

Extracting functional components of neural dynamics with Independent Component Analysis and inverse Current Source Density

Szymon Łęski · Ewa Kublik · Daniel A. Świąjkowski · Andrzej Wróbel · Daniel K. Wójcik

Received: November 25, 2009/ Accepted: date

Abstract Local field potentials have good temporal resolution but are blurred due to the slow spatial decay of the electric field. For simultaneous recordings on regular grids one can reconstruct efficiently the current sources (CSD) using the inverse Current Source Density method (iCSD). It is possible to decompose the resultant spatiotemporal information about the current dynamics into functional components using Independent Component Analysis (ICA). We show on test data modeling recordings of evoked potentials on a grid of $4 \times 5 \times 7$ points that meaningful results are obtained with spatial ICA decomposition of reconstructed CSD. The components obtained through decomposition of CSD are better defined and allow easier physiological interpretation than the results of similar analysis of corresponding evoked potentials in the thalamus. We show that spatiotemporal ICA decompositions can perform better for certain types of sources but it does not seem to be the case for the experimental data studied.

Having found the appropriate approach to decomposing neural dynamics into functional components we use the technique to study the somatosensory evoked potentials recorded on a grid spanning a large part of the forebrain. We discuss two example components associated with the first waves of activation of the somatosensory thalamus. We show that the proposed method brings up new, more detailed information on the time and spatial location of specific activity conveyed through various parts of the somatosensory thalamus in the rat.

Keywords local field potentials (LFP) · inverse Current Source Density (iCSD) · Independent Component Analysis (ICA) · somatosensory evoked potentials (EP) · thalamic processing

1 Introduction

One of the obstacles in understanding how the brain works is that all of its parts work simultaneously. The paradigm of evoked activity (averaging responses with respect to the chosen marker) is helpful in extracting dynamics relevant for a given stimulus. To locate the anatomical structures processing a representation of the stimulus and to pinpoint the order of their activation one must be able to simultaneously probe many regions of the brain. A number of techniques assist the researcher in this enterprise, including PET, MRI, voltage or calcium sensitive dyes. However, all of them have limitations, either in terms of available time-scales, interpretation of obtained results, or the extent of the tissue that can be accessed simultaneously.

A convenient choice when it comes to temporal resolution is the local field potential (LFP): low-frequency part of the extracellular electric potential, which reflects brain dynamics at the level of neural populations (Nunez and Srinivasan 2005). The extracellular potential is generated by currents (mostly resulting from synaptic inputs) crossing the cell membranes. The coarse-grained distribution of trans-membrane currents leaving or entering cells is called the current source density (CSD).

In the simplest model, which assumes quasi-static approximation of electrodynamic equations (Nunez and Srinivasan 2005), the density of currents C is related to the extracellular voltage ϕ by the Poisson equation:

$$\nabla(\sigma \nabla \phi) = -C, \quad (1)$$

where σ is the electrical conductivity tensor. Mathematically, this is the same equation as the one which describes

This work was partly financed from the Polish Ministry of Science and Higher Education grants PBZ/MNiSW/07/2006/11 and 46/N-COST/2007/0.

S. Łęski · E. Kublik · D. A. Świąjkowski · A. Wróbel · D. K. Wójcik
Department of Neurophysiology
Nencki Institute of Experimental Biology
3 Pasteur St.
02-093 Warsaw, Poland
E-mail: d.wojcik@nencki.gov.pl

the electric field of static charge distribution. The fundamental solution of this equation has the form of $\phi \sim \frac{1}{r}$. One consequence of the slow decay is that even for localized CSD distributions, for dipole or closed fields, which decay much faster, the resulting potential ϕ is non-local. In other words, the electric signal recorded in one place may reflect activity of a distant neural population¹. This spatial blurring is the first challenge in analysis of LFP. If several time-aligned recordings of LFP on a line, 2D or 3D grid, are available, one can try to address this problem by estimating the underlying CSD from the measured voltages. This is called the CSD method (Nicholson and Freeman 1975; Freeman and Nicholson 1975; Mitzdorf 1985).

Another challenge follows from simultaneous activity of different populations which may contribute to each of the recorded signals. This is the ‘cocktail party’ problem: how to hear what a single population is saying if many populations are speaking at the same time. Independent component analysis (ICA) (Comon 1994; Bell and Sejnowski 1995) is a technique that proved to be of use for this kind of blind source separation problem in many fields. In neuroscience it was used, among others, in the analysis of fMRI data (McKeown et al. 1998; Stone et al. 2002), optical imaging recordings (Schiessl et al. 2000; Reidl et al. 2007), in characterization of receptive fields (Saleem et al. 2008), and in analysis of local field potentials (Tanskanen et al. 2005). Our application of ICA to volume CSD is similar in spirit to fMRI and optical imaging analysis as in all of these cases the tissue is probed regularly on fine spatial scales and one can assume the studied signals to be generated by sources of well-defined spatial location.

Neither of these techniques (CSD, ICA) alone seems to be sufficient for analysis of LFP recorded in a geometrically and morphologically complex brain region, such as the rat thalamus. For example, even if we knew exactly the time-course of CSD, it could still happen that sources and sinks occupying the same spatial localization represent activity of different populations of cells, performing different functions. On the other hand, independent components calculated from LFP would ideally represent single functional populations, but the localization of that populations would be not so precise. While this may seem technical, better localization of CSD versus LFP translates into substantially better resolution of the method, as we discuss further. In this paper we show that the combination of CSD and ICA provides a powerful and stable method for extraction of well-localized functional components from multi-electrode recordings of LFP. We discuss several variants of performing the decomposition and point out the optimal approach for our data. Finally, we study selected components obtained in analysis of data from three dimensional

($4 \times 5 \times 7$ locations) mapping of the local field potentials in the rat forebrain during early stages of tactile information processing.

The paper is organized as follows. In Section 2 we discuss the computational methods. Specifically, in Subsection 2.1 we describe the inverse CSD method, in 2.2 we present the spatiotemporal ICA method we use, and in 2.3 we describe how we use clustering to study stability of ICA decomposition. We test different combinations of spatial and temporal decompositions on volumetric test data in Section 3 and show that depending on the structure of sources different variants of ICA give best reconstructions. In Section 4 we provide details about the experimental procedure. The insights obtained in the decomposition of test data are used in the analysis of experimental data presented in Section 5 and discussed in Section 6.

2 Computational methods

2.1 Inverse Current Source Density

The simplest technique of CSD estimation which was used traditionally was through the finite-difference approximation to Eq. (1), often smoothed for stability (Freeman and Nicholson 1975). It is based on the replacement of the second derivative with appropriate difference quotient. Let us focus on a one-dimensional example and assume that CSD varies only along z axis and is constant in perpendicular plane. Equation (1) becomes then

$$\sigma \frac{\partial^2 \phi}{\partial z^2} = -C(z). \quad (2)$$

Here we also assume that the medium is isotropic and homogeneous so that σ is a constant scalar. Let us consider signals from a laminar electrode with multiple equidistant contacts positioned at $z_k = z_0 + kh$, where h is the inter-contact distance. The CSD is then approximated using the three-point formula for second derivative:

$$C(z_k) = -\sigma \frac{\phi(z_k + h) - 2\phi(z_k) + \phi(z_k - h)}{h^2}. \quad (3)$$

This method has been applied extensively over the years to electrophysiological signals from various parts of the brain (Schroeder et al. 1992; Ylinen et al. 1995; Shimono et al. 2000; Lakatos et al. 2005; Lipton et al. 2006; de Solages et al. 2008; Rajkai et al. 2008; Stoelzel et al. 2009). One can also use other numerical approximations of $\frac{\partial^2 \phi}{\partial z^2}$, for example using not only immediate neighbors but also next neighbors or use different weights which may lead to better stability of the estimates at the cost of spatial filtering of the results (see for example Freeman and Nicholson (1975); Tenke et al. (1993)).

An advantage of the traditional CSD method is its simplicity. Moreover, in simple one dimensional geometry for recordings in laminar structures it often gives meaningful

¹ See (Hunt, Falinska, Łęski, Wójcik, Kasicki (2009) unpublished. Available at <http://www.neuroinf.pl/Members/danek/homepage/preprints/Article.2009-10-22.4312>

results. However, this straightforward approach has a number of drawbacks. For instance, as shown in Pettersen et al. (2006), traditional CSD method may introduce errors in estimating the CSD when the sources vary considerably in the plane perpendicular to the electrode at the scale of a few intercontact distances. This is the case for the recordings from the somatosensory (barrel) cortex, for example. Also, in traditional CSD it is hard to account for conductivity jumps, e.g. at cortical surface. Another drawback is that the three-point formula, Eq. (3), can not be applied at the boundary. This means that we lose boundary data. In one dimension we lose two points out of, say, twenty, but for two- or three-dimensional data the boundary may comprise most of the recording points. For example, in two dimensional recordings on an 8 by 8 grid there are 24 boundary points out of all 64, and in the three-dimensional ($4 \times 5 \times 7$) data we study here there are 110 points at the boundary out of 140 (more than 75%). Despite these facts even two dimensional CSD estimates through discrete derivation were studied for example in Novak and Wheeler (1989), Shimo-no et al. (2000) and Lin et al. (2002). For one dimensional cortical recordings, to overcome the problem of the boundaries, Vaknin et al. (1988) proposed to add virtual recording points at both ends of the line of contacts and to assume that ϕ at these points is the same as at the nearest end. This procedure allows us to estimate the CSD at the ends, but the assumption that voltage does not vary outside the recording electrode is not always right.

Recently another approach to CSD estimation called inverse Current Source Density method was proposed by Pettersen et al. (2006) for one dimensional recordings and generalized to three-dimensional recordings by Łeński et al. (2007). Here, unlike the traditional CSD, one does not try to estimate C directly from Eq. (1). Instead, one constructs a class of CSD distributions parameterized with as many parameters as the number of recorded signals. Then one has to establish a one-to-one relation F between measured voltages ϕ and the parameters of the CSD distribution. For a specific example assume N recording points on one-, two- or three-dimensional Cartesian grid. Consider a family of CSD distributions parameterized by N numbers, $I_{c_1, \dots, c_N}(x, y, z, t)$. By this we mean that knowing the values of the N parameters we can assign a value of CSD to each spatial position. Then the values of the potential, ϕ , on the grid can be obtained by solving Eq. (1), which yields the relation F . A direct solution to the Poisson equation leads to the formula

$$\phi_{c_1, \dots, c_N}(x, y, z, t) = \int \frac{I_{c_1, \dots, c_N}(x', y', z', t) dx' dy' dz'}{4\pi\sigma \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}}.$$

Therefore, the N measured voltages

$$\phi_n(t) = \phi_{c_1, \dots, c_N}(x_n, y_n, z_n, t)$$

are functions of the N CSD parameters

$$(\phi_1(t), \dots, \phi_N(t))^T = F(c_1, \dots, c_N)(t) \quad (4)$$

If the family and the parameterization are chosen appropriately, one can invert F and from the N measured potentials get the N parameters of CSD. One particularly useful parameterization which we are using here, is to use CSD values on the measurement grid as the N parameters c_n and obtain the CSD within the grid $I_{c_1, \dots, c_N}(x_0, y_0, z_0, t)$ via interpolation (for example nearest neighbor, linear or with splines). Then one can reduce equation (4) to an invertible linear relation between the recorded potentials and the CSD values at the grid. One can also consider other CSD distributions and parameterizations. For example, Pettersen et al. (2006) showed that the traditional one-dimensional CSD is equivalent to inverse CSD if we assume that the CSD is localized on infinitely thin, infinite planes located at $z = z_i$.

Assuming that all the sources are located within the electrodes grid may lead to noticeable reconstruction errors close to the boundary when there are active sources located outside the probed region (Łeński et al. 2007). This is because the inverse CSD method tries to imitate the influence of these sources by adjusting the distribution within the grid. A way to overcome this difficulty is to consider a family of CSD distributions extending one layer beyond the original grid. One can set the CSD values at the additional nodes to zero or duplicate the value from the nearest node of the original grid. We denote these two approaches by B and D boundary conditions, in line with notation of (Łeński et al. 2007). Note that new CSD family on the extended grid is still parameterized with N parameters. In this paper we use spline interpolation between the nodes and the D boundary conditions.

2.2 Independent Component Analysis

We follow the spatiotemporal ICA scheme as described in detail by Stone and Porrill (1999) and Stone et al. (2002). The data we aim to analyze are $m = 4 \times 5 \times 7 = 140$ signals $f_1(t), \dots, f_m(t)$, with $1 \leq t \leq n$. We can arrange the data in an $m \times n$ matrix, $X = (f_1, \dots, f_m)^T$, so that the rows of X are our signals. Equivalently we can view our data as a sequence of n snapshots $X = (s_1, \dots, s_n)$.

First we use principal component analysis to reduce dimensionality of the problem. That is, we approximate our data matrix X with $\tilde{X} = UV^T$, where U is an $m \times k$ matrix of k spatial components ('coefficients' in PCA), and V is an $n \times k$ matrix of temporal components (principal component 'scores'). The choice of k is an important issue. Especially taking k too small is to be avoided, as that may result in losing information contained in the signal. Also, for k not large enough one may end up with components which could be further separated for larger k . On the other hand, taking k too large may result in unnecessarily large data sets, but is less dangerous than the opposite. To choose k for the experimental data we performed

the analysis for $3 \leq k \leq 32$. Using clustering (Section 2.3) we found that the number of components we could extract saturates around $k = 20$ (see Figure 1 and 2 in the supplementary material). Finally, we took $k = 24$ for analysis of experimental data. For model data (Section 3) we also took $k = 24$, despite the fact that the true dimensionality of the data was lower (between 5 and 8). In such cases we get several large components corresponding to the original ones, and several components which are close to zero. This backs the claim that choosing too large k is safer than choosing k too small.

The idea behind spatiotemporal ICA is to take into account information about independence carried by the temporal signals and the spatial distributions and decompose $\bar{X} = ST^T$ into maximally independent spatial and temporal components S and T . This decomposition is found by maximization of a linear combination of entropies obtained separately for spatial, H_S , and temporal, H_T , probability distribution functions:

$$h_{ST} = \alpha H_S + (1 - \alpha) H_T. \quad (5)$$

Parameter $\alpha \in [0, 1]$ quantifies how much weight we attribute to the temporal and spatial independence. $\alpha = 0$ is purely temporal decomposition, when one assumes independent temporal sources with dual spatial distributions unconstrained. $\alpha = 1$ is purely spatial decomposition, when one assumes independent sources localized in space and dual unconstrained temporal signals. Intermediate α take into account possible independence of both parts. In the article we use $\alpha = 0.5$ or $\alpha = 0.8$ although we tested a wider range of values.

Both the temporal and spatial ICA decompose the original signals into pairs of temporal signals and associated spatial profiles. The difference is in the assumptions underlying the decomposition. The temporal ICA tries to find signals which are independent over time, while the spatial ICA looks for signals which are independent over space. Physiologically, this means that the temporal ICA assumes independent processes in the brain taking place simultaneously, while spatial ICA assumes that there exist independent modules, or spatially organized nuclei, which perform distinct functions and are generators of independent signals. One should expect that the most reasonable assumptions will depend on the experimental situation. For example, for analysis of ongoing EEG signal the temporal ICA may work better, as the resulting signals are produced by many independent generators performing different functions simultaneously. On the other hand, the spatial ICA may be better suited for the analysis of averaged stimulus-evoked potentials, as in such case the signals are related to a single brain computation, which is performed by a network with fixed anatomical connections.

The details of the algorithm are presented in (Stone and Porrill 1999; Stone et al. 2002). We used the MAT-

LAB code² by J.V. Stone and J. Porrill. We required that the spatial independent components have high kurtosis (model probability density function $x \mapsto 1 - \tanh^2 x$), and that the temporal components have low kurtosis (model pdf $x \mapsto e^{-x^4}$). This requirement is satisfied for example by localized spatial components and oscillatory time courses. Any other combination of kurtosis assumptions led to physiologically meaningless components.

2.3 Clustering of independent components

All variants of ICA methods rely on minimization of a certain goal function. In our case this goal function, Eq. (5), is negative of entropy of estimated components (Bell and Sejnowski 1995; Stone et al. 2002). One problem that may arise here is that the minimization algorithm typically finds only a local minimum of the goal function. *A priori* there is no guarantee that this local minimum corresponds to a valid solution of blind source separation problem, although it may be the case. To make sure that we obtain correct solutions is to repeat the ICA algorithm several times, using for example different starting points for minimization, and to see if the components are stably extracted (Himberg et al. 2004).

For each rat we repeated the ICA procedure 30 times with random initial mixing matrix. Every time we looked for 24 components, which produced the total of 720 components. We clustered these components using hierarchical agglomerative clustering (group average linkage), as in (Himberg et al. 2004). We used the `hcluster.m` MATLAB script available as a part of `Icasso` software package³. As the dissimilarity matrix we used a matrix of (squared) L^2 distances between components. Specifically, for two components $f_i(t)$, $f_j(t)$, associated with spatial profiles (values at 140 measurement points) $\mathbf{s}_i(x, y, z)$, $\mathbf{s}_j(x, y, z)$, we took $D(i, j)$ as dissimilarity, where

$$D(i, j) = \frac{D_T(i, j)}{\langle D_T(i, j) \rangle_{i, j}} + \frac{D_S(i, j)}{\langle D_S(i, j) \rangle_{i, j}},$$

$$D_T(i, j) = \min \left(\sum_t (f_i(t) - f_j(t))^2, \sum_t (f_i(t) + f_j(t))^2 \right),$$

$$D_S(i, j) = \min \left(\sum_{x, y, z} (\mathbf{s}_i(x, y, z) - \mathbf{s}_j(x, y, z))^2, \sum_{x, y, z} (\mathbf{s}_i(x, y, z) + \mathbf{s}_j(x, y, z))^2 \right).$$

We take the minimum over measures using $f_i(t) - f_j(t)$ and $f_i(t) + f_j(t)$, analogously for $\mathbf{s}(x, y, z)$, to avoid spurious dissimilarities which would occur when signs are dif-

² STICA software available at <http://jim-stone.staff.shef.ac.uk/>

³ The package is available at <http://www.cis.hut.fi/jhimberg/icasso/>

ferently distributed between $f_i(t)$ and $s_i(t)$ for two runs of ICA.

We also used clustering to choose the number of extracted components k . We pooled all 525 components obtained for $3 \leq k \leq 32$ (one run for each k) and applied the clustering procedure described above. We divided the data into 32 clusters and counted how many clusters are populated by components obtained for selected k in the studied range. For our experimental data this number saturates slightly above $k = 20$, see the supplementary material available online.

3 Analysis of test data

To test the proposed ICA / iCSD method we used model data of the same form as the experimental data studied in the next section. Thus we generated 140 LFP signals from a number (between $M = 5$ to $M = 8$) of spatially localized sources with time-courses $g_i(t)$ and spatial profiles $c_i(x, y, z)$. Specifically, we assumed the current-source density of the form

$$C(x, y, z, t) = \sum_{i=1}^M c_i(x, y, z) g_i(t). \quad (6)$$

We used three types of the time courses $g_i(t)$: 1) oscillatory signals (Gaussian white noise low-pass filtered below 300Hz), 2) simulated “evoked potentials” — these signals consisted of two step functions of opposite sign filtered below 300Hz, 3) experimental evoked potentials chosen from the recordings described in Section 4. As the spatial components we took CSD distributions for which the inverse CSD method gives exact results. The reason for choosing such sources was to focus on comparing different variants of ICA. The reconstruction error of the iCSD method itself has already been studied in Łeński et al. (2007); Wójcik and Łeński (2009). The sources were constructed as follows: on a grid of $4 \times 5 \times 7$ nodes we set non-zero CSD values at a small number of nodes. These values were then interpolated using splines in the same way as used for construction of F (see Section 2.1). Again, we used several kinds of spatial components: a) two non-zero nodes, for which we took the same absolute value of CSD and opposite signs (a dipole); b) as in a), but with the same signs; c) two non-zero nodes with the CSD values of opposite signs and non-equal amplitude; d) a non-zero value V_{central} at one node surrounded by non-zero values in 6 neighboring nodes (the number of neighboring nodes was smaller if the central node was located at the boundary), the values at the neighboring nodes were equal to $-V_{\text{central}}/6$; e) single non-zero CSD node. Due to spline interpolation every distribution was non-zero in some region surrounding initially set CSD nodes. These choices of spatial and temporal profiles were motivated by the structures observed in the analysis of experimental data. All components from

two of the model datasets are presented in supplementary material available online.

For these different sets of sources and generated potentials we performed the inverse CSD analysis, followed by spatial, temporal, and spatiotemporal ICA decomposition (Section 2.2) with different values of α . This framework allowed us to compare the components obtained for different choices of α with the generating sources known exactly. It was shown in Stone and Porrill (1999) on test data modeling synthetic fMRI images that the best decomposition of signal was obtained when both spatial and temporal independence of the data were taken into account (for spatiotemporal ICA with $\alpha = 0.5$). We expected the same to be true in case of three-dimensional LFP data. However, in case of our simulated signals we found that the values of α which give good results for some combinations of temporal and spatial components do not necessarily lead to valid reconstructions for different classes of components. One general observation was that purely temporal decompositions usually led to meaningless components. Otherwise, depending on test case, either spatial or spatiotemporal decompositions were performing better. For every test case at least one approach allowed to recover the original sources with high precision.

Here is the summary of our observations:

1. For spatial components of type a) the best choice is the spatiotemporal ICA (for example $\alpha = 0.5$ although the results were robust against changes of α), which allows for faithful extraction of the original components. Spatial ICA ($\alpha = 1$) extracted some components faithfully, but other components were not unmixed properly. For example, two of the original components were extracted as three reconstructed components, with some non-trivial cancellations occurring.
2. For type b) and c) both spatial ($\alpha = 1$) and spatiotemporal ($\alpha = 0.5$) ICA methods recover the original components.
3. For d) spatial components the spatial ICA works better than the spatiotemporal variant, which tends to produce several components with overlapping spatial parts and similar time-courses; however, the spatial ICA sometimes puts some part of the surrounding current sources into a different component than the central source.
4. For e) spatial components (single non-zero CSD node) and oscillating signals (type 1) the spatial ICA is again a better choice.

To illustrate these effects Fig. 1 shows example components obtained in spatiotemporal and spatial ICA decompositions for test sources of the form d), which corresponds to the closed field we observe in the experimental data, with part of the temporal signals given by the filtered step potentials, 2), and another part being the experimental EPs, 3). The dataset was a linear combination of eight sources of such form with different spatial placement. Fig. 1 A and

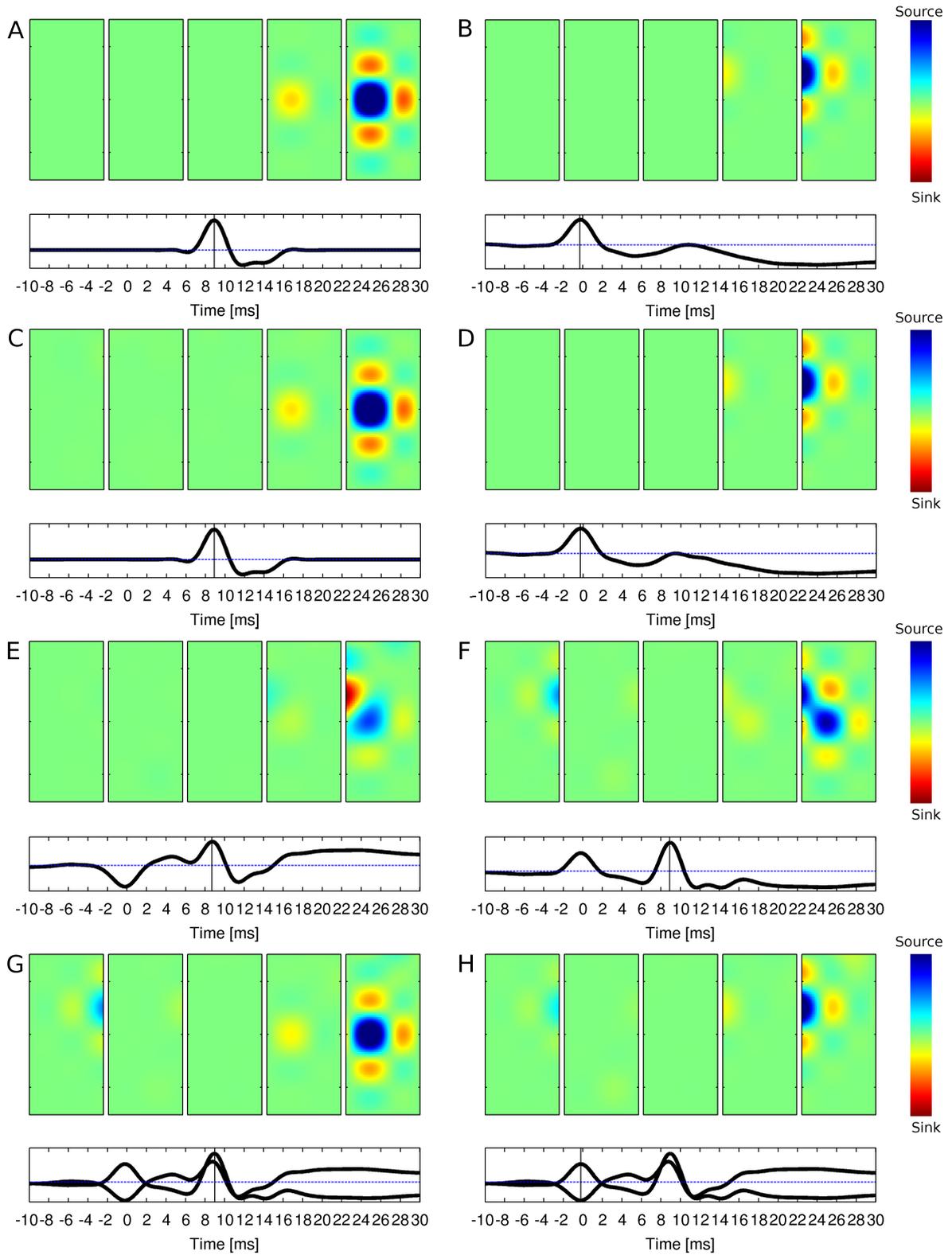


Fig. 1 Tests of spatial ($\alpha = 1$) and spatiotemporal ($\alpha = 0.5$) ICA on model data described in Section 3, point 3. Each panel shows one (A–F) or two (G, H) components in 3D. Five images in every panel are parallel, consecutive slices through the volume spanned by the $4 \times 5 \times 7$ grid and show the spatial distribution of the current source density. Below each planar representation of the spatial CSD distribution we plot the time course of the components, $f_i(t)$, the vertical line marks the time for which the spatial distribution is drawn. A), B) two of the components present in model data. Spatial ICA recovers these components faithfully (C, D). The spatiotemporal ICA generates two components E), F), overlapping in space. The sum of these two components at $t = 9\text{ms}$, G), is similar to the original component A), while the same sum at $t = -0.2\text{ms}$ yields the component B). This means that in this case the spatiotemporal ICA produces mixtures of original components.

B show two of the original model components. The spatial ICA leads to well separated components, Fig. 1 C and D. For spatiotemporal decomposition these two components become mixed, Fig. 1 E and F. Fig. 1 G and H show their sum at two different time points. This dataset, processed using several variants of the procedure, is available online together with a suitable MATLAB viewer⁴.

Clearly, the choice of model data may favor one method over the other, therefore these results, to some extent, are a consequence of the assumed structure of sources. Specifically, while the spatial distributions of two components could overlap, there was usually substantial difference between any pair. On the other hand, we considered sources with substantially correlated time courses. Our intention was to model similar activation of these spatially separated sources following the dynamics observed in the experiment (see Section 4, especially Fig. 5 and 6). It was also inspired by the simultaneous divergent signal propagation commonly taking place in the neural system.

One issue that should be addressed is why decompose iCSD reconstructions instead of raw potentials. When we perform the two procedures, ICA decomposition and iCSD reconstruction, in principle they can be done in either order. In practice, however, it turns out that ICA decomposition of potentials followed by reconstruction of the underlying CSD did not produce good results. This is caused by the long range of electric potential which smears the current sources in the LFP picture. Therefore, the same source in the CSD picture will be much more localized than its potential. This leads to higher kurtosis for the distribution of CSD values. As the algorithm we used is unmixing sources looking for high kurtosis in the spatial domain, it is a major issue. Figure 2 shows components obtained through spatial ICA on LFP followed by iCSD (Fig. 2 C) with components obtained by ICA performed on iCSD reconstructed signals (Fig. 2 B) corresponding to a specific original source in the signal (Fig. 2 A). The kurtosis of the distribution of the CSD values of the original source (Fig. 2 A) was 113, which was almost twice the kurtosis of the distribution of LFP values for the same source (60). Clearly, the ICA on LFP leads to very poor decomposition. Curiously, the time course of the source was correctly recovered by both approaches.

4 Animal experiments and data

We used the ICA / iCSD method described above to understand the dynamics of complex spatiotemporal activation pattern evoked in the rat's thalamus by somatosensory stimulation. The local field potential (LFP) responses to vibrissal displacement were recorded from deep forebrain structures of the anaesthetized (urethane; 1.5 g/kg b.w.)

⁴ http://www.neuroinf.pl/Members/szleski/iCSD_data/ICviewer

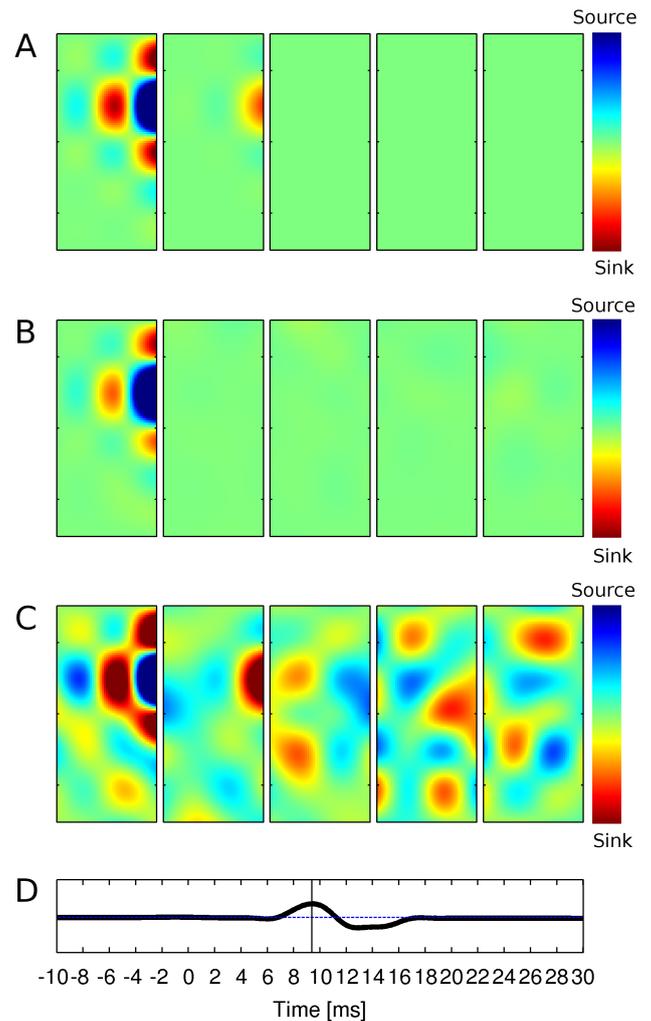


Fig. 2 Comparison of spatial components obtained with ICA applied to either LFP signals or reconstructed CSD. A) Original (model) spatial component. B) Spatial component recovered using spatial ICA on CSD. C) Spatial component recovered using spatial ICA on LFP (decomposition followed by iCSD). D) Time-course of the original component is faithfully recovered with both methods.

adult male Wistar rats ($N = 7$). Local anesthetic (EMLA 5% cream) was applied into external auditory meatus and animals placed in a stereotaxic apparatus. Skin on the head was injected with 0.5% lidocaine prior to incision. The body temperature was kept at 37°C by a thermostatic blanket, liquid requirements were fulfilled by subcutaneous injections of 0.9% NaCl (2 ml every 2 hrs). The depth of anesthesia was monitored by observation of presence of corneal reflexes and whisker twitching and adjusted with supplementary doses of urethane (10% of the initial dose). All experimental procedures followed the 86/609/EEC Directive and were accepted by the 1st Warsaw Local Ethics Committee.

To access the deep forebrain structures most of the right parietal bone was removed and the resulting opening was covered with agar dissolved in saline. A set of 2-5 stainless

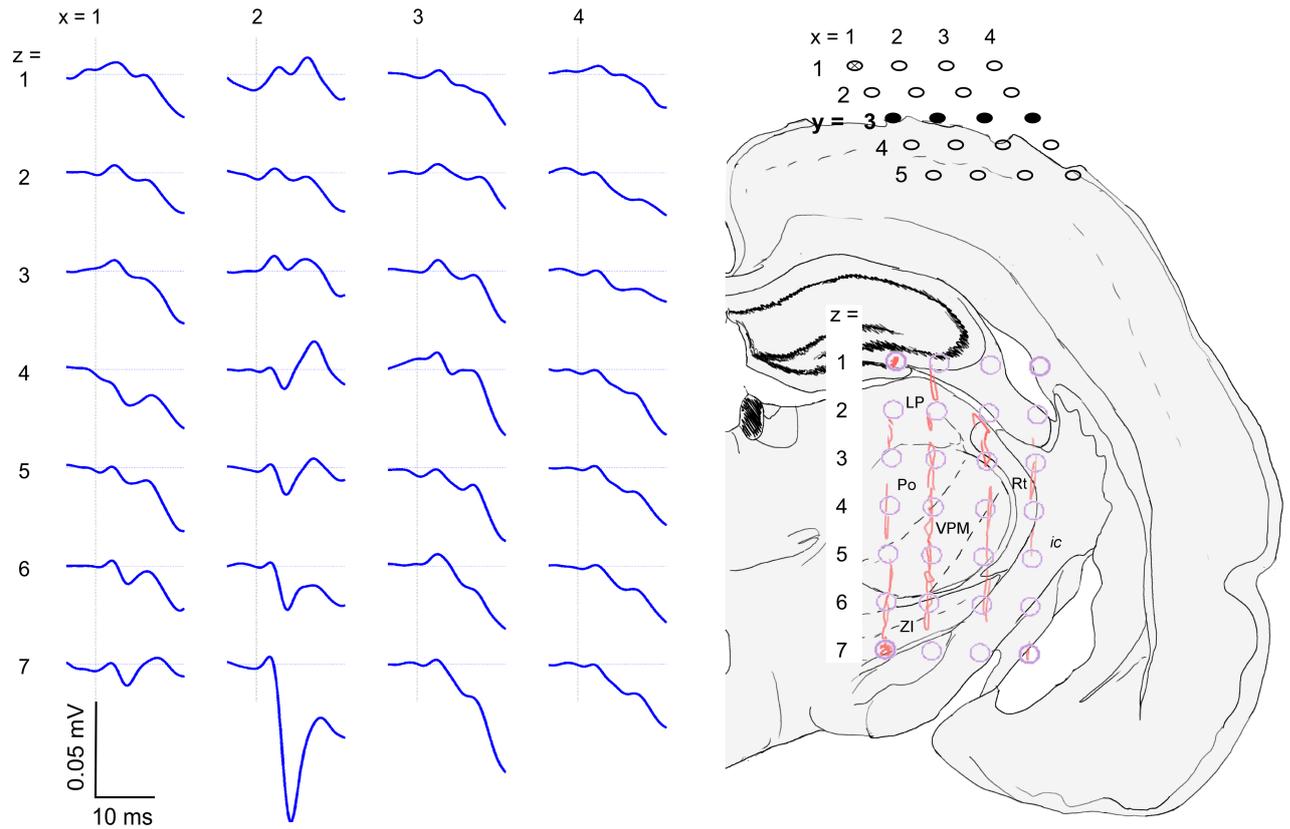


Fig. 3 Potentials recorded in Rat 1 in one of the coronal planes. A) Evoked potentials recorded at positions marked by full circles in the coronal plane shown in B; B) Drawing of structures and electrode tracks in one of the investigated coronal planes. Ellipses on the top outline the approximate surface positions of the insertions of the electrode grid (filled ellipses correspond to the drawn plane, crossed ellipse marks the surface position where the stereotaxic manipulator was set to zero; the coordinates of the surface grid are 2.1, 2.8, 3.5 and 4.2 mm to the right from midline, 1.9, 2.6, 3.3, 4.0 and 4.7 mm posterior to Bregma). Light gray lines mark electrodes traces and light gray circles indicate recording positions at seven depths at the deep forebrain. The notation of the recording positions and respective EPs corresponds to X_Y_Z coordinates, where X deciphers lateral position, Y anterior-posterior position and Z codes the recording depth. Accordingly, the coronal plane shown in B corresponds to Y = 3 and the biggest potential can be traced at location 2_3_7. For the names of the structures see section Abbreviations at the end of the article.

steel micro-electrodes (FHC, Bowdoin, USA; impedance of 1-1.5 M at 1 kHz) was mounted in parallel on the manipulator holder with 0.7 mm horizontal spacing between the tips. Electrodes were lowered vertically through the opening and stopped each 0.7 mm at seven depths in the brain tissue, starting at 3.4 mm from the cortical surface (zero being set at the most medial-anterior position; cross in Fig. 3B). At each stop the group of left whiskers (attached to piezoelectric stimulator) was deflected by $100\mu\text{m}$, 60 times and then averaged to obtain the evoked potentials (EPs) from a specific point. After completion of one penetration the electrodes were moved to the other recording positions on the 4 by 5 surface grid (Fig. 3B). Thus, a 3D Cartesian grid of 140 ($4 \times 5 \times 7$) recording points comprising a slab of forebrain tissue with portions of the thalamus, pre-ectum, hippocampus, and cerebral white matter was completed. The length of the interelectrode distance was based on the average size of the studied thalamic nuclei (e.g. the posterior nucleus, Po, $1.4 \times 1 \times 2$ mm), and the expected

large-scale activation evoked by the strong multi-whisker input.

The LFPs were recorded monopolarly (with Ag-AgCl reference electrode under the skin on the neck), amplified (1000 times) and filtered (0.1–5 kHz). Epochs (of about 1.4 s) containing EPs were digitized on-line (10 kHz) with 1401plus interface and Spike-2 software (CED, Cambridge, England). All stored data were examined for integrity and epochs with artifacts were excluded from further analysis. To obtain clear LFP, not contaminated by multi-unitary activity, the recorded signal was low-pass filtered below 300Hz (a linear-phase FIR digital filter of order 100 with Kaiser window, $\beta = 5$).

After completion of the experiment the electrolytic lesions were made in corner positions of the recording grid; rats received overdose of pentobarbital (150 mg/kg b.w.) and were perfused transcardially with saline followed by 10% formalin. Removed brains were stored in fixative, then cryoprotected in sucrose solutions, sectioned on freezing

of components obtained for different α , on the clustering analysis, and on the results from tests of different variants of ICA on several sets of model LFPs, we chose to use the spatial ICA ($\alpha = 1$) to process the experimental data.

Observing the recorded potentials one can see stimulus evoked waves at almost each of 140 positions (Fig. 3). This activity pattern encompassed several distinct regions of the thalamus due to massive divergence of the ascending fibers and heterogeneous inputs from various vibrissal afferents (for a review, see Waite (2004)) and the timing order of their activation might shape specificity of information which is passed on to the cortex. Figure 3A shows EPs recorded at one coronal plane (-3.7 mm from Bregma) in Rat 1. There is only one early EP that clearly stands out in this view suggesting the short latency activation (signal from the electrode just below the zona incerta, position 2_3_7). Other locations share many similar deflections of recorded traces which may partially result from electrotonic spread of field potentials.

To decompose the complex spatiotemporal activation pattern of voltage responses evoked in the rat's thalamus by sensory stimulation we applied the ICA/iCSD method. Figure 5 exemplifies two components (out of 24) which constituted the current sinks with the shortest latencies evoked specifically in the first order somatosensory thalamic nucleus, i.e. the ventral postero-medial nucleus (VPM) in Rat 7. The spot of activation differentiated by the first component in Fig. 5A was located at the border of VPM and the posterior thalamic nucleus (Po) and reached its maximum at 3.6 ms after deflection of a bunch of contralateral whiskers. With a little delay, the second, stronger component (Fig. 5B) activated more lateral part of the dorsal VPM and reached its maximum at 4.7 ms after stimulation, i.e. about 1 ms later than the first one. The detailed analysis confirmed that the locations of the two sinks overlapped the expected positions of so called 'heads' (first component) and 'cores' (second component) of barreloids, i.e. the neuronal groupings representing the stimulated bundle of large vibrissae (Pierret et al. 2000; Urbain and Deschênes 2007).

Analogous components recorded in Rat 1 are presented in Fig. 6. The faster, weaker component peaks at 4.0ms after stimulation (black trace in Fig. 6E); the second component dominates about 1ms later (gray trace in Fig. 6E). To show the different dynamics of the two components we plot their sum in Fig. 6B and D. As time goes on, the first component originally dominates the other (shown at the latency of 3.6 ms at Fig. 6B), then the second component takes over (shown at the peak of the second component, latency of 5.2 ms, Fig. 6D). Fig. 6A and C contrast the complete pattern of tissue activation 3.6 and 5.2 ms after the onset of stimulus with the sum of the currents given by the two components under discussion shown in Fig. 6B and D.

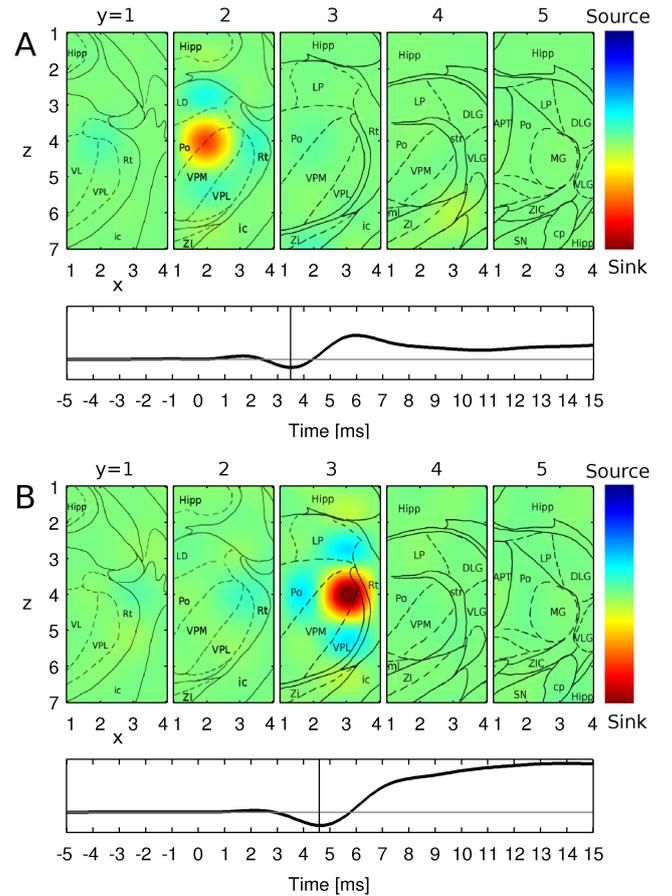


Fig. 5 Experimental data from Rat 7. Two independent components (A, B) with current sinks in somatosensory thalamus. The first component (A) is located at the border of VPM and Po, the second, stronger component (B) is located in more lateral part of the dorsal VPM. The locations match the expected positions of 'heads' and 'cores' of barreloids, see text.

	'heads'	'cores'
1	4.0	5.2
2	4.1	5.0
3	4.6	5.8
4	5.1	5.9
5	4.2	5.1
6	—	4.8 (small amplitude)
7	3.6	4.7

Table 1 Peak latencies for 'core' and 'head' components in all rats.

The sequence of described activation could be traced in data from all rats (Table 1) but the position of the first component was accurately placed at the VPM/Po border only in five animals due to the matched locations of recording electrodes (as in shown examples of Rat 1 and 7). In another rat (Rat 4) the first component was weaker and in Rat 6 it was included into less weighted, mixed components obtained in the analysis. The experimental dataset

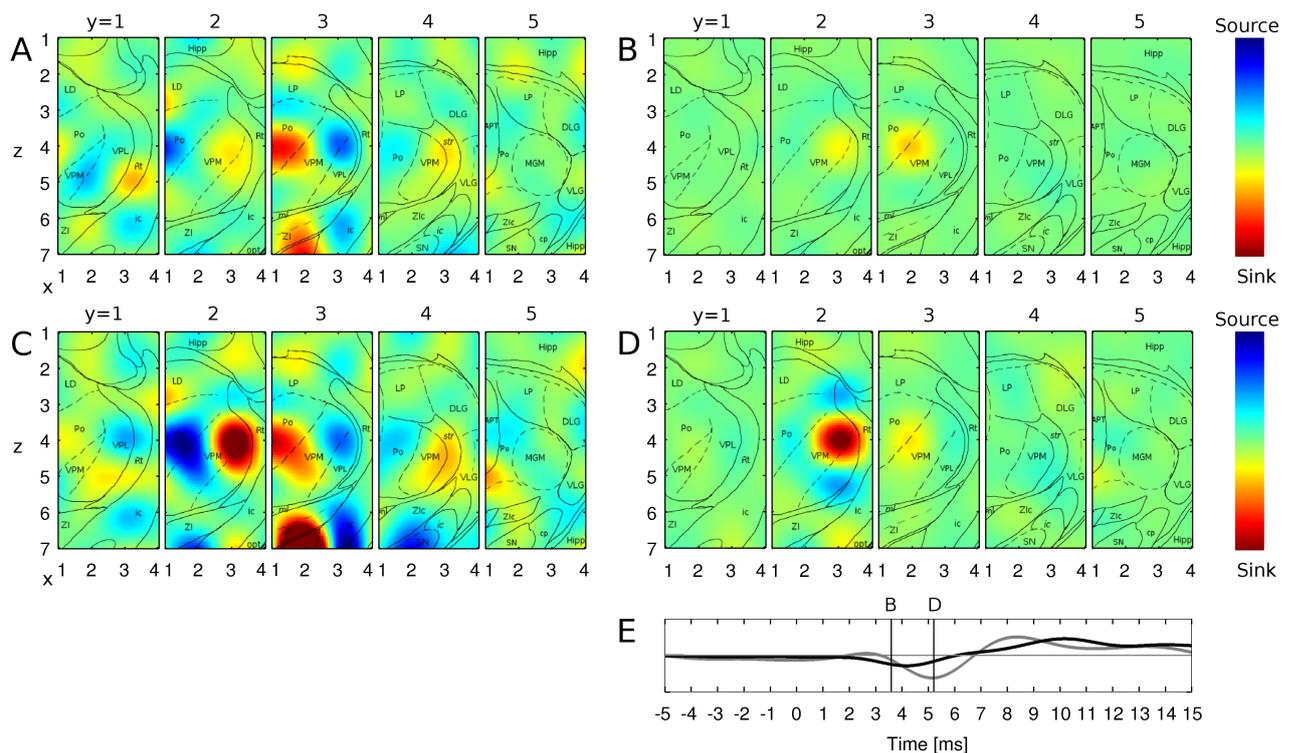


Fig. 6 Experimental data from Rat 1. A), C) Full CSD reconstruction and B), D) sum of the CSD contained in the 'head' and 'core' components extracted using spatial ICA. E) The time course of the two extracted components, black line represents the component placed at the border of VPM and Po, gray line — at VPM core). The time after stimulation is 3.6ms in A), B), and 5.2 ms in C), D), marked with vertical lines on the temporal plot, E). See text for details.

analyzed above (Rat 1), is available together with a MATLAB viewer⁵.

6 Summary and discussion

In this paper we have shown that the combination of ICA following iCSD reconstruction from EP recordings on a regular grid is a powerful technique for extraction of functional components of neural dynamics. The method described here brought up new, more detailed information on the latency and spatial location of specific activity conveyed through various parts of the rat somatosensory thalamus. This was possible due to two characteristics of electrophysiological data utilized by the method. First, averaging of EPs makes it possible to extract accurate anatomical position and time dependence (dynamics) of small current sources accompanying the different volleys of conveyed spike activity. Second, utilizing both the physiological and anatomical results from the same animal allows for extraction of highly specific, individual relationships between the two data sets.

We have tested different variants of spatiotemporal ICA applied to CSD and LFP, from purely temporal to purely

spatial decompositions. For the cases tested the temporal ICA usually gave meaningless results. This is probably the result of observed temporal correlation between sets of sources receiving common input. For some sources spatiotemporal ICA worked best, however, for our experimental data spatial ICA seems to extract the functional components most efficiently. In principle one could perform ICA decomposition and CSD reconstruction in either order. However, as the ICA procedure is nonlinear the results obtained are very different. Indeed, as we showed, it turns out that ICA of LFP results in very poor decomposition. This is an effect of the long range of electric potential which smears current sources in the LFP picture. As a result, the distribution of CSD values of a source will have a higher kurtosis than the distribution of values of LFP generated by this source. The ICA algorithm we used assumes high kurtosis of the distributions of spatial profiles, so the better localization of reconstructed current sources versus potentials is an advantage. This is why we chose to use iCSD first, and then spatial ICA, to decompose the neural dynamics obtained from extracellular multi-electrode recordings into meaningful, functional components.

Application of any ICA algorithm to data always leads to some "independent" components in the context of the given framework. It is non-trivial to verify that they are the

⁵ http://www.neuroinf.pl/Members/szleski/iCSD_data/ICviewer

meaningful, or functional as we say, components one requires. Our conviction that the components obtained through ICA on iCSD in our experiments are indeed meaningful comes from the positive results of extensive testing on different types of data and from the structure of the spatial profiles and time courses of the obtained components which allow for meaningful physiological interpretation. We believe that further confirmation can be obtained by realistic modeling of the system generating the measured signals. Having control over different nuclei and their dynamics we will be able to verify if indeed the components obtained through our procedure are generated in different functional structures.

The physiological validity of the proposed method is supported by several features of the resulting components. First, the components are clearly separated (cf. Fig. 5 and 6), with neither of the components calculated for a given animal having similar location in space and time simultaneously. Second, most of the activated sources are tightly anchored to the somatosensory pathway structures known from the literature as relaying the vibrissal input (Waite 2004).⁶ Specifically, among few strongest components we typically found the ones with short latencies located in the first order thalamic nucleus VPM and zona incerta, structures known to receive strong and fast excitatory vibrissal input (Diamond et al. 1992; Nicolelis et al. 1992).

The results presented in Fig. 5, 6 based on the proposed ICA/iCSD method deliver the first functional confirmation of the separation between two pathways conveying the somatosensory information via the principal trigeminal nucleus and that part of VPM, which contains the cytochrome oxidase-rich barreloids (Land et al. 1995; Veinante and Deschênes 1999; Urbain and Deschênes 2007). The anatomical locations of both pathways were recently well established. The stratum of VPM at its dorsomedial margin ('heads' of barreloids) contains cells with obligatorily multi-whisker input whereas neurons from more lateral part of dorsal VPM ('cores' of barreloids), at certain levels of anesthesia, are capable of conveying information from single vibrissae (Urbain and Deschênes 2007). Thus, judging from their anatomical locations, the two components differentiated in our experiment would represent the multi- and single-whisker pathways, correspondingly. The fact that the cells from the principal trigeminal nucleus giving off the axons of the barreloid 'head' pathway have larger somata than those of the 'core' pathway (Veinante and Deschênes 1999; Lo et al. 1999) suggests that they may also conduct spikes with larger speed which is in line with our findings that the first component, encompassing 'heads' of barreloids reaches its maximum faster.

⁶ The single exception is Rat 6, where the recording grid encompassed the border with cortical tissue. For this rat we found three strong cortical components with long latencies accompanied by moderate strength components typically located at somatosensory thalamus.

It is worth emphasizing that our method which averages the volleys of activity with specific spatiotemporal characteristics is much more accurate in this respect than analyzes based on single unit recordings, as the spiking of single neurons in anaesthetized preparation is capricious and the required number of cells necessarily leads to averaging data from many animals. In contrast, our method allows for reliable analysis of data from a single animal. For example, the results shown in Fig. 6 quite well differentiate the two parallel volleys traversing the thalamus with different time dynamics and via different thalamic subregions. Obviously, one would need a group of animals to identify all sources of smaller amplitude.

Abbreviations

APT	anterior pretectal nucleus
DLG	dorsal lateral geniculate nucleus
Hipp	hippocampus
LD	latero-dorsal thalamic nucleus
LP	lateral posterior thalamic nucleus
MGM	medial geniculate nucleus, medial part
Po	posterior thalamic nuclear group
Rt	reticular thalamic nucleus
SN	substantia nigra
VPM	ventral postero-medial thalamic nucleus
VLG	ventral lateral geniculate nucleus
ZI	zona incerta
ZIc	caudal part of zona incerta
cp	cerebral peduncle
ic	internal capsule
ml	medial lemniscus
opt	optic tract

References

- Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation* 7(6), 1129–1159.
- Comon, P. (1994). Independent Component Analysis, a new concept? *Signal Processing* 36(3), 287–314.
- Diamond, M. E., Armstrong-James, M., & Ebner, F. F. (1992). Somatic sensory responses in the rostral sector of the posterior group (pom) and in the ventral posterior medial nucleus (vpm) of the rat thalamus. *The Journal of Comparative Neurology*, 318(4), 462–476.
- Freeman, J. A., & Nicholson, C. (1975). Experimental optimization of current source-density technique for anuran cerebellum. *Journal of Neurophysiology* 38(2), 369–382.
- Himberg, J., Hyvärinen, A., & Esposito, F. (2004). Validating the independent components of neuroimaging time series via clustering and visualization. *Neuroimage* 22(3), 1214–1222.
- Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology* 94(3), 1904–1911.

- Land, P. W., Buffer, S. A., & Yaskosky, J. D. (1995). Barreloids in adult rat thalamus: three-dimensional architecture and relationship to somatosensory cortical barrels. *The Journal of Comparative Neurology* 355(4), 573–588.
- Lin, B., Colgin, L. L., Brücher, F. A., Arai, A. C., & Lynch, G. (2002). Interactions between recording technique and AMPA receptor modulators. *Brain Research* 955(1-2), 164–173.
- Lipton, M. L., Fu, K. M. G., Branch, C. A., & Schroeder, C. E. (2006). Ipsilateral hand input to area 3b revealed by converging hemodynamic and electrophysiological analyses in macaque monkeys. *Journal of Neuroscience* 26(1), 180–185.
- Lo, F. S., Guido, W., & Erzurumlu, R. S. (1999). Electrophysiological properties and synaptic responses of cells in the trigeminal principal sensory nucleus of postnatal rats. *Journal of Neurophysiology* 82(5), 2765–2775.
- Łęski, S., Wójcik, D. K., Tereszczuk, J., Świejkowski, D. A., Kublik, E., & Wróbel, A. (2007). Inverse Current-Source Density method in 3D: reconstruction fidelity, boundary effects, and influence of distant sources. *Neuroinformatics* 5(4), 207–222.
- McKeown, M. J., Makeig S., Brown G. G., Jung T., Kindermann S. S., Bell A. J., & Sejnowski T. J. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping* 6 (3), 160–188.
- Mitzdorf, U. (1985). Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiological Reviews* 65(1), 37–100.
- Nicholson, C., & Freeman, J. A. (1975). Theory of current source-density analysis and determination of conductivity tensor for anuran cerebellum. *Journal of Neurophysiology* 38(2), 356–368.
- Nicolelis, M. A., Chapin, J. K., & Lin, R. C. (1992). Somatotopic maps within the zona incerta relay parallel gabaergic somatosensory pathways to the neocortex, superior colliculus, and brainstem. *Brain Research*, 577(1), 134–141.
- Novak, J. L., & Wheeler, B. C. (1989). Two-dimensional current source density analysis of propagation delays for components of epileptiform bursts in rat hippocampal slices. *Brain Research* 497(2), 223–230.
- Nunez, P. L., & Srinivasan, R. (2005). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford: University Press.
- Paxinos, G., & Watson, C. (1996). *The Rat Brain in Stereotaxic Coordinates, Compact Third Edition*. Academic Press.
- Paxinos, G., & Watson, C. (2007). *The Rat Brain in Stereotaxic Coordinates, 6th Edition*. Academic Press.
- Pettersen, K. H., Devor, A., Ulbert, I., Dale, A. M., & Einevoll, G. T. (2006). Current-source density estimation based on inversion of electrostatic forward solution: effects of finite extent of neuronal activity and conductivity discontinuities. *Journal of Neuroscience Methods* 154(1-2), 116–133.
- Pierret, T., Lavallée, P., & Deschênes, M. (2000). Parallel streams for the relay of vibrissal information through thalamic barreloids. *Journal of Neuroscience* 20(19), 7455–7462.
- Rajkai, C., Lakatos, P., Chen, C. M., Pincze, Z., Karmos, G., & Schroeder, C. E. (2008). Transient cortical excitation at the onset of visual fixation. *Cerebral Cortex* 18(1), 200–209.
- Reidl, J., Starke, J., Omer, D. B., Grinvald, A., & Spors, H. (2007). Independent component analysis of high-resolution imaging data identifies distinct functional domains. *Neuroimage*, 34(1), 94–108.
- Saleem, A. B., Krapp, H. G., & Schultz, S. R. (2008). Receptive field characterization by spike-triggered independent component analysis. *Journal of Vision*, 8(13), 2.1–216.
- Schiessl, I., Stetter, M., Mayhew, J. E., McLoughlin, N., Lund, J. S., & Obermayer, K. (2000). Blind signal separation from optical imaging recordings with extended spatial decorrelation. *IEEE Transactions on Biomedical Engineering*, 47(5), 573–577.
- Schroeder, C. E., Tenke, C. E., & Givre, S. J. (1992) Subcortical contributions to the surface-recorded flash-vep in the awake macaque. *Electroencephalography and Clinical Neurophysiology* 84(3), 219–231.
- Shimono, K., Brucher, F., Granger, R., Lynch, G., & Taketani, M. (2000) Origins and distribution of cholinergically induced beta rhythms in hippocampal slices. *Journal of Neuroscience* 20(22), 8462–8473.
- de Solages, C., Szapiro, G., Brunel, N., Hakim, V., Isope, P., Buisseret, P., Rousseau, C., Barbour, B., & Léna, C. (2008). High-frequency organization and synchrony of activity in the Purkinje cell layer of the cerebellum. *Neuron* 58(5), 775–788.
- Stoelzel, C. R., Bereshpolova, Y., & Swadlow, H. A. (2009). Stability of thalamocortical synaptic transmission across awake brain states. *Journal of Neuroscience* 29(21), 6851–6859.
- Stone, J. V., & Porrill, J. (1999). Regularisation using spatiotemporal independence and predictability. *Computational Neuroscience Report 201*, Psychology Department, Sheffield University.
- Stone J. V., Porrill J., Porter N. R., & Wilkinson I. D. (2002). Spatiotemporal independent component analysis of event-related fmri data using skewed probability density functions. *Neuroimage* 15(2), 407–421.
- Tanskanen, J. M. A., Mikkonen, J. E., & Penttonen, M. (2005). Independent component analysis of neural populations from multi-electrode field potential measurements. *Journal of Neuroscience Methods* 145(1-2), 213–232.
- Tenke, C. E., Schroeder, C. E., Arezzo, J. C., & Vaughan, H. G. (1993). Interpretation of high-resolution current source density profiles: a simulation of sublaminar contributions to the visual evoked potential. *Experimental Brain Research* 94(2), 183–192.
- Urbain, N., & Deschênes, M. (2007). A new thalamic pathway of vibrissal information modulated by the motor cortex. *Journal of Neuroscience* 27(45), 12407–12412.
- Vaknin, G., DiScenna, P. G., & Teyler, T. J. (1988). A method for calculating current source density (CSD) analysis without resorting to recording sites outside the sampling volume. *Journal of Neuroscience Methods* 24(2), 131–135.
- Veinante, P., & Deschênes, M. (1999). Single- and multi-whisker channels in the ascending projections from the principal trigeminal nucleus in the rat. *Journal of Neuroscience* 19(12), 5085–5095.
- Waite, P. (2004). Trigeminal sensory system. In G. Paxinos (Ed.), *The Rat Nervous System* (pp. 817–851). Elsevier.
- Wójcik, D. K., & Łęski, S. (2009). Current source density reconstruction from incomplete data. *Neural Computation*, doi:10.1162/neco.2009.07-08-831.
- Ylinen, A., Bragin, A., Nádasdy, Z., Jandó, G., Szabó, I., Sik, A., & Buzsáki, G. (1995). Sharp wave-associated high-frequency oscillation (200 Hz) in the intact hippocampus: network and intracellular mechanisms. *Journal of Neuroscience* 15(1), 30–46.