Improved rejection of artifacts from EEG data using high-order statistics and independent component analysis

Arnaud Delorme[1,2], Tzyy-Ping Jung[2], Terrence Sejnowski[1], Scott Makeig[2]

[1] Computational Neurobiology Laboratory, Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92107 USA
[2] Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego
9500 Gilman Drive, La Jolla, CA 92093-0961 USA

Abstract-**While it is now generally accepted that independent component analysis (ICA) is a good tool for isolating both artifacts and cognition-related processes in EEG data, there is little definite proof that data preprocessed using ICA is more effective than artifact rejection on raw channel data, especially when more subtle signal processing methods are used to detect artifacts. Here we applied five statistical signal processing methods for detecting artifactual data epochs from either the raw data containing simulated artifacts or from the ICA decomposition of these data, and tested their performance for different sizes of introduced artifacts. The most efficient rejection method used threshold limits applied to the single trial data spectra. We show that for this or other methods ICA preprocessing can improve the detection of data epochs containing eye, muscle, and electrical artifacts by 10-20%.**

I. INTRODUCTION

In event-related experiments, each data epoch normally represents a single experimental trial time locked to one or more experimental events of interest. Usually, EEG software first subtracts a baseline – e.g., the average potential before the stimulus occurs – from each trial, then finds and eliminates electrodes at which potential values exceed some defined threshold. The retained electrodes usually include central scalp placements which may contain record brain activity, parietal placements that may contain temporal muscle artifacts and frontal electrodes that may contain blinks and eye movement artifacts. It is critical to detect such artifact contaminating event-related EEG data for several reasons. First, artifactual signals often have high amplitudes relative to brain signals. Thus, even if their distribution in the recorded EEG is sparse, they can bias evoked potential or other averages constructed from the data and, as a consequence, bias results of an experiment.

In most current EEG software packages, single data trials that contain out-of-bounds potential values at single electrodes are rejected from analysis. A problem with the simple thresholding criterion is that it only takes into account low-order signal statistics (minimum and maximum). For example, this type of metric may not detect muscle activity, which typically involves rapid electromyographic (EMG) signals of small to moderate size, nor will it detect artifacts generated by small eye movements. Higher order statistical properties of the EEG signals might contain more relevant information about these and other types of artifacts.

Independent Component Analysis (ICA) [1-3] applied to a concatenated collection of single-trial EEG data has proven to be efficient for separating distinct artifactual processes, ranging from eye artifacts to muscle and electrical artifacts [4-12]. Although different ICA algorithms in different implementations have been used to detect artifacts in EEG and MEG data, they all are derived using related mathematical principles [13]. While ICA is now considered an important technique for removing artifacts, there is still little quantitative results of what advantage for artifact detection results from ICA decomposition.

Here we develop a framework for comparing between rejection methods and determine if preprocessing EEG data using ICA can help in detecting artifactual data epochs. We first apply a set of statistical and spectral analysis methods to detect artifacts in the raw data, and optimized a free parameter for each rejection method to detect optimally known artifactual data epochs. Then, we applied the same procedure to the data decomposed using ICA. Finally, we quantitatively compared rejection results of all rejection methods applied either to the EEG data or to the data preprocessed using ICA.

II. METHODS FOR ARTIFACT REJECTION

We applied five different methods for detecting trials containing artifacts [5, 11].

Extreme values.. First, we used standard thresholding of the raw EEG data. Here, data trials were labeled as artifactual if the absolute value of any data point in the trial exceeded a fixed threshold. This method is currently the most widely used artifact detection method in the EEG community, and is mostly effective for detecting gross eye blinks or eye movement artifacts.

Linear trends. Marked linear trends at one electrode typically indicate transient recording-induced current drifts. To detect such events, we measured the goodness of fit of EEG activity to an oblique straight line within a sliding time window. We then either accepted or rejected the data trial depending on the
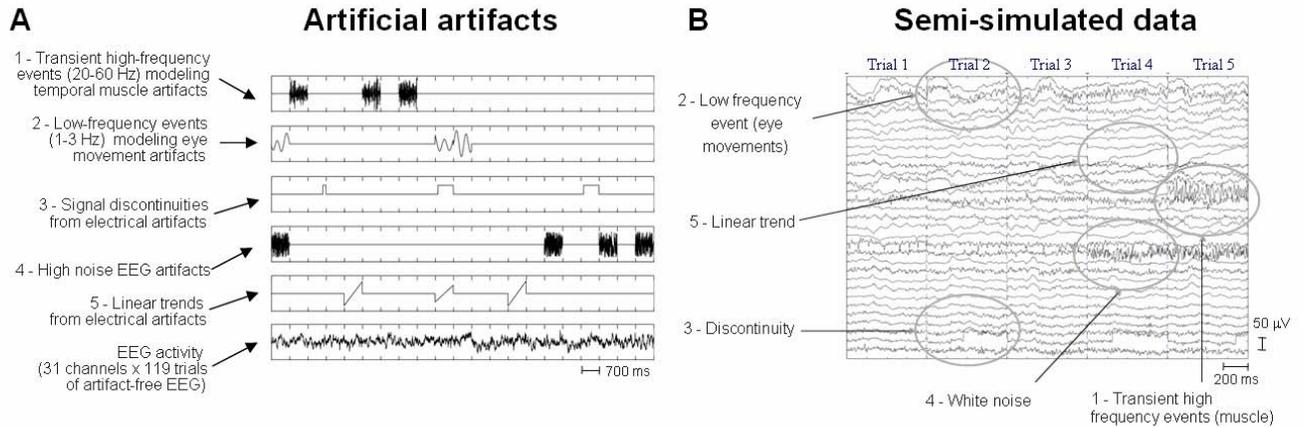
Fig. 1. A. Types of artificial artifacts introduced into actual EEG data. Data epochs are concatenated in this representation and graduations indicate epoch boundaries every 700 ms. Since artifact amplitude varies, plotting limits for each curve are arbitrary (except for the background EEG where the plotting range is 100 μV). B. Semi-simulated data obtained by adding simulated artifacts shown in A to actual and apparently artifact-free EEG data. Plotted artifacts correspond approximately to 0 dB signal/noise.

minimum slope of this straight line and its goodness to fit (in terms of r^2).

Improbability. Most artifacts have "unusual" time courses, e.g., they appear as transient 'odd' or unexpected events, and may be so identified by their outlying statistics. Here we used the distribution of data values and its kurtosis to detect such artifacts. To estimate the relative probability of each trial from the raw data, we first computed the observed probability density function (D_e) of data values for each electrode e. Then, we computed the joint log probability $J_e(i)$ of the activity values (A_i) in each data trial i at electrode e by

$$J_e(i) = -\log(\prod_{x \in A_i} p_{D_e}(x))$$ (1)

where $p_{D_e}(x)$ is the probability of observing the value x in the probability distribution D_e of activity at electrode e. We used the joint log probability for more effective graphic presentation of very low joint probability values. The joint probability was computed for every data trial at each electrode.

Although this probability measure allowed us to detect some artifactual outlier trials, we used another measure to detect unusually peaky distributions of potential values. For this we used the kurtosis (K) of the activity values in each trial.

$$K = m_4 - 3m_2^2$$ (2)

$$m_n = E\{(x - m_1)^n\}$$ (3)

where m_n is the nth central moment of all activity values of the trial, m_1 the mean, and E an expectation function (in our case the average). If the kurtosis is highly positive, the activity distribution is highly peaked (usually around zero) with a sparse appearance of extreme values, and the identified data is likely to contain an artifact. If all activity values are similar, or the values alternate between two or more extreme, the kurtosis will be highly negative. Once more, this type of activity is not

typical of true EEG signals and may reflect non-stationary processes, so negative kurtosis values indicate the possible presence of artifacts. Strong negative kurtosis values usually reflect AC (alternating current) or DC (direct current) artifacts, for example those induced by screen currents, loose electrode contacts, and/or strong induced line noise from electrical machinery or fixtures.

Before defining rejection thresholds for joint probability and kurtosis, we first normalized these measures to have zero mean and unit standard deviation. We were thus able to define (z) thresholds in terms of standard deviations from expected mean values.

Spectral pattern. Finally, some EEG artifacts have specific activity and scalp topographies that are more easily identifiable in the frequency domain. For instance, temporal muscle activations typically induce relatively strong 20-60 Hz activity at temporal electrodes, while saccadic eye movements produce unusually strong (1-3 Hz) low frequency activity at frontal electrodes. To detect these artifacts, we computed the Slepian multitaper spectrum [14] for each single trial and each single channel, then subtracted the data mean spectrum, and finally applied positive thresholds to the remaining trial spectral differences.

III. DATA SIMULATION

To test and optimize the rejection process, we used event related EEG data from a 'Go/Nogo' visual categorization task [15]. EEG was recorded at a 1000 Hz sampling rate using a 32-electrode scalp montage with all channels referenced to the vertex electrode (Cz). The montage did not include specific eye artifact channels, but did include channels above the eyes (FPz; FP1, FP2). Responses to target and non-target stimulus presented about every 2 seconds were recorded for each subject. Data epochs were extracted extending from 100 before to 600 ms after stimulus onsets, and the mean value in the pre-stimulus baseline (-100 to 0 ms) was subtracted from each individual epoch. Data were then visually pruned of

noticeable eye and muscle artifacts by experimenter AD, resulting in 119 "clean" data epochs.

We then simulated five types of artifacts (Fig. 1): (1) We modeled eye blink time courses using random noise band-pass filtered (FIR) between 1 and 3 Hz. Eye blinks have stereotyped scalp topographies that can be isolated using ICA [4], so, to obtain topographical maps for these simulated eye blinks, we applied ICA decomposition to data from *another* subject and visually identified eye blink component by its time course and scalp topography (high gains on the most frontal electrodes; small gains everywhere else). (2) We modeled temporal muscle artifacts using random noise band-pass filtered (FIR) between 20 and 60 Hz and multiplied by a typical muscle scalp map, again isolated by ICA from another subject (high gains at a few temporal electrodes and near zero gains for other electrodes). (3) We then modeled electrical shift artifacts by implementing discontinuities at one randomly selected data channel. (4) We also modeled unfiltered white noise at another randomly selected data channel. (5) Finally, we modeled linear trends (with randomly selected slopes from 100 to 300 µV per epoch at the lowest level of noise) at another randomly selected data channel.

In the test data depicted above, each data channel could only have one type of artifact, excepting the first two artifact types, which projected with varying strengths to all the electrodes. We took care that the randomly selected channels for each artifact type differed from each other and did not coincide with channels where the two first topographical artifacts had maximum amplitude.

Since our goal was to test the sensitivity of each method to detect artifacts, we varied simulated artifact amplitude to find the smallest artifacts that each method could detect. Artifacts at the smallest amplitudes were so small that no method was able to detect them. For each artifact type, amplitude was gradually increased from -50 dB to 0 dB. To compute signal (artifact) to noise (background EEG) ratio (SNR), we computed the SNR at each frequency by dividing the spectrum for each type of artifact (not mixed yet with data) at each frequency with the data spectrum at the same frequency. We then found the frequency with the largest SNR and converted it to dB scale ($10*\log_{10}(SNR)$). Note that, prior to computing SNR for the first two (topographic) artifacts, we scaled their amplitudes by the highest channel gain in the applied scalp map.

IV. AUTOMATIC ARTIFACT REJECTION

Since we knew which data trials contained simulated artifacts, we could determine the most efficient artifact rejection method for each type of artifact. For each method, we chose a unique free parameter that we optimized to make the method best able to detect artifacts of a given type. The optimization procedure minimize the total number of trials misclassified (both misses and false alarms). To optimize the method parameter, we used a procedure that recursively divided its value range until a minimum was reached. (This assumes that there is a single local minimum for each

parameter; using more powerful non-linear optimization methods was not computationally feasible since they required too many iterations.)

Optimization was performed independently for each rjection method and for each type of artifact. We assumed voltage thresholds to be symmetrical in polarity, so only one parameter had to be optimized in the standard thresholding method. Linear trend detection had two parameters (minimum slope and goodness of fit). We set the slope to be equal or higher than the minimal artifactual slope at the maximum level of noise (0.5 µV/700 ms) and then optimized the goodness-of-fit parameter. Since it was time consuming and was specifically aimed at detecting trends, we applied this method only to detecting linear trend artifacts. For the probability and kurtosis methods, we optimized the standard deviation threshold. Finally, for the spectral measure, we optimized the dB limits independently for three frequency bands (0 to 3 Hz, 20 to 60 Hz, and 60 to 125 Hz) and then used the frequency band that returned the best results.

ICA-based rejection.

ICA separates EEG processes whose time waveforms are maximally independent of each other. The separated processes may be generated either within the brain or outside it. For instance, eye movements and muscle activities produce ICA components with specific activity patterns and component maps [4, 16]. However, scalp EEG activity as recorded at different electrodes is highly correlated and thus contains much redundant information. Also, several artifacts might project to overlapping sets of electrodes. Thus it would be useful to isolate and measure the overlapping projections of the artifacts to all the electrodes and this is what ICA does [2, 17]. To build intuition about how ICA works, one might imagine an *n*-electrode recording array as an *n*-dimensional space. The recorded signals can be projected into a more relevant coordinate frame than the single-electrode space: e.g. the independent component space. In this new coordinate frame, the projections of the data on each basis vector – i.e. the independent components – are maximally independent of each other. Intuitively, by assessing the statistical properties of the data in this space, we might be able to isolate and remove the artifacts more easily and efficiently.

Multiplication of the scalp data, *U,* by the unmixing matrix, *W,* found for example by infomax ICA represents a linear change of coordinates from the electrode space to the independent component space, or

$$S = W * U \qquad (4)$$

where *S* is the 'activation' matrix of the components across time. Each component is a linear weighted sum of the activity recorded at all of the electrodes. Each independent components comprises an activation time course and an associated scalp map (the corresponding column of W^{-1}) that gives the relative projection strengths (and polarities) of the component to each of the electrodes.

All the rejection methods described in the previous section were also applied to raw potential values decomposed using

different ICA algorithms. We used three different algorithms most commonly used to process EEG data: Infomax ICA, SOBI and fastICA. For Infomax, we used default parameters (since we checked that the EEG data did not contain any sub-Gaussian component, we did not use the extended version of Infomax for faster computation). For the SOBI algorithm [18], since we were processing data epochs, we slightly modified the algorithm to average covariance matrices processed for each data epochs. We also forced the number of correlation matrices (or time steps) to be equal to 100, which is a better default for EEG data than the original SOBI function default (Akaysha Tang, personal communication). For the fastICA algorithm [19], we forced the decomposition to be symmetrical and to estimate all components in parallel which is believed, though at the sacrifice of any speed advantage, to be a better mode for EEG data analysis (Aapo Hyvärinen, personal communication). Since we could not determine which component contained relevant artifacts, for each artifact type, each rejection method was applied to all components and

the single component returning the best results was retained.

We generated a total set of 20 data sets with different artifact and noise implementations, and then used a Linux cluster of 36 processors (1.4 GHz or higher) to optimize parameters for each method and each dataset. Final results presented here correspond to results of about 24 hours of computation on this cluster.

V. RESULTS

Results for each rejection method and each artifact type are presented in Fig. 2, which shows results for one artifact type in each row and one rejection method in each column. At all detectable artifact levels, all rejection methods performed better applied to ICA components (isolated using Inofmax ICA) than applied to raw data. Frequency thresholding performed the best; the joint probability method was second best, and standard thresholding third. Kurtosis thresholding performed the poorest, though it was partly successful in detecting large discontinuity and trend artifacts. Finally, the
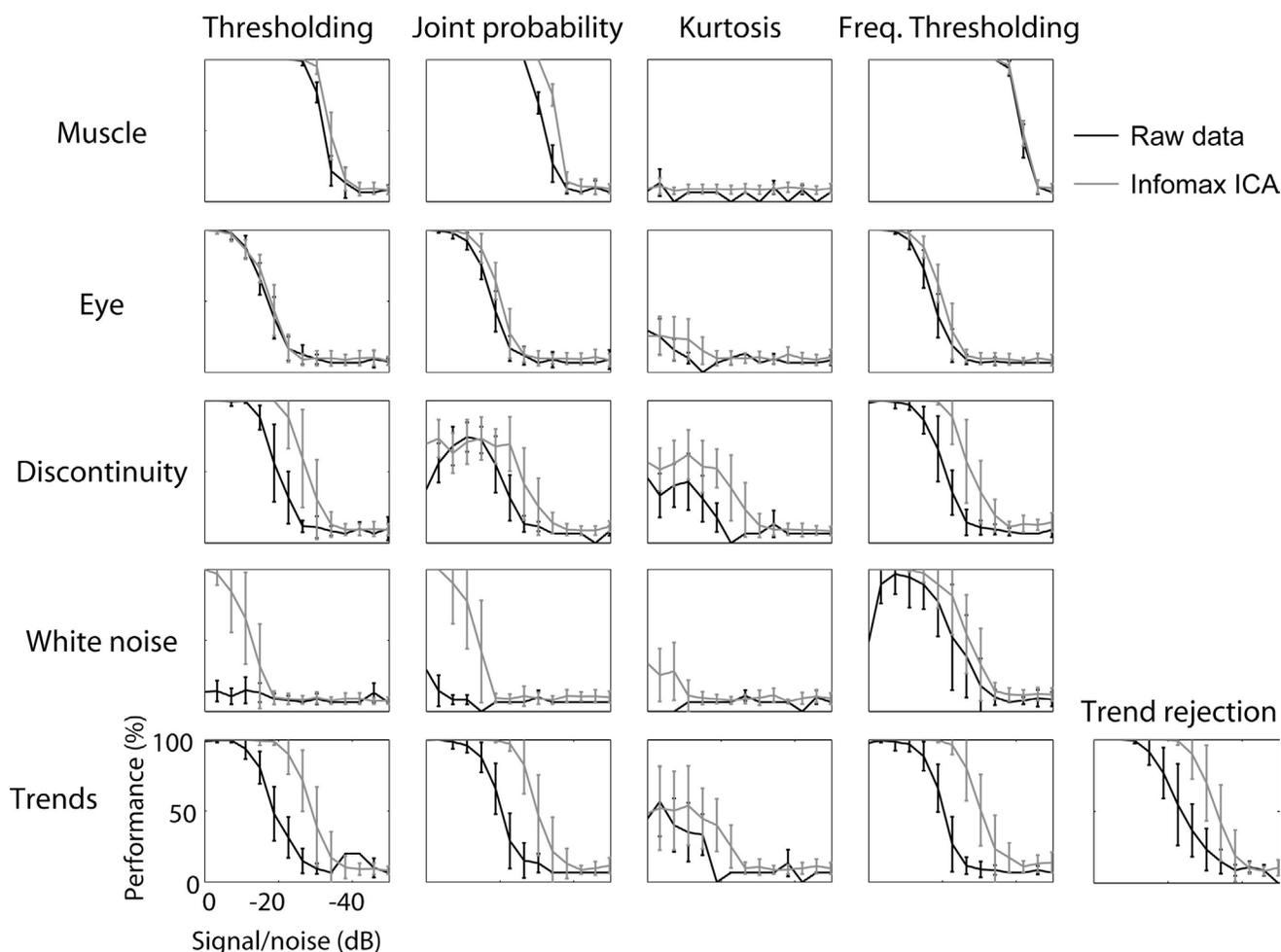
Fig. 2. Artifact rejection performance (y-axis, 0% to 100%) by five methods (columns) for five types of simulated artifacts (rows) of decreasing strength (x axis, 0 to -50 dB). Here, rejection performance was measured by correctly detected artifacts minus incorrectly rejected artifacts, divided by the total number of artifacts. The five methods were first optimally applied to the raw channel data (dark traces). (Since the trend detection method was time consuming it was only applied to the trend artifacts). Error bars indicate standard deviations across the 20 replications. Results of the same methods applied to the same simulated artifact data after decomposition by infomax ICA are shown as grey traces. Overall, spectral threshold methods worked best, and rejection performance was better when applied to the 'best' ICA component than when applied to the 'best' single scalp channel.
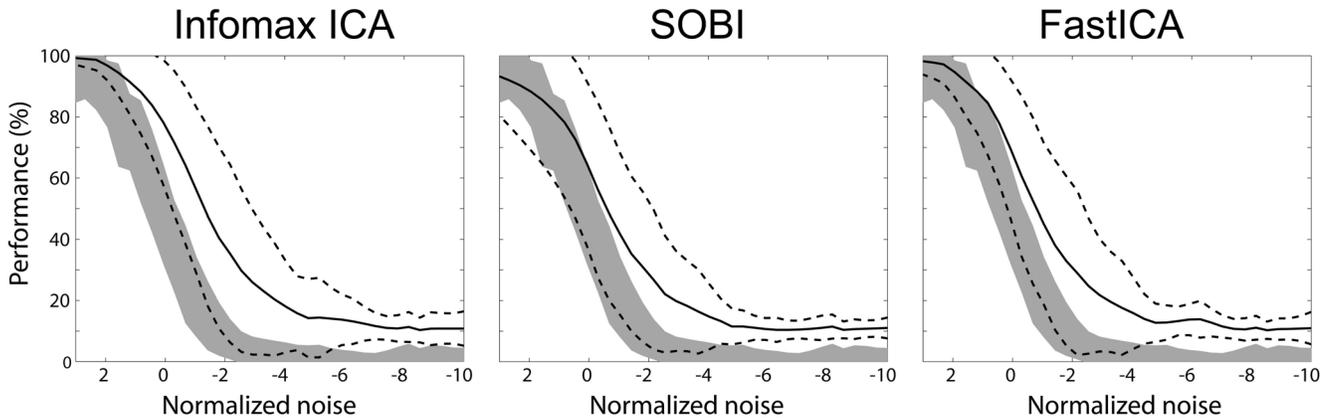
Fig. 3. Comparison of mean artifact rejection performance by spectral threshold methods applied to the raw channel data (grey band) and the same data decomposed using three well-known ICA algorithms (solid and dashed traces). Shaded areas show mean performance, plus and minus one standard deviation, of the spectral threshold methods applied optimally to raw best-channel data. To obtain these mean results, rejection performance for each type of artifact (shown in Fig. 2) were fitted to a logistic function then averaged (the normalized noise unit thus corresponds to 1 to 3 dB depending on artifact type). Mean performance of the same spectral threshold methods applied to the best independent component returned by three widely-used ICA algorithms are shown by thick traces; flanking dashed traces show these means plus or minus one standard deviation. For all but the highest artifact amplitudes, rejection performance of the spectral threshold method applied to the ICA component was about 10-20% higher than when applied to the single channel data.

trend rejection method was the most efficient method for detecting linear trends in the data, although slightly lower performance could be achieved using the frequency thresholding method (1-3 Hz band).

We then attempted to estimate the global performance of artifact rejection on the raw data and on data preprocessed using ICA. To do so, we only considered the frequency thresholding method since it outperformed most methods for all types of artifacts. Since the performance trend, as artifact size decreased, was different for each artifact type, we normalized each performance curve to the logistic function before averaging. In Fig. 3, we plot the average performance for three ICA algorithms commonly used to process EEG data, extended infomax ICA, SOBI and FastICA (see methods). We expected that frequency thresholding would be more efficient when applied to data preprocessed by any of the tested ICA algorithms than when applied to raw data. This is indeed what we observed (Fig. 3). Data preprocessing by ICA led to a 10-20% increase in artifact detection performance for all ICA algorithms we tested.

VI. DISCUSSION

We have shown that optimally applying spectral methods to isolate artifacts in 32-channel EEG data epochs allowed more reliable detection of smaller artifacts than optimally applying standard thresholding methods on the data. We have also shown that preprocessing the data using ICA allows more effective artifact detection.

In our simulated data, mixing of artifacts with data was perfectly linear. Might this not be the case for real data? In fact, instantaneous mixing via EEG volume conduction of artifacts and EEG processes is linear. By Ohm's law, externally imposed electrical artifacts (DC trends, discontinuities, white noise) also mix linearly with EEG data at scalp electrodes. On this basis, at least, linear ICA

decomposition algorithms are not inappropriate for separating artifacts from other data processes. On the other hand, these simulations may have disadvantaged the ICA approach since, for example, the simulated muscle artifacts were most probably mixed with actual muscle activity occurring in the background EEG. This might explain why muscle artifact detection applied to the ICA decompositions did not outperform artifact detection applied to the raw data. Moreover, threshold optimization might be of most benefit applied to the raw channel data, where frequency domain thresholds must be finely tune to best separate artifacts from the background noise. Optimized tuning of frequency domain thresholds for ICA component activity might be less important in the (typical) case in which ICA largely isolates artifacts to a single ICA component.

Beyond these practical results, there are several theoretical advantages to using ICA in EEG analysis. First, several major assumptions of ICA seem to be fulfilled in the case of EEG recordings (for a detailed justification, see [6, 17]). As mentioned previously, the first assumption is that the ICA component projections are summed linearly at scalp electrodes. The second assumption is that sources are independent. This is not strictly realistic but even if the appearance of artifacts might be related to brain activity – muscle contractions, for example, triggered by activity in the motor cortex – the time courses of the resulting artifacts and the triggering brain events are typically different across all or some trials. Thus, they should be accounted for by different independent components [4]. A third assumption concerns the non-gaussianity of the source activity distributions. This last condition is quite plausible for artifacts, which are usually sparsely active and thus far from Gaussian in distribution.

The EEGLAB toolbox [20], a user-friendly graphic interface developed in an open source environment running under Matlab (The Mathworks, Inc.), allows users to apply

ICA to their data, to combine and compare all types of rejection methods, and then manually review and edit artifact rejections. We have found that high-order statistics and spectral properties of independent component activities, available in this software and described in this article, may be strongly indicative of artifactual activity as defined and accepted in our laboratory. To use the EEGLAB software, we generally recommend (1) setting thresholds such that roughly 10% of data trials are detecting using a specified method, (2) visually inspecting data trials marked for rejection, and (3) optimizing the thresholds manually. For instance, for the joint data probability measure (which in our tests here performed better than standard thresholding yet was much faster to compute than spectral thresholding), we usually use thresholds higher than 5 standard deviations above the mean. (For a Gaussian distribution, the probability that a tagged artifact trial belongs to the 'ordinary' trial distribution would then be less than $1.6*10^{-12}$). After finally rejecting the marked and checked data artifacts, the cleaned data may be decomposed by ICA for studying brain source dynamics and/or processed by other analysis methods.

REFERENCES

[1] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput*, vol. 7, pp. 1129-59., 1995.

[2] S. Makeig, A. J. Bell, T. P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," in *Advances in Neural Information Processing Systems*, vol. 8, D. Touretzky, M. Mozer, and M. Hasselmo, Eds., 1996, pp. 145-151.

[3] T. P. Jung, S. Makeig, M. J. McKeown, A. J. Bell, T. W. Lee, and T. J. Sejnowski, "Imaging brain dynamics using Independent Component Analysis," *Procesedings of the IEEE*, vol. 89(7), pp. 1107-22, 2001.

[4] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects," *Clin Neurophysiol*, vol. 111, pp. 1745-58., 2000.

[5] A. Delorme, S. Makeig, and T. J. Sejnowski, "Automatic artifact rejection for EEG data using high-order statistics and independent component analysis," presented at International workshop on ICA, San Diego, CA, 2001.

[6] L. Zhukov, D. Weinstein, and C. Johnson, "Independant Component Analysis for EEG source separation," *IEEE engineering in medecine and biology*, vol. 19, pp. 87-96, 2000.

[7] J. Iriarte, E. Urrestarazu, M. Valencia, M. Alegre, A. Malanda, C. Viteri, and J. Artieda, "Independent component analysis as a tool to eliminate artifacts in EEG: a quantitative study," *J Clin Neurophysiol*, vol. 20, pp. 249-57, 2003.

[8] C. J. James and O. J. Gibson, "Temporally constrained ICA: an application to artifact rejection in electromagnetic brain signal analysis," *IEEE Trans Biomed Eng*, vol. 50, pp. 1108-16, 2003.

[9] Y. Tran, A. Craig, P. Boord, and D. Craig, "Using independent component analysis to remove artifact from electroencephalographic measured during stuttered speech," *Med Biol Eng Comput*, vol. 42, pp. 627-33, 2004.

[10] E. Urrestarazu, J. Iriarte, M. Alegre, M. Valencia, C. Viteri, and J. Artieda, "Independent component analysis removing artifacts in ictal recordings," *Epilepsia*, vol. 45, pp. 1071-8, 2004.

[11] G. Barbati, C. Porcaro, F. Zappasodi, P. M. Rossini, and F. Tecchio, "Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals," *Clin Neurophysiol*, vol. 115, pp. 1220-32, 2004.

[12] C. A. Joyce, I. F. Gorodnitsky, and M. Kutas, "Automatic removal of eye movement and blink artifacts from EEG data using blind component separation," *Psychophysiology*, vol. 41, pp. 313-25, 2004.

[13] T. W. Lee, M. Girolami, A. J. Bell, and T. J. Sejnowski, "A Unifying Information-theoretic Framework for Independent Component Analysis," *Comput. Math. Appl.*, vol. 31, pp. 1-21, 2000.

[14] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proceeding of the IEEE*, vol. 20, pp. 1055-1096, 1982.

[15] A. Delorme, G. A. Rousselet, M. J. Mace, and M. Fabre-Thorpe, "Interaction of top-down and bottom-up processing in the fast visual analysis of natural scenes," *Brain Res Cogn Brain Res*, vol. 19, pp. 103-13, 2004.

[16] S. Makeig, T. P. Jung, A. J. Bell, D. Ghahremani, and T. J. Sejnowski, "Blind separation of auditory event-related brain responses into independent components," *Proc Natl Acad Sci U S A*, vol. 94, pp. 10979-84., 1997.

[17] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, pp. 163-78., 2000.

[18] A. Belouchrani and A. Cichocki, "Robust whitening procedure in blind source separation context," *Electronics Letters*, vol. 36, pp. 2050-2053, 2000.

[19] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Netw.*, vol. 13, pp. 411-30, 2000.

[20] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J Neurosci Methods*, vol. 134, pp. 9-21, 2004.