UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Data-driven re-referencing of intracranial EEG based on independent component analysis (ICA)

Michelmann, Sebastian; Treder, Matthias; Griffiths, Benjamin; Kerren, Casper; Roux, Frederic; Wimber, Maria; Rollings, David; Sawlani, Vijay; Chelvarajah, Ramesh; Gollwitzer, Stephanie; Kreiselmeyer, Gernot; Hamer, Hajo; Bowman, Howard; Staresina, Bernhard; Hanslmayr, Simon

DOI:
10.1016/j.jneumeth.2018.06.021

License:
Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version
Peer reviewed version

Citation for published version (Harvard):
Michelmann, S, Treder, M, Griffiths, B, Kerren, C, Roux, F, Wimber, M, Rollings, D, Sawlani, V, Chelvarajah, R, Gollwitzer, S, Kreiselmeyer, G, Hamer, H, Bowman, H, Staresina, B & Hanslmayr, S 2018, 'Data-driven re-referencing of intracranial EEG based on independent component analysis (ICA)', *Journal of Neuroscience Methods*. https://doi.org/10.1016/j.jneumeth.2018.06.021

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 19. Nov. 2026

UNIVERSITY OF BIRMINGHAM

Research at Birmingham

Data-driven re-referencing of intracranial EEG based on independent component analysis (ICA)

Hanslmayr, Simon; Chelvarajah, Ramesh; Sawlani, Vijay; Michelmann, Sebastian; Bowman, Howard; Wimber, Maria; Staresina, Bernhard; Kerren, Casper; Roux, Frederic

Citation for published version (Harvard):
Hanslmayr, S, Chelvarajah, R, Sawlani, V, Michelmann, S, Bowman, H, Wimber, M, Staresina, B, Kerren, C & Roux, F 2018, 'Data-driven re-referencing of intracranial EEG based on independent component analysis (ICA)' Journal of Neuroscience Methods.

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 28. Jun. 2018

Data-driven re-referencing of intracranial EEG based on independent component analysis (ICA)

Sebastian Michelmann[1]*, Matthias S. Treder[1], Benjamin Griffiths[1], Casper Kerrén[1], Frédéric Roux[1], Maria Wimber[1], David Rollings[2], Vijay Sawlani[2], Ramesh Chelvarajah[2], Stephanie Gollwitzer[3], Gernot Kreiselmeyer[3], Hajo Hamer[3], Howard Bowman[1,4], Bernhard Staresina[1], Simon Hanslmayr[1]

1. School of Psychology, University of Birmingham, UK
2. Complex Epilepsy and Surgery Service, Neuroscience Department, Queen Elizabeth Hospital Birmingham, UK
3. Epilepsy Center, Department of Neurology, University Hospital Erlangen, Germany
4. Centre for Cognitive Neuroscience and Cognitive Systems and the School of Computing, University of Kent at Canterbury, UK

*corresponding author. Contact: sebastian_michelmann@hotmail.de, School of Psychology, University of Birmingham, Edgbaston, B15 2TT

Keywords:

Independent Component Analysis

Intracranial EEG

Referencing

Preprocessing

Bipolar Reference

Depth Electrodes

Neuroscience

Electroencephalography

Abstract

-Background

Intracranial recordings from patients implanted with depth electrodes are a valuable source of information in neuroscience. They allow for the unique opportunity to record brain activity with high spatial and temporal resolution. A common pre-processing choice in stereotactic EEG (S-EEG) is to re-reference the data with a bipolar montage. In this, each channel is subtracted from its neighbor, to reduce commonalities between channels and isolate activity that is spatially confined.

-New Method

We challenge the assumption that bipolar reference effectively performs this task. To extract local activity, the distribution of the signal source of interest, interfering distant signals, and noise need to be considered. Referencing schemes with fixed coefficients can decrease the signal to noise ratio (SNR) of the data, they can lead to mislocalization of activity and consequently to misinterpretation of results.

We propose to use Independent Component Analysis (ICA), to derive filter coefficients that reflect the statistical dependencies of the data at hand.

-Results

We describe and demonstrate this on human S-EEG recordings. In a simulation with real data, we quantitatively show that ICA outperforms the bipolar referencing operation in sensitivity and importantly in specificity when revealing local time series from the superposition of neighboring channels.

-Comparison with Existing Method(s)

We argue that ICA already performs the same task that bipolar referencing pursues, namely undoing the linear superposition of activity and will identify activity that is local.

-Conclusions

When investigating local sources in human S-EEG, ICA should be preferred over re-referencing the data with a bipolar montage.

Introduction

An increasingly popular tool in modern cognitive neuroscience is to investigate local field potentials recorded from electrodes implanted into the brain of epileptic patients. Data recorded from these patients offer a unique and powerful opportunity to understand how local neural populations implement cognitive operations. As is the case in all EEG recordings, electrical potentials are recorded as differences between two sites, in the simplest case an "active" site and a "passive" reference. Therefore a delicate decision has to be made as to what is the best reference for a given recording site (e.g. Nunez & Srinivasan, 2006). Online recording of stereotactic EEG (S-EEG) can for example be performed with a mastoid or a subdermal reference; however the offline analysis is often preceded by a re-referencing operation. While ICA has been shown to be a useful tool for the analysis of Electrocorticographic (EcoG) recordings in the past (Hu, Stead, & Worrell, 2007; Whitmer, Worrell, Stead, Lee, & Makeig, 2010), it is still standard to re-reference the data with a bipolar montage (e.g. Staudigl, Vollmar, Noachtar, & Hanslmayr, 2015; Tallon-Baudry, Bertrand, & Fischer, 2001). Other re-referencing operations that can be used for S-EEG recordings include working with scalp referenced (e.g. linked mastoid) data (Staresina et al., 2016), re-referencing to average (Liu, Coon, De Pesters, Brunner, & Schalk, 2015; Ludwig et al., 2009), laplacian referencing (see Tenke & Kayser, 2012), neutral referencing via two adjacent white matter leads (P. Avanzini, Pelliccia, Lo Russo, Orban, & Rizzolatti, 2018; Pietro Avanzini et al., 2016) and infinity reference (Yao, 2001).

In bipolar re-referencing, each channel is subtracted from its neighbor. Bipolar referencing is seen as advantageous since it removes contamination from activity at the online reference electrode, and highlights activity that is local (J. P. Lachaux, Rudrauf, & Kahane, 2003).

Especially in a recent paper Mercier et al. (Mercier et al., 2017) investigate the diffusion of cortical field potential in S-EEG and empirically compare the effect of commonly used reference choices on the recorded signal. They find that using both neighboring electrodes as a reference in a local referencing scheme reduces correlations between channels. Based on this reference, the authors analyze the activity recorded in cortical white matter and conclude that it contains a mixture of signal sources spreading from nearby gray matter and signal from other sources that sometimes correlates with distant gray matter activity. Their findings highlight the importance of considering anatomical structure when choosing the reference. They further render the use of a white matter reference somewhat suboptimal and put a spotlight on the fact that recorded activity at a given electrode reflects a mixture of activity from different sources.

The authors further highlight that *"re-referencing electrophysiological data is a critical preprocessing choice that could drastically impact signal content and consequently the results of any given analysis"*

(Mercier et al., 2017 p. 219, l.4; see also: Shirhatti, Borthakur, & Ray, 2016) . We fully agree with this statement and think that the problem of referencing in the realm of intracranial studies needs more discussion, especially because S-EEG is becoming an increasingly popular tool in cognitive neuroscience.

Historically, a bipolar reference has offered a coarse but efficient standard way to make epileptic spikes visible in a clinical setting. Today, it is still a starting point to identifying the source of epileptic spikes and additional reference montages are routinely included in recordings (Flink et al., 2002). The intricacies of interpreting a bipolar montage are well known among clinicians. The analysis of S-EEG recordings for research purposes, however, often discards epileptic spikes and focuses on the function of healthy brain tissue (J.-P. Lachaux, Axmacher, Mormann, Halgren, & Crone, 2012). This requires rigorous procedures that not only maximize the signal to noise ratio (SNR) of the data, but also carefully consider the nature of the recorded signal.

Especially when differential functions of local structures are investigated with S-EEG, a wrong choice of reference can lead to drastic misinterpretations of results; that is, mislocalization of effects. This is highly relevant in cognitive neuroscience, since results from intracranial recordings act somewhat as a ground truth for the localization of electrophysiological activity.

In the current paper, we therefore want to add to this discussion. Specifically, we believe that it is helpful to point out disadvantages of bipolar and local referencing schemes, especially with regard to activity recorded in subcortical structures (notably the Hippocampus), which were not considered by Mercier et al., due to "*distinct characteristics of the corresponding signal*" (Mercier et al., 2017 p. 221, l. 21). By spelling out the mathematical details of the referencing operation, we more explicitly discuss its effects on signal-to-noise-ratio (SNR) and identify conditions in which bipolar referencing can lead to mislocalization of signal activity. As an alternative, we propose to use Independent Component Analysis (ICA) in order to derive a data-driven referencing scheme that is based on the statistical dependencies between channels. Previously, ICA has been used in free moving animals, to remove contamination from the reference electrode and from volume conducted noise (Whitmore & Lin, 2016) and its application to intracranial EEG in order to identify activity at the scalp reference electrode has been previously suggested (Hu et al., 2007). In this context, we here want to introduce ICA as a general re-referencing framework for human s-EEG data. We champion a perspective on referencing schemes that describes them as spatial filtering operations. We argue that ICA, as a spatial filtering operation with coefficients derived from the statistical dependencies in the data at hand, should effectively replace any re-referencing operation during preprocessing. For simplicity, in our comparisons, we will mainly focus on the bipolar referencing scheme that is often the preferred

preprocessing choice; however most of our analysis and simulations also generalize to other referencing schemes such as the local referencing scheme proposed by Mercier et al. (Mercier et al., 2017). In our basic simulations, we manipulate the superposition of real data under varying levels of noise. Importantly these simulations show that ICA, which assumes a linear superposition of sources in the same way that bipolar re-referencing does, not only results in higher sensitivity when it comes to localizing a simulated source at the electrode on which it was originally placed, but it also outperforms a bipolar referencing scheme in specificity, when activity from distant sources is to be discarded.

Material and Methods

Participants

For the analyses, data from 9 patients (5 female) were used who were on average 35.78 years old (range: 24-53). Seven patients were recorded at the Queen Elizabeth Hospital Birmingham (QEHB), 2 patients were recorded at the University Hospital Erlangen (UKE).

Patients suffered from drug-resistant epilepsy and were implanted with intracranial depth electrodes. They underwent pre-surgical monitoring purely for diagnostic purposes. All patients volunteered to participate in a memory study. Written informed consent was obtained in accordance with the Declaration of Helsinki.

Task

During the memory task, patients were watching 3 different 4-second-long movie sequences that each consisted of 2 distinct scenes. In one of the scenes, a word appeared in the center of the screen and patients were instructed to vividly associate the word with the exact scene within the movie. After performing a short distractor task in which they categorized numbers as odd or even, patients were shown the words in an arbitrary order. For every word they first decided whether they had previously learned it in scene 1 or scene 2 of a movie, and then identified the movie it was associated with. To improve performance and increase the number of successfully remembered trials, each association was learned 3 times and later recalled 3 times.

Data recording and preprocessing

The data were continuously recorded at a sampling rate of either 1024 Hz or 1000 Hz, with an online linked-mastoid reference. Data were imported into MATLAB 2014a (MathWorks) using FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) for data from QEHB and using Brainstorm (Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011) for data from UKE. Subsequently, data were downsampled to 1000 Hz and epochs of 7 seconds were created, beginning 2 seconds before video onset at encoding and word onset at retrieval. Epochs ended 5 seconds after video/word onset.

All channels that displayed frequent epileptic spiking or electrical noise were excluded from further analysis. Remaining trials that still contained artifacts were manually removed at a later stage.

ICA computation

ICA was computed using the EEGLab implementation ‘runica’ (Delorme & Makeig, 2004). Before computation of the unmixing matrix, all data were filtered with a high-pass filter of 1.5 Hz and trials

containing artifacts were heuristically removed based on statistical characteristics (i.e. variance, kurtosis and maximum value). This was done to improve the estimation of filter coefficients that should reflect the relationship between channels in the absence of coarse artifacts. In this, whole trials were excluded, ensuring the same number of sampling points at every channel. The unmixing-matrix was later applied to the unfiltered data including all trials; trials containing artifacts were then removed based on visual inspection. For the illustration in Fig. 1, only those trials that remained after visual inspection were included and ICA computation was limited to the three channels displayed in the figure.

Electrode localization

For the example-channels displayed in Fig. 1 and Fig. 2, the pre-surgical MRI was first segmented using Freesurfer (Fischl, 2012; Reuter, Schmansky, Rosas, & Fischl, 2012) and anatomical labels were derived from the Desikan-Killiany Atlas (Desikan et al., 2006). Electrode locations were manually determined based on the center of signal drop-out in the post-surgical MRI-scan.

The post-surgical MRI-scan was then co-registered with the pre-operational MRI using robust coregistration (Reuter, Rosas, & Fischl, 2010) as implemented in Freesurfer. Lastly, the labeled segmentation was overlaid with the post-operational MRI and the highlighted electrode positions to determine the approximate structure in which the electrode was located. This was done to derive labeling in a partly automated and standardized way.

The locations of the three electrodes for simulations were manually labeled based on the anatomical characteristics of the surrounding tissue in the post-operational MRI.

Power Spectra

To derive the power spectra in the encoding block, Fourier transformed data were multiplied with a Hanning taper of 4 cycles for a given frequency. Power spectra of every full frequency between 2 and 39 Hz were computed. This was done every 10ms in a sliding window from 0.5 seconds prior to movie onset to 1.5 seconds during the movie (e.g. Jokisch & Jensen, 2007) .

All processing of the data was done using the FieldTrip toolbox for EEG/MEG-analysis (Oostenveld et al., 2011).

Simulations

For all simulations, three arbitrary electrodes from three different patients were selected, which were later used to define independent sources based on real data. Electrodes were located in the right Parahippocampal Cortex, in the right Perirhinal Cortex and in the left Middle Temporal Gyrus. Those

channels were selected because they were particularly free of epileptic spiking. For the simulations, ICA was run on unfiltered visually inspected data. After Fisher Z-transformation of correlation coefficients, a dependent samples t-test was computed to derive analytical statistics across 100 repetitions of the simulation.

Bipolar and local (re-)referencing

Bipolar referencing is a popular preprocessing choice for several reasons (J. P. Lachaux et al., 2003; Shirhatti et al., 2016; Trongnetrpunya et al., 2015):

Firstly, bipolar re-referencing removes the reference activity. Activity from the online reference channel is expressed on all channels, since it is subtracted into all electrodes during recording. By later re-referencing one channel against another, the reference activity is removed.

Secondly, bipolar referencing removes noise sources with a broad spatial distribution (Trongnetrpunya et al., 2015); if all electrodes are picking up an external noise source (e.g. 50Hz line noise), this noise will likewise disappear in the subtraction of a channel from its neighbor. In practice however, line noise and other noise sources often do not affect all electrodes to the same extent. Therefore, if the data is not sufficiently inspected before re-referencing, this can for example lead to the subtraction of noise into an otherwise clean channel; in other words, the distribution of external noise sources in the data needs to be considered.

Finally, bipolar referencing emphasizes spatially local activity. By subtracting one channel from its neighbor, only signal that is unique to this channel will supposedly remain unaffected. Common (broad) activity is attenuated or even fully removed. This postulation is somewhat problematic, since it relies on a number of assumptions (see below). For the extraction of local signal, the distribution of the signal source of interest, as well as the distribution of interfering distant signals and noise needs to be considered, which can have a variable spatial extent (Kajikawa & Schroeder, 2011).

The referencing operation

In general, a referenced channel, whether it was computed via bipolar, local, or average referencing will be a simple linear combination of electrodes and therefore a mixture of activity recorded from several electrodes. Whether this linear combination reduces the dependencies between the channels or even introduces new dependencies and mixes separate sources even more, depends on the spatial location and extent of the underlying signal of interest, as well as factors like noise distribution, online reference and electrode location. The importance of considering anatomical structure in the reference choice has been extensively analyzed by Mercier et al. (Mercier et al., 2017). We now want to consider the implicit assumptions about the distribution of underlying source-signals in bipolar referencing and their impact on the resulting channel-activity.

In the bipolar referencing scheme, the new re-referenced channel is a linear mixture of two electrodes: the time series on the neighboring channel is subtracted from the time series on the electrode of interest.

Using one neighbor as a reference (e.g. Staudigl et al., 2015; Tallon-Baudry et al., 2001), we define the bipolar reference in accordance with the nomenclature proposed by Mercier el al. (Mercier et al., 2017) (eq. 1) as

$$V'_k = (V_k - V_{k-1}) \quad (1)$$

where V_k refers to the electrode in the kth position and V'_k is the re-referenced time series.

Considering the time series on electrode $V_{k,}$ we can write it as a linear combination of activity that we would take for signal S_k and activity that is noise N_k:

$$V_k = \alpha_k S_k + \beta_k N_k \quad (2)$$

Here, we conceive of noise on a given electrode V, as both true sensor noise and general brain activity that does not emanate from the immediate vicinity of the electrode of interest; that is, brain activity either spread through passive conduction or transported via white matter tracts. If we now compute

$$V'_k = \alpha_k S_k + \beta_k N_k - V_{k-1} \quad (3)$$

we implicitly expect that activity that is shared between the neighboring channels will constitute mostly noise and that the channel V_{k-1} contributes little or no unique activity to the re-referenced signal. We now want to define S^*_{k-1} as signal that is unique to contact (k-1) and N^*_{k-1} as noise that is unique to contact (k-1). Together they account for the remaining activity that is not shared with electrode V_k.

Since we defined V_k as the electrode of interest, the activity at electrode V_{k-1} would be written as

$$V_{k-1} = \alpha_{k-1} S_k + \beta_{k-1} N_k + S^*_{k-1} + N^*_{k-1} \quad (4)$$

that is, a combination of activity S that would be considered signal at the electrode V_k, activity N that would be considered Noise at the electrode V_k, activity S^*_{k-1} which is unique to the electrode V_{k-1} and would be considered part of the signal at V_{k-1} and noise N^*_{k-1} which is unique to the electrode V_{k-1}. Combining these terms with the previous formula (3), we obtain the activity at the new re-referenced channel V'_k as

$$V'_k = (\alpha_k S_k - \alpha_{k-1} S_k) + (\beta_k N_k - \beta_{k-1} N_k) - S^*_{k-1,} - N^*_{k-1} \quad (5)$$

Consequently, the best conditions for bipolar referencing are, if the common noise term N_k is of equal magnitude on both channels ($\beta_k = \beta_{k-1}$), so it can cancel out, and if the local signal of interest S_k is considerably stronger on the contact V_k than on the contact V_{k-1} ($\alpha_k >> \alpha_{k-1}$) and therefore not dampened much. Additionally, little or no unique signal S^*_{k-1} and unique noise N^*_{k-1} should be present on the neighboring channel V_{k-1} which could otherwise overshadow the signal S_k.

In practice, we accept that the re-referenced channel V'_k will consist of a linear mixture of dampened signal and residual noise from V_k and inverted unique signal and unique noise from V_{k-1}.

After bipolar referencing, the spatial resolution of the data will be reduced: we can maximally locate the signal of interest between the two electrodes V_k and V_{k-1}. However, a signal of interest emanating in the close vicinity of V_k or V_{k-1}, will not automatically have its maximal strength between these electrodes after re-referencing. Since subtracting one channel from its neighbor is essentially taking the spatial derivative, the strongest signal after re-referencing will appear at the point where the signal drops/increases most between two channels. This means that for a correct localization, the difference in signal strength α between electrode k and (k-1) must be bigger than between (k-1) and (k-2) and bigger than between (k+1) and k (i.e. $|\alpha_{k+1} - \alpha_k| < |\alpha_k - \alpha_{k-1}| > |\alpha_{k-1} - \alpha_{k-2}|$).

This requirement can be especially problematic in cases where some electrodes on a shaft are in a discrete structure. The two most mesial electrodes on a shaft can for example fall in the Hippocampus and pick up a strong signal there. A third electrode in the Parahippocampal Cortex can pick up very little of this activity. In the re-referenced channels, it will then appear like the peak of the underlying hippocampal signal is between Hippocampus and Parahippocampal Cortex (or incorrectly interpreted on the re-referenced channel V'_3 (= V_3-V_2) which falls in the Parahippocampal Cortex), even though

only the two electrodes in the Hippocampus strongly picked up the characteristic signal in the first place.

In another hypothetical scenario, consider a signal that has a very broad distribution and is nearly equally strong on the first (n-1) electrodes of the shaft. It will only appear in inverted form at the end of the electrode shaft after bipolar re-referencing and could easily be mistaken for activity from a local source. These are just two examples of how the standard approach of bipolar referencing can lead to drastic mislocalization of a source signal.

Impact on SNR

An important quantity for assessing signal quality after re-referencing is signal-to-noise-ratio (SNR). In the bipolar referencing scheme, SNR can either increase or decrease, depending on the distribution of signal of interest (S), the distribution of signal that is not of interest (S*) and the distribution of noise sources (N and N*).

If we quantify the signal to noise ratio before re-referencing as

$$\frac{Var(S_k)}{Var(N_k)} \tag{6}$$

we can write the signal to noise ratio of the re-referenced channel as

$$\frac{var(\alpha_k S_k - \alpha_{k-1} S_k)}{\mathrm{var}(\beta_k N_k - \beta_{k-1} N_k) + \mathrm{var}(S^*_{k-1}) + \mathrm{var}(N^*_{k-1})} \tag{7}$$

for which we used that the unique signal S^*_{k-1} and the unique noise N^*_{k-1} are by definition independent of each other and independent of the signal and noise time series S_k and N_k.

Generally, only noise that is shared between a selected channel and the neighboring channel will have a positive impact on SNR with re-referencing. Shared signal will lead to a dampening of the signal, i.e. decrease the numerator in the Signal/Noise fraction and therefore decrease the signal to noise ratio on the re-referenced electrode, whereas unique noise and signal that is not of interest (e.g. stemming from white matter) will add to its denominator. Under favorable conditions, the bipolar referencing scheme can therefore improve SNR, however it can also decrease SNR and signals can appear stronger on neighboring or even distant channels than on the source-channel (see above).

Notably, reduction in SNR will also decrease correlations between channels, i.e. a reduction in correlation between channels as observed by Mercier et al. (Mercier et al., 2017), can occur partly because signal is lost in the re-referencing process. In other words: noise can be uncorrelated as well, so a lack of correlation alone is not a reliable quality measure for successful extraction of a local signal of interest.

Referencing is spatial filtering

A useful perspective in assessing referencing schemes is to consider them as spatial filters (e.g. Bleichner & Debener, 2017). In analogy to temporal filters, spatial filters can sometimes be characterized by the spatial frequencies that are selected or suppressed.

The bipolar referencing operation is an approximation of the first spatial derivative. Since the derivative and the Fourier transform are both linear operations, we can estimate the effect of bipolar referencing on the spatial frequency spectrum by taking the derivative of the Fourier-coefficients, in which each coefficient's complex conjugate is multiplied by the frequency itself. This means that high frequencies are amplified and low frequencies are dampened.

If we assume that low temporal frequencies will show a broader spread to neighboring electrodes than high temporal frequencies (i.e. also have a lower spatial frequency), the shared signal between neighboring electrodes will be stronger in the low frequencies and therefore low temporal frequencies will be disproportionally dampened by the referencing operation, in other words the spectral properties of the time series can change. This is further complicated because neural sources are best described as current dipoles. In addition to the spatial frequency of sources, their orientation will therefore have an impact on the effects of re-referencing.

Importantly, all these confounding variables will be modulated by the distance between neighboring electrodes and their location, impeding the comparability between studies and even patients.

Furthermore, a general problem with local/bipolar referencing is the inevitable loss of information. Activity on the *n* electrodes on an electrode shaft cannot be sufficiently explained by the linear combination of (n-1) electrodes that is obtained via bipolar referencing ((n-2) electrodes in local referencing) unless there are already linear dependencies between the channels beforehand (i.e. the re-referenced data will have a reduced rank). For this reason, routinely re-referencing the data can mean to throw away information. This problem is additionally aggravated in datasets in which only few electrodes are present on each shaft.

Towards a data driven reference

The purpose of this manuscript is to demonstrate that a standard ICA algorithm (Comon, 1994; Hyvärinen & Oja, 2000) already outperforms bipolar referencing (and other referencing schemes) in extracting the signal of interest. This applies specifically to cases in which the goal of referencing is to extract sources that are local and have high spatial specificity (i.e. do not pick up activity that originates in distant structures).

The goal of ICA is very similar to that of bipolar referencing, namely to apply spatial filters to the recorded activity in order to undo the linear superposition of sources and find the underlying signal time series.

By separating the activity into the same number of sources as there are channels, no information is lost in this process. The filters that are computed via ICA not only optimize the statistical independence between underlying components (which is a stronger requirement than merely reducing correlations (Rodgers, Nicewander, & Toothaker, 1984)), they also allow for hidden sources with high and low spatial frequency and this distribution can be inspected in the columns of the mixing matrix. The resulting filter coefficients will be static across time and ICA therefore acts in the same way as referencing operations: each channel is replaced with a linear combination of channels. In contrast to the classical referencing schemes, ICA is a data-driven approach. That is, its coefficients adapt to the data at hand, whereas bipolar referencing uses the fixed coefficients (1,-1) on neighboring channels. Thus bipolar referencing is but one implementation of a much larger space of coefficients afforded by ICA.

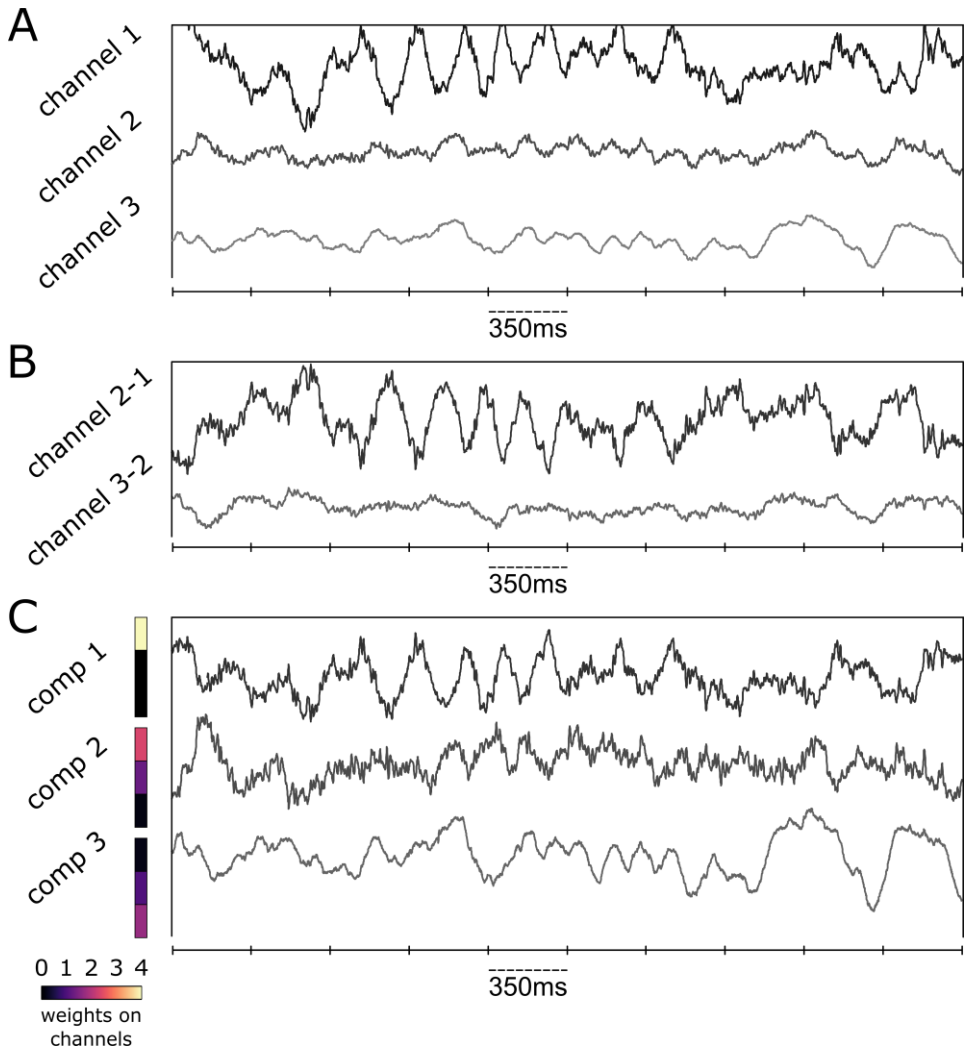


Fig. 1. Example of how bipolar re-referencing can impair the interpretability of the data. A: Activity during a single trial on 3 selected channels in the medial temporal lobe of a patient implanted with intracranial electrodes. Channels 1 and 2 fall in the Hippocampus, channel 3 is located in the nearby white matter. B: The same channels were re-referenced with a bipolar scheme. The first channel appears inverted and the shared activity between channel 2 and 3 is almost cancelled out. C: The same trial using an ICA decomposition of the data. Inspecting the weights in the columns of the 3x3 mixing matrix (i.e. the topography), we can see how the hidden components mix into the channels observed in A. Instead of subtracting the channels from each other and therefore mixing them even more, ICA splits the 3 channels into activity that is unique to channel 1 (component 1), activity that is shared between channel 1 and 2 (component 2) and activity that is shared between channel 2 and 3 (component 3). Interestingly component 2 is mostly present in channel 1, however it appears strongest (and inverted) on the combined bipolar channel 3-2. This is because the difference between its contribution to channel 1 and to channel 2 is smaller than the difference between its contribution to 2 and to 3, which can be observed in C.

ICA can reconstruct local sources

Independent component analysis (ICA) is a method for blind source separation. It aims at explaining a random vector by a linear combination of underlying components (sources) that are statistically independent (Comon, 1994).

ICA is extensively used in scalp EEG/MEG in order to separate brain signals from artifacts (Delorme & Makeig, 2004).

In the underlying framework, the observed data can be sufficiently explained by a linear mixture of underlying sources:

$$X = A*S \tag{8}$$

where X is the data (channels x time points), S are the underlying sources (components x time points) and A is a transformation matrix, which is usually called the mixing matrix.

In practice, the mixing matrix A and the underlying sources are both unknown. ICA aims to undo the linear mixing – algorithms find the inverse of the mixing matrix A^{-1} (i.e. the unmixing matrix), such that the resulting components are maximally statistically independent:

$$A^{-1}*X = S \tag{9}$$

See (Hyvärinen & Oja, 2000) for a review. Since A^{-1} is a square matrix of full rank, it is invertible and we can switch between channel representation and component representation of the data without losing information. Importantly, the rows of the unmixing matrix act as a set of spatial filters, i.e. they are linear transformations of the electrodes that highlight the activity of an independent source. When we inspect the columns of the mixing matrix A, we can see how much an underlying source contributes to each channel in the observed data representation.

Usually, a column of the mixing matrix can be used to visualize the topography of a component, which is helpful in detecting artifacts in scalp M/EEG (Delorme & Makeig, 2004). For intracranial data, we can still use this information in order to identify local sources and separate them from hidden sources that affect all channels to an almost identical extent.

Intuitively, ICA feels like a drastic transformation of the data and it appears less justifiable then simple re-referencing. In particular, the ICA framework relies on assumptions like a non-Gaussian distribution of underlying sources and a linear mixing model. In practice, however, the output of the ICA is a set of linear filters that optimally complies with these assumptions and unmixes the linear superposition on the channels accordingly. The strong benefit from this is that no a priori assumptions about the filter-

coefficients are made. We can further show via simulations that ICA yields higher sensitivity and specificity than bipolar re-referencing which only makes the assumption of a linear superposition and allows for no flexibility in filter-coefficients (see below).

If the goal of the analysis is to draw conclusions about local sources, we propose to do an ICA on the data and then compute a measure of uniformity for every column in the mixing matrix. This has the advantage that no prior assumptions about the distribution of signal and noise on the electrodes are made; rather the actual signal spread is estimated from the statistical properties of the data. Later, the distribution of underlying sources can be inspected and broad sources can be discarded from further analysis.

In particular, we propose the following steps:

1) Compute ICA on a cleaned version of the data, excluding channels and trials that are contaminated by epileptic spikes.
2) Calculate a measure of uniformity ('broadness') for the absolute (since a component can have positive and negative polarity on different channels) of every column of the mixing matrix. As a measure of uniformity, we propose a simple chi^2-value; degrees of freedom correspond to the number of channels. Alternatively, one could consider the Kullback–Leibler divergence (Kullback & Leibler, 1951) from a uniform distribution or a related measure (see also: Whitmore & Lin, 2016).
3) Discard all sources that are broad in topography (e.g. the column has a chi^2 value past a threshold corresponding to $p = 0.2$) from further analysis.
4) One of a, b, c:
    a) Project the remaining components back in order to keep working on channel-data. The channels in the cleaned dataset will have linear dependencies; effects can be interpreted as being present in the structure they are measured in and not originating from broad sources or the reference channel.
    b) Continue working with the local independent components. Since the largest weight in the column of the mixing matrix determines which channel picks up most of the activity from that component, component-labels can be changed to the label of their peak-weight. The reasoning behind this is that signals should be strongest where they originate. Additionally, all other weights can be kept. They can be useful when interpreting the spatial extent of an effect and when averaging across several subjects.
    c) Selectively build channels from a linear combination of only those independent components that have their largest weight (peak in the column of the mixing matrix) on the corresponding

electrode. This special case of option 'b' deals with multiple sources localized to the same contact. Specifically, one adds together those sources that peak on the same contact, weighted by their strength. This way, only the contributions of identified sources at an electrode are considered, when a channel is built as a linear combination. This probably produces the most realistic representation of the signal that would be measured by a reference-free, uncontaminated electrode in the respective structure, without interference from distant sources.

Importantly, the three options at step four have different advantages and disadvantages. The back-projection of components that are local (4a) is certainly a robust way to obtain data that has been stripped of broad (i.e. shared) components. For very local sources, the channel-representation can also provide a more stable solution (see simulations). Furthermore, most researchers are probably more comfortable working with the familiar channel representation of the data. A disadvantage of the back-projection is that channels will again represent a mixture of neighboring sources, even though the spread of these sources is more confined. A further aspect to keep in mind is that the channel-data will be highly rank deficient, which however, is not necessarily a disadvantage.

The main advantage of working with the local components (4b) is that the overlap of neighboring sources is optimally reduced. This approach is therefore best suited if differential conclusions about neighboring structures are to be drawn. To disentangle a hippocampal and a parahippocampal source is a prime example in which two local components that express a clear peak in one of the respective structures would be compared. The main disadvantage of this approach is that working with component-data instead of the channel representation can be unintuitive at first.

The selective linear combination (4c) of several components that express a peak at a certain electrode should be a special case that is only relevant, if a researcher encounters multiple sources that are localized to the same contact. In this scenario one might want to combine those sources into a single channel, in order to have e.g. a single hippocampal channel that reflects all sources that were localized in the Hippocampus.

An example of steps 1-3 is given in Fig. 2. Data from one patient is presented who had electrodes implanted in the left and right hemisphere. ICA was computed on the channels from the right hemisphere, whereas channels from the left were discarded due to epileptic spiking. Three shafts were located in the right anterior, medial and posterior temporal lobe; some channels on the medial and posterior shaft were located in the Hippocampus (Fig. 2A, left). A chi^2 test was then computed on the columns of the absolute of the mixing matrix. Columns of the mixing matrix were then sorted by the chi^2 value in descending order (Fig. 2A, right). The resulting component 1 and component 3

showed a local distribution and peaked in the Hippocampus. It can be seen that component 1 is expressed most strongly on the posterior hippocampal channel R POST 1, but it also affects the channel in the medial Hippocampus to a larger extent than any other source. Another local source could be revealed on the medial hippocampal channel R MID 2. Importantly, the component which explained the second most mean projected variance in the data (i.e. component 2) had an almost uniform distribution. It can therefore not be considered local activity and was probably due to activity on the reference mastoid electrodes. All components 1-3 expressed an evoked response upon stimulus onset (Fig. 2B). They also showed distinct responses in the power spectrum upon stimulus presentation, namely a strong theta power increase was observed in the local hippocampal sources, whereas the broad source showed mostly alpha power decreases, consistent with this source picking up scalp EEG from the mastoids (Fig. 2C). This illustrates the importance of separating local and broad activity when analyzing intracranial EEG and the power of ICA in effectively performing this task. In order to demonstrate the usefulness of an ICA based approach for several datasets, ICA was computed for 8 further patients. The mixing matrix of each patient was sorted from local (left) to broad (right). Fig. 3 shows these examples of the separation of local and broad sources on datasets with up to 110 electrodes.

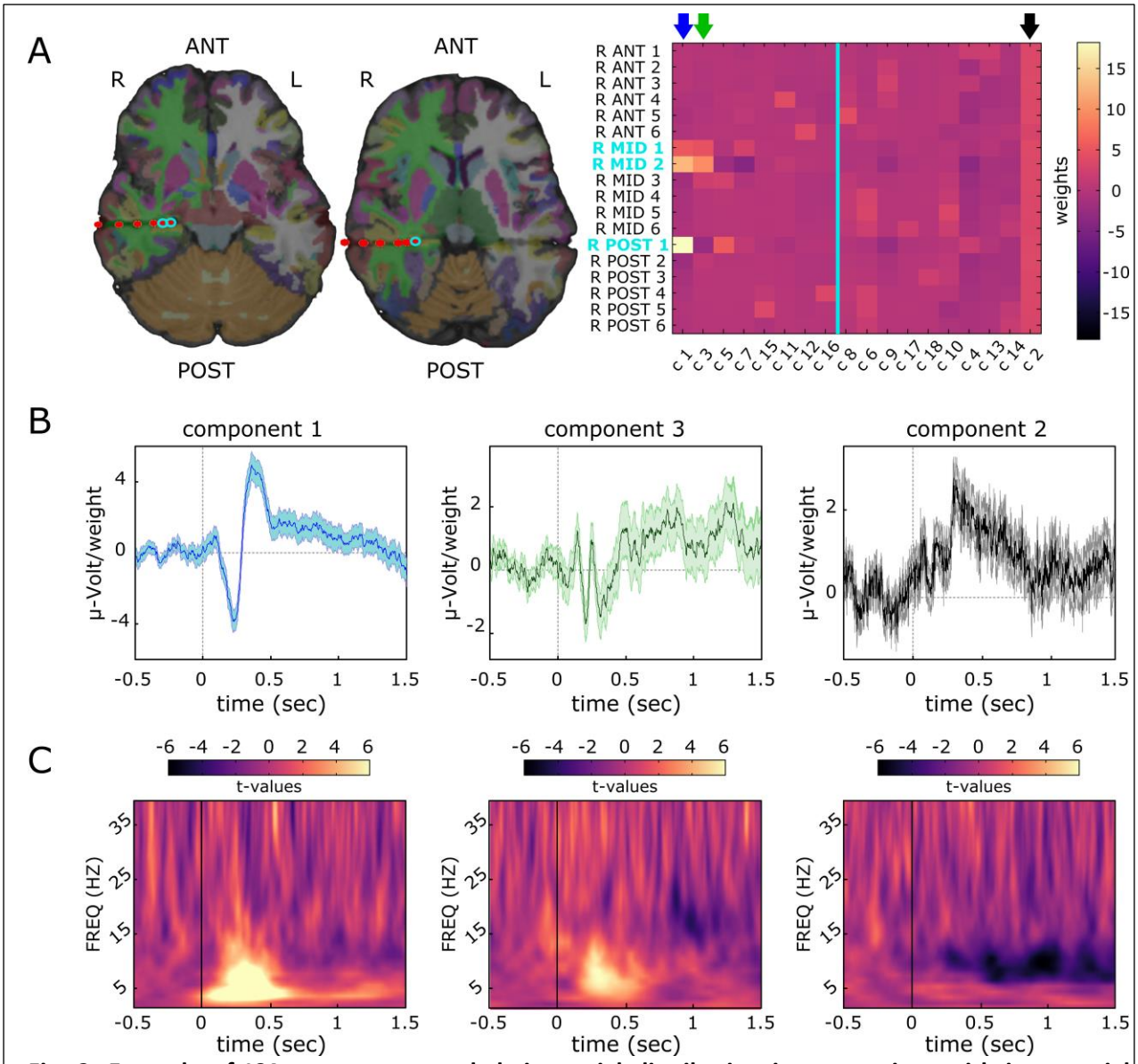


Fig. 2. Example of ICA-components and their spatial distribution in one patient with intracranial electrodes.

The patient was performing a memory task in which they repeatedly associated short video clips with words and recalled them later. A: two out of three used electrode shafts are depicted on the left. The brain is superimposed by labels corresponding to the Desikan-Killiany Atlas, which were derived via Freesurfer parcellation. R-Mid 1-6 and R-Post 1-6 are located in the right medial and right posterior temporal lobe respectively (counting from inside to outside). ICA was computed and the inverse of the unmixing-matrix (i.e. the topography) was sorted by the spatial "broadness" of the components. Columns of this mixing-matrix (right) correspond to the distribution of the components. The turquois vertical line separates local components on the left from components that have a broad distribution across channels (corresponding to a chi^2- value of $p > 0.2$), which one could consider discarding from further analysis. Importantly, the second component (explaining the second most projected variance in the dataset) has an almost uniform distribution across channels, i.e. it is mixed into every channel to an equal extend. This component probably picks up activity from the mastoid-reference. The electrodes labelled in turquois (R-Post 1, R-Mid 1-2) fall inside the right Hippocampus (R-Mid 3 falls partly in the nearby white matter). Component 1 and 3 are independent sources that each peak inside the Hippocampus, with activity related to component 1 originating from the posterior HC, even though it is picked up strongly on R-Mid 2 and R-Mid 1 as well. B-C: Evoked responses and standard error (B) of component 1-3, during the first 1.5 seconds of the movie clips and event locked changes in power spectral density (PSD) during this time (C). Evoked responses were baseline corrected to -500 to -100 ms before the onset of the movie. To show changes in PSD, the power-spectra were rank transformed across 192 trials and a dependent-sample t-test was computed between the PSD rank at each time point and the average PSD rank between -500 and -100 ms. Interestingly, all components show event related changes and the profiles of the broad component and the HC-sources have very distinct properties.

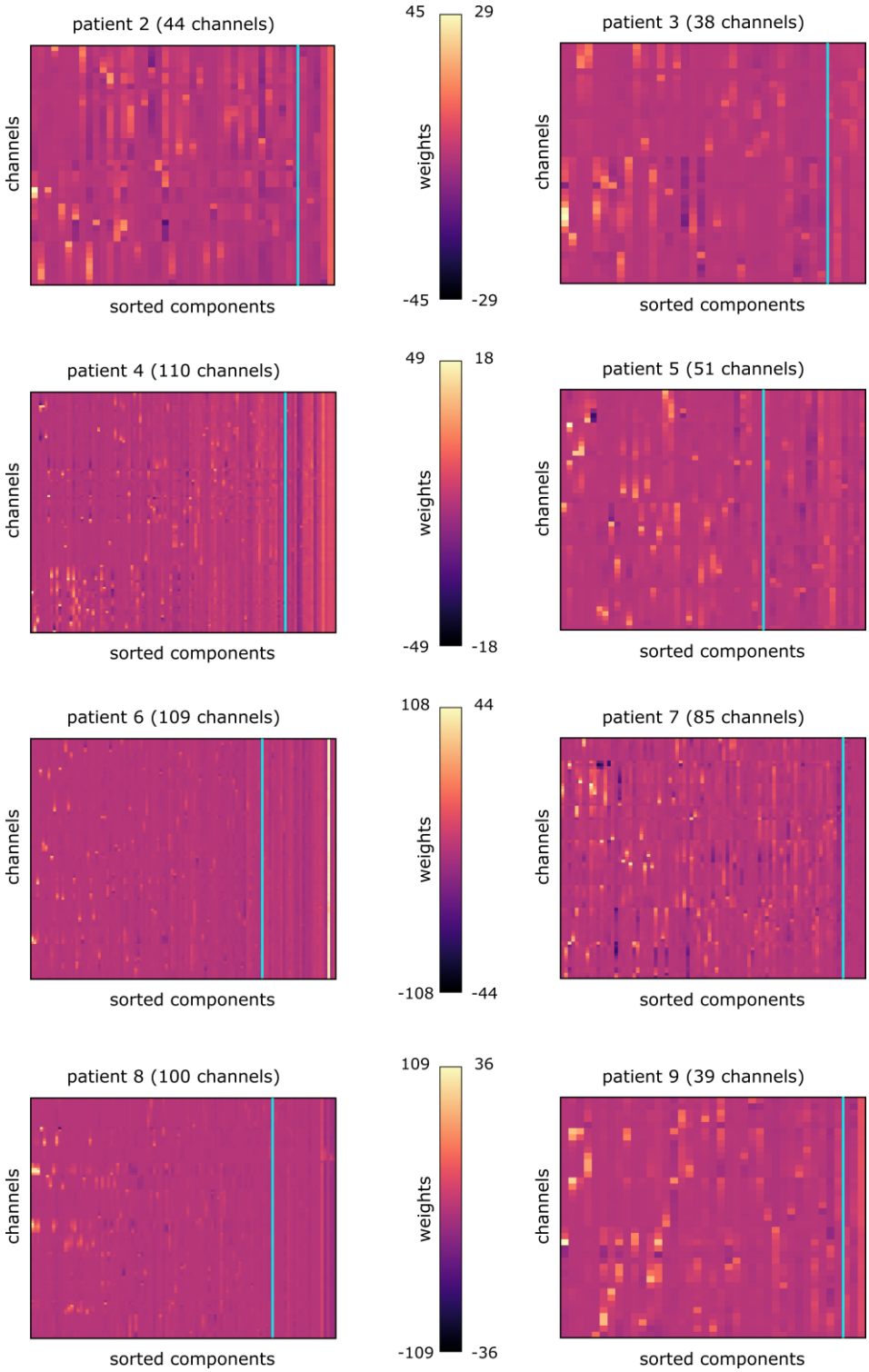


Fig. 3. Further examples of the separation between local and broad components.
ICA was computed for each of the patients and the inverse of the unmixing-matrix (i.e. the topography) was sorted by the spatial "broadness" of the components. Columns of this mixing-matrix correspond to the distribution of the components. The turquois vertical line separates local components on the left from components that have a broad distribution across channels (corresponding to a chi2- value of $p > 0.2$) on the right, which one would consider discarding from further analysis.

Simulation of the influence of spatial frequency and noise on signal recovery in ICA and bipolar reference

To quantitatively compare the abilities of ICA and local referencing in uncovering the activity of local sources, we performed simulations using real iEEG data.

For simplicity, we compared ICA to a bipolar referencing scheme using only three electrodes. However, the observed principles generalize to other setups with a larger number of electrodes.

Firstly, we defined three latent sources. To this end, 3 different channels were selected using datasets from 3 different patients. This allowed us to use real data but avoided signal spread or other unwanted statistical dependencies. Note that using real channel data in isolation will present a challenging scenario for ICA since this recorded activity from each patient is itself a superposition of other sources and will therefore likely violate the assumption of a non-gaussian distribution, which means that in a realistic scenario, ICA should perform a lot better. A total of 480,000 sampling points was now used on each channel, which (at a sampling rate of 1000 Hz) corresponds to an 8-minute recording. These three channels served as source 1, source 2 and the source which acted as underlying reference. All channels were demeaned and scaled to unit variance; the reference source was then scaled to 10 % of the strength of the other sources. This is intended to reflect a common recording scenario of intracranial EEG before re-referencing: The online reference is realized, for instance, with a scalp electrode attached to the mastoid and should pick up little activity. Next we defined a linear mixture model: Sources 1 and 2 were placed on the outer electrodes 1 and 3 (i.e. the source-channels affected these electrodes with strength of 1) and spread into the neighboring electrodes with a varying strength of $1/a$ and $1/a^2$ respectively (Fig. 4A). These parameters were chosen because the decay of the local field potential with distance from the source can be modelled with exponentials (Bedard & Destexhe, 2012; e.g. Bédard, Kröger, & Destexhe, 2004; Herreras, 2016; Teleńczuk et al., 2017). The base allows for a convenient manipulation of the "broadness" of the spread. The mixing parameter a was consequently decreased in 40 logarithmic steps from 10 to 1.02 in order to increase the "broadness" of the underlying sources, such that they affected neighboring channels to a greater extent.

Additionally, the two sources were referenced against the reference signal by subtraction. These linear mixing and referencing operations can be succinctly summarized in the following mixing matrix:

$$A = \begin{pmatrix} 1 & 1/a^2 & -1 \\ 1/a & 1/a & -1 \\ 1/a^2 & 1 & -1 \end{pmatrix} \quad (10)$$

The mixing matrix was then used to define observed data by left-multiplying it with the matrix of source signals as X = A*S, where S are the underlying sources (3 components x time points) with source 1 in the first row, source 2 in the second row and the reference source in the third row. Each row in the mixing matrix specifies how these sources are combined to form the electrode activities. For instance, inspecting row 1 of A, we see that source 1 is added with factor 1, source 2 is added in attenuated form with a factor of $1/a^2$, and the reference is subtracted by adding it with factor -1.

The ICA was now run 100 times on the observed data X for every level of the parameter *a*. Likewise, the bipolar reference was computed by subtracting the second electrode in X (i.e. the second row) from electrode 1 and from electrode 3. We then added increasing levels of pink noise to each individual observed channel in X. In this, we adjusted the variance of the noise signal in 40 linear steps from 0 to 100% of the variance of the signal-sources. Again, we repeated the ICA and the bipolar referencing 100 times for every level of *a* and every level of noise.

We defined the sensitivity of ICA to recover source-channel 1 and 2 respectively, by selecting the recovered components that had their peak weight on one of the outer electrodes; we then took the absolute correlation of each of these components with the respective underlying source signal that had been placed on that channel.

Similarly, we defined sensitivity of bipolar reference by taking the absolute correlation of the re-referenced outer electrodes with their underlying sources.

To define specificity of the recovered source, we computed the correlation with the underlying source that was placed on the opposite end of the three electrodes and used 1 – correlation as a metric. In a combined measure, we then multiplied sensitivity and specificity to account for the fact that specificity can be due to a loss of signal altogether and therefore is only informative in the presence of sensitivity. Average sensitivity, average specificity and the combined measure were compared between ICA and bipolar referencing with a dependent samples t-test of Fisher Z-transformed correlation coefficients.

Results showed that ICA performed constantly better in separating the two sources (Fig. 4B). This was most pronounced when noise levels were high and spatial mixing was high, but interestingly ICA could also recover signals better when noise levels were low and mixing was little. Importantly, a crucial result is that ICA shows constantly better specificity than bipolar referencing. This is a critical point, since the proclaimed goal of bipolar referencing is to extract activity that is local, in other words to be highly spatially specific.

In another analysis, we projected the two identified components back to a channel representation. Sensitivity was consequently defined as the absolute correlation of the outer channels after back-projection, with the underlying sources. Specificity was again defined as 1 – correlation with the opposite source. Again, the ICA solution consistently outperformed bipolar re-referencing in sensitivity and specificity, when recovering local activity (see Supplemental Fig. 1).

Finally, we compared these two approaches of working with local components (step 4 b, see above) and back-projecting local activity (step 4 a, see above) directly. When mixing between sources was high, the local components had higher sensitivity and specificity than the back-projected channels (Fig.5, top and middle row). This is not surprising, since the back-projection mixes the contribution of opposite sources back into each outer channel. Surprisingly, however, projecting the data back resulted in a slightly better performance for little mixing under low levels of noise. This may be due to slight inaccuracies in the data representation introduced by ICA unmixing that can be resolved via back-projection. Regarding the fourth step, we would therefore generally recommend to back-project the remaining data (step 4a), if multiple distributed channels are analyzed in a dataset. The back-projection yields a stable representation of the signal if mixing is little and can resolve ambiguities in interpretation. If the goal of analysis is, however, to draw differential conclusions about nearby structures, it may be best to disentangle activity via ICA and take the component representation of the data as a new synthetic (i.e. re-referenced) channel.

There are some limitations to this simulation. Firstly, we only used an 8-minute recording, whereas in practice an experimental session with a patient may last longer. However, computation on a full recording should result in an even better estimation of the ICA filters, whereas the bipolar reference operation remains the same. Secondly, we only simulated a linear mixture of three channels, whereas in practice a dataset can be formed of up to several hundred electrodes (see Supplemental Information for analyses regarding the stability of ICA decomposition with >100 channels, as a function of time).

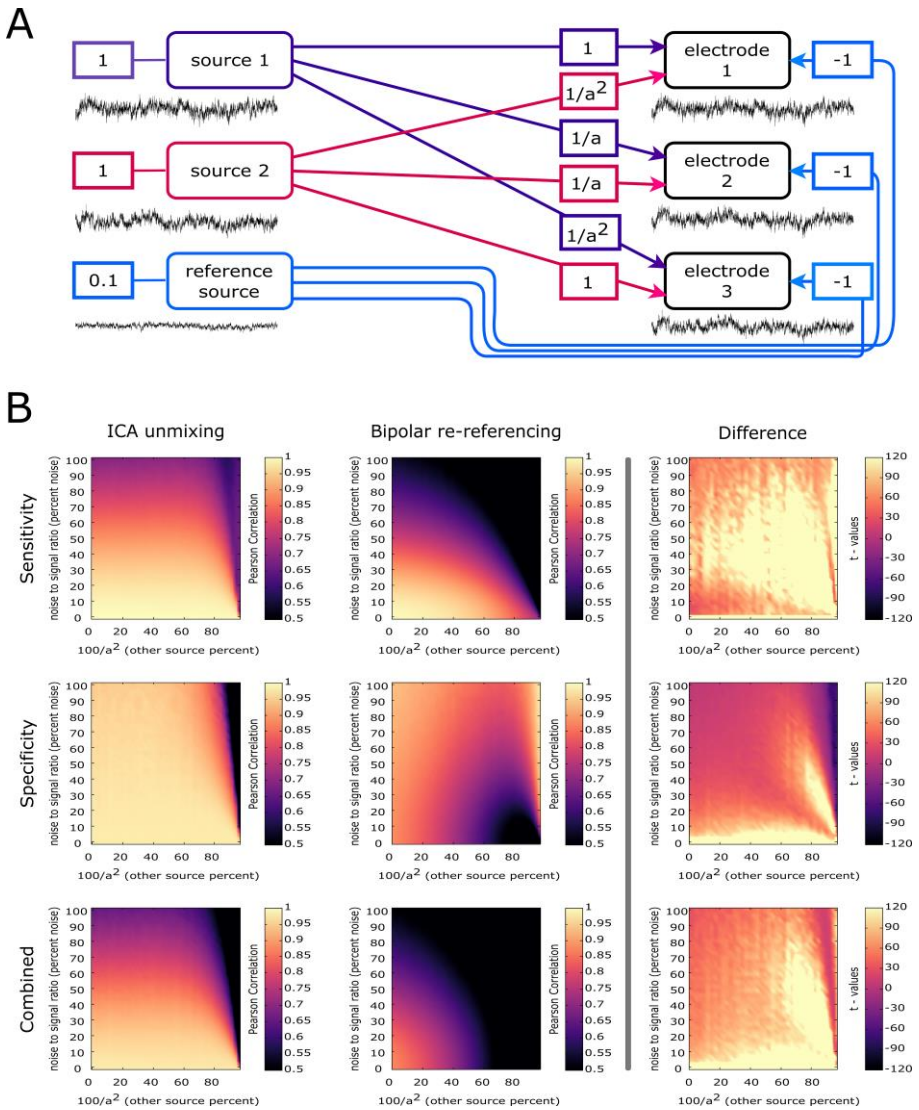


Fig. 4. Simulation of the influence of spatial frequency and noise on signal recovery, comparing ICA and bipolar reference.

A: A linear mixture model was defined in which two hidden sources and a reference mixed into three neighbouring electrodes. Each source channel was located at one end of the simulated electrode shaft and affected the neighboring electrodes with a decreasing strength of 1/a and $1/a^2$. The factor 'a' was decreased in logarithmic steps in order to modulate the "broadness" of the signal spread. B: The recovery performance was evaluated for ICA and bipolar reference for different levels of broadness. Additionally, different levels of noise were simulated by adding independent pink noise of increasing amplitude to the electrodes. In order to determine sensitivity (upper row), the ICA component that peaked on electrode 1 (electrode 3) was correlated with source 1 (source 3). Likewise for bipolar reference, the timecourse of electrode 1-2 (electrode 3-2) was correlated with source 1 (source 2). Specificity (middle row), refers to 1-correlation with the opposite source. Sensitivity and Specificity in the top and middle row are average absolute correlation averaged across both electrodes and 100 repetitions. The combined measure in the bottom row is sensitivity*specificity. The three panels on the right (right column) show the t-statistic of difference between ICA-unmixing and bipolar re-referencing acoss 100 runs. ICA showed better performance in recovering the original source (sensitivity) and importantly was more spatially specific than bipolar reference. Bipolar reference performed well when the spatial frequency of underlying sources was high and noise levels were low. The increase of bipolar reference in specificity under high "broadness" of sources and high levels of noise is due to a loss of signal altogether (i.e. noise correlations).

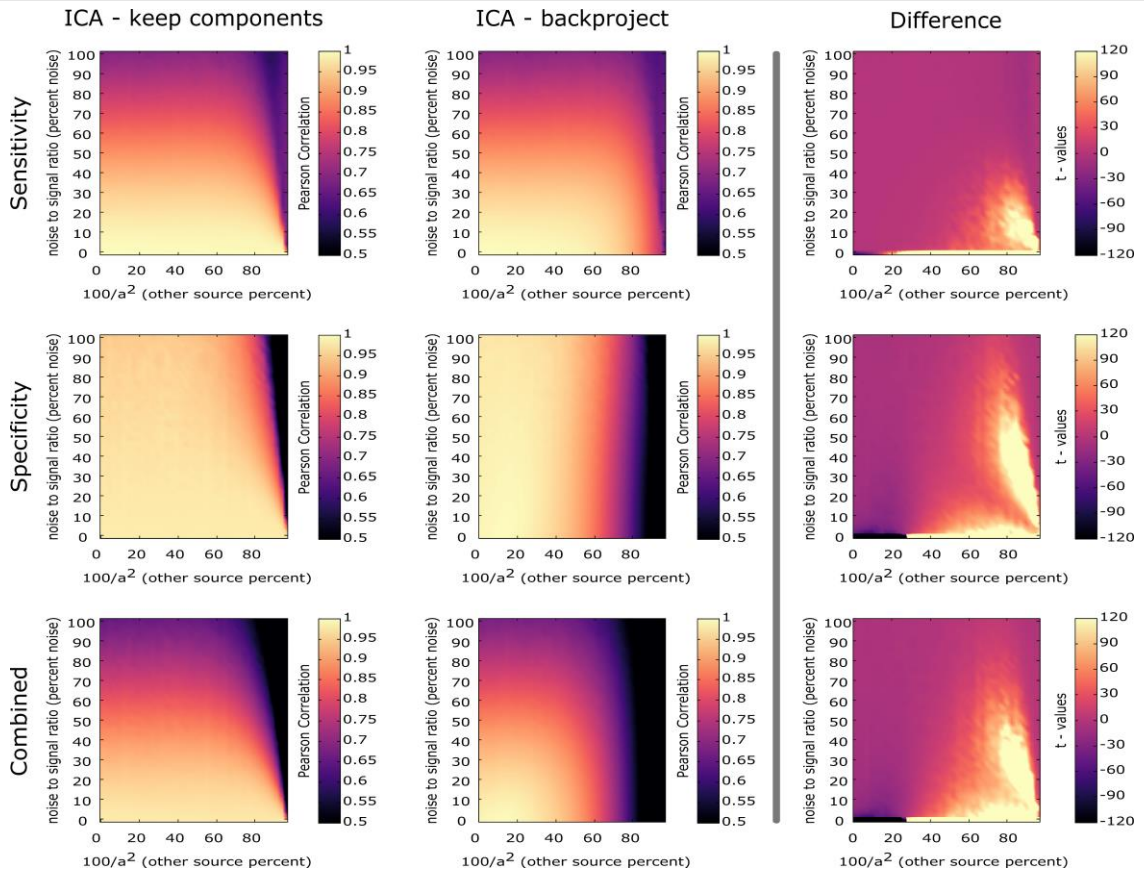


Fig. 5. Simulation of the influence of spatial frequency and noise on signal recovery, comparing ICA compontents and backprojected ICA components.

The recovery performance of ICA was evaluated under the same mixture model from Fig. 4A. Different levels of broadness and different levels of noise were simulated. In order to determine sensitivity of components (upper row, left), the ICA component that peaked on electrode 1 (electrode 3) was correlated with source 1 (source 3). Likewise to derive the sensitivity of the backprojection approach (upper row, middle) the backprojected channel 1 (channel 2) was correlated with source 1 (source 2). Specificity (middle row), refers to 1-correlation with the opposite source. Sensitivity and Specificity in the first and second row are absolute correlation, averaged across both electrodes and 100 repetitions. The two panels below (bottom row) show the combined measure of sensitivity*specificity. The three panels on the right (right column) show the t-statistic of difference acoss 100 runs. The isolated components showed better performance in recovering the original source (sensitivity) and were more spatially specific than the backprojected channels, when mixing was strong. Backprojection performed better, when mixing was little and noise levels were low.

Discussion

Fixed (re-)referencing schemes such as bipolar reference are standard solutions for preprocessing. They are powerful tools, but they address every dataset in exactly the same way. Given the large variability of S-EEG electrode spacing and location, an adaptive solution seems necessary.

We here show that a conventional ICA algorithm performs the same task that is usually addressed via re-referencing, in an elegant and intuitive way. Importantly it outperforms the widely used approach of bipolar re-referencing in extracting the local signal of interest (sensitivity) and discarding activity from distant brain regions and noise (specificity).

Even though careful preprocessing may entail the hand-picking of reference electrodes based on anatomical structure, a recent paper by Mercier et al. (Mercier et al., 2017) demonstrates that channels located in white matter contain signal from distant gray matter too; re-referencing can therefore lead to the superposition of distant sources in channels. Nevertheless, the reason why local and bipolar referencing schemes are often used in preprocessing is, to extract signal that is local. We here explicitly argue that bipolar referencing does not automatically convey spatial specificity, especially when applied without considering anatomical information and the data at hand (see also: Zaveri, Duckrow, & Spencer, 2006).

Applying ICA on the other hand is a data driven approach in which the statistical dependencies between channels are exploited. We want to promote the perspective that referencing is a spatial filtering operation and suggest using ICA, which automatically derives spatial filters that optimally isolate hidden sources.

Hence, we consider the spatial filters computed via ICA as a referencing scheme that was derived from the data. In practice, the spatial filters can then actually resemble a bipolar or local referencing scheme, meaning that ICA finds coefficients close to (1, -1) on neighboring electrodes, however the particular advantage of ICA is its adaptiveness; that is, filter coefficients are in no way restricted to prior assumptions about signal propagation in the brain. The “broadness” of the mean projected variance across channels should then be inspected for each IC time series in order to exclude (or at least separate) broad activity from the analysis. We propose to use a chi^2 test in order to quantify the broadness of components.

Certainly we champion a considerate application of ICA as a re-referencing choice. Arguably, setting a fixed criterion for broadness can be problematic since the electrode coverage and density will affect those metrics. A potential solution to this could be to incorporate information about the spread of a component in relation to the covered brain volume when deciding how broad the spatial extent of a

component is. Furthermore, anatomical information still needs consideration. A source could appear statistically broad but show a clear peak in cortical gray matter, whereas all other weights are on electrodes in the surrounding white matter; accordingly, one might consider this source as a local one. These ambiguities in the ICA-based approach should inform a considerate application of the method that mirrors the caution that can be employed other referencing approaches (e.g. hand-picking of reference channels). We want to state however that it is possible to apply these steps in a mechanical way with a fixed threshold for broadness; ultimately its shortcomings are hugely surpassed by the problematic assumptions of fixed re-referencing schemes.

Even though a complete comparison with all fixed re-referencing schemes is beyond the scope of this paper, the implications of fixed coefficients for SNR and mislocalization of activity will be similar. In the common approach of re-referencing the data with an average reference, for instance, strong activity from gray matter contacts or from noisy electrodes may be subtracted in attenuated form into otherwise silent channels. One can easily think of a pathological scenario for this approach: If half of the contacts pick up a strong theta rhythm in the hippocampus and the other half are located in cortical white matter, re-referencing will result in a near equal distribution of theta amplitude across all eight channels, i.e. no spatial specificity. In the ideal condition for an average referencing scheme, on the other hand, the channel activity sums to zero and only reference activity and broad noise will be subtracted. This condition is probably never met with intracranial recordings. ICA can therefore similarly address the limitations of this approach.

The ICA based approach can also help to deal with recordings that are separated by hours or days. In general, the mixing of channels may change when, for example, electrodes move (especially the reference) or when the noise in the patient's room changes. This will generate a problem for all referencing schemes, since the filter coefficients remain the same, whereas the composition of the data changes. If the goal of re-referencing is merely to exclude broad activity from the recording, then a separate ICA for each session will provide a better solution, since coefficients adapt to the mixture of channels within each session. If the goal is, however, to keep the coefficients between two separate sessions consistent, then ICA can be computed on the concatenated data from both sessions.

An additional advantage of ICA is that it can help to identify which signal an electrode is picking up (compare Fig. 2). An electrode that is located between two neighboring structures could pick up the same source as other electrodes that are located in only one of the structures. This way, the statistical dependencies between channels can be informative of electrode location, when the anatomical information is ambiguous.

A general criticism, targeting ICA itself, argues that there are more sources than electrodes in a dataset. This has been addressed in the EEG community with overcomplete ICA models (Grosse-Wentrup & Buss, 2007; Lee, Lewicki, Girolami, & Sejnowski, 1999) and could potentially be extended to work in a referencing context. Importantly however, we here show that a standard ICA algorithm that is easily available and has been established in the EEG community (Delorme & Makeig, 2004), outperforms bipolar referencing in the very task that justifies the application of re-referencing in the first place; that is, extracting local signal.

Another broad criticism about ICA in general is that components may not always be local and could be hard to interpret. This may be the case when distant sources are correlated because there is communication between the regions. In practice, the filters would still extract time courses that capture the linear dependencies between the distant sources. Either way, if the components are of interest, projecting them back will resolve ambiguities in their interpretation and bring the researcher back to a familiar channel representation of the data that have been cleansed of broad activity; another potential solution to this might be to apply ICA separately to different electrode shafts.

Researchers who want to apply connectivity analysis (Greenblatt, Pflieger, & Ossadtchi, 2012) to their data might be worried about the implications of ICA for their analysis. The general problem is that ICA decomposition aims at reducing statistical dependencies between channels. If one accepts that independent components reflect the activity of separate generators, it is still possible to assess connectivity between these components by assessing, for instance, statistical relationships in band-specific power (e.g. Chen, Ros, & Gruzelier, 2013). Yet, it is also possible to simply exclude broad components and back-project the remaining data to a channel-representation. This will reduce spurious correlations between channels that are due to shared noise or reference activity and will therefore benefit subsequent analyses of connectivity.

Regardless of the difficulties one can experience when interpreting independent components, it is always possible to inspect how a linear superposition of time courses mixes into the observed data at hand. Together with knowledge about the anatomical locations of electrode contacts and the type of surrounding tissue, ICA therefore allows for an informed decision about anatomical sources of recorded activity. Filtering the data with coefficients that were derived via ICA can therefore be a crucial advantage in determining anatomical sources of effects and certainly it is a more informative preprocessing choice than the application of a fixed predetermined re-referencing scheme which does not account for the complex dependencies that may exist between channels.

Acknowledgements

S.H. is supported by a grant from the European Research Council (Consolidator Grant Agreement 647954) and further supported by the Wolfson Society and the Royal Society.

B.S. is funded by the Wellcome Trust/Royal Society Sir Henry Dale Fellowship (107672/Z/15/Z).

M.W. is supported by the ERC Starting Grant STREAM (ERC-2016-STG-715714).

C.K. is funded by Stiftelsen Olle Engkvist Byggmästare.

References

Avanzini, P., Abdollahi, R. O., Sartori, I., Caruana, F., Pelliccia, V., Casaceli, G., … Orban, G. A. (2016). Four-dimensional maps of the human somatosensory system. *Proceedings of the National Academy of Sciences*, *113*(13), E1936–E1943. http://doi.org/10.1073/pnas.1601889113

Avanzini, P., Pelliccia, V., Lo Russo, G., Orban, G. A., & Rizzolatti, G. (2018). Multiple time courses of somatosensory responses in human cortex. *NeuroImage*, *169*, 212–226. http://doi.org/10.1016/j.neuroimage.2017.12.037

Bedard, C., & Destexhe, A. (2012). *Modeling local field potentials and their interaction with the extracellular medium*. *Handbook of Neural Activity Measurement*.

Bédard, C., Kröger, H., & Destexhe, A. (2004). Modeling Extracellular Field Potentials and the Frequency-Filtering Properties of Extracellular Space. *Biophysical Journal*, *86*(3), 1829–1842. http://doi.org/10.1016/S0006-3495(04)74250-2

Bleichner, M. G., & Debener, S. (2017). Concealed, Unobtrusive Ear-Centered EEG Acquisition: cEEGrids for Transparent EEG. *Frontiers in Human Neuroscience*, *11*. http://doi.org/10.3389/fnhum.2017.00163

Chen, J.-L., Ros, T., & Gruzelier, J. H. (2013). Dynamic changes of ICA-derived EEG functional connectivity in the resting state. *Human Brain Mapping*, *34*(4), 852–868. http://doi.org/10.1002/hbm.21475

Comon, P. (1994). Independent component analysis, A new concept? *Signal Processing*, *36*(3), 287–314. http://doi.org/10.1016/0165-1684(94)90029-9

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. http://doi.org/10.1016/j.jneumeth.2003.10.009

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. http://doi.org/10.1016/j.neuroimage.2006.01.021

Fischl, B. (2012). FreeSurfer. *NeuroImage*. http://doi.org/10.1016/j.neuroimage.2012.01.021

Flink, R., Pedersen, B., Guekht, A. B., Malmgren, K., Michelucci, R., Neville, B., … Özkara, C. (2002). Guidelines for the use of EEG methodology in the diagnosis of epilepsy. *Acta Neurologica Scandinavica*, *106*(1), 1–7. http://doi.org/10.1034/j.1600-0404.2002.01361.x

Greenblatt, R. E., Pflieger, M. E., & Ossadtchi, A. E. (2012). Connectivity measures applied to human brain electrophysiological data. *Journal of Neuroscience Methods*, *207*(1), 1–16. http://doi.org/10.1016/j.jneumeth.2012.02.025

Grosse-Wentrup, M., & Buss, M. (2007). Overcomplete Independent Component Analysis via Linearly Constrained Minimum Variance Spatial Filtering. *The Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, *48*(1–2), 161–171. http://doi.org/10.1007/s11265-006-0028-3

Herreras, O. (2016). Local Field Potentials: Myths and Misunderstandings. *Frontiers in Neural Circuits*, *10*. http://doi.org/10.3389/fncir.2016.00101

Hu, S., Stead, M., & Worrell, G. A. (2007). Automatic identification and removal of scalp reference signal for intracranial eegs based on independent component analysis. *IEEE Transactions on*

Biomedical Engineering, *54*(9), 1560–1572. http://doi.org/10.1109/TBME.2007.892929

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks*, *13*(4–5), 411–430. http://doi.org/10.1016/S0893-6080(00)00026-5

Jokisch, D., & Jensen, O. (2007). Modulation of Gamma and Alpha Activity during a Working Memory Task Engaging the Dorsal or Ventral Stream. *The Journal of Neuroscience*, *27*(12), 3244 LP-3251. Retrieved from http://www.jneurosci.org/content/27/12/3244.abstract

Kajikawa, Y., & Schroeder, C. E. (2011). How local is the local field potential? *Neuron*, *72*(5), 847–858. http://doi.org/10.1016/j.neuron.2011.09.029

Kullback, S., & Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86. http://doi.org/10.1214/aoms/1177729694

Lachaux, J.-P., Axmacher, N., Mormann, F., Halgren, E., & Crone, N. E. (2012). High-frequency neural activity and human cognition: Past, present and possible future of intracranial EEG research. *Progress in Neurobiology*, *98*(3), 279–301. http://doi.org/10.1016/j.pneurobio.2012.06.008

Lachaux, J. P., Rudrauf, D., & Kahane, P. (2003). Intracranial EEG and human brain mapping. *Journal of Physiology Paris*, *97*(4–6), 613–628. http://doi.org/10.1016/j.jphysparis.2004.01.018

Lee, T. W., Lewicki, M. S., Girolami, M., & Sejnowski, T. J. (1999). Blind source separation of more sources than mixtures using overcomplete representations. *IEEE Signal Processing Letters*, *6*(4), 87–90. http://doi.org/10.1109/97.752062

Liu, Y., Coon, W. G., De Pesters, A., Brunner, P., & Schalk, G. (2015). The effects of spatial filtering and artifacts on electrocorticographic signals. *Journal of Neural Engineering*, *12*(5). http://doi.org/10.1088/1741-2560/12/5/056008

Ludwig, K. A., Miriani, R. M., Langhals, N. B., Joseph, M. D., Anderson, D. J., & Kipke, D. R. (2009). Using a Common Average Reference to Improve Cortical Neuron Recordings From Microelectrode Arrays. *Journal of Neurophysiology*, *101*(3), 1679–1689. http://doi.org/10.1152/jn.90989.2008

Mercier, M. R., Bickel, S., Megevand, P., Groppe, D. M., Schroeder, C. E., Mehta, A. D., & Lado, F. A. (2017). Evaluation of cortical local field potential diffusion in stereotactic electro-encephalography recordings: A glimpse on white matter signal. *NeuroImage*, *147*, 219–232. http://doi.org/10.1016/j.neuroimage.2016.08.037

Nunez, P. L., & Srinivasan, R. (2006). *Electric Fields of the Brain*. Oxford University Press. http://doi.org/10.1093/acprof:oso/9780195050387.001.0001

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9. http://doi.org/10.1155/2011/156869

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, *53*(4), 1181–1196. http://doi.org/10.1016/j.neuroimage.2010.07.020

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, *61*(4), 1402–1418. http://doi.org/10.1016/j.neuroimage.2012.02.084

Rodgers, J. L., Nicewander, W. A., & Toothaker, L. (1984). Linearly Independent, Orthogonal, and Uncorrelated Variables. *The American Statistician*, *38*(2), 133. http://doi.org/10.2307/2683250

Shirhatti, V., Borthakur, A., & Ray, S. (2016). Effect of Reference Scheme on Power and Phase of the Local Field Potential. *Neural Computation*, *28*(5), 882–913.

http://doi.org/10.1162/NECO_a_00827

Staresina, B. P., Michelmann, S., Bonnefond, M., Jensen, O., Axmacher, N., & Fell, J. (2016). Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection. *ELife*, *5*(AUGUST). http://doi.org/10.7554/eLife.17397.001

Staudigl, T., Vollmar, C., Noachtar, S., & Hanslmayr, S. (2015). Temporal-Pattern Similarity Analysis Reveals the Beneficial and Detrimental Effects of Context Reinstatement on Human Memory. *Journal of Neuroscience*, *35*(13), 5373–5384. http://doi.org/10.1523/JNEUROSCI.4198-14.2015

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: A User-Friendly Application for MEG/EEG Analysis. *Computational Intelligence and Neuroscience*, *2011*, 1–13. http://doi.org/10.1155/2011/879716

Tallon-Baudry, C., Bertrand, O., & Fischer, C. (2001). Oscillatory synchrony between human extrastriate areas during visual short-term memory maintenance. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(20), RC177. http://doi.org/20015744 [pii]

Teleńczuk, B., Dehghani, N., Le Van Quyen, M., Cash, S. S., Halgren, E., Hatsopoulos, N. G., & Destexhe, A. (2017). Local field potentials primarily reflect inhibitory neuron activity in human and monkey cortex. *Scientific Reports*, *7*, 40211. http://doi.org/10.1038/srep40211

Tenke, C. E., & Kayser, J. (2012). Generator localization by current source density (CSD): Implications of volume conduction and field closure at intracranial and scalp resolutions. *Clinical Neurophysiology*. http://doi.org/10.1016/j.clinph.2012.06.005

Trongnetrpunya, A., Nandi, B., Kang, D., Kocsis, B., Schroeder, C. E., & Ding, M. (2015). Assessing Granger Causality in Electrophysiological Data: Removing the Adverse Effects of Common Signals via Bipolar Derivations. *Frontiers in Systems Neuroscience*, *9*(January), 189. http://doi.org/10.3389/fnsys.2015.00189

Whitmer, D., Worrell, G., Stead, M., Lee, I. K., & Makeig, S. (2010). Utility of Independent Component Analysis for Interpretation of Intracranial EEG. *Frontiers in Human Neuroscience*, *4*. http://doi.org/10.3389/fnhum.2010.00184

Whitmore, N. W., & Lin, S.-C. (2016). Unmasking local activity within local field potentials (LFPs) by removing distal electrical signals using independent component analysis. *NeuroImage*, *132*, 79–92. http://doi.org/10.1016/j.neuroimage.2016.02.032

Yao, D. (2001). A method to standardize a reference of scalp EEG recordings to a point at infinity. *Physiological Measurement*, *22*(4), 693–711. http://doi.org/10.1088/0967-3334/22/4/305

Zaveri, H. P., Duckrow, R. B., & Spencer, S. S. (2006). On the use of bipolar montages for time-series analysis of intracranial electroencephalograms. *Clinical Neurophysiology*, *117*(9), 2102–2108. http://doi.org/10.1016/j.clinph.2006.05.032

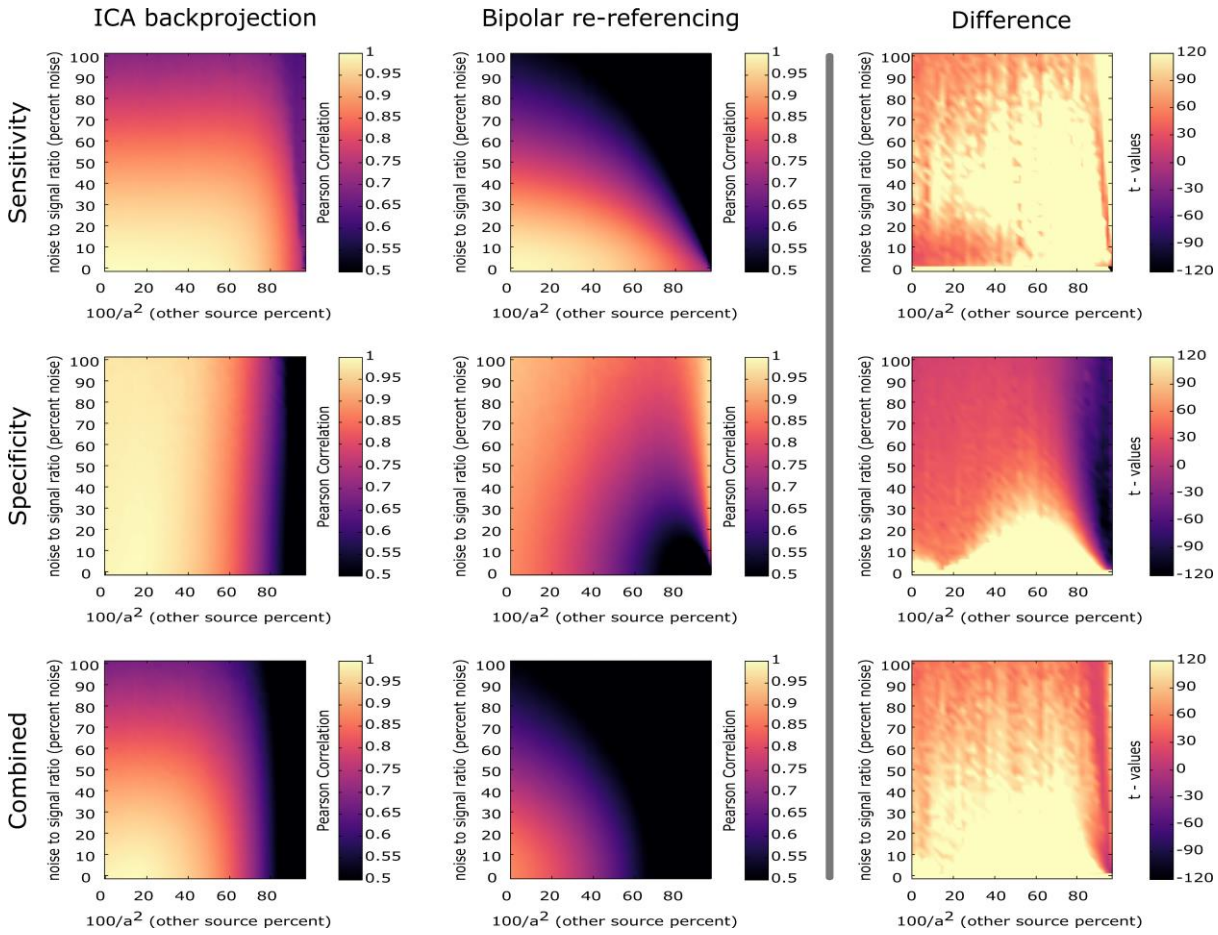
Supplemental Information

Effect of the “broadness” criterion

To address the issue of an arbitrary criterion of what constitutes a spatially “broad” component, we assessed the amount of excluded components for various thresholds of “broadness”. I.e. we used a chi^2 value that corresponds to an alpha of 0.25, 0.2, 0.15, 0.1 and 0.05. Additionally we assessed how much the mean projected variance (MPV) of the data was reduced by excluding the corresponding components. The absolute amount of excluded variance differed substantially between subjects, due to different levels of noise; the amount of excluded components and MPV, however, was fairly invariant to the change of criterion (see Supplemental Tab. 1).

Chi^2 criterion	**5%**	**10%**	**15%**	**20%**	**25%**
MPV reduction					
patient 1	37.69	35.16	33.75	33.75	33.75
patient 2	23.52	23.52	23.52	23.52	23.52
patient 3	9.02	6.66	6.25	5.73	5.26
patient 4	43.34	43.34	43.34	43.34	43.34
patient 5	20.11	19.23	18.27	17.98	17.98
patient 6	89.52	89.45	89.43	89.43	89.43
patient 7	2.12	0.35	0.35	0.35	0.35
patient 8	17.94	17.57	17.57	17.57	17.57
patient 9	9.38	9.38	9.38	9.38	9.38
N components rejected					
patient 1	14	12	11	11	11
patient 2	5	5	5	5	5
patient 3	10	7	6	5	4
patient 4	28	28	28	28	28
patient 5	20	19	18	17	17
patient 6	28	27	26	26	26
patient 7	8	6	6	6	6
patient 8	23	20	20	20	20
patient 9	3	3	3	3	3

Supplemental Tab. 1. Effect of "broadness" criterion on the amount of data rejected.
Effect of various thresholds of broadness on the reduction of mean projected variance (MPV) of the data (top) and on the amount of components rejected for each patient (bottom).

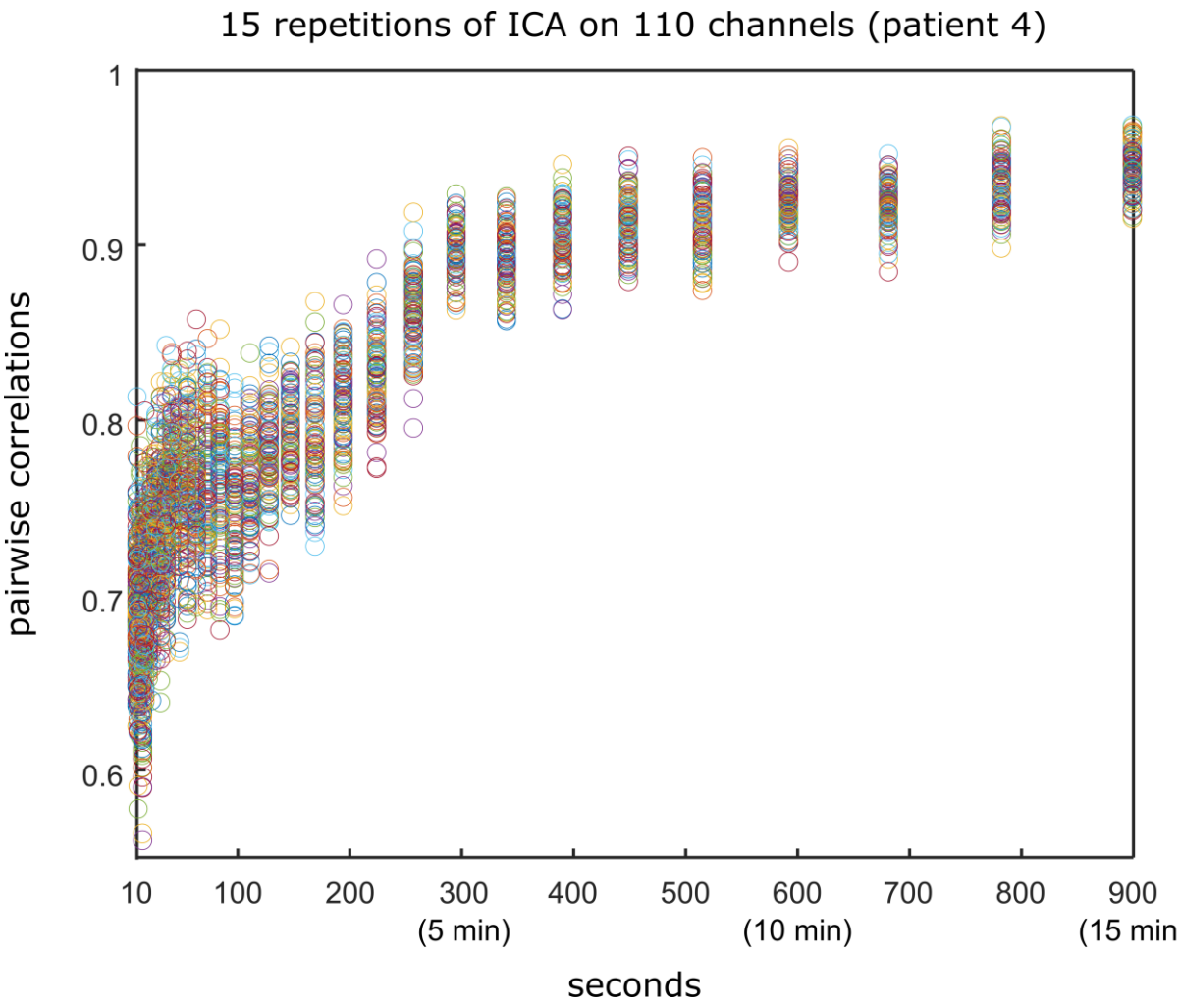


Supplemental Fig. 1. Simulation of the influence of spatial frequency and noise on signal recovery, comparing backprojected ICA-components and bipolar reference.

The recovery performance to the linear mixture model from Fig. 4A was evaluated for ICA and bipolar reference for different levels of broadness. Additionally, different levels of noise were simulated by adding independent pink noise of increasing amplitude to the electrodes. After computation of ICA, the components that were localized to one of the outer electrodes were backprojected. In order to determine sensitivity of the ICA-backprojection approach (upper row) the backprojected channel 1 (channel2) was correlated with source 1 (source 2). Specificity (middle row), refers to 1-correlation with the opposite source. Likewise for bipolar reference, the timecourse of electrode 1-2 (electrode 3-2) was correlated with source 1 (source 2). Specificity (middle row), refers to 1-correlation with the opposite source. Sensitivity and Specificity in the first and second row are absolute correlation, averaged across both electrodes and 100 repetitions. The two panels below (bottom row) show the combined measure of sensitivity*specificity. The three panels on the right (right column) show the t-statistic of difference acoss 100 runs. ICA showed better performance in recovering the original source (sensitivity) and importantly was more spatially specific than bipolar reference. Bipolar reference performed well when the spatial frequency of underlying sources was high and noise levels were low. The increase of bipolar reference in specificity under high “broadness” of sources and high levels of noise is due to a loss of signal altogether (i.e. noise correlations).

Stability of ICA as a function of time in a large dataset

An important consideration for applying ICA to intracranial data is how much data is needed for a stable solution of the ICA filter coefficients. In order to estimate when the broad components become stable, we repeated ICA with different random seeds on a large dataset (i.e. with 110 electrodes, patient 4). We excluded the broad components (chi^2 value past a threshold of $p = 0.2$), and projected the remaining data back to a channel representation. We repeated this process 15 times for increasingly longer time-segments. Segments started with a length of 10s and were increased in logarithmic steps until a length of 15 minutes was reached. All pairwise correlation between the 15 solutions, are shown in Supplemental Fig. 2. It is visible that the solution stabilizes as a function of time and that using less than 5 minutes of clean data should best be avoided. Importantly, this is only a general orientation. In practice the stability of the solution not only depends on the amount of observations and the amount of channels but also on the amount of noise in the data and the stability of signal and noise over time. If the researcher is in doubt about the stability of their own ICA solution, pairwise correlations between the backprojected remaining data under different random seeds can provide an orientation.



Supplemental Fig. 2. Stability of ICA as a function of time in a large dataset
ICA was repeated 15 times for increasingly longer time segments. Broad components were rejected and the remaining data were projected back to a channel representation. All pairwise correlations between backprojected datasets are plotted as a function of the amount of data used for computation. It can be seen that ICA stabilizes as a function of time. Correlations also suggest that the use of less than 5 minutes of clean data should be avoided for ICA computation on intracranial EEG.

Figure legends

Fig. 1. Example of how bipolar re-referencing can impair the interpretability of the data.

A: Activity during a single trial on 3 selected channels in the medial temporal lobe of a patient implanted with intracranial electrodes. Channels 1 and 2 fall in the Hippocampus, channel 3 is located in the nearby white matter. B: The same channels were re-referenced with a bipolar scheme. The first channel appears inverted and the shared activity between channel 2 and 3 is almost cancelled out. C: The same trial using an ICA decomposition of the data. Inspecting the weights in the columns of the 3x3 mixing matrix (i.e. the topography), we can see how the hidden components mix into the channels observed in A. Instead of subtracting the channels from each other and therefore mixing them even more, ICA splits the 3 channels into activity that is unique to channel 1 (component 1), activity that is shared between channel 1 and 2 (component 2) and activity that is shared between channel 2 and 3 (component 3). Interestingly component 2 is mostly present in channel 1, however it appears strongest (and inverted) on the combined bipolar channel 3-2. This is because the difference between its contribution to channel 1 and to channel 2 is smaller than the difference between its contribution to 2 and to 3, which can be observed in C.

Fig. 2. Example of ICA-components and their spatial distribution in one patient with intracranial electrodes.

The patient was performing a memory task in which they repeatedly associated short video clips with words and recalled them later. A: two out of three used electrode shafts are depicted on the left. The brain is superimposed by labels corresponding to the Desikan-Killiany Atlas, which were derived via Freesurfer parcellation. R-Mid 1-6 and R-Post 1-6 are located in the right medial and right posterior temporal lobe respectively (counting from inside to outside). ICA was computed and the inverse of the unmixing-matrix (i.e. the topography) was sorted by the spatial "broadness" of the components. Columns of this mixing-matrix (right) correspond to the distribution of the components. The turquois vertical line separates local components on the left from components that have a broad distribution across channels (corresponding to a chi^2- value of $p > 0.2$), which one could consider discarding from further analysis. Importantly, the second component (explaining the second most projected variance in the dataset) has an almost uniform distribution across channels, i.e. it is mixed into every channel to an equal extend. This component probably picks up activity from the mastoid-reference. The electrodes labelled in turquois (R-Post 1, R-Mid 1-2) fall inside the right Hippocampus (R-Mid 3 falls partly in the nearby white matter). Component 1 and 3 are independent sources that each peak inside the Hippocampus, with activity related to component 1 originating from the posterior HC, even though it is picked up strongly on R-Mid 2 and R-Mid 1 as well. B-C: Evoked responses and standard error (B) of component 1-3, during the first 1.5 seconds of the movie clips and event locked changes in power spectral density (PSD) during this time (C). Evoked responses were baseline corrected to -500 to -100 ms before the onset of the movie. To show changes in PSD, the power-spectra were rank transformed across 192 trials and a dependent-sample t-test was computed between the PSD rank at each time point and the average PSD rank between -500 and -100 ms. Interestingly, all components show event related changes and the profiles of the broad component and the HC-sources have very distinct properties.

Fig. 3. Further examples of the separation between local and broad components.

ICA was computed for each of the patients and the inverse of the unmixing-matrix (i.e. the topography) was sorted by the spatial “broadness” of the components. Columns of this mixing-matrix (right) correspond to the distribution of the components. The turquois vertical line separates local components on the left from components that have a broad distribution across channels (corresponding to a chi2- value of $p > 0.2$), which one would consider discarding from further analysis.

Fig. 4. Simulation of the influence of spatial frequency and noise on signal recovery, comparing ICA and bipolar reference.

A: A linear mixture model was defined in which two hidden sources and a reference mixed into three neighbouring electrodes. Each source channel was located at one end of the simulated electrode shaft and affected the neighboring electrodes with a decreasing strength of 1/a and $1/a^2$. The factor 'a' was decreased in logarithmic steps in order to modulate the "broadness" of the signal spread. B: The recovery performance was evaluated for ICA and bipolar reference for different levels of broadness. Additionally, different levels of noise were simulated by adding independent pink noise of increasing amplitude to the electrodes. In order to determine sensitivity (upper row), the ICA component that peaked on electrode 1 (electrode 3) was correlated with source 1 (source 3). Likewise for bipolar reference, the timecourse of electrode 1-2 (electrode 3-2) was correlated with source 1 (source 2). Specificity (middle row), refers to 1-correlation with the opposite source. Sensitivity and Specificity in the top and middle row are average absolute correlation averaged across both electrodes and 100 repetitions. The combined measure in the bottom row is sensitivity*specificity. The three panels on the right (right column) show the t-statistic of difference between ICA-unmixing and bipolar re-referencing acoss 100 runs. ICA showed better performance in recovering the original source (sensitivity) and importantly was more spatially specific than bipolar reference. Bipolar reference performed well when the spatial frequency of underlying sources was high and noise levels were low. The increase of bipolar reference in specificity under high "broadness" of sources and high levels of noise is due to a loss of signal altogether (i.e. noise correlations).

Fig. 5. Simulation of the influence of spatial frequency and noise on signal recovery, comparing ICA compontents and backprojected ICA components.

The recovery performance of ICA was evaluated under the same mixture model from Fig. 4. Different levels of broadness and different levels of noise were simulated. In order to determine sensitivity of components (upper row, left), the ICA component that peaked on electrode 1 (electrode 3) was correlated with source 1 (source 3). Likewise to derive the sensitivity of the backprojection approach (upper row, middle) the backprojected channel 1 (channel2) was correlated with source 1 (source 2). Specificity (middle row), refers to 1-correlation with the opposite source. Sensitivity and Specificity in the first and second row are absolute correlation, averaged across both electrodes and 100 repetitions. The two panels below (bottom row) show the combined measure of sensitivity*specificity. The three panels on the right (right column) show the t-statistic of difference acoss 100 runs. The isolated components showed better performance in recovering the original source (sensitivity) and were more spatially specific than the backprojected channels, when mixing was strong. Backprojection performed better, when mixing was little and noise levels were low.

Supplemental Fig. 1. Simulation of the influence of spatial frequency and noise on signal recovery, comparing backprojected ICA-components and bipolar reference.

The recovery performance to the linear mixture model from Fig. 4A was evaluated for ICA and bipolar reference for different levels of broadness. Additionally, different levels of noise were simulated by adding independent pink noise of increasing amplitude to the electrodes. After computation of ICA, the components that were localized to one of the outer electrodes were backprojected. In order to determine sensitivity of the ICA-backprojection approach (upper row) the backprojected channel 1 (channel2) was correlated with source 1 (source 2). Specificity (middle row), refers to 1-correlation with the opposite source. Likewise for bipolar reference, the timecourse of electrode 1-2 (electrode 3-2) was correlated with source 1 (source 2). Specificity (middle row), refers to 1-correlation with the opposite source. Sensitivity and Specificity in the first and second row are absolute correlation, averaged across both electrodes and 100 repetitions. The two panels below (bottom row) show the combined measure of sensitivity*specificity. The three panels on the right (right column) show the t-statistic of difference acoss 100 runs. ICA showed better performance in recovering the original source (sensitivity) and importantly was more spatially specific than bipolar reference. Bipolar reference performed well when the spatial frequency of underlying sources was high and noise levels were low. The increase of bipolar reference in specificity under high “broadness” of sources and high levels of noise is due to a loss of signal altogether (i.e. noise correlations).

Supplemental Fig. 2. Stability of ICA as a function of time in a large dataset.

ICA was repeated 15 times for increasingly longer time segments. Broad components were rejected and the remaining data were projected back to a channel representation. All pairwise correlations between backprojected datasets are plotted as a function of the amount of data used for computation. It can be seen that ICA stabilizes as a function of time. Correlations also suggest that the use of less than 5 minutes of clean data should be avoided for ICA computation on intracranial EEG.