Supplementary Material for: Global and local complexity of intracranial EEG decreases during NREM sleep

Michael M. Schartner*, Andrea Pigorini, Steve A. Gibbs, Gabriele Arnulfo, Simone Sarasso, Lionel Barnett, Lino Nobili, Marcello Massimini, Anil K. Seth, Adam B. Barrett
*m.schartner@sussex.ac.uk (correspondence)

November 30, 2016

Contents

1 Supplementary measures for global analysis

In addition to ACE, SCE and LZc, we computed other measures for the global analysis of 18 channels per subject. These were normalized spectral power bands, LZsum (the mean Lempel-Ziv complexity of single channels) and sumCov (the mean of the absolute values of correlation coefficients between all channels). Tab. S1 summarizes the results. Normalized spectral power bands were obtained by fast Fourier transform of the segment, averaged over the channels and grouped into frequency bands, normalized such that all bands' scores together sum to one. Frequency bands are defined - following convention - as δ =1-4Hz, θ=4-8Hz, α =8-13Hz, β =13-30Hz, γ =30-70Hz. As expected, normalized delta power is consistently substantially higher in NREMe than WR, while beta and gamma power are in almost all cases substantially lower. LZsum behaved very similarly to LZc for state pairs WR/NREMe and REM/NREMe, yet was slightly less consistent at discriminating between WR or REM and NREMl. sumCov's behavior slightly resembled that of delta power, yet was overall less consistent and with weaker effect sizes across subjects.

	ACE	SCE	LZc	LZsum	sumCov	δ	θ	α	β	γ
WR/NREMe	**10 0 0**	**10 0 0**	**10 0 0**	**10 0 0**	0 4 6	**0 0 10**	5 1 4	7 2 1	9 1 0	9 0 1
WR/NREMl	8 2 0	8 2 0	8 2 0	6 4 0	0 5 5	1 0 9	4 4 2	7 3 0	7 2 1	5 4 1
WR/REM	4 5 1	2 8 0	1 9 0	4 4 2	0 7 3	0 2 8	4 3 3	7 3 0	3 5 2	2 3 5
REM/NREMe	**10 0 0**	**10 0 0**	9 1 0	9 1 0	1 2 7	0 1 9	4 2 4	5 2 3	**10 0 0**	8 2 0
REM/NREMl	9 1 0	8 2 0	8 2 0	7 2 1	1 1 8	1 3 6	2 7 1	3 3 4	7 2 1	5 3 2
NREMl/NREMe	7 2 1	8 2 0	2 8 0	8 2 0	1 8 1	0 2 8	2 5 3	6 1 3	8 2 0	7 3 0

Table S1: **Comparisons of broader set of measures per state pair, using 18 channels per subject.** For each measure and state pair, the three numbers correspond to how many subjects out of 10 had higher score for the left state with Cohen's $d > 0.8$ (left digit), no substantial difference, $d < 0.8$, (middle) and higher score for the right state with Cohen's $d > 0.8$ (right). The results were obtained from applying the measures to the broadband signal of 18 channels.

2 Correlation between measures

Correlations between ACE, SCE, LZc and spectral power in the 5 canonical frequency bands were analysed for each state separately and the complete results summarised in Tab. S2. Scatter plots for each subject and selected measure pairs are shown in Fig. S1, illustrating variation across subjects, with subject 7 as a visible outlier. Correlations between all three complexity measures exist, showing that ACE, SCE and LZc are similarly sensitive to certain signal features, yet variations across states are also apparent. These imperfections in correlations indicate that the three measures capture not entirely identical types of spatio-temporal signal diversity. Strong correlations were also found between complexity measures and spectral power in certain bands - especially delta - showing that the measures are affected by the spectral profile of the signal, although as we show elsewhere the decrease in measured complexity during NREMe compared to WR does not only arise due to spectral changes (see Fig. S8 and Fig. 5c).

Measure pair	WR	REM	NREMl	NREMe
ACE/LZc	0.71	0.6	0.86	0.72
ACE/SCE	0.72	0.73	0.84	0.71
ACE/delta	-0.56	-0.6	-0.76	-0.66
ACE/theta	-0.27	n.s.	n.s.	n.s.
ACE/alpha	n.s.	n.s.	n.s.	0.29
ACE/beta	0.5	0.43	0.62	0.5
ACE/gamma	0.57	0.66	0.65	0.54
LZc/SCE	0.51	0.39	0.62	0.37
LZc/delta	-0.32	-0.29	-0.56	-0.31
LZc/theta	n.s.	n.s.	n.s.	n.s.
LZc/alpha	n.s.	n.s.	n.s.	n.s.
LZc/beta	0.23	0.18	0.4	n.s.
LZc/gamma	0.28	0.26	0.45	0.26
SCE/delta	-0.46	-0.62	-0.82	-0.75
SCE/theta	n.s.	n.s.	n.s.	n.s.
SCE/alpha	n.s.	n.s.	n.s.	0.33
SCE/beta	0.44	0.48	0.69	0.65
SCE/gamma	0.36	0.5	0.7	0.62

Table S2: **Pearson correlation for all measure pairs, pooled across subjects.** For a given state and subject the Pearson correlation of two different measures' scores across 10s segments was computed. For any state and pair of measures (choosen from ACE, SCE, LZc and spectral power of the 5 canonical frequency bands, excluding pairs of spectral power alone) the value of the Pearson correlation averaged across subjects, r, is displayed if $p < 0.05$ (obtained by a FDR corrected t-test across the values for the 10 subjects, setting 0 correlation as null hypothesis). If the p-value was above 0.05, the table entry was set to not significant (n.s.).

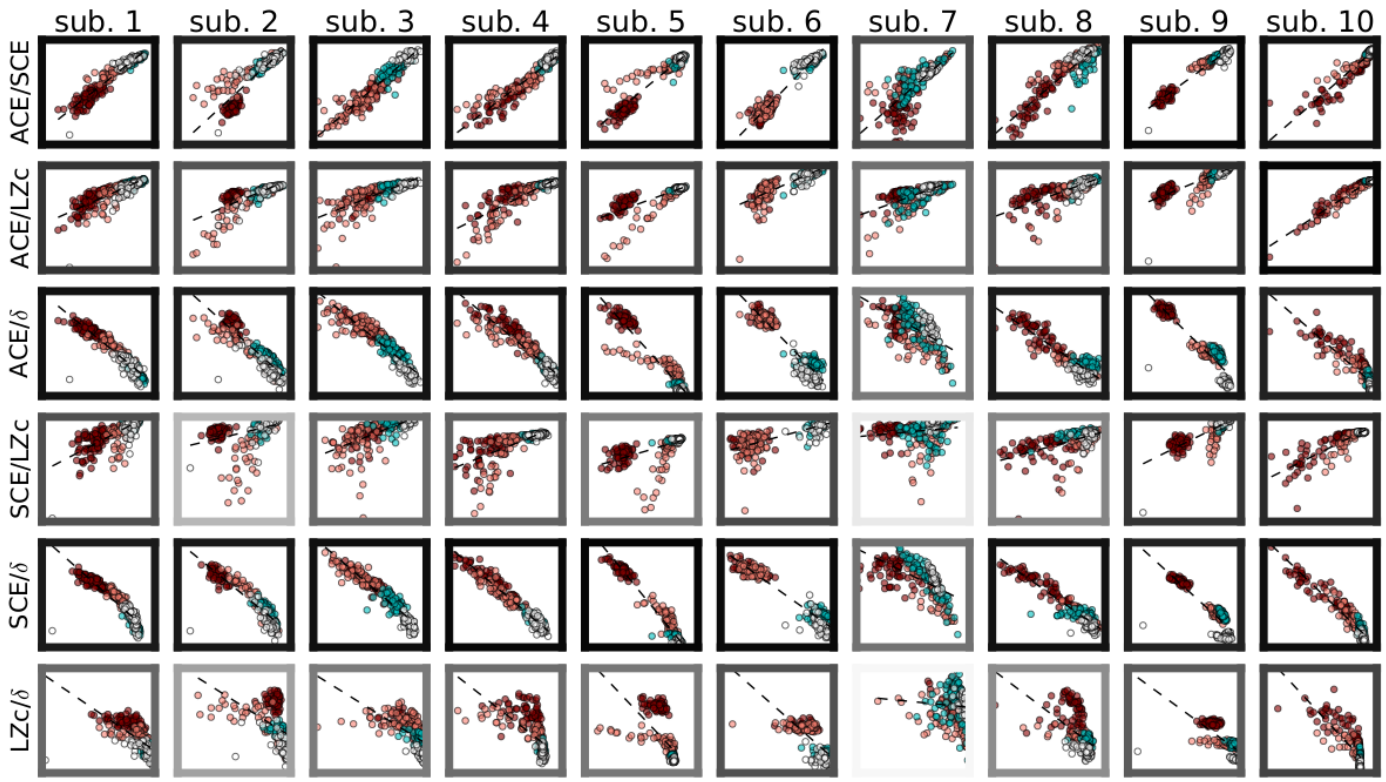

Figure S1: **Scatter plots illustrating correlation of pairs of measures**. Each panel shows a scatter plot for a pair of measures (chosen from ACE, LZc, SCE and normalized delta power) and one of the ten subjects. Each dot - coloured by state as in main text figures - has as x value the score for a 10s segment of the first measure of the pair indicated for each row and as y value the score of the second measure. The best linear fit is indicated with a dotted line. The frame of each panel is colored according to the absolute value of the Pearson correlation coefficient; white signifying 0, black signifying 1 and linearly scaled grey tones showing values in between. Measure pairs ACE/SCE, ACE/LZc and SCE/delta show high correlation across subjects. Subject 7 shows low correlation for all displayed measure pairs, demonstrating that the complexity measures are not all sensitive to the same dynamics. (compare with Tab. S2)

3 Local analysis

As described in the main text, Lempel-Ziv complexity of single channels (LZs, same computation as LZc with trivial concatenation) scored predominantly lower during NREMe than WR, with some outliers. For each subject, a panel in Fig. S2 displays LZs' scores for all 4 states (WR, REM, NREMl, NREMe), for all channels. Across all subjects, a total of 247 channels were analysed. For 206 of those LZs scored higher for WR than NREMe with Cohen's $d > 0.8$, for 35 $d < 0.8$ and for 6 LZs scored lower for WR than NREMe with $d > 0.8$. The 6 outliers were found in subject 8 (Hi, In)[1], subject 9 (Hi, Oc) and subject 10 (Fr, In), i.e. different regions and different subjects, thus they do not indicate a consistent local deviation from the global trend. When considering LZs scores per subject and region, averaged across channels in the same region, values were lower during NREMe than WR for 48 out of 50 subject region pairs (see Tab. S3).

We repeated the above analysis for normalised delta power, see Fig. S3. For 210 of the 247 channels, delta(WR)<delta(NREMe) with Cohen's $d > 0.8$, for 26 $d < 0.8$ and for 11 delta(WR)>delta(NREMe) with $d > 0.8$. All the 11 channels for which delta behaved against the main trend were in subject 7, and were located in the regions Hi, Oc and Te. We note that for all of these 11 channels LZs by contrast did behave with the trend, LZs(WR)>LZs(NREMe) (compare Fig. S2 with Fig. S3), further adding evidence for complexity not just reflecting the power spectrum. Considering delta power per subject averaged across all channels from the same region in a given subject, delta power is higher for NREMe than WR for 49 out of 50 region/subject pairs.

Further local analyses involved exploring the complexity, as measured by $SCE^{(i)}$, of the interaction of a seed channel in one region with a group of target channels in another (see Fig. S4). We call the so-obtained measure SCE1N for the case of N target channels and computed it for $N = 2$ and $N = 3$ for all possible region pairs (seed channel in one region, N target channels in another region) for each subject. If there were multiple channel choices for a given region pair 10 different random combinations were chosen per $10s$ segment and results averaged across them. We also computed complexity of synchrony (CS) as being Lempel Ziv compression of the synchrony time series of two channels, thus capturing temporal signal diversity thereof. The scores of SCE13, SCE12 and CS can be compared between pairs of regions for single subjects with results shown in Figs. S5 to S7, respectively.

For each of these three measures of the complexity of the interaction of a channel in one region with channels in another region, we observed lower values during NREMe than WR for almost all choices of seed and target regions for all subjects. When using Cohen's $d > 0.8$ as a criterion for strong effect size, SCE13, computed for 129 pairs of regions across all subjects, scored higher for WR with $d > 0.8$ for 24, for 103 $d < 0.8$ and for 2 lower for NREMe than WR with $d > 0.8$. This overall result across subjects for SCE13 can be expressed as a triplet: $\{24, 103, 2\}$. The analogous triplet for SCE12 is $\{11, 130, 2\}$ and that for CS is $\{98, 33, 8\}$ (for CS we included also reflexive region pairs). The high fraction of cases with small effect size for SCE13 and SCE12 indicates that the small number of possible synchrony coalitions for the case of just 2 or 3 target regions is insufficient to render these measures effective indices of complexity. There were no exceptional choices of seed and target region that went against the main trend for more than one subject. Hence this SCE1N analysis did not yield evidence for a change in score against the main trend involving some region consistently in more than one subject, further supporting that the signal complexity decrease during sleep occurs both globally and locally.

Tab. S3 summarises the consistency across regions of all the measures analysed locally.

[1]Macro-region label abbreviations: Fr (frontal), Pa (parietal), Te (temporal), Oc (occipital), Ci (cingulate), In (insula), Hi (Hippocampus).

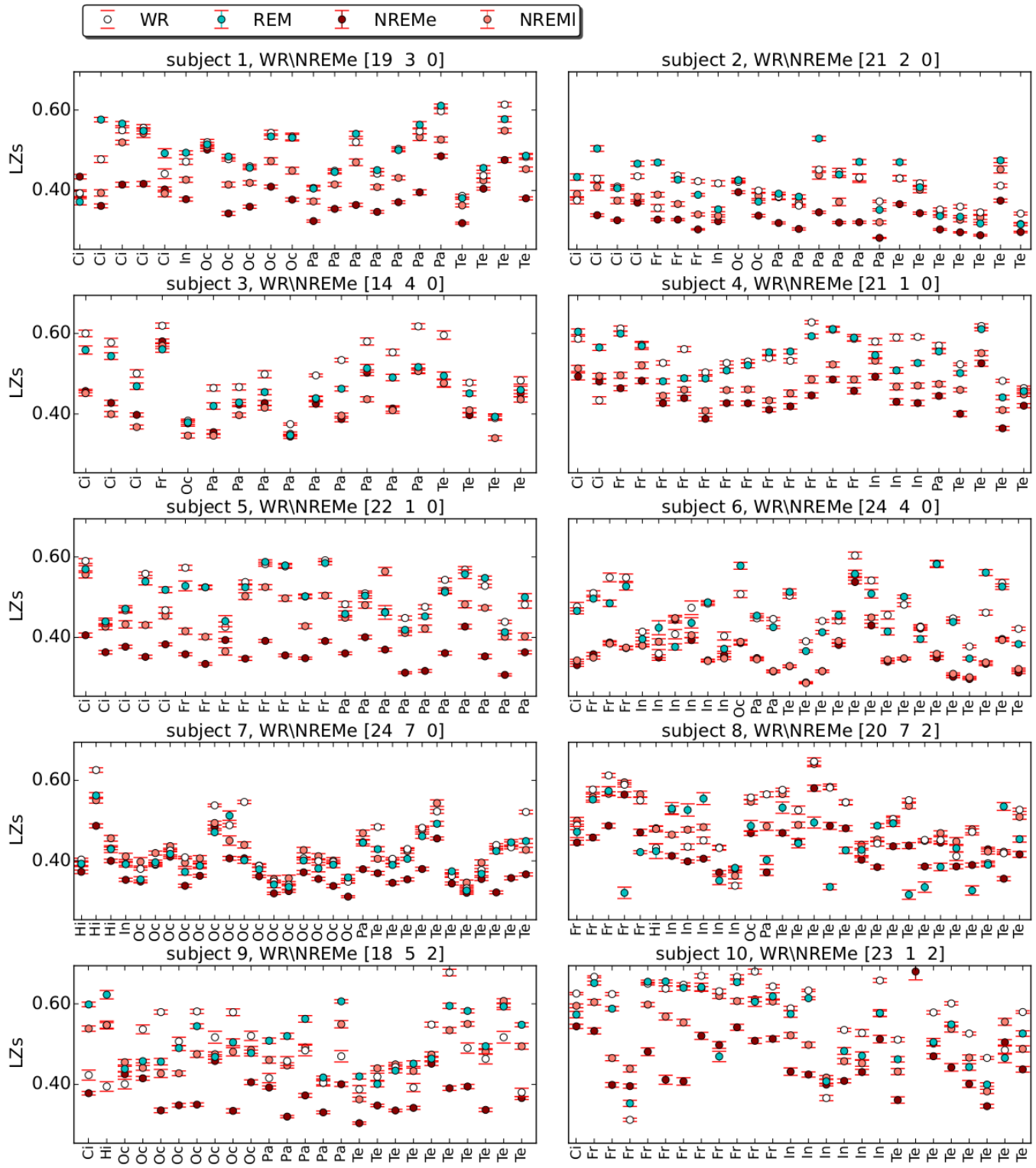

Figure S2: **Lempel Ziv complexity of single channels (LZs) for each subject and state.** States shown are WR, REM, NREMl and NREMe. LZs was computed across all 18-31 channels per subject, the channel's region is indicated as an x-tick label (first 2 letters of the anatomical regions as listed in Fig. 2). Error bars indicate standard error across 10s segments. All 10 panels share the same y-axis. The numbers in the triplet in the title of each panel indicate respectively how many channels scored higher (with Cohen's $d > 0.8$) for WR than NREMe (left digit), non-substantially different (Cohen's $d < 0.8$, middle digit) or lower (with Cohen's $d > 0.8$, right digit). For the clear majority of channels, LZs scored substantially higher for WR as compared to NREMe. The few outliers are not consistently in one region, showing that there is no local effect against the global trend (unlike for delta power, compare with Fig. S3).

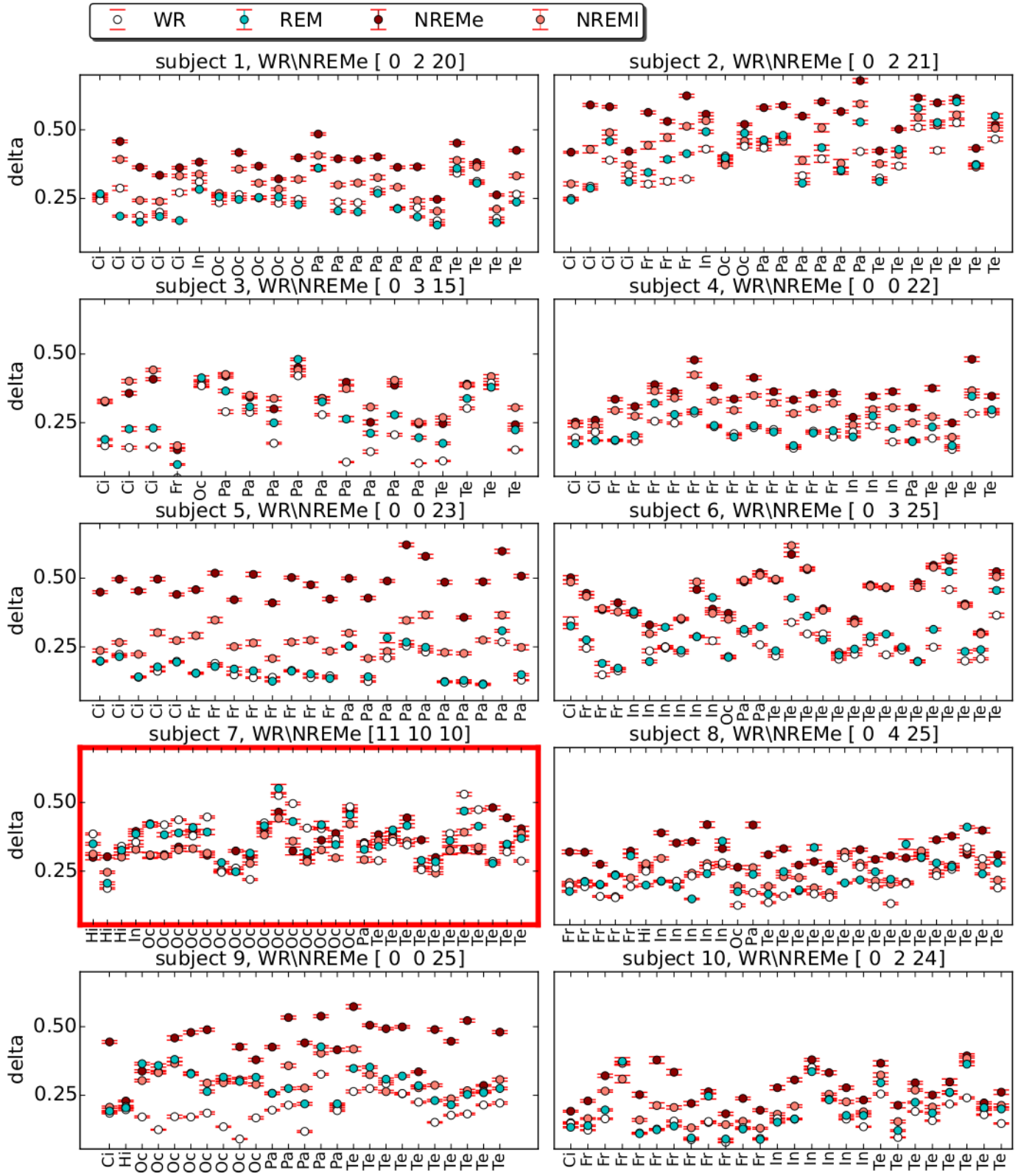

Figure S3: **Normalized delta power of single channels for each subject and state.** States shown are WR, REM, NREMl and NREMe. Delta power was computed across all 18-31 channels per subject, the channel's region is indicated as an x-tick label (first 2 letters of the anatomical regions as listed in Fig. 2). Error bars indicate standard error across 10s segments. All 10 panels share the same y-axis. The numbers in the triplet in the title of each panel indicate respectively how many channels scored higher (with Cohen's $d > 0.8$) for WR than NREMe (left digit), non-substantially different (Cohen's $d < 0.8$, middle digit) or lower (with Cohen's $d > 0.8$, right digit). For all regions in all subjects (except occipital lobe in subject 7 - 10 channels, panel in red frame), delta power is lower for WR as compared to NREMe.

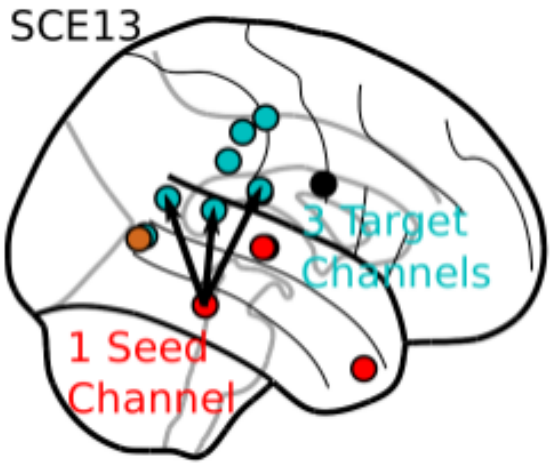

Figure S4: **Schematic of the computation of SCE13**. This measure quantifies the complexity of the interaction between a seed channel in one region with three target channels in another region using the entropy $SCE^{(i)}$ of the constitution of the set of target channels in synchrony with the seed channel.

State pair	ACE	SCE	LZc	LZs	SCE13	SCE12	CS	delta	gamma
WR/NREMe	26 1	20 7	27 0	48 2	107 22	109 34	122 17	1 49	42 8
WR/NREMl	26 1	16 11	27 0	41 9	93 36	96 47	103 36	6 44	33 17
WR/REM	14 13	19 8	17 10	36 14	70 59	77 66	101 38	14 36	18 32
REM/NREMe	25 2	17 10	27 0	46 4	102 27	114 29	105 34	2 48	42 8
REM/NREMl	26 1	16 11	24 3	36 14	90 39	98 45	86 53	11 39	35 15
NREMl/NREMe	18 9	19 8	25 2	42 8	94 35	90 53	122 17	6 44	47 3

Table S3: **Summary of local results (without applying statistics).** Considering all subject and region combinations, the left-hand number in each cell gives the total occurrences of the state on the left scoring higher, and the right-hand number gives the total occurrences of the state on the right scoring higher. ACE, LZc and SCE were computed across 4 channels (thus for these measures there were 27 subject and region combinations in total, see Fig. 7). For SCE12 and SCE13 there were respectively 129 and 143 combinations of subject, target region and seed region, and for CS there were 139 combinations. LZs, delta power and gamma power were computed for each single channel in each subject (247 channels in total, resulting in 50 region-subject combinations when averaging per subject across channels in the same region).

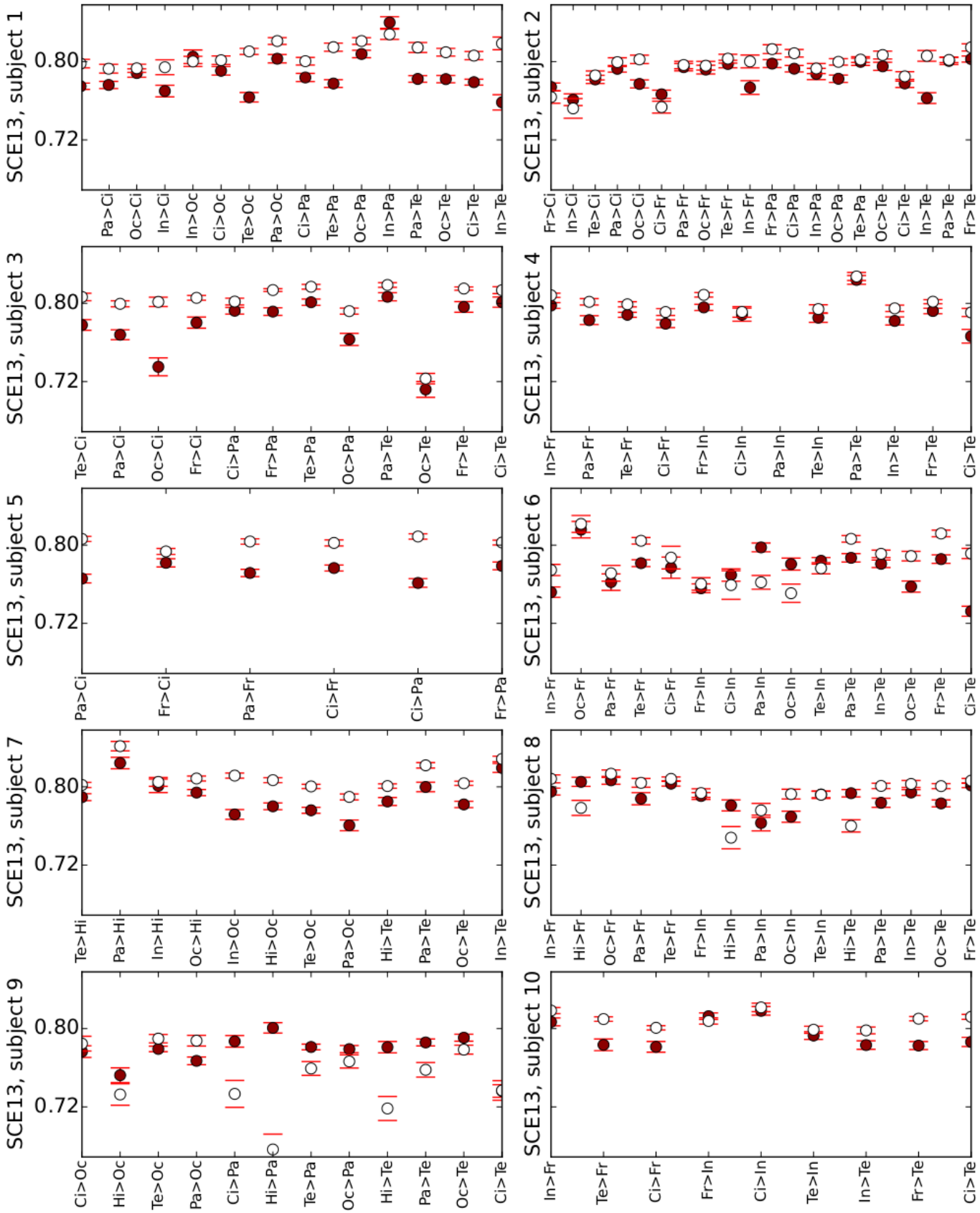

Figure S5: **SCE13 computed for each possible combination of seed and target regions for each subject.** The labels along the x-axis give the pairs of anatomical regions, abbreviated as in the previous figures, with e.g. Oc>Fr denoting that the seed is located in the occipital lobe and the targets are located in frontal lobe.

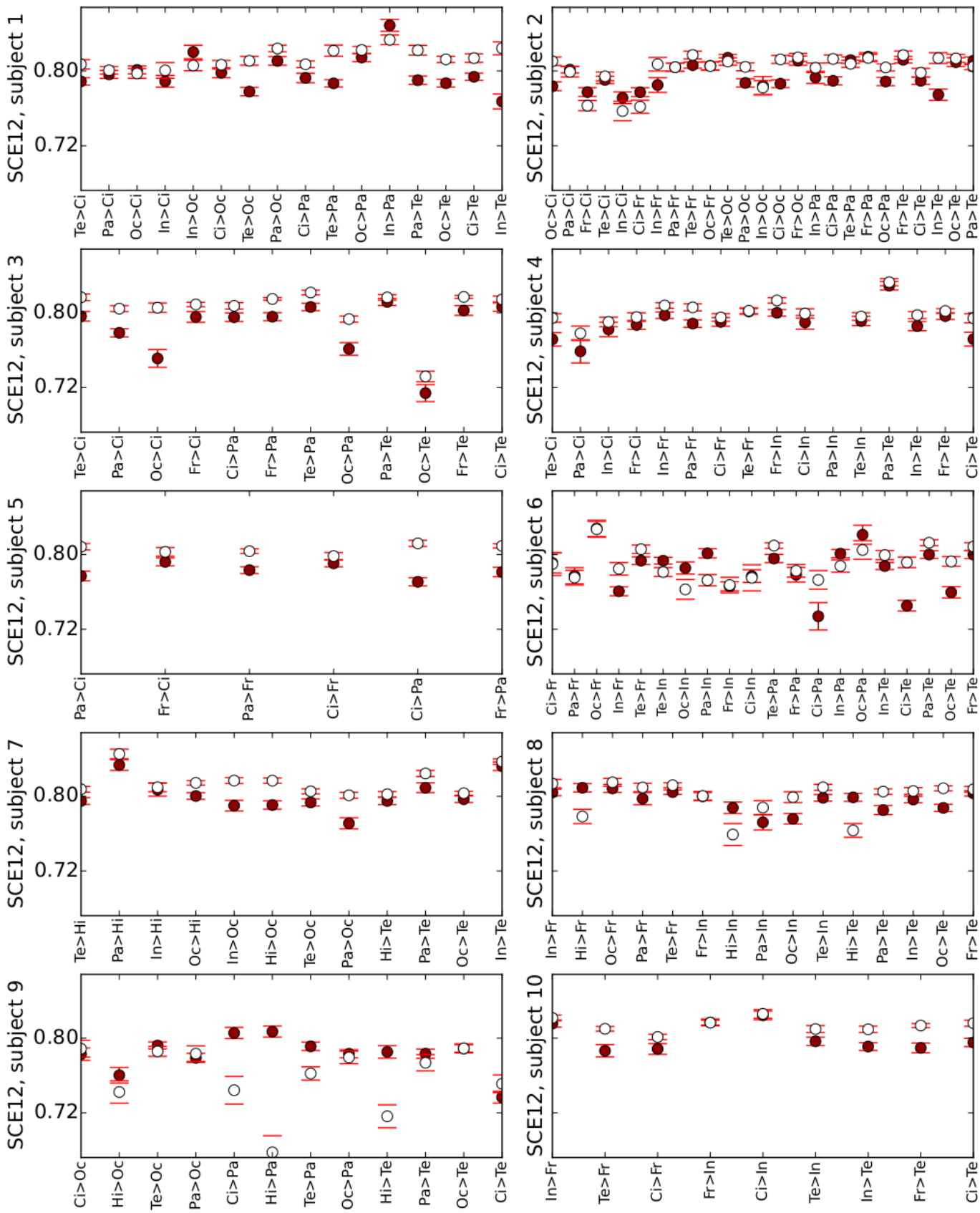

Figure S6: **SCE12 computed for each possible combination of seed and target regions for each subject.** The labels along the x-axis give the pairs of anatomical regions, abbreviated as in Fig. S5.

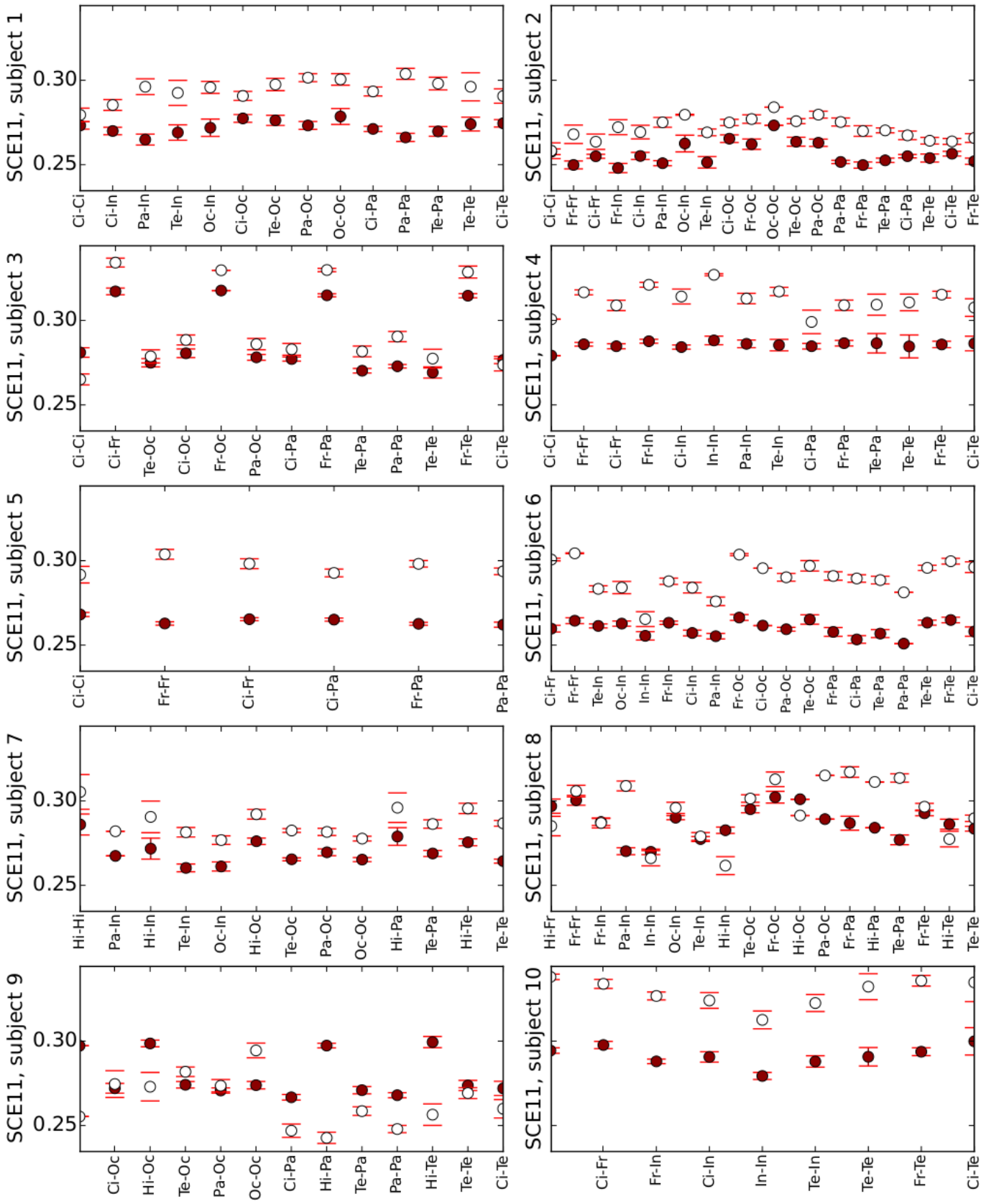

Figure S7: **CS computed for each non-ordered pair of regions (including reflexive pairs) for each subject.** The labels along the x-axis give the pairs of anatomical regions, abbreviated as in previous figures.

4 Further Controls

Fig. S8 compares the behaviour of the complexity measures computed in the default way (using normalisation by score for time-shuffled data) with their behaviour after phase-randomised surrogate data renormalisation (as described in Section 3.1, main text), subject by subject.

Fig. S9 shows ACE_N, SCE_N and LZc_N scores computed for channel quartets restricted to single regions, i.e., a replot of the graphs shown in Fig. 8 but using the phase-randomised surrogate renormalisation of the complexity measures.

Fig. S10 compares the measures' scores computed on data where bipolar referencing was applied (identical to Fig. 5) with scores obtained when omitting the bipolar referencing step. This shows that bipolar referencing as a pre-processing step of the data is not crucial for the detection of complexity changes between WR and NREMe.

Fig. S11 shows the results of our tests of the behaviour of the complexity measures for data with different sampling rates.

Fig. S12 shows the results of our tests of the behaviour of the complexity measures for different length data segments.

Tab. S4 shows the results from repeating the analysis of the complexity measures across globally distributed sets of channels as in Fig. 5a for different random selections of channels.

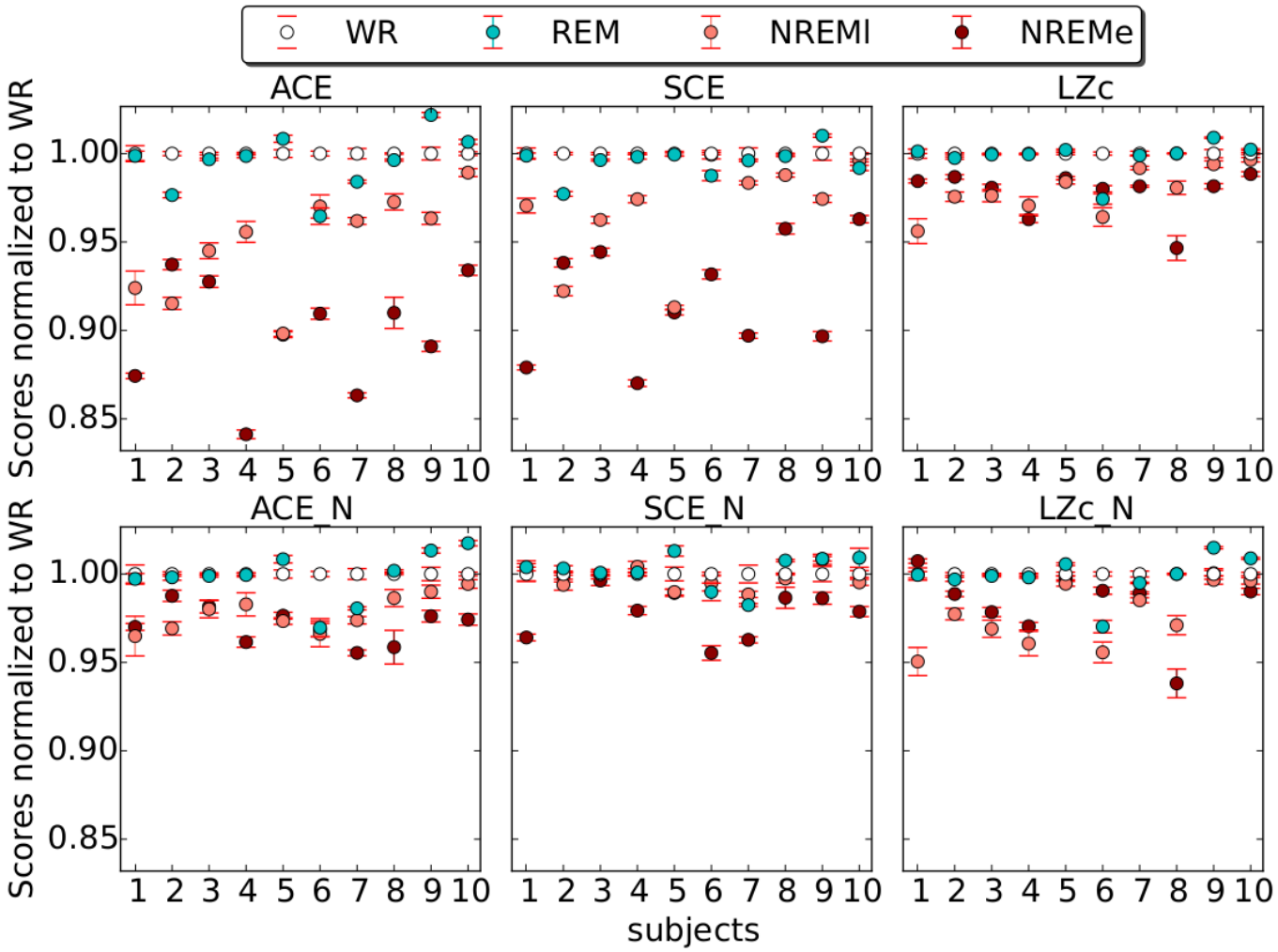

Figure S8: **Control for changes in spectral power profile.** (Top) ACE, SCE and LZc computed across 18 channels, normalized using their scores for time-shuffled data, as in Fig. 5. (Bottom) ACE_N, SCE_N and LZc_N, which are the same measures but normalized using their scores for phase-shuffled data. The differences in scores between WR and NREMe are only partially attenuated by applying the phase-shuffling normalisation, and thus can be attributed to changes in signal properties that go beyond changes to the power spectrum.

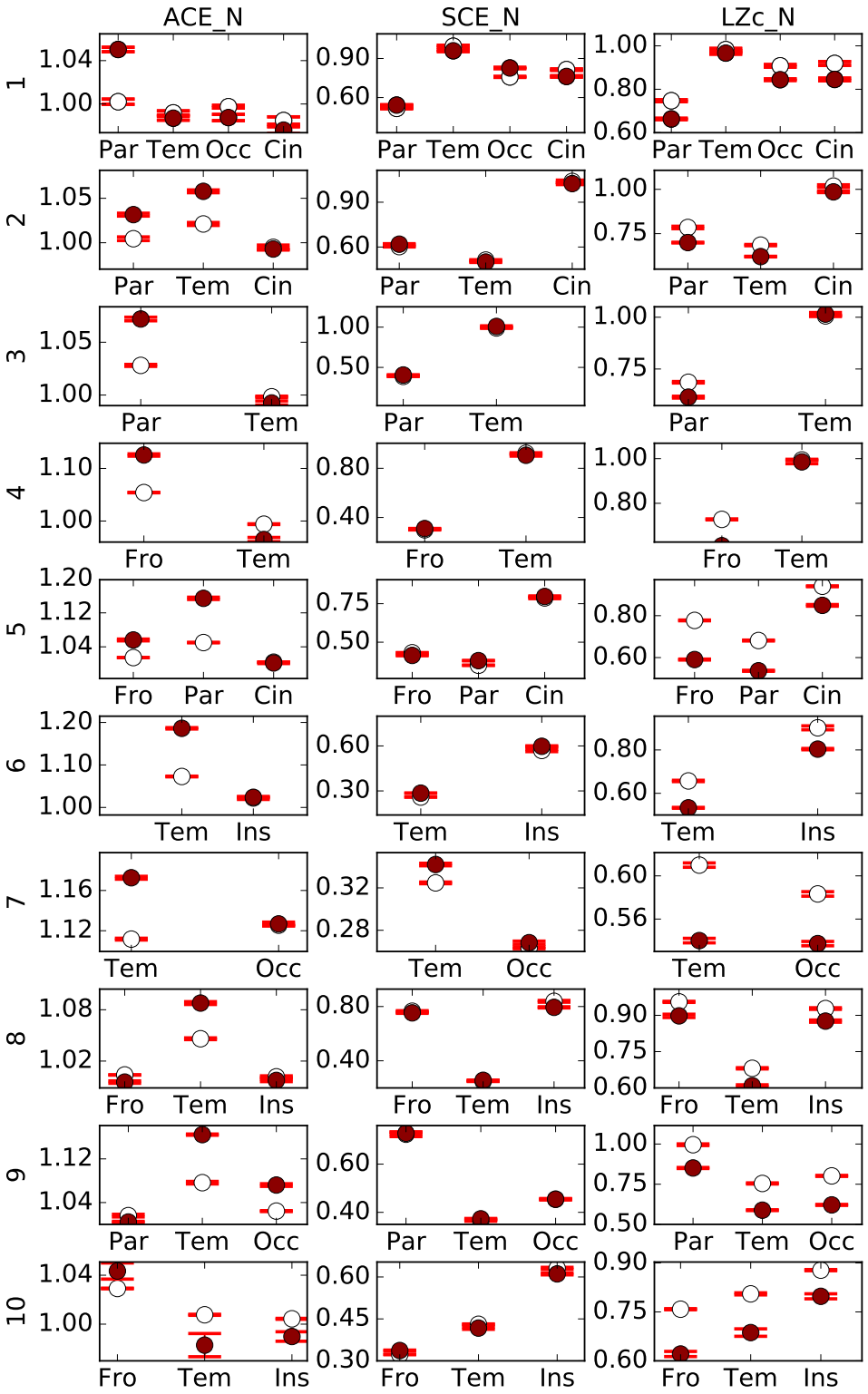

Figure S9: ACE_N, SCE_N **and** LZc_N **scores computed for channel quartets restricted to single regions.** The same analysis as shown in Fig. 7 was performed, yet normalising the measures by their average score across 100 phase-randomised data segments. Note that here the scales vary across subjects and measures.

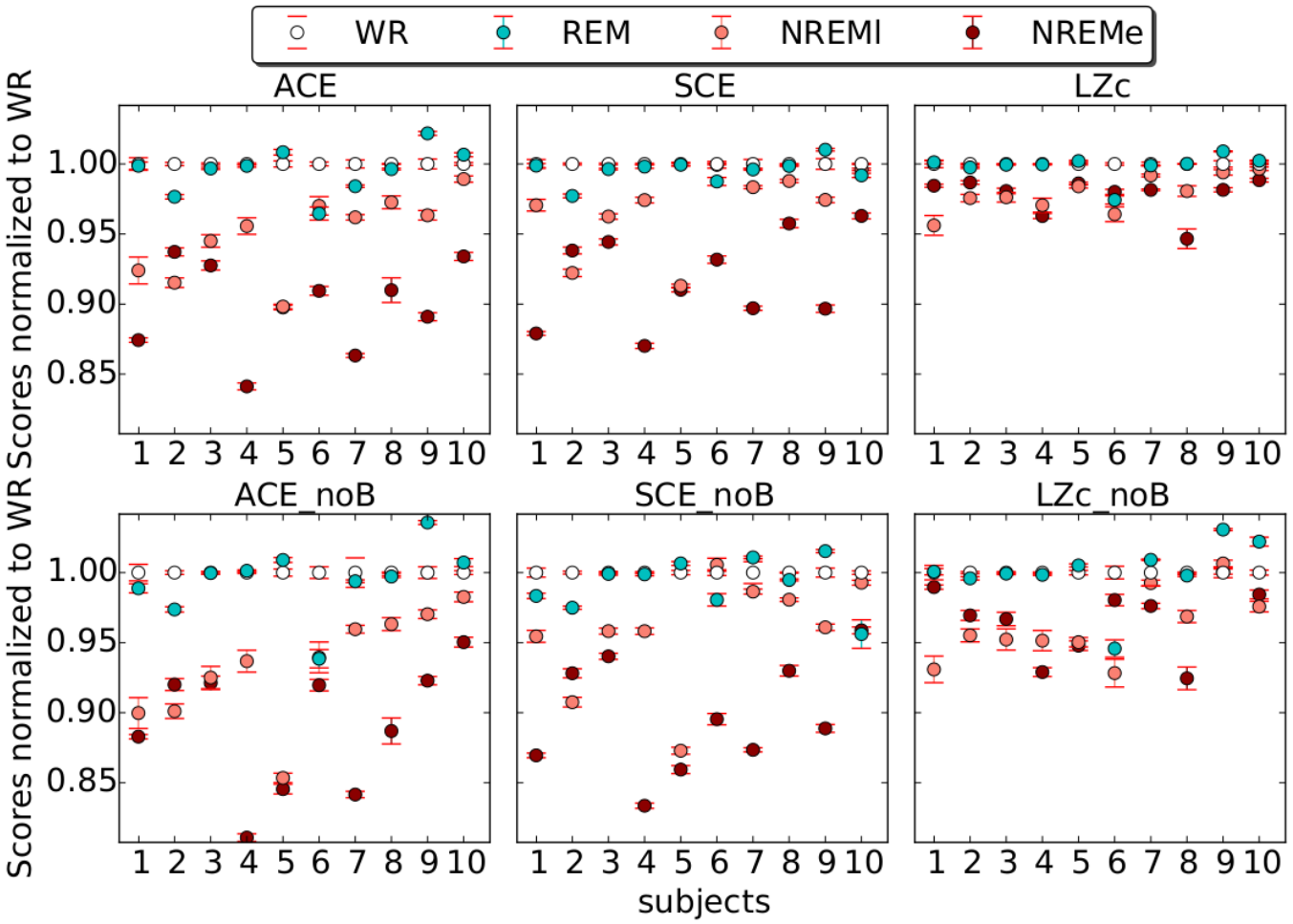

Figure S10: **Comparison of bipolar and monopolar montage.** ACE, SCE and LZc computed across 18 channels using (top) the bipolar montage used for the main analyses; (bottom) monopolar referencing (same data omitting the bipolar re-referencing). Plotted points show mean across 10s segments, and error bars show standard error.

Table S4: Channel choice control: 10 different sets of 10 channels

	1	2	3	4	5	6	7	8	9	10
ACE	10 0 0	10 0 0	10 0 0	10 0 0	10 0 0	10 0 0	10 0 0	10 0 0	10 0 0	10 0 0
SCE	9 1 0	8 2 0	9 1 0	8 2 0	10 0 0	9 1 0	9 1 0	9 1 0	9 1 0	10 0 0
LZc	10 0 0	10 0 0	10 0 0	10 0 0	9 1 0	10 0 0	10 0 0	10 0 0	10 0 0	10 0 0

For each measure and set of channels, the three numbers correspond to how many subjects out of 10 had higher score for WR than NREMe with Cohen's $d > 0.8$ (left digit), no substantial difference, $d < 0.8$, (middle) and higher score for NREMe than WR with Cohen's $d > 0.8$ (right), computed across trials. For each of 10 different sets of 10 channels, the results were obtained from applying the measures to 10s segments. The scores of ACE, SCE and LZc were higher for WR than NREMe for all subjects and sets of channels, with $d > 0.8$ for all but at most 2 subjects.

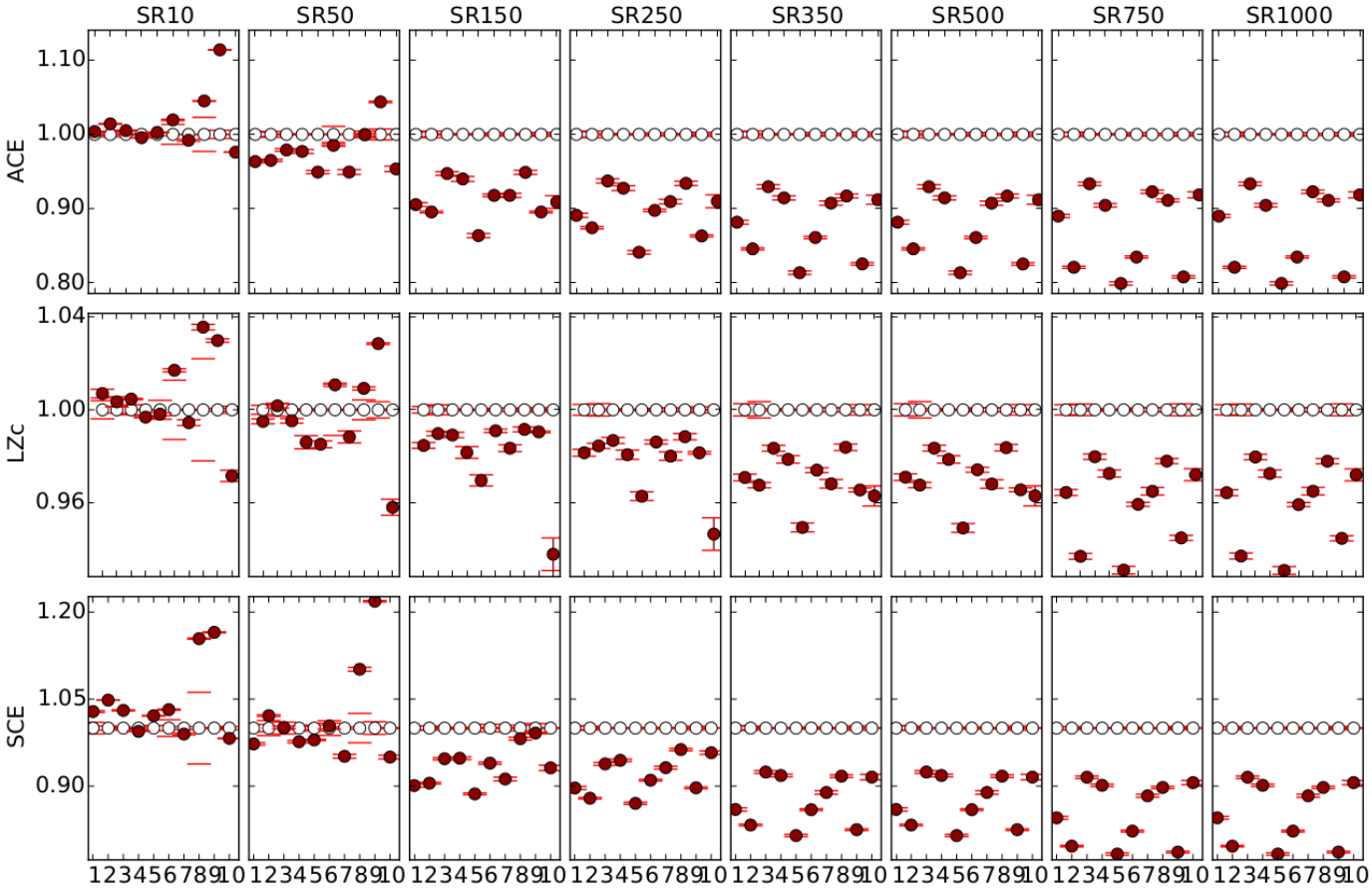

Figure S11: **ACE, SCE and LZc for 18 channels and 2500 observations for different sampling rates**. Each column of panels displays the three measures' scores for all subjects and the sampling rate indicated in the column's title ($SR = 10Hz, 50Hz, 150Hz, 250Hz, 350Hz, 500Hz, 750Hz, 1000Hz$). Each measure's score is higher for WR than NREMe for every subject and sampling rate of at least $150Hz$. For sampling rate smaller or equal to $50Hz$ consistency of the scores decreases strongly for all 3 measures. Scores are normalized by the result for WR.

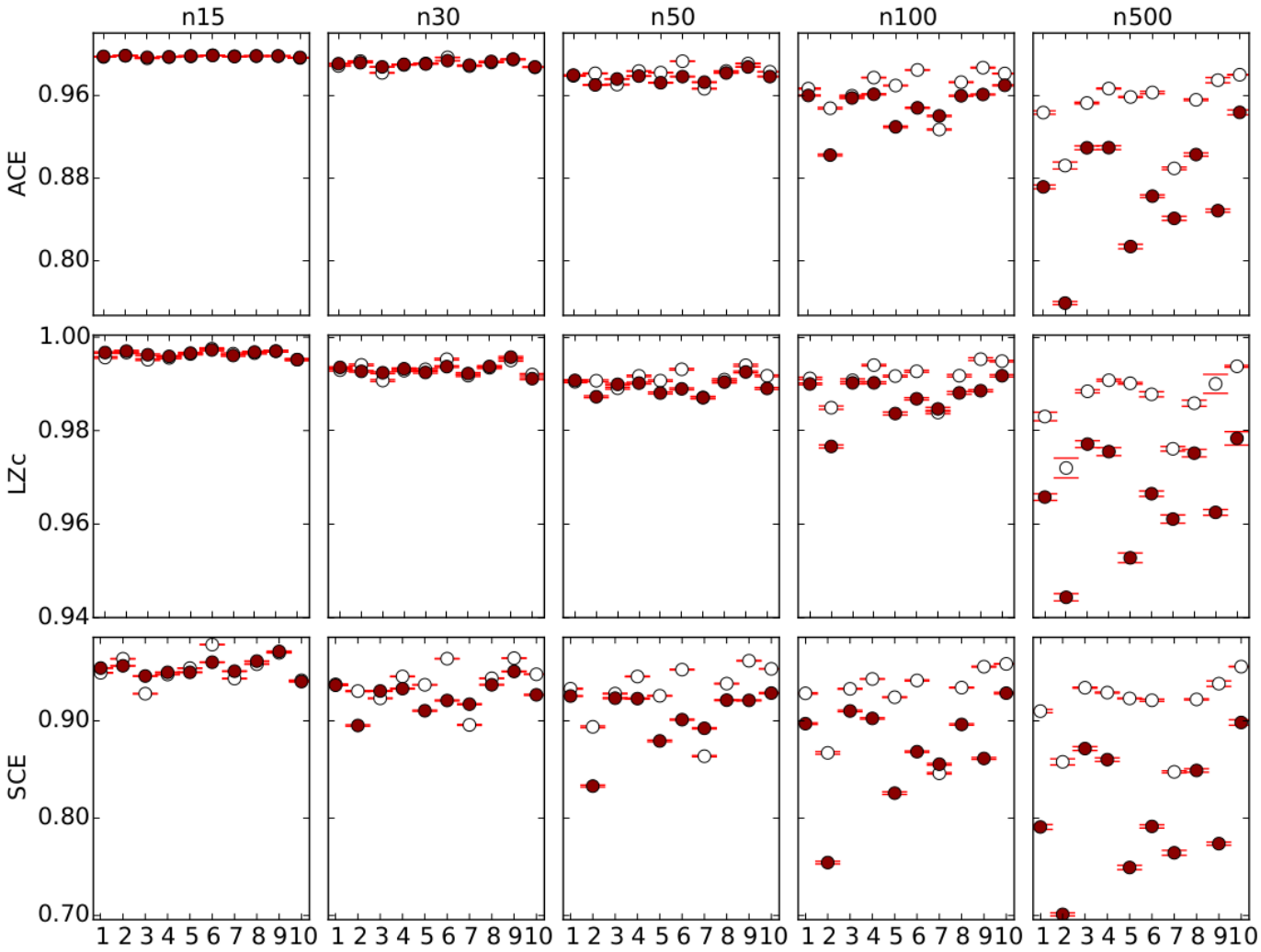

Figure S12: **ACE, SCE and LZc for 18 channels for different numbers of observations at** $250Hz$**.** Each column of panels displays the three measures' scores for all subjects and the number of observations indicated in the column's title. For segment lengths below $2s$ (corresponding to 500 observations at sampling rate $250Hz$), the measures' scores are no longer consistently higher for WR (white) than NREMe (red). Scores are normalized by the result for WR.

References

[1] Schartner M, Seth A, Noirhomme Q, Boly M, Bruno M-A, et al. (2015) Complexity of Multi-Dimensional Spontaneous EEG Decreases during Propofol Induced General Anaesthesia. PLoS ONE 10(8): e0133532. doi: 10.1371/journal.pone.013353