

# EEG SPATIAL DECODING WITH SHRINKAGE OPTIMIZED DIRECTED INFORMATION ASSESSMENT

Xu Chen, Zeeshan Syed and Alfred Hero

University of Michigan, Ann Arbor, MI 48109

## ABSTRACT

This paper proposes an approach to infer neural interactions from EEG data using a James-Stein estimator of directed information called shrinkage optimized directed information assessment (SODA). SODA uses shrinkage regularization on empirical histograms to deal with the high dimensionality of multi-channel EEG signals and the small sizes of many real-world datasets. It is designed to make few *a priori* assumptions, and can handle both non-linear and non-Gaussian flows across electrode sites. The use of James-Stein shrinkage allows the SODA algorithm to achieve higher sensitivity to directed neural interactions for a given specificity. We augment this through a central limit theorem-based approach that can assess the statistical significance of each discovered interaction. When evaluated on brain computer interface EEG motor activity data the neural decoding obtained using SODA outperformed several state-of-the-art approaches including Granger causality, MI, unregularized directed information, and spatial coherence. Our results show that SODA localizes 30% more directed interactions in regions that are consistent with Brodmann functional areas of motor activity.

**Index Terms**— James Stein estimators, information flow, small sample size, directional interaction graph

## 1. INTRODUCTION

An extensive body of research focuses on the goal of identifying and classifying brain activity using EEG data. Central to these efforts is developing an understanding of how the brain coordinates information processing to achieve specific tasks. Multiple studies have shown that neural activations in certain regions of the human brain have strong interactions [1][2][3]. In recent years, directed information (DI) has grown to be a popular approach to study these interactions. DI provides a decomposition of the mutual information (MI) between EEG signals into causal and anti-causal components. It can model non-linear and non-Gaussian dependencies between different signals [4], and differs in particular from MI by providing an asymmetric function of the time-aggregated feature densities extracted from pairs of measurement sites.

While DI has been demonstrated to be superior to other approaches such as Granger causality and MI [4], it is diffi-

cult to estimate in the presence of high dimensionality and small sample sizes; both issues commonly associated with real-world multi-channel EEG datasets. These issues can introduce severe bias into unregularized empirical estimates of MI [5] (and correspondingly DI [6]). In this paper, we address this issue by introducing an optimally regularized DI estimator. Our approach, shrinkage optimized directed information assessment (SODA), uses a shrinkage regularization that minimizes estimator mean square error and provides asymptotic expressions for estimator bias and variance. We describe a central limit theorem (CLT) that can be used in conjunction with our SODA approach to assess the statistical significance of putative interactions across EEG electrode sites, and to reduce false discoveries (i.e., false positives). To the best of our knowledge, this is the first time that such an approach has been applied for interaction detection in EEG signals.

When evaluated on brain computer interface EEG motor activity data, the directed information graph discovered by SODA was consistent with activation of Brodmann areas of the brain associated with motor function. Compared to unregularized DI by Quinn et al [4], and other state-of-the-art approaches for interaction discovery, SODA had a substantially higher sensitivity to statistically significant information flows under identical false positive rate constraints.

## 2. SODA FRAMEWORK FOR EEG

We introduced the general framework underlying SODA in a recent study on audio-video indexing in multi-modal databases [6]. Here we cast these ideas into the context of neural interaction discovery from EEG data, focusing in particular on the question of how to extend the SODA framework to assess the statistical significance of each discovered neural interaction. We start by recalling the fundamental definitions and properties of DI. Consider two EEG electrodes,  $E_x$  and  $E_y$ , placed at positions  $x$  and  $y$  on the scalp or intracranially, with  $M_x$  and  $M_y$  time points respectively. Denote by  $X_m$  and  $Y_m$  the temporal feature variables extracted at time  $m$  for  $E_x$  and  $E_y$  respectively, and define  $X^{(m)} = \{X_k\}_{k=1}^m$  and  $Y^{(m)} = \{Y_k\}_{k=1}^m$ . The DI from electrode  $E_x$  to electrode  $E_y$

is then a non-symmetric generalization of the MI defined as

$$\text{DI}(E_x \rightarrow E_y) = \sum_{m=1}^M I(X^{(m)}; Y_m | Y^{(m-1)}) \quad (1)$$

where  $M = \min\{M_x, M_y\}$ , and  $I(X^{(m)}; Y_m | Y^{(m-1)})$  is the conditional MI between  $X^{(m)}$  and  $Y_m$  given the past  $Y^{(m-1)}$ . The conditional MI can be expressed in terms of joint entropies as  $H(X^m, Y^{m-1}) - H(Y^{m-1}) - H(X^m, Y^m) + H(Y^m)$ . The maximum likelihood (ML) estimator of the DI is obtained by quantizing the  $M_x + M_y$  dimensional features  $\{X^{M_x}, Y^{M_y}\}$  and computing estimates of these joint entropies from the empirical histogram  $\{z_1, \dots, z_{p^m}\}$ , where  $p$  is the number of levels per dimension,  $z_i$  represents the frequency within the  $i$ th histogram, and  $m = M_x + M_y$ . Under general conditions, the histogram count vector  $Z$  is multinomial distributed  $P_\theta(Z_1 = n_1, \dots, Z_{p^m} = n_{p^m}) = \frac{n!}{\prod_{k=1}^{p^m} n_k!} \prod_{k=1}^{p^m} \theta_k^{n_k}$ , where  $\theta = E[Z]/n = [\theta_1, \dots, \theta_{p^m}]$  is a unknown vector of cell probabilities and  $\sum_{k=1}^{p^m} n_k = n$  with  $n$  corresponding to the number of samples and  $\sum_{k=1}^{p^m} \theta_k = 1$ . Since the number  $p^m$  of quantization cells is larger than the number of trials  $n$ , we must compensate for the overfilling error of the ML estimates of  $\theta$ . To do this, we apply a James-Stein shrinkage approach based on shrinking the ML estimator of  $\theta$  towards a target distribution  $t = [t_1, \dots, t_{p^m+1}]$ , which in our work is chosen to be the uniform distribution consistent with [5]. This results in the regularized ML estimator  $\hat{\theta}_k^\lambda = \lambda t_k + (1 - \lambda)\hat{\theta}_k^{ML}$ , where  $\lambda \in [0, 1]$  is a shrinkage coefficient. Following [5], for a given  $\lambda$ , the entropy estimator for a single sample  $S$  of an EEG electrode  $E_X$  is defined as  $\hat{H}_{\hat{\theta}^\lambda}(S) = -n \sum_{k=1}^p \hat{\theta}_k^\lambda \log(\hat{\theta}_k^\lambda)$ . However, unlike the James-Stein entropy estimator [5], for the proposed James-Stein DI estimator,  $\lambda$  is selected to minimize the MSE of the estimated DI, denoted by  $\widehat{\text{DI}}^\lambda = \text{DI}_{\hat{\theta}^\lambda}(E_x \rightarrow E_y)$ , which is expressed as sums and differences of estimated entropies. The following is proven in [6].

**Definition:** The optimal DI shrinkage parameter  $\lambda$  that minimizes estimator MSE is:  $\lambda^\circ = \arg \min_\lambda \left\{ C_1^2 + (2C_1C_2 + T_2\Sigma_2T_2')/n + O(n^{-2}) \right\}$ .

In the above,  $C_1, C_2, T_2$  and  $\Sigma_2$  are constants that can be estimated from the EEG data. The resultant James-Stein DI estimator,  $\widehat{\text{DI}}^{\lambda^\circ}(X^M \rightarrow Y^M)$ , is called the SODA estimator. The SODA estimator satisfies a CLT:

**Theorem:** Let  $\Phi(x)$  be the standard normal distribution function. Then asymptotically in the number of samples, the standard normal random variable

$$Pr\left(\frac{\widehat{\text{DI}}^\lambda - \mathbf{E}[\widehat{\text{DI}}^\lambda]}{\sqrt{\text{Var}(\widehat{\text{DI}}^\lambda)}} \leq \alpha\right) \rightarrow \Phi(\alpha). \quad (2)$$

Due to space limitations, the expressions for the mean and variance of  $\widehat{\text{DI}}^\lambda$  in the theorem are not given here. The theo-

rem will be used to perform significance testing of discovered directional interactions in the EEG.

In [6], a local version of DI was introduced for temporal interaction localizations. This is an important step while studying physiological signals such as the EEG, due to issues related to time warping inherent in these data. Here we describe this algorithm in the context of EEG. The local DI is defined similarly to the DI except that for a pair of EEG signals  $X$  and  $Y$ , the signals are time shifted and windowed prior to DI computation. Specifically, let  $\tau_x \in [0, M_x - T]$  and  $\tau_y \in [0, M_y - T]$  be the respective time shift parameters, where  $T$  is the sliding window width, and denote by  $X_{\tau_x}^{M_x}$  and  $Y_{\tau_y}^{M_y}$  the time shifted sequences. Then the local DI,  $\text{DI}(X_{\tau_x}^{M_x} \rightarrow Y_{\tau_y}^{M_y})$ , defines a surface over  $\tau_x$  and  $\tau_y$ , and summation indices range over smaller sets of  $T$  time samples. We use peaks of the local DI surface to detect and localize the spatial-temporal interactions in pairs of EEG signals.

The overall process for SODA-based interaction discovery in EEG data is as follows: (1) **Temporal Alignment:** Align the EEG signals temporally by segmenting the EEG data according to local DI peak locations. (2) **Pairwise DI and p-value computation:** After alignment, calculate the  $K \times K$  matrices of SODA estimated DIs and p-values  $1 - \Phi\left(\frac{\widehat{\text{DI}}_{ij} - \mu_{ij}}{\sigma_{ij}}\right)$  on these DI estimates. (3) **False Discovery Rate Control:** Threshold the DI and p-value matrices to find interaction regions exhibiting large and statistically significant DI. The construction of the interaction graph over the  $K$  EEG electrodes is performed by testing the  $K \times (K - 1)$  hypotheses that there is a significant interaction (both directions) between pairs of electrodes. This is a multiple hypothesis testing problem and we control false discovery rates using the corrected Benjamini-Hochberg (BH) procedure [7].

### 3. EXPERIMENTAL RESULTS

The SODA algorithm was applied to a publicly available BCI EEG motor activity dataset [8]. The EEG consisted of random movements of the left and right hand recorded with eyes closed. The data consisted of multiple data matrices corresponding to multiple activities, where each column of a data matrix represented one electrode and there were a total of 19 electrodes and 3008 samples in each row. The motor activity lasted about 6 seconds. The sample rate of the recording was 500Hz. The subject executed 10 classes of movements where each class contained different trials of the same movement (e.g., three trials of left hand forward movement, three trials of left hand backward movement, etc.) The performance of the SODA-based interaction detected was compared to four state-of-the art approaches: Granger causality [3], coherency measure [2], MI [1] and unregularized DI by Quinn et al. [4]. In [2], coherency was defined as normalized cross-spectrum between two EEG signals, where only the imaginary part of the signal was employed. In [1], mutual information was

applied to feature selection for EEG signal classification. In [4], Quinn et al. utilized unregularized directed information to capture the non-linear and non-Gaussian dependency structure of spike train recordings. The time sequences were first divided into segments of 200ms length for feature extraction. There was 100ms overlap between neighboring segments. We estimated the joint probability density functions for each segment of EEG data by first mapping the features to a codebook by quantization as described in Sec.2. Then we applied the proposed shrinkage method to the ML estimator. Here the number of samples  $n$  was the total number of trials for all the subjects performing the same task, and the Lloyd-max quantization level in the scalar quantizer was selected to be 10 for multinomial distribution.

**Interaction Detection and Comparison:** Fig. 1 presents a visual illustration of the DI matrices (expressed as heatmaps) obtained through SODA for left hand forward movement, left hand backward movement, and right hand forward movement respectively. Colors in the heatmaps indicate the magnitudes for different strengths of interactions between the 19 electrodes. In the interest of space, we discuss the results for left hand forward movement in more detail below (these results are representative of other movements). We utilized a heat kernel to transform the (symmetrized) pairwise DI matrix into the distance matrix and applied K-means clustering with the number of clusters (3) chosen by setting a threshold to within-cluster sums of point-to-centroid distances. For left hand forward movement, the three clusters discovered were  $(C_3, C_4, T_3, T_5)$ ,  $(F_4, FP1, FP2)$  and the rest of electrodes. Mapping the EEG channels into Brodmann areas [9], we identified cluster  $(C_3, C_4, T_3, T_5)$  as reflecting auditory processing such as that associated with detecting a cue to start motion (Brodmann area 21) and the execution of motor function (Brodmann area 4). Similarly, we identified cluster  $(F_4, FP1, FP2)$  as corresponding to the planning of complex movements (Brodmann area 8) and cognitive branching (Brodmann area 10). The third cluster corresponded to electrodes that were not very active. The localization accuracy was highlighted for each heatmap with red numbers in Fig. 1. In Fig. 1, the order of the electrodes by clustering is  $T_3, T_5, C_3, C_4, T_4, F_8, CZ, F_7, F_3, T_6, PZ, P_3, P_4, O_1, O_2, FZ, F_4, FP_1, FP_2$  from left to right and from top to bottom.

Fig. 2 compares the neural interaction graph obtained through SODA with other state-of-the-art approaches. The interactions shown for SODA are significant at the 5% level after accounting for multiple hypotheses as described above. The regions in the brain for Fig. 1 that exhibited the highest activity match perfectly with the regions in Fig. 2 that have the densest number of links. Fig. 2 indicates that SODA discovers significantly more new interactions than other approaches, and that these interactions are significant according to statistical tests. For instance, the edges between  $(FP1 \rightarrow F_4)$  and  $(FP1 \rightarrow F_7)$  corresponded to adjusted  $p$ -values of 0.026 and 0.036 respectively. The results of applying SODA to the

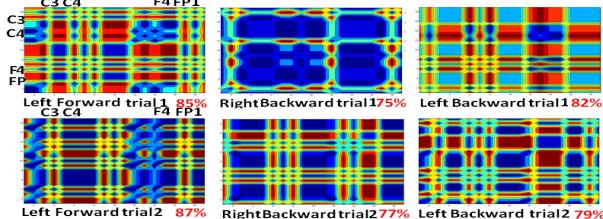
replicates of EEG signals with different trials also indicated that during these periods, the electrodes  $FP1, F_4$  and  $F_7$  are highly interactive and therefore can serve as strong evidence that the activity was indeed being localized to these electrodes in the brain associated with motor control. Compared to the unregularized DI [4], SODA has the advantage that it can control false positive rate more accurately with optimal shrinkage regularization and its predictions are validated by neural pathway locations as determined by Brodmann areas.

**Consistency Measure:** To study the ability of SODA to uncover interactions that were consistently observed during the same class of activity, we randomly divided the data into equal sized training and test sets. SODA was applied to the training set and the localization consistency was computed by mapping the detected interactions to the testing set. Table 1 compares the different algorithms for varying false discovery rates (FDR) including MI, Granger causality where the covariance matrix was regularized with Ledoit Wolf shrinkage method, coherence measure, unregularized DI and SODA. A localization consistency of 100% means that all interactions discovered on the training set were observed in the test set. As shown in Table 1, SODA consistently outperformed other methods over a range of practical FDR thresholds (0.1 to 0.05), with this improvement increasing in particular as the FDR threshold was lowered. Since the weak dependencies were filtered out and strong dependencies remained with a lower threshold, false positives were significantly reduced with SODA. This explains the fundamental reason that SODA achieved the best performance with the threshold for declaring an edge present corresponding to FDR of level 0.05. Compared to the next best performing method for interaction localization (unregularized DI), SODA improved the localization consistency by about 5%.

We also assessed localization consistency in terms of interactions within regions of the brain coordinating similar task related behavior (i.e., as opposed to between all EEG electrode sites). For an FDR of 0.1, SODA discovered 36 interactions, while 23, 27 and 31 interactions are detected using MI, CM and unregularized DI. Together, the set of results presented here demonstrate that SODA is able to both discover more statistically significant neural interactions from EEG data, and to discover interactions that are consistent across multiple observations of the same behavior.

## 4. CONCLUSIONS

We proposed applying a James-Stein estimator of directed information, called SODA, for EEG signal interaction detection and classification based on directed information. We illustrated the SODA estimator for EEG signals interaction detection/localization using a publicly available brain computer interface EEG motor activity database. Our results, relative to other state-of-the-art algorithms, show that SODA provides interaction estimates that are consistent with neural pathway



**Fig. 1.** Visual illustration of SODA heatmaps with Project BCI dataset for left hand forward/backward movement and right hand forward movement, where colors indicate the magnitudes for different strengths of interactions calculated by SODA between 19 electrodes. SODA is able to detect more interactions than MI, such as the interactions between *FP1* and *F<sub>7</sub>*, *FP1* and *FZ*. The red number indicates localization accuracy computed by mapping the detection results using SODA on different replicates of the same class of activity.

FDR	0.1	0.07	0.05
MI	0.641	0.657	0.676
GC	0.653	0.705	0.728
CM	0.657	0.694	0.726
uDI	0.669	0.721	0.743
SODA	<b>0.698</b>	<b>0.755</b>	<b>0.809</b>

**Table 1.** Comparison of EEG localization consistency for different levels of significance (FDR thresholds). The consistency is computed by mapping the detection results using SODA on different replicates and the number of these electrodes that are connected in the MI, GC, CM, uDI and SODA interaction graphs, determined by thresholding these quantities at the same FDR level. uDI, CM and GC represents unregularized DI, coherence measure and Granger causality.

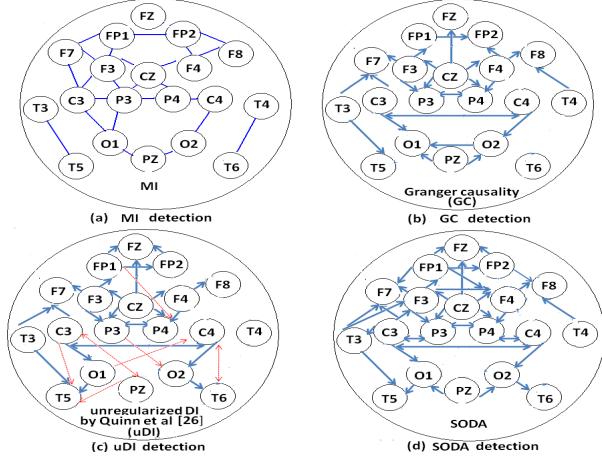
locations of motor activities as determined by Brodmann area. We demonstrate that SODA provides better performance as compared to unregularized DI, MI and coherence measure.

## 5. ACKNOWLEDGEMENT

This work was partially supported by ARO grant W911NF-09-1-0310 and NSF grant IIS-1054419.

## 6. REFERENCES

- [1] L. Tian, D. Erdogmus, A. Adama, and M. Pavel, “Feature selection by independent component analysis and mutual information maximization in eeg signal classification,” in *International Joint Conference on Neural Network*, 2005.
- [2] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett, “Identifying true brain interaction from eeg data using the imaginary part of coherency,” in *Clinical Neurophysiology*, 2004, vol. 115.
- [3] W. Hesse, E. Moller, M. Arnold, and B. Schack, “The use of time-variant EEG granger causality for inspecting directed interdependencies of neural assemblies,” in *Journal of Neuroscience Methods*, 2003, vol. 124.
- [4] C.J. Quinn, T.P. Coleman, N. Kiyavash, and N.G. Hat-sopoulos, “Estimating the directed information to infer causal relationships in ensemble neural spike train recordings,” in *computational neuroscience*, 2011.
- [5] J. Haussler and K. Strimmer, “Entropy inference and the james-stein estimator, with application to nonlinear gene association networks,” in *JMLR*, 2009, vol. 10.
- [6] X. Chen, A. Hero, and S. Savarese, “Multimodality video indexing and retrieval using directed information,” in *IEEE Transactions on Multimedia*, 2001.
- [7] Y. Benjamini and D. Yekutieli, “The control of the false discovery rate in multiple testing under dependency,” in *Ann Stat*, 2001, vol. 29.
- [8] “<http://sites.google.com/site/projectbci/>,” in *EEG motor activity data set*.
- [9] L. J. Garey, “Brodmann’s localisation in the cerebral cortex,” in *Springer*, 2006.



**Fig. 2.** Visual illustration of the dependencies between different electrodes reconstructed using MI, Granger causality, unregularized DI by Quinn et al. [4] and SODA averaged over the total 6 seconds. The red arrows are false positives with unregularized DI by Quinn et al. [4] which do not have motor-related functions validated by neural pathway locations as determined by Brodmann areas.