# Quantifying Evoked Responses through Encoded Information

Julian Fuhrer*[,a], Kyrre Glette[a], Tor Endestad[b,d], Anne-Kristin Solbakk[b,c,d], and Alejandro Blenkmann[d]

[a]*RITMO, Department of Informatics, University of Oslo, 0373 Oslo, Norway*
[b]*Department of Neurosurgery, Oslo University Hospital, Rikshospitalet, 0372 Oslo, Norway*
[c]*Department of Neuropsychology, Helgeland Hospital, 8657 Mosjøen, Norway*
[d]*RITMO, Department of Psychology, University of Oslo, 0373 Oslo, Norway*

**Abstract**

Being the most sophisticated information processing apparatus known to us, understanding the brain offers great possibilities. A viable candidate to advance the understanding of this cortical information processing machine is information theory. With its universal applicability, it enables the modeling of complex systems, is free of any requirements about the data structure and can infer the present underlying brain mechanisms. Algorithmic information theory seems particularly suitable since it can estimate the information contained in individual brain responses. Here, we propose a measure grounded in algorithmic information theory termed Encoded Information as a novel approach to analyze neurophysiological recordings. Specifically, it enables an assessment of the encoded information mean responses share with one another by compressing the respective signals. By applying the approach to five cortical activity types originating from intracranial electroencephalography recordings of humans and marmoset monkeys, we demonstrate that the information-based encoding can compete with conventional approaches such as the t-test or mutual information. Information-based encoding is attractive whenever one is interested in detecting where in the brain the neural responses differ across experimental conditions.

*EEG | SEEG | ECOG | Information Content | Kolmogorov Complexity | t-test | Frequency Tagging*

## Introduction

Efficient processing of *information* is a core capacity of the brain. It enables us to rapidly perceive, comprehend, identify changes, and (if needed) engage with our environment – all seemingly without conscious effort. Accordingly, in the attempt to understand the respective underlying brain mechanisms responsible for this capacity, it is plausible to employ approaches that originate from *information* theory. This mathematical theory provides multivariate analysis tools, is not bound to a single type of data, is model-independent (i.e., does not require assumptions about the data itself) and can capture nonlinear interactions (1–4). Hence it is apparent to utilize those principles for the analysis of neurophysiological recordings originating from electroencephalography (EEG), intracranial EEG (iEEG), magnetoencephalography (MEG) or functional magnetic resonance imaging (fMRI).

Specifically, by measuring the degree of redundancy, *algorithmic* information theory (AIT) estimates the absolute information contained in individual responses. The higher the information content, the more complex its structure. Accordingly, the less compressible or more random the response. Assessing the absolute information content of responses recorded from different contact sites enables to infer on the complexity of the activity recorded at the respective sensors and potentially allows to identify underlying dynamics.

While AIT has been applied to analyze cognitive operations or to discriminate between states of consciousness measured with EEG, iEEG, MEG or fMRI recordings (5–9), its use in neuroscience is limited – despite its potential. Here we demonstrate the measure of *encoded information* (EI, 10) as a worthwhile

*E-mail address: julianpf@uio.no

tool to directly quantify the level of similarity between responses of neurophysiological data. By use of iEEG recordings stemming from 22 humans (11) and three marmoset monkeys (12, 13), we validate this approach of encoded information considering theta (5 to 7 Hz), alpha (8 to 12 Hz), beta (12 to 24 Hz), high-frequency (HFA; 75 to 145 Hz) and ERP activity by comparing it to that of a conventional t-test, to Mutual Information (MI), Gaussian Copula Mutual Information (GCMI; 3) and to the method of Neural Frequency Tagging (NFT; 14, 15).

# Material & Methods

## Neurophysiological Recordings

### Human

We analyzed experimental iEEG data obtained from intracranial electrodes implanted in (self-reported) normal-hearing adults with drug-resistant epilepsy (11). Participants (n=22, mean age 31 years, range 19 to 50 years, 6 female) performed a passive auditory oddball paradigm where a standard tone alternated with random deviant tones. The tones had a duration of 75 ms and were presented every 500 ms in blocks of 5 min consisting of 300 standards and 300 deviants. At the beginning of each block, 15 standards were played. To capture automatic, stimulus-driven processes, participants were asked not to pay attention to the sounds while reading a book or magazine. They completed 3 to 10 blocks, providing at least 1800 trials (for details, see 10, 11). From the 22 participants, a total of 1078 channels (mean: 48, range: 12 to 104) were available after cleaning. Theta (5 to 7 Hz), alpha (8 to 12 Hz), beta (12 to 24 Hz), high-frequency (HFA; 75 to 145 Hz) and evoked related potential (ERP) activity was then extracted from the data. For each activity type, differences between standard and deviant tone responses were evaluated in the 400 ms time window following the sound onset across channels and subjects. Additionally, the neural oscillation synchronization to the tone onset frequency (2 Hz) was examined by considering the raw data.

### Marmoset

Further, we considered experimental iEEG data employing a roving oddball paradigm (see 12, 13). Based on three awake adult male common marmosets (*callithrix jacchus*), iEEG recordings were captured while they were passively exposed to a stream of sounds. The paradigm consisted of sound trains of three, five or 11 repeated single tones of 20 different frequencies (250 to 6727 Hz with intervals of 1/4 octaves). All tones were pure sinusoidal tones, lasted 64 ms (7 ms rise/fall) and there was stimulus onset asynchrony 503 ms between them. For each sound-train, all tones were identical but varied across tone-trains. Consequently, the mismatch occured between the transition from one train to another in the form of a frequency change. According to that, the last tone of a train is defined as a standard tone while the first tone of a new train is considered as a deviant tone. Standards and deviants transitioned then 240 times during a recording session.

The number of implanted electrodes varied from monkey to monkey. For monkey ”Fr” 32 (located in the left hemisphere), 64 for ”Go” (in the right hemisphere) and 64 electrodes for monkey ”Kr” (in the right hemisphere) were implanted in the epidural space (Fig. 6c; see 12, 13, for more detailed information). Recorded data was re-referenced through an average reference montage and epoched into −950 to 2000 ms segments relative to the standard or deviant tone onset. Theta, alpha and beta frequency bands were then extracted from the demeaned signals using wavelet time frequency transformation (Morlet wavelets) based on convolution in the time domain (16). To obtain the HFA, the Hilbert transform was applied to the signal. To eliminate any residual artifacts, responses with an amplitude larger than five standard deviations from the mean for more than 25 consecutive ms, or with a power spectral density above five standard deviations from the mean for more than six consecutive Hz were excluded. For the analysis, all recordings were shortened to the −100 to 350 ms interval relative to sounds onset.

## Encoded Information

Given the mean responses, we estimated the EI across conditions as described in (10). In short, by employing algorithmic or Kolmogorov Complexity (K-complexity), we estimated the EI between conditions through the measure of Normalized Compression Distance (NCD) (2, 17). For a pair of strings $(x, y)$, it is defined as

$$\mathrm{NCD}(x, y) = \frac{C(xy) - \min(C(x), C(y))}{\max(C(x), C(y))},$$

with $C(xy)$ denoting the compressed size of the concatenation of $x$ and $y$, and $C(x)$ and $C(y)$ their respective size after compression (2, 17). Further, the NCD is non-negative, that is, it is $0 \leq \mathrm{NCD}(x, y) \leq 1 + \epsilon$, where the $\epsilon$ accounts for the imperfection of the employed compression technique. Small NCD values suggest similar objects, and high values suggest rather different objects.

To obtain a compressed version of a signal it was simplified by grouping its values into 128 discrete steps (bins). The bins covered equal distances and in a range between the global extrema of all the signals considered. The compressor received the indices of the bins that contained the elements of the signal (6, 7). Compression then proceeded through a compression routine based on Python's standard library with gzip. The statistical hypothesis testing was then performed through a permutation-based approach as described below.

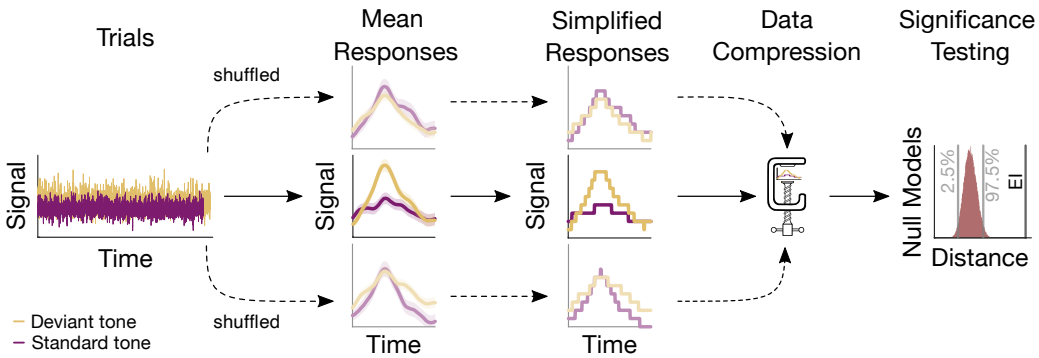



**Figure 1:** Sketch of the procedure for an example channel. Based on the trials, a mean response for each condition is computed. Trials are then shuffled, resulting in surrogate mean responses. Subsequently, these signal undergo a simplification procedure, followed by their compression. The output of the compression routine is the EI, quantifying the similarity between responses. Resulting values are then evaluated, leading to a null model distribution. This distribution serves to assess the significance of the actual EI value.

## t-test

We compared our approach to that of a conventional t-test. For each sample and channel, a two-sided t-test was performed, where the resulting t-value represented the activity difference between the two conditions (i.e., standard or deviant). To correct for multiple comparisons across samples and channels, False Discovery Rate (FDR) adjustment was applied with an FDR of 0.05.

## Neural Frequency Tagging

We used NFT to identify the brain's capability to automatically segment the continuous auditory stream (14, 15). In our sequence, the two main segments are represented by the frequency of sound onsets (2 Hz) or by the frequency of transitions between standards to standards or deviants to deviants (i.e., half the frequency or 1 Hz). If the neural activity of a recording site shows such "tagging" of frequency-specific properties within the stimuli, we define it as "responsive". If it tags half the frequency we identify it as sensitive to the pattern of standard-deviant alternation.

To assess this tagging, we computed the power-spectral density (PSD) of the epoched raw data using Welch's method (Fig. 2). Subsequently, we estimated the signal-to-noise ratio (SNR) of the PSD (18).

Here, SNR defined the ratio of power in a given frequency (signal) to the average power in the surrounding frequencies (noise). By doing so, we normalized the spectrum and accounted for the 1/f power decay (19). We then identified significant peaks through a lower threshold consisting of two times the standard deviation. The latter was estimated by computing the median absolute deviation, which was obtained by taking the median SNR multiplied by the constant distribution-dependent scale factor (Fig. 2). In the case of normally distributed observations it reflects the 50 % of the standard normal cumulative distribution function, leading to a scale factor of 1.4826 (20, 21).

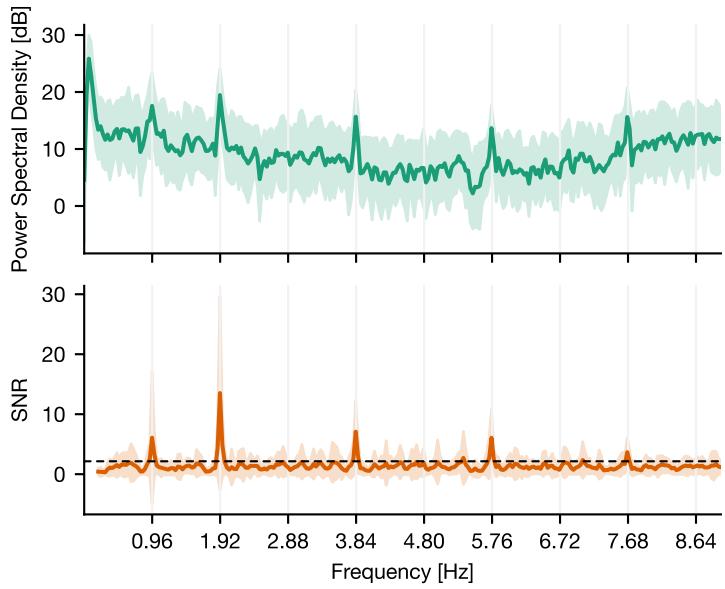



**Figure 2:** PSD and SNR of a responsive channel located in the superior temporal sulcus. The activity of the channel shows synchrony to the main frequency of tone onsets, which is around 2 Hz. The displayed deviation is due to a constant lag of tone presentation during recording (resulting theoretical main sequence is at 1.919 Hz). Further, it shows synchrony to half the main frequency. Accordingly, the underlying region of this channels tags the presentation rate of solely standard or deviant tones. Thus, discriminates between the standard and deviant condition.

## Mutual Information

Besides the t-test and NFT, EI was compared to the measure of MI. While it also grounds in information theory, in contrast to EI it draws on the concept of Shannon entropy (i.e., classic information theory). For a discrete random variable x with N outcomes, the entropy can be defined as

$$H(x) = -\sum_{k=1}^{\mathrm{N}} \mathrm{p}(x_{\mathrm{k}}) \log \mathrm{p}(x_{\mathrm{k}}) \,, \tag{1}$$

with $\mathrm{p}(x_{\mathrm{k}})$ being the occurrence probability for each element $x_{\mathrm{k}}, \ldots, x_{\mathrm{N}}$ of $x$. Given this definition, the MI between two discrete random variables $(x, y)$ with N or M outcomes can be defined as

$$\begin{aligned} \mathrm{MI}(x; y) &= H(x) - H(x|y) \\ &= H(y) - H(y|x) \\ &= \sum_{\mathrm{k}=1}^{\mathrm{N}} \sum_{\mathrm{j}=1}^{\mathrm{M}} p(x_{\mathrm{k}}, y_{\mathrm{j}}) \log \frac{p(x_{\mathrm{k}}, y_{\mathrm{j}})}{p(x_{\mathrm{k}}) p(y_{\mathrm{j}})} \,, \end{aligned} \tag{2}$$

with the joint probability $p(x_{\mathrm{k}}, y_{\mathrm{j}})$ and the marginal probabilities $p(x_{\mathrm{k}})$ and $p(y_{\mathrm{j}})$.

This estimation of MI proceeds through estimating the probability distribution of Eq. 2 by binning the signal into discrete steps. This is followed by a maximum likelihood estimation, yielding the probability distribution estimates. In our analysis, the respective signals were binned into four bins.

Besides this binning approach, a novel estimation technique of MI after Ince et al. (3) was employed. With this approach, MI is estimated via Gaussian Copula. In short, each univariate marginal distribution is transformed into a standard normal. Subsequently, a Gaussian parametric MI estimate is applied. That yields a lower bound estimate of MI, named GCMI. Statistical significance testing was then estimated through surrogate data testing.

### Surrogate Testing

The statistical significance of the information-based measures (EI, MI and GCMI) were assessed through surrogate data testing. Accordingly, p-values were obtained by evaluating the observed information-based quantity in terms of a null distribution (Fig. 1). For EI and binned MI, null distributions were created by repeatedly shuffling the trials (i.e., single evoked responses) between conditions (e.g. standard and deviant) and then re-computing the information-based measure. For the GCMI, this proceeded for each sample of the time series signal. Single-sided p-values lower or equal to 0.05 were considered statistically significant. To correct for multiple comparisons across channels, FDR adjustment was applied with an FDR of 0.05.

### Significance Ratio

To compare the different methods with each other, the ratio of significant channels to the total amount of channels was determined. This was either computed for each subject, where the total significance ratio was the mean ratio across all subjects accompanied by a bootstrapped 95 % confidence interval. Or, by taking the total ratio by collapsing across all channels (regardless of subjects). The latter was used when the number of channels for a specific region was limited, leading to distorted ratios with great confidence intervals. For the second data set, the ratio of each monkey was considered individually.

## Results

Based on two neurophysiological data sets, we compared our method's performance on discriminating standard to deviant tone responses to that of a conventional t-test, MI and GCMI across five different cortical activity types. Additionally, we compared it to NFT on the ERP activity.

Overall, EI and t-test showed the greatest significance ratios (Fig. 3a&6a). Across the 1078 channels stemming from the human brain recordings, for alpha, beta and HFA activity, EI's significance ratio was significantly greater than all methods considered (Fig. 3a), while the t-test showed the greatest significance ratio for the ERP activity (EI and t-test comparison on a subject level: two-sided paired t-tests, p=[1.43e−3, 3.61e−2, 2.87e−3 and 6.44e−3]). MI showed the lowest ratios, although it was only significantly lower on a subject level for HFA (two-sided paired t-test, $p_{min}$=3.02$e$−2). Worth pointing out is that for human iEEG, the number of channels and their distribution is not constant across subjects, and can thus greatly vary. Consequently, the less channels for a subject, the less robust the estimated individual significant ratio. For the human iEEG data set, the mean number of channels were 49 per subject with a minimum of 12 and a maximum of 105 channels.

The observations from this data set are in line with with the results from the marmoset recordings, where 128 channels in total exist. Based on the significance of each individual marmoset monkey, for theta, alpha and beta activity, EI detects more channels than the other approaches. For HFA, GCMI and EI perform similarly well, while for ERP, GCMI and t-test perform best (Fig. 6a). Additionally, all methods in common is the variance of detected channels across monkeys.

To assess channels the methods commonly detect, we put the absolute numbers of significant channels relative to the total amount of significant channels for each pairwise method combination (Fig. 6c, i.e. each method's specific unique set of significant channels plus their intersection), which was then defined as the intersection ratio. For the human recordings, the intersection ratio between t-test and GCMI was

the greatest across all activity types, reaching its maximum for the ERP activity (Fig. 3b). For beta and HFA activity, there is also a high intersection ratio between the t-test and GCMI. However, by considering Fig. 3a, it appears that the number of significant channels is low for these activities, leading to higher percentages. This can also be observed by correlating the method specific channel distributions across subjects (Fig. 3c), i.e., by correlating the significant ratios across subjects for each method to method combination. Aside from that, the approaches correlate the least for beta activity, which is also the activity with the lowest significance ratios. Collapsing all activities together into one model reveals that overall a high correlation across the methods exists (Fig. 3c).

The relatively high intersection ratio between t-test and GCMI across activities is not observable for the roving oddball paradigm. For this data set, the intersection ratios are the highest for HFA and ERP activity (Fig. 6b). Further, the same effect as for the human recordings occurs. For HFA, the monkey Fr shows maximum intersection ratios between t-test, GCMI or MI. However, the respective significance ratios are rather low (Fig. 6c). Moreover, when comparing Kr to Go it appears that the latter has a smaller number of significant and overlapping channels for EI and t-test. However, because Kr exhibits more significant channels for both methods, the intersection ratio is about the same for both monkeys.

For the human iEEG data set, we further evaluated all contact sites through the approach of NFT. The amount of responsive channels, i.e., the amount of channels that solely tag the frequency of standards or deviants is lower in comparison to all other methods (two-sided paired t-test between NFT and MI, $p=6.87e-3$).

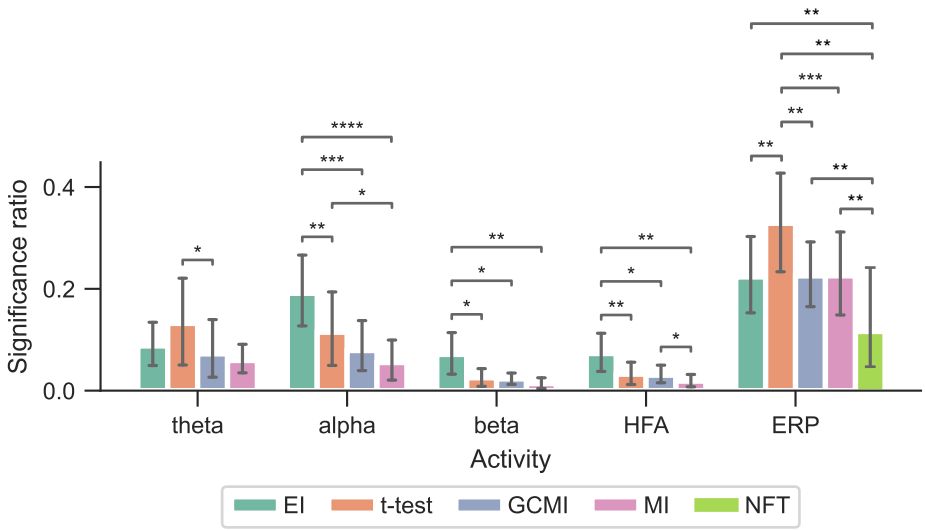
Given the differing sensitivity to deviating sounds across individual brain regions, we then compared the methods' significance ratios across individual regions. For this reason, we evaluated the methods' performances across different areas comprising temporal, frontal, insular, peri-central sulci, and ACC cortices, as well as the hippocampus (Fig. 4; see 5, for exact definitions). Overall, the significance ratios of EI aligns with the other methods across regions and activity types. That is to say, the ratio for higher hierarchical areas such as the superior or middle frontal cortex is low, while being higher in areas such as the superior temporal plane (which includes the Heschl's gyrus). The NFT method together with MI showed the lowest sensitivities across brain regions. However, there were higher or equally high ratios in sensitive brain areas such as the superior temporal plane or posterior insular cortices.

To examine the capabilities of the two best performing methods, we further applied the approach of EI and t-test to scenarios where only a reduced number of trials is available. The number of trials varied from 1 to 100 % of all available trials ($759.50 \pm 360.85$ trials for deviant responses and $715.14 \pm 388.12$ trials for standards across all channels). For each percentage and condition, trials were randomly picked. This step was repeated 50 times for each trial increment to obtain an average p-value as a function of the number of trials (Fig. 5). This was evaluated based on the human iEEG recordings for channels with differing sensitivity to deviating tones. The channels where chosen in terms of the 25, 50, 75, and 97.5 % percentiles of the t-value distribution emerging from HFA. As can be seen in Fig. 5, the two methods only marginal differ in reaching statistical significance, although EI tends to reach this threshold slightly earlier.

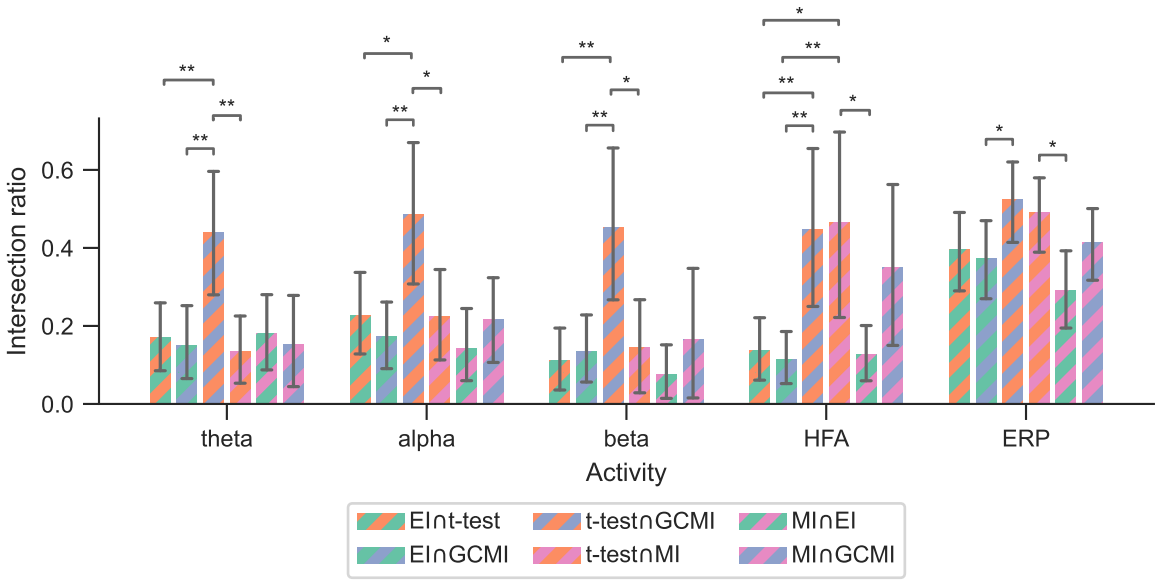
## Discussion

With the measure of EI, we proposed information-based condition discrimination as an alternative to the classical t-test approach. Using compression as the core principle, this procedure quantifies the similarity between recordings stemming from different conditions. By applying this procedure to five cortical activity types, and comparing it to that of t-test, MI, GCMI and NFT, we demonstrate the high sensitivity of EI.

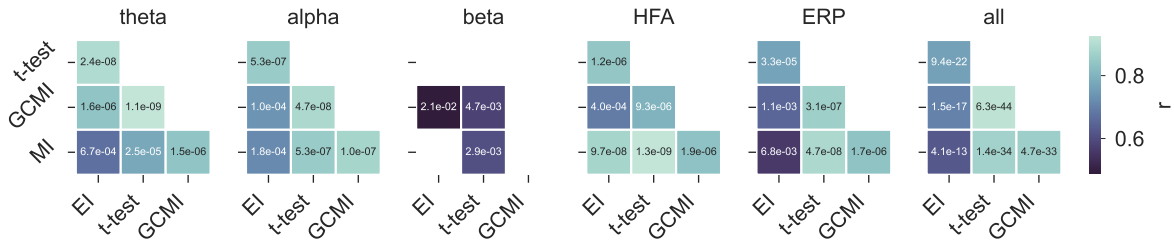
Besides their individual procedure in estimating test statistics, the examined methods also differ in how the data is processed. While for the t-test and GCMI approach, each sample or time point across trials is considered independently, EI and also MI in its current implementation make use of all time samples by computing the respective mean responses across trials. Consequently, for each channel, the first two methods result in test statistics along the time axis, while the last two output one concrete channel-dependent test statistic. Such a time course is advantageous when the core interest is within the time




**(a)** Significance ratios across activities and methods.




**(b)** Intersection of significant channels across methods.




**(c)** Correlation of methods across subjects.

**Figure 3:** Performance of the different methods. The error bars indicate the 95% CIs across subjects. Importantly, each subject has a unique electrode distribution such that the range of significant channels can greatly vary. **a**: Significance ratio across activity types. Statistical significance is indicated with * $p \leq 5e{-}2$, ** $p \leq 1e{-}2$, *** $p \leq 1e{-}3$ and **** $p \leq 1e{-}4$. **b**: Intersection of the significant channels for each method combination. Each number is put relative to the total number of significant channels. **c**: Correlation matrices comparing the subject specific significant ratios. In each square, the respective p-value is indicated.

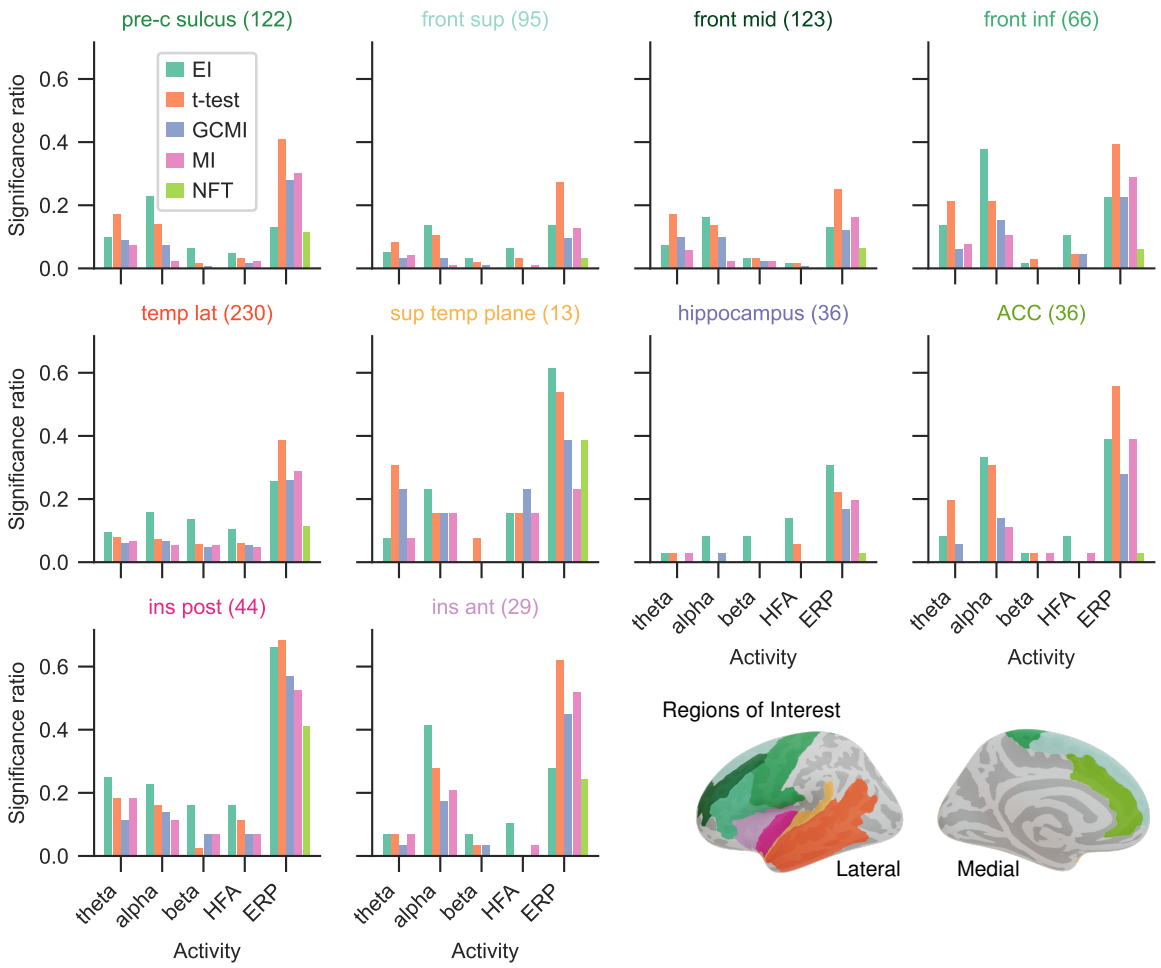



**Figure 4:** Significance ratio across regions for the different methods. The number of channels per ROI is indicated in brackets (For more details on the ROIs, see 5).

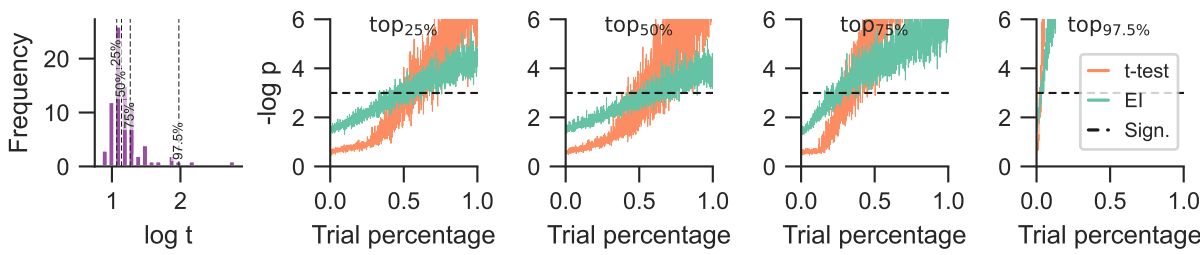
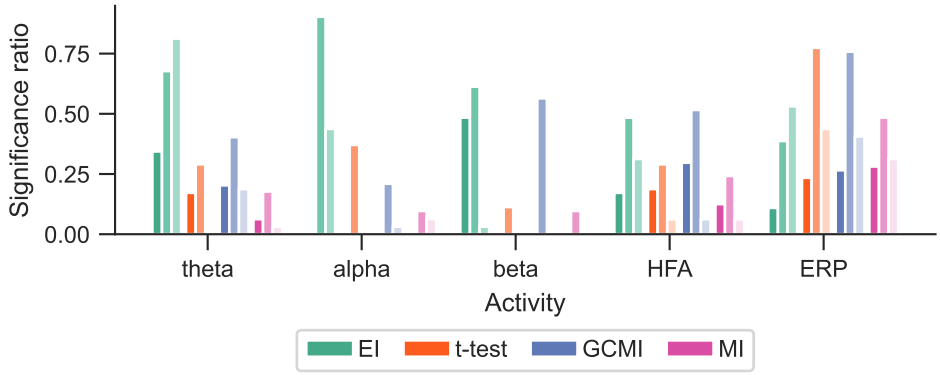
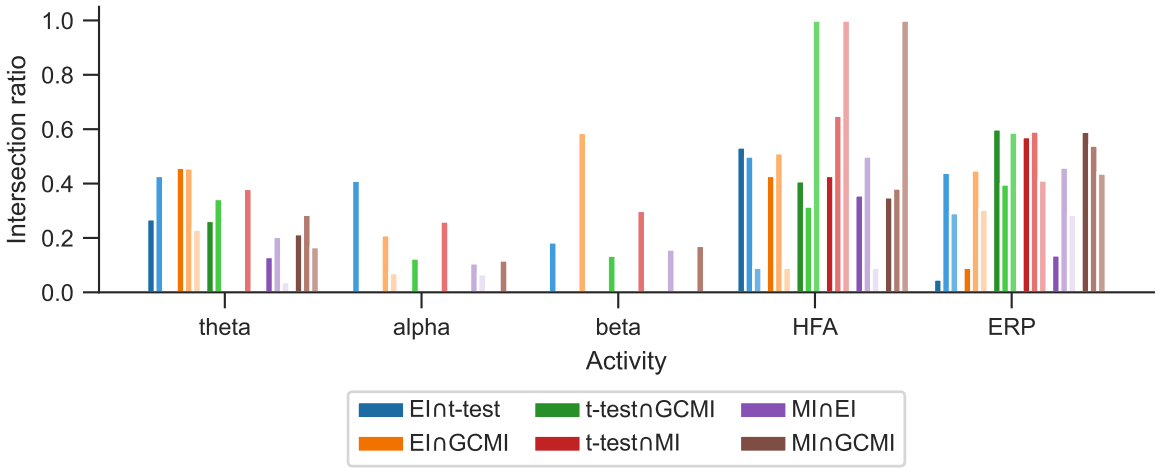


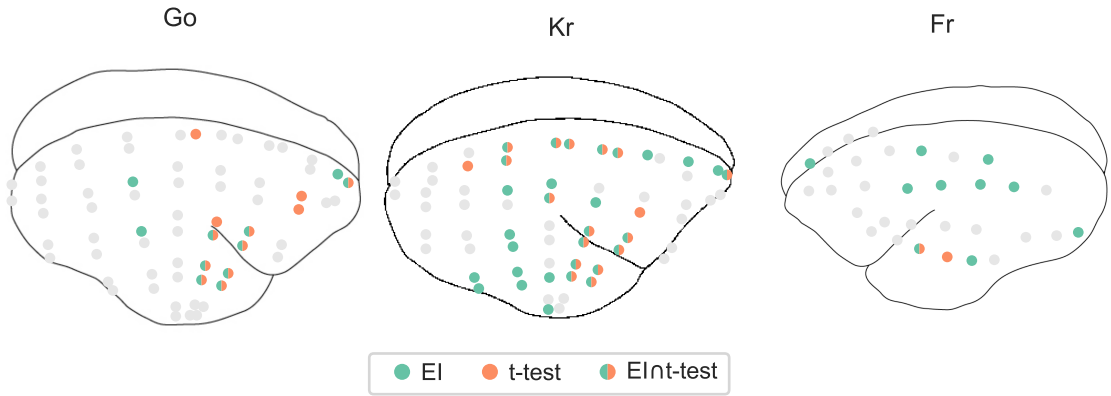

**Figure 5:** Results of numerical investigation of the performance for limited amount of data. The number of trials is reduced by randomly picking a varying number of trials for each condition. This is followed by applying the measures and was repeated 50 times. The dashed line represents the significant threshold of p=0.05. Note that the y-range is limited to 6 (corresponding to a p-value of $2.5e{-}3$). 100 % of the trials was around $759.50 \pm 360.85$ trials for deviant responses and $715.14 \pm 388.12$ trials for standards for each channel.




**(a)** Significance ratios across activities for the three monkeys.




**(b)** Intersection of significant channels across methods.




**(c)** Location of the significant channels based on HFA across the marmosets.

**Figure 6:** Performance of the different methods across the three marmoset monkeys ”Go”, ”Kr” and ”Fr”. **a**: Global significance ratio for each marmoset. **b**: Intersection of the significant channels for each method combination. Each number is put relative to the total number of significant channels for each method combination. **c**: Location of the significant channels for EI and t-test for HFA (monkey ”Go” has 64, ”Kr” 62 and ”Fr” exhibits 32 channels).

domain or latency of channels, i.e. at which time point responses differ most. However, when it comes to assessing how sensitive contact sites are towards different conditions, this time course is of a minor role. By masking a channel's significance, it is necessary to reduce the time course of the test statistics to one concrete value. Additionally, the complete time series of the respective statistic also needs to be corrected for multiple comparisons across channels and trials (i.e., correction of a 2D-array that covers samples and channels versus a 1D array that only includes channels).

Adopting the sample-based approach mentioned above is not beneficial for EI. The measure of EI is a compression-based approach and thus compresses the entire time series while making use of structures along the time axis. This mechanism is a key feature of this approach and is possibly responsible for the higher significance ratios across activities (Fig.3a, 4 and 6). In using the trails for each sample, this feature would be removed. Especially for responses that are close in magnitude (Fig. 1), this mechanism proves beneficial. By exploiting complementary structures within both responses, subtle information-grounded differences that extend beyond the width of one sample can be identified in such scenarios. Additionally, the binning parameter or resolution along the time-axis is of minor importance.

Especially for beta and HFA activity, EI appears to be a worthwhile tool for detecting active channels (Fig.3a). For both data sets, EI exhibits the highest significance ratios compared to the t-test, MI and GCMI. HFA presumably carries mismatch or prediction error signals (22). In that regard, detecting a high number of channels in HFA discriminating between standard and deviating sounds in regions such as the hippocampus appears to be especially interesting (Fig. 4).

For the ERP activity, the t-test shows the greatest significance ratios. One possible reason for this could be the generally low variance (or standard error) in EPR compared to the other activity types. This low variance leads to statistically strong differences in amplitudes resulting from the different conditions. Further, the number of samples that significantly differ in amplitude after correcting for multiple comparison can be low for a channel (i.e., the whole trial versus only a few sample points). EI, on the other hand, considers the full time course and only assesses channels as significant when the information content of the entire mean responses differs.

Overall the sample based approaches of t-test and GCMI correlate the most (Fig. 3c). Besides that, there is no clear outlier apparent for both data sets. This also applies when considering the significance ratios across individual brain regions (Fig. 4). This can be seen as an validation of the EI measure. While showing higher significance ratios in some ROIs (e.g., hippocampus or pre-central sulcus), it performs similarly well as the other methods across brain areas and does not indicate implausible results.

Compared to EI, t-test, MI and GCMI, the approach of NFT lags behind. However, when considering individual ROIs, it performs equally well on regions close to or belonging to the temporal cortex. Worth pointing out is that NFT is based on a rather different procedure. By making use of the full frequency spectrum and the entire sequence, it exploits the static presentation rates of the stimuli. Fig. 2 shows such a "tagging" effect, where the brain synchronizes to the exact frequency the stimuli were presented. The highest SNR or power is contained at 1.92 Hz which is the presentation rate of the standard or deviant tones. In addition, it also shows a peak at half this frequency, which is the combination of either standard to standard or deviant to deviant tones. This implies that the brain discriminates between standard and deviant tones. One prerequisite here is that the presentation rate is static. Any temporal lab during the presentation disturbs this "tagging" effect. One possible reason for the relatively low performance could be traced back to this unique approach of NFT together with the employed multidimensional oddball paradigm. It is possible that it performs better in case of a paradigm with fewer deviant types (as compared to eight deviant tones).

When it comes to performance, MI is close to NFT. It detects less channels than the remaining methods, especially for alpha, beta and HFA activity. Important to notice is that for comparison reasons, we implemented MI in the same way as EI. That is to say, the entire time series of a mean response is considered. It is also possible to implement it in the sample-based fashion of GCMI, where it has been shown to operate similarly well (3).

Further, the number of bins during the binning procedure across the data sets was held constant. We chose 128 bins for EI and four bins for MI (3, 5, 10), because the trials of both data sets had a similar length. It is possible to adapt this parameter, which might be needed for longer lasting trials, although

both EI and MI should be robust to this parameter choice to some extent. For example, EI showed similar results when varying the number of bins from 128 to 64, but showed a decrease when using only 32 bins.

Lastly, in the case of limited data availability, the two best-performing methods show a relatively similar performance (Fig. 5). For lower t-values (below the 75 %-percentile), both methods need around 50 % of the maximally available trials to identify a channel as sensitive. Although EI appears to require slightly less trials than the t-test. Worth pointing out is that for channels with relatively distinct responses to standards and deviants (above the 97.5 %-percentile), both measures only require around 3 % of the available trails to identify the respective channel as significant.

## Conclusion

Taken together, our proposed procedure of EI proves as a worthwhile procedure to discriminate differences in neurophysiological recordings of evoked responses. Among the methods considered, EI can compete in detecting channels sensitive to deviating sound types across activities. Especially for beta and HFA activity, EI proves as a worthwhile approach, detecting a higher number of sensitive channels in comparison to the other procedures. Future studies could explore this further by extending EI through a hybrid approach, i.e., using Shannon information theory aside to AIT. Another possibility is to focus on the employed compression method. Modern compressors such as LZMA, Brotli or neural network compressors might be able to increase its performance. In sum, our proposed information-based encoding measure proved to be useful for assessing where in the brain neural responses differ across experimental conditions.

## References


1. N. M. Timme and C. Lapish, "A tutorial for information theory in neuroscience," *eNeuro*, vol. 5, no. 3, 2018.
2. M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and Its Applications*, 3rd ed., ser. Texts in Computer Science. Springer New York, 2008.
3. R. A. Ince, B. L. Giordano, C. Kayser, G. A. Rousselet, J. Gross, and P. G. Schyns, "A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula," *Human Brain Mapping*, vol. 38, no. 3, pp. 1541–1573, 2017.
4. E. Piasini and S. Panzeri, "Information theory in neuroscience," *Entropy*, vol. 21, no. 1, 2019. [Online]. Available: https://www.mdpi.com/1099-4300/21/1/62
5. J. Fuhrer, K. Glette, J. Ivanovic, P. G. Larsson, T. Bekinschtein, S. Kochen, R. T. Knight, J. Tørresen, A.-K. Solbakk, T. Endestad, and A. Blenkmann, "Direct brain recordings reveal continuous encoding of structure in random stimuli," *bioRxiv*, 2021.
6. J. D. Sitt, J.-R. King, I. El Karoui, B. Rohaut, F. Faugeras, A. Gramfort, L. Cohen, M. Sigman, S. Dehaene, and L. Naccache, "Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state," *Brain*, vol. 137, no. 8, pp. 2258–2270, 06 2014.
7. A. Canales-Johnson, A. J. Billig, F. Olivares, A. Gonzalez, M. d. C. Garcia, W. Silva, E. Vaucheret, C. Ciraolo, E. Mikulan, A. Ibanez, D. Huepe, V. Noreika, S. Chennu, and T. A. Bekinschtein, "Dissociable Neural Information Dynamics of Perceptual Integration and Differentiation during Bistable Perception," *Cerebral Cortex*, vol. 30, no. 8, pp. 4563–4580, 03 2020.
8. M. Schartner, A. Seth, Q. Noirhomme, M. Boly, M.-A. Bruno, S. Laureys, and A. Barrett, "Complexity of multi-dimensional spontaneous eeg decreases during propofol induced general anaesthesia," *PLOS ONE*, vol. 10, no. 8, pp. 1–21, 08 2015.
9. M. Schartner, A. Pigorini, S. A. Gibbs, G. Arnulfo, S. Sarasso, L. Barnett, L. Nobili, M. Massimini, A. K. Seth, and A. B. Barrett, "Global and local complexity of intracranial EEG decreases during NREM sleep," *Neuroscience of Consciousness*, vol. 2017, no. 1, 01 2017.
10. J. Fuhrer, A. Blenkmann, T. Endestad, A.-K. Solbakk, and K. Glette, "Complexity-based encoded information quantification in neurophysiological recordings," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2022, pp. 2319–2323.
11. A. O. Blenkmann, S. Collavini, J. Lubell, A. Llorens, I. Funderud, J. Ivanovic, P. G. Larsson, T. R. Meling, T. Bekinschtein, S. Kochen, T. Endestad, R. T. Knight, and A.-K. Solbakk, "Auditory deviance detection in the human insula: An intracranial eeg study," *Cortex*, vol. 121, pp. 189 – 200, 2019.
12. M. Komatsu, K. Takaura, and N. Fujii, "Mismatch negativity in common marmosets: Whole-cortical recordings with multi-channel electrocorticograms," *Scientific Reports*, vol. 5, no. 1, p. 15006, 10 2015.
13. A. Canales-Johnson, A. F. Teixeira Borges, M. Komatsu, N. Fujii, J. J. Fahrenfort, K. J. Miller, and V. Noreika, "Broadband dynamics rather than frequency-specific rhythms underlie prediction error in the primate auditory cortex," *Journal of Neuroscience*, vol. 41, no. 45, pp. 9374–9391, 2021.

14. A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: A review," *Journal of Vision*, vol. 15, no. 6, pp. 4–4, 05 2015.
15. T. W. Picton, M. S. John, A. Dimitrijevic, and D. Purcell, "Human auditory steady-state responses: Respuestas auditivas de estado estable en humanos," *International Journal of Audiology*, vol. 42, no. 4, pp. 177–219, 2003.
16. R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data," *Computational intelligence and neuroscience*, vol. 2011, p. 156869, 01 2011.
17. M. Li, X. Chen, X. Li, B. Ma, and P. Vitanyi, "The similarity metric," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3250–3264, 2004.
18. T. Meigen and M. Bach, "On the statistical significance of electrophysiological steady-state responses," *Documenta Ophthalmologica*, vol. 98, no. 3, pp. 207–232, Jul 1999.
19. A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. S. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.
20. R. Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised Spike Detection and Sorting with Wavelets and Superparamagnetic Clustering," *Neural Computation*, vol. 16, no. 8, pp. 1661–1687, 08 2004.
21. D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
22. A. M. Bastos, J. Vezoli, C. Bosman, J.-M. Schoffelen, R. Oostenveld, J. Dowdall, P. De Weerd, H. Kennedy, and P. Fries, "Visual areas exert feedforward and feedback influences through distinct frequency channels," *Neuron*, vol. 85, no. 2, pp. 390 – 401, 2015.