

The EEG multiverse of schizophrenia

Authors: Dario Gordillo^{1,a*}, Janir Ramos da Cruz^{1,2,a}, Eka Chkonia^{3,4}, Wei-Hsiang Lin¹, Ophélie Favrod¹, Andreas Brand¹, Patrícia Figueiredo², Maya Roinishvili^{4,5}, and Michael H. Herzog¹

Affiliations:

¹ Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

² Institute for Systems and Robotics – Lisboa (LARSyS), Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

³ Department of Psychiatry, Tbilisi State Medical University, Tbilisi, Georgia

⁴ Institute of Cognitive Neurosciences, Free University of Tbilisi, Tbilisi, Georgia

⁵ Laboratory of Vision Physiology, Beritashvili Centre of Experimental Biomedicine, Tbilisi, Georgia

^a These authors contributed equally

***Corresponding author:**

Dario Gordillo, Laboratory of Psychophysics, Brain Mind Institute, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Phone number: +41 76 633 40 47

Email: dario.gordillolopez@epfl.ch

Keywords: Electroencephalography, Schizophrenia, Psychosis, Neuroimaging, Resting-State, Psychiatry

20 Abstract

21 Research on schizophrenia typically focuses on one paradigm, for which clear-cut differences between
22 patients and controls are established. Great care is taken to understand the underlying genetical,
23 neurophysiological, and cognitive mechanism, which eventually may explain the clinical outcome. One
24 tacit assumption of these *deep rooting* approaches is that paradigms tap into common and
25 representative aspects of the disorder. Here, we analyzed the resting-state electroencephalogram
26 (EEG) of 121 schizophrenia patients and 75 controls. Using multiple signal processing methods, we
27 extracted 194 EEG features. Sixty-nine out of the 194 EEG features showed a significant difference
28 between patients and controls indicating that these features detect an important aspect of
29 schizophrenia. Surprisingly, the correlations between these features were very low, suggesting that
30 each feature picks up a different aspect of the disorder. We propose that complementing *deep* with
31 *shallow* rooting approaches, where many roughly independent features are extracted from one
32 paradigm (or several paradigms), will strongly improve diagnosis and potential treatment of
33 schizophrenia.

34 1. Introduction

35 Schizophrenia patients show strong abnormalities in many domains including personality, cognition,
36 perception, and even immunology. In many experimental paradigms, the differences between patients
37 and controls have large effect sizes, indicating that important aspects of the disease are detected. This
38 provokes two questions: what do these abnormalities have in common and how representative are they
39 for the disease? For example, patients exhibit strong deficits in cognition, such as in working memory
40 tasks (Meyer-Lindenberg et al., 2001), which are attributed to abnormalities of cortico-cerebellar-
41 thalamic-cortical circuits (Andreasen et al., 1998). Patients show also diminished skin flushing with the
42 niacin skin test (Rybakowski and Weterle, 1991), which is attributed to dysfunctional phospholipase A2

43 arachidonic acid signaling (Messamore, 2012). How do the working memory deficits correspond to
44 deficits in skin functioning? Very few studies have correlated deficits with each other (Braff et al., 2007a,
45 2007b; Dickinson et al., 2011; Price et al., 2006; Seidman et al., 2015; Toomey et al., 1998). The
46 Consortium on the Genetics of Schizophrenia studied neurocognitive and neurophysiological
47 abnormalities in schizophrenia patients with a battery of 15 paradigms (Seidman et al., 2015). They
48 found that neurocognitive measures shared a significant amount of variance while neurophysiological
49 measures were almost entirely independent. Price and colleagues (2006) studied four candidate
50 electrophysiological endophenotypes of schizophrenia (mismatch negativity, P50, P300, and
51 antisaccades). Even though patients *and* their family members showed deficits in each of these
52 endophenotypes, the features were largely uncorrelated.

53 Here, we took another road. Instead of comparing different paradigms, we analyzed the very same data
54 of the very same patients with different electroencephalogram (EEG) analysis methods, including many
55 that have shown strong atypical patterns in patients (Andreou et al., 2015; Boutros et al., 2008; da Cruz
56 et al., 2020; Di Lorenzo et al., 2015; Kim et al., 2000, p. 200; Nikulin et al., 2012; Sun et al., 2014;
57 Uhlhaas and Singer, 2010). Resting-state EEG features revealing significant differences between
58 schizophrenia patients and controls are thought to reflect brain mechanisms linked to important aspects
59 of the disorder. For example, schizophrenia patients exhibit reduced long-range temporal correlations
60 (LRTC) in the alpha and beta frequency bands (Nikulin et al., 2012). These responses were suggested
61 to reflect excessive switching of neuronal states in patients. Patients also have shown atypical patterns
62 in the dynamics of EEG microstates classes C and D (da Cruz et al., 2020; Rieger et al., 2016).
63 Microstates abnormalities have been proposed to correspond to imbalances in attentional and
64 information processing in schizophrenia. Schizophrenia patients have shown increased power in the
65 delta, theta, and beta frequency bands (Venables et al., 2009). Increased beta power was suggested
66 to reflect cortical hyperexcitability, increased power in the delta and theta bands were proposed to
67 relate to atypical dopaminergic function, to name a few examples. All these results, individually, suggest

68 that each EEG feature captures important aspects of schizophrenia. But how representative are these
69 abnormalities of the disorder? Does a patient showing abnormal microstate dynamics also show deficits
70 in LRTC, or in other EEG features? Even though all these atypical patterns in patients are obtained
71 from the same EEG data, no study has evaluated how these EEG features relate to each other. This is
72 not only the case for resting-state EEG studies, but constitutes the conventional approach in
73 schizophrenia research. This approach is centered on drawing general conclusions about
74 schizophrenia based on one paradigm, assumed to unveil common and representative aspects of the
75 disorder.

76 Aiming to shed light on this EEG *multiverse* of schizophrenia, in this work, we analyzed the resting-
77 state EEG data of 121 schizophrenia patients and 75 healthy controls with multiple methods. This
78 allowed us to extract 194 EEG features, such as time-domain features, frequency-domain and
79 connectivity features both in electrode and source space, and nonlinear dynamical features. Then, we
80 correlated the features revealing group differences to evaluate how these abnormalities/deficits relate
81 to each other. We also examined whether EEG features show adequate predictive power to clinical
82 scales measuring key symptoms of schizophrenia. We propose that future studies in schizophrenia
83 research should consider multiple features extracted from the same and/or different paradigms in order
84 to improve diagnosis and potential treatment.

85 **2. Material and Methods**

86 **2.1. Participants**

87 Two groups of participants joined the experiment: schizophrenia patients ($n = 121$) and healthy controls
88 ($n = 75$). All participants took part in a battery of tests comprising perceptual and cognitive tasks as well
89 as EEG recordings. Data of 101 patients and 75 controls have already been published in different
90 contexts (da Cruz et al., 2020a, 2020b; Favrod et al., 2018; Garobbio et al., 2021). Patients were

91 recruited from the Tbilisi Mental Health Hospital or the psycho-social rehabilitation center. Patients were
92 invited to participate in the study when they had recovered sufficiently from an acute psychotic episode.
93 Thirty-five were inpatients and 86 were outpatients. Patients were diagnosed using the Diagnostic and
94 Statistical Manual of Mental Disorders Fourth Edition (DSM-IV) by means of an interview based on the
95 Structured Clinical Interview for DSM-IV, Clinical Version, information from staff, and study of patients'
96 records. Psychopathology of patients was assessed by an experienced psychiatrist using the Scale for
97 the Assessment of Negative Symptoms (SANS) and the Scale for the Assessment of Positive
98 Symptoms (SAPS). Out of the 121 patients, 106 were receiving neuroleptic medication.
99 Chlorpromazine (CPZ) equivalents are indicated in **Table 1**. Controls were recruited from the general
100 population in Tbilisi, aiming to match patients' demographics as closely as possible. All controls were
101 free from psychiatric axis I disorders and had no family history of psychosis. General exclusion criteria
102 were alcohol or drug abuse, severe neurological incidents or diagnoses, developmental disorders
103 (autism spectrum disorder or intellectual disability), or other somatic mind-altering illnesses, assessed
104 through interview by certified psychiatrists. All participants were no older than 55 years. Group
105 characteristics are presented in **Table 1**. All participants signed informed consent and were informed
106 that they could quit the experiment at any time. All procedures complied with the Declaration of Helsinki
107 (except for pre-registration) and were approved by the Ethical Committee of the Institute of
108 Postgraduate Medical Education and Continuous Professional Development (Georgia). Protocol
109 number: 09/07. Title: "Genetic polymorphisms and early information processing in schizophrenia".

110 Table 1 - Group average statistics (\pm SD)

	Patients	Controls	Statistics
Gender (F/M)	22/99	39/36	$\chi^2(1) = 24.702, p = 6.690e-7^b$
Age (years)	35.8 ± 9.2	35.1 ± 7.7	$t(194) = 0.519, p = 0.604^c$
Education (years)	13.3 ± 2.6	15.1 ± 2.9	$t(194) = -4.418, p = 1.657e-5^c$
Handedness (L/R)	6/115	4/71	$\chi^2(1) = 0.013, p = 0.908^b$
Illness duration (years)	10.8 ± 8.7		
SANS	10.1 ± 5.2		

SAPS	8.6 ± 3.2		
CPZ equivalent ^a	561.1 ± 389.4		
SANS - Scale for the Assessment of Negative Symptoms, SAPS - Scale for Assessment of Positive, CPZ - chlorpromazine			
^a Average CPZ equivalents calculated over the 106 Patients receiving neuroleptic medication			
^b Pearson's chi-squared test			
^c Two-sided independent samples <i>t</i> -test			

111

112

2.2. EEG recording and data processing

113

Participants were sitting in a dim lit room. They were instructed to keep their eyes closed and to relax for 5 minutes. Resting-state EEG was recorded using a BioSemi Active Two Mk2 system (Biosemi B.V., The Netherlands) with 64 Ag-AgCl sintered active electrodes, referenced to the common mode sense electrode. The recording sampling rate was 2048 Hz. Offline data were downsampled to 256 Hz and preprocessed using an automatic pipeline (da Cruz et al., 2018). Preprocessed EEG data were analyzed using multiple signal processing methods in the electrode and source space. In total, 194 EEG features were extracted (See [Supplementary Table 1](#)). Out of the 194 EEG features, 50 were obtained in the source space and 144 in the electrode space. For source space analysis, we defined 80 brain regions (40 per hemisphere) according to the AAL atlas (See [Supplementary Table 4](#)). See [Supplementary Methods](#) for a detailed description of the analysis methods.

114

115

116

117

118

2.3. Group comparisons

119

We compared patients' and controls' scores for each of the 194 EEG features. For each of the *J* variables (i.e., 64 electrodes, 80 brain regions, or 12 microstate parameters, depending on the number of variables of each EEG feature) of a given feature, we performed a two-way ANCOVA, with Group (patients and controls) and Gender (male and female) as factors and Education as a covariate. *P*-values for the effect of Group were corrected for *J* comparisons using False Discovery Rate (FDR; with an error rate of 5%). Group effects' η^2 were converted to Cohen's *d*.

120

121

122

123

124

2.4. Pearson and partial least squares correlations

First, for each EEG feature that contained at least one variable showing a significant difference between patients and controls (after correcting for multiple comparisons), we selected the variable (i.e., electrode, brain region, or microstate parameter) with the biggest effect size to be the representative variable for that feature. Then, for patients and controls separately, we computed pairwise Pearson correlations between the representative variables of each significant EEG feature. Second, to quantify the overall relationship, i.e., the amount of *shared information*, between pairs of multivariate EEG features, we used Partial Least Squares Correlation (PLSC). PLSC generalizes correlations between two variables to two matrices (McIntosh et al., 1996; Tucker, 1958). The *shared information* can be quantified as the *inertia* common to the two features (Krishnan et al., 2011). The statistical significance of the inertia was assessed using a permutation test (Abdi and Williams, 2013; McIntosh et al., 2004). The inertia values were normalized. Hence, the normalized inertias ($\mathfrak{I}_{relative}$) ranged from 0 (the two EEG features are completely unrelated) to 1 (the two EEG features contain the same information). PLSC analysis was done for patients and controls separately. See [Supplementary Methods](#) for details.

2.5. Regression analysis

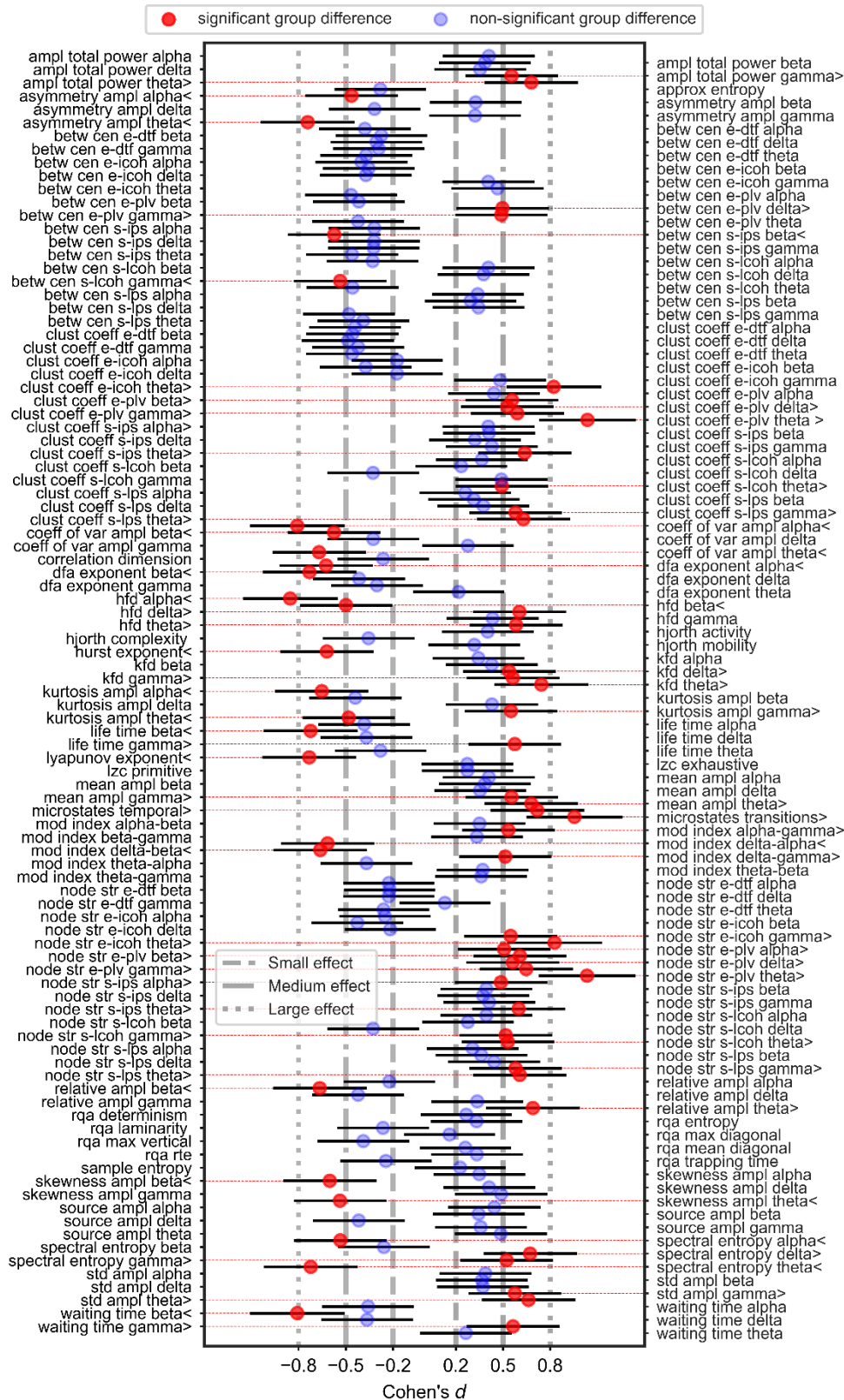
To evaluate whether EEG features predict psychopathology scores (SAPS and SANS) adequately, we used elastic net regression models (Zou and Hastie, 2005). Elastic nets can handle regression problems where the number of predictors is relatively large compared to the number of samples as well as multicollinearity (i.e., the predictors are not linearly independent) by combining the L1 and L2 penalties to achieve regularization. For each of the 194 EEG features (with all its variables), we built two regression models, one to predict SAPS scores and one to predict SANS scores. We performed 20 repetitions of a 3-fold nested cross-validation procedure. First, one third of the data (1 fold) was left out for validation (test set), while the remaining data (2 folds; train set) were used to find the optimal

parameters, namely the amount of penalization and the compromise between L1 and L2 penalties, using 3-fold cross-validation. The model with the parameters leading to best performance in the train set was tested on the left-out data (test set). The entire procedure was repeated 20 times, with different allocations of the patients in the train and test sets. Using the same cross-validation procedure, i.e., 20 repetitions of a 3-fold cross-validation, we also evaluated predictive performance using a nonlinear random forest regression model, setting the maximum depth of the tree to 10 and the number of trees to 100. Random forests are meta estimators that average several decision trees trained on subsets of the dataset to improve accuracy and avoid overfitting. Prediction performance was calculated using the coefficient of determination (R^2) and the root mean squared error ($RMSE$). The distribution of the prediction performance values was obtained from the 60 aggregated $RMSE$ and R^2 , across repetitions of the procedure.

3. Results

3.1. Multiple EEG features reveal significant group effects

For 121 patients (22 females, 35.8 ± 9.2 years old, 13.3 ± 2.6 years of education) and 75 age-matched healthy controls (39 females, 35.1 ± 7.7 years old, 15.1 ± 2.9 years of education; Table 1), we extracted in total 194 features from the resting-state EEG recordings, including time-domain, frequency-domain, connectivity, and nonlinear dynamical features (Supplementary Table 1). Among the 194 EEG features, 69 (35.57%) showed significant differences between patients and controls with medium to large effect sizes (Cohen's d varied from 0.463 to 1.037, Figure 1). Patients showed significantly reduced values in 24 out of the 69 EEG features revealing significant group differences (illustrated as negative effect size in Figure 1). Patients exhibited significantly higher values than controls in 45 EEG features.



174

175

176

Figure 1. Effect size (Cohen's *d*) of the group differences between patients and controls for each of the 194 EEG features. We took the values of the electrode, brain region, or

177 microstate parameter with the largest effect size according to Cohen's d (η^2 values were
178 converted to Cohen's d) to be the representative variable for each feature. Significant
179 group differences, after correction for multiple comparisons (using FDR), are depicted in
180 red, with dotted red vertical lines serving as a guide to their labels. $>$ and $<$ were added
181 to the feature labels to indicate if patients had significantly higher or lower values than
182 controls, respectively. The non-significant effects are shown in blue. Error bars represent
183 95% confidence intervals (C.I.). A list with the abbreviations and the corresponding name
184 of each feature is presented in [Supplementary Table 1](#).

185 **3.2. Correlations between EEG features**

186 To evaluate to what extent features that showed significant group differences are sensitive to the same
187 aspects of the disorder we computed Pearson's correlations between pairs of features (**Figure 2**). As
188 the representative variable for each feature, we took the values of the electrode, brain region, or
189 microstate parameter that showed the largest group difference according to Cohen's d (**Figure 1**).
190 Surprisingly, we found that in the patients group only 36.49% of the pairwise correlations were
191 significant at a level of 0.05 (without correcting for multiple comparisons). For the control group, only
192 26.73% of the correlations were significant. Since significance depends on the sample size, here, we
193 focus on the magnitude of the correlation coefficients (r -values). In general, the magnitudes of the r -
194 values were very low in both patients (0.055, 0.122, 0.251, for the 25th, 50th, and 75th percentiles,
195 respectively) and controls (0.059, 0.129, 0.242, for the 25th, 50th, and 75th percentiles, respectively;
196 **Figure 2**). Strong correlations were found mainly for pairs of very closely related features, such as
197 between waiting-time statistics of gamma bursts (*waiting time gamma*) and life-time statistics of gamma
198 bursts (*life time gamma*; $r=0.836$ and $r=0.926$, in patients and controls, respectively).

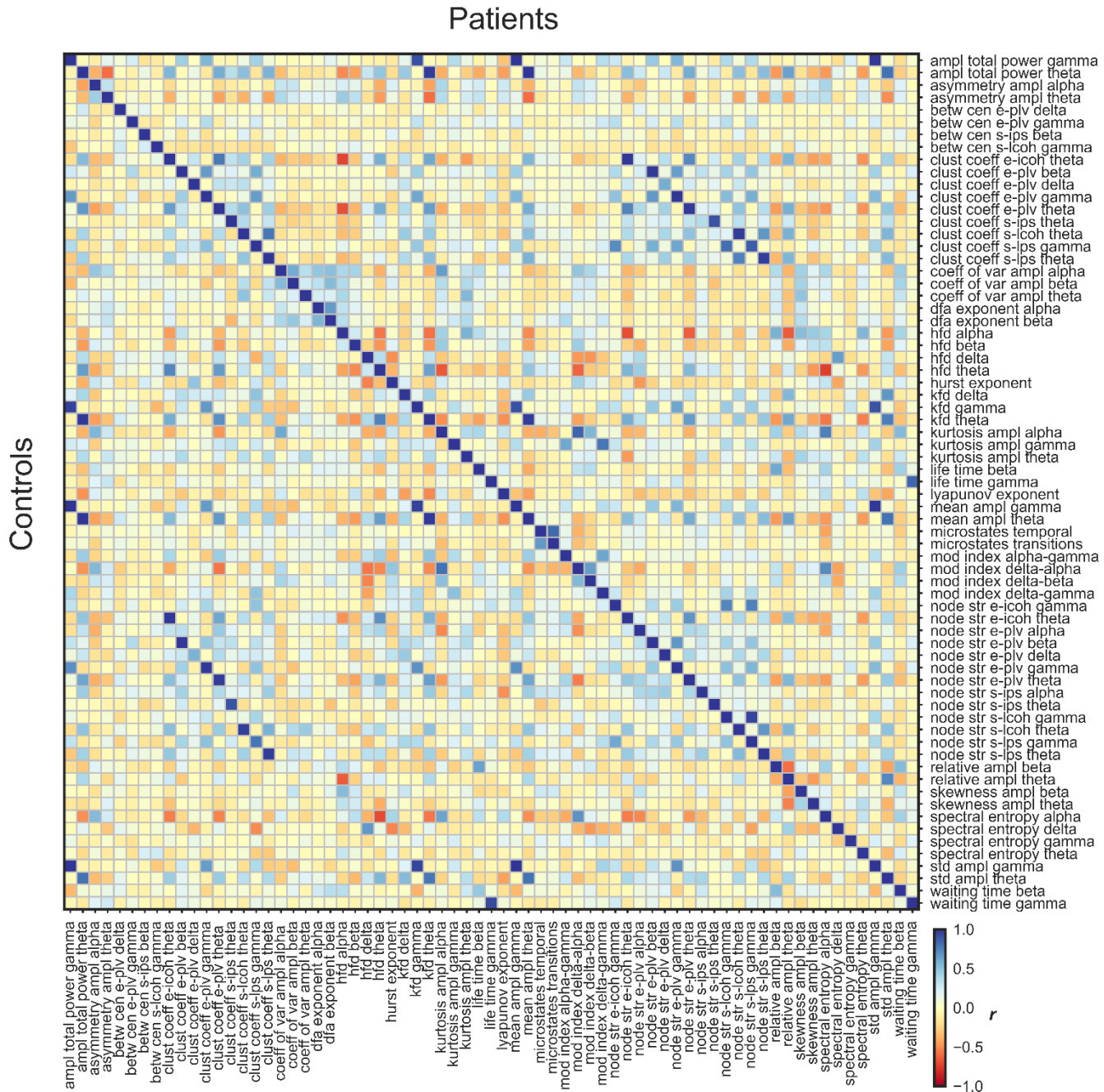


Figure 2. Pairwise correlations between the 69 EEG features that showed significant group differences between patients and controls. Patients' r -values are presented in the upper triangle and controls' r -values are shown in the lower triangle. Strong negative and positive r -values are depicted in red and blue, respectively, and r -values around 0 in yellow. For each feature, we used the values of the electrode, brain region, or microstate parameter that showed the largest effect size as the representative variable for the

206 correlations. A list with the abbreviations and corresponding name of each feature is
207 shown in [Supplementary Table 1](#).

208 To quantify the overall *shared information* between pairs of EEG features, which showed significant
209 group differences, by taking not only variables with the largest effect size into account but all variables
210 of the features we used partial least squares correlation (PLSC). For the patients, 55.92% of the
211 pairwise inertias were significant (without correcting for multiple comparisons) and for controls, 40.28%.
212 In general, relative inertias were not very high in both patients (0.254, 0.329, 0.409, for the 25th, 50th,
213 and 75th percentiles, respectively) and controls (0.305, 0.387, 0.472, for the 25th, 50th, and 75th
214 percentiles, respectively; **Figure 3**). As in the Pearson's correlation results, features that showed strong
215 associations were mainly similar features, such as the same network statistics for different connectivity
216 measures in the theta band, for example, at the electrode level: clustering coefficient connectivity
217 estimated with phase locking value (*clust coeff e-plv theta*) and with imaginary part of coherence (*clust*
218 *coeff e-icoh theta*; $\mathfrak{S}_{relative}=0.804$ and $\mathfrak{S}_{relative}=0.826$, in patients and controls, respectively).

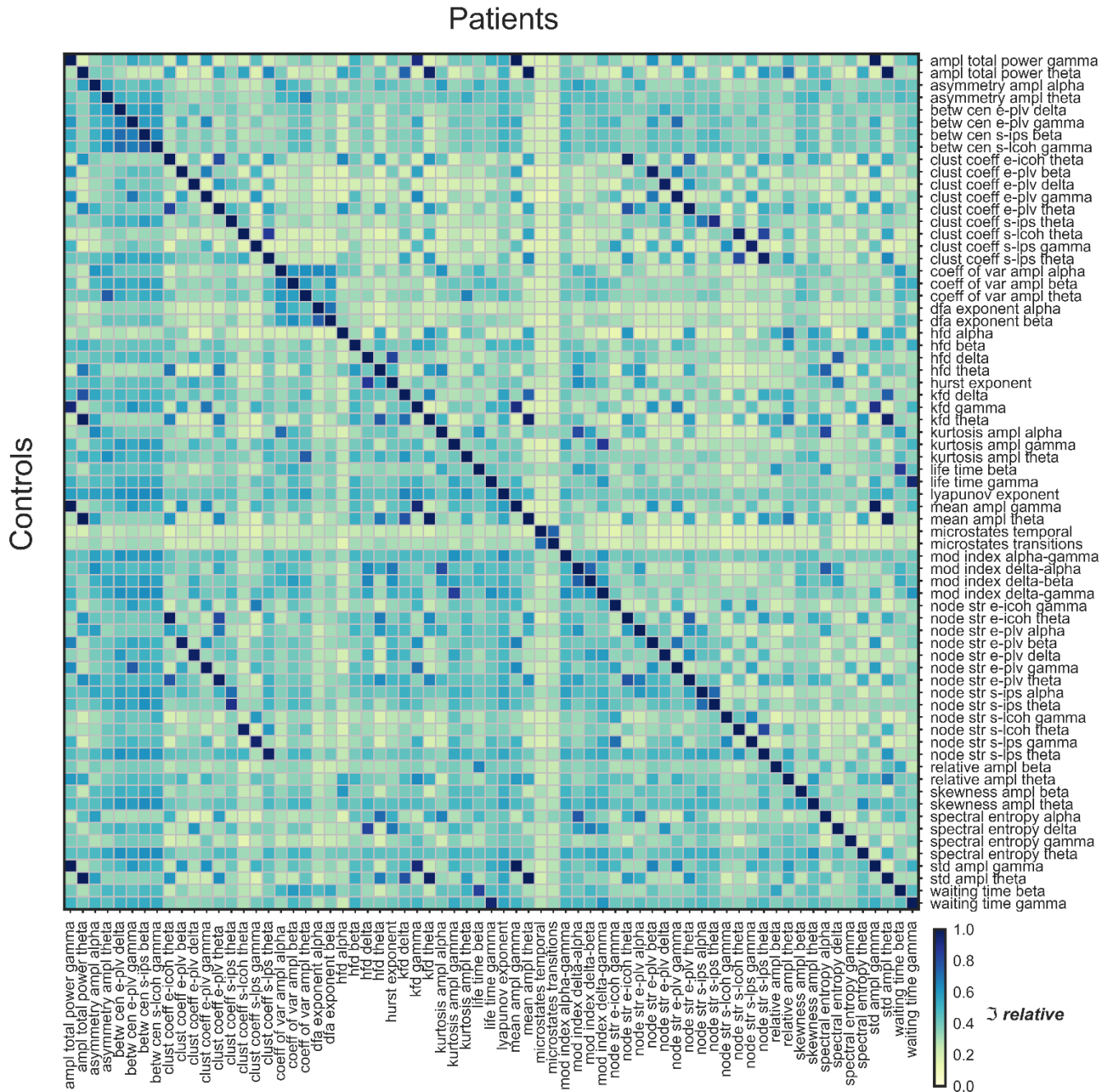


Figure 3. Shared information between the 69 EEG features that showed significant group differences, as measured by the relative inertia ($\mathfrak{S}_{relative}$) computed with partial least squares correlations (PLSC). The relative inertia ranges from 0 (two features are completely unrelated) to 1 (the two features' values move together by the exact same percentage). Patients' relative inertias are presented in the upper triangle and controls'

225 relatives are shown in the lower triangle. A list with the abbreviations and corresponding
226 name of each feature is shown in [Supplementary Table 1](#).

227 **3.3. Prediction of psychopathology scores**

228 We evaluated whether EEG features were adequate predictors of psychopathology scores determined
229 by the Scale for the Assessment of Positive Symptoms (SAPS) and the Scale for the Assessment of
230 Negative Symptoms (SANS), which target positive (hallucinations, delusions, bizarre behavior, and
231 positive formal thought disorder) and negative (affective flattening, alogia, apathy, anhedonia, and
232 attention) symptoms, respectively. All 194 EEG features exhibited very weak out-of-sample predictive
233 ability to both the SANS and SAPS scores. Results were very similar for both the linear (i.e., elastic
234 net) and nonlinear (i.e., random forest) models. See [Supplementary Table 2](#) and [Supplementary Table](#)
235 [3](#) for details.

236 **4. Discussion**

237 Traditionally, schizophrenia research focuses on a single experimental paradigm and analysis method
238 showing significant differences between patients and controls, and then tries to derive the underlying
239 genetic or neurophysiological causes of the disorder. This approach has been quite successful in the
240 formulation of hypotheses, such as the dopamine hypothesis (Howes and Kapur, 2009), the social brain
241 hypothesis (Burns, 2006), the glutamate hypothesis (Hu et al., 2015), or the dysconnection hypothesis
242 (Friston et al., 2016), just to name a few. Here, we examined to what extent abnormalities quantified
243 by different EEG features correlate with each other. Many of the investigated features were previously
244 linked to different abnormalities of brain processes in schizophrenia, and, here, we reproduced many
245 of these results, such as imbalance in microstates dynamics (da Cruz et al., 2020a; Rieger et al., 2016),
246 decreased long-range temporal correlations in the alpha and beta bands (Nikulin et al., 2012),
247 decreased life- and waiting-times in the beta band (Sun et al., 2014), increased spectral amplitude in

248 the theta band (Boutros et al., 2008), increased connectivity in the theta band at the source level
249 (Andreou et al., 2015; Di Lorenzo et al., 2015), decreased Lyapunov exponent (Kim et al., 2000), among
250 others. With our systematic analysis, we also found abnormalities in EEG features, which, to the best
251 of our knowledge, have not been reported yet, namely, delta-phase gamma-amplitude coupling, range
252 EEG coefficient of variation and asymmetry in the theta and alpha bands, etc. In some way, deeper
253 analysis of each feature may have warranted an in-depth study and a potential publication. However,
254 we did not want to elaborate on these methods individually because we wanted to understand how all
255 EEG features relate to each other in their entirety. The surprising insight from our analysis is that, even
256 though we are probing the same signals from the same participants, we found only weak correlations
257 between the 69 significant features. The only strong correlations were between features that are similar
258 from the outset, thereby resembling test-retests. This suggests that, even though EEG features reveal
259 clear-cut and reproducible atypical patterns in patients, none of the features is truly representative for
260 the disease, but rather that all these features pick up more or less independent aspects of
261 schizophrenia. Hence, the traditional approach of focusing on a single experimental paradigm and
262 analysis method has its limitations. These results remind us that schizophrenia is indeed a very
263 heterogeneous disease, a well-known fact, which is however not always taken seriously enough
264 because, as mentioned above, most research tries to find the one or a few causes of schizophrenia
265 within one well described paradigm by digging as deep as possible into the underlying
266 neurophysiological and genetic mechanisms. In analogy to botany, one may call these approaches
267 “deep rooting” approaches.

268 We propose that it may be useful to complement these deep rooting approaches with “shallow rooting”
269 approaches, representing schizophrenia within a high-dimensional space, where many tests and
270 analysis outcomes are the basis variables. The outcomes should ideally have large effect sizes, low
271 mutual correlations, and a “flat” factor structure. Whether this is possible is an open question and
272 depends very much on the underlying causes of schizophrenia. On the lowest complexity level, there

273 may be only a few independent causes (or even only one), which were not found yet. Given the
274 heterogeneity of the disease, including abnormalities in the cognitive (Andreasen et al., 1998) but also
275 the skin functioning domain (Messamore, 2012), the causes need to be on a rather general level, likely
276 subcellular, present in all human functioning. Alternatively, schizophrenia may be an approximately
277 “additive” disease, where many small abnormalities add up to severe symptoms. In an even more
278 complex scenario, only certain combinations of redundant functions, each coming with at least two
279 variants, cause the disease. For instance, if one function is up-regulated and another one down-
280 regulated in an individual, there are no abnormalities. Deficits manifest only when all or most functions
281 are either up- or down-regulated. In such a combinatorial scenario, it would be difficult to find the
282 underlying causes since each variant itself does not lead to a deficit; only certain combinations do. Our
283 correlation analysis provides evidence supporting the more complex/multifactorial scenario, where
284 each feature makes up its own factor and manifests differently in different patients. Since our data
285 shows that each EEG feature is sensitive to roughly independent aspects of schizophrenia, each brain
286 process captured by the analysis methods might be neither necessary nor sufficient to explain important
287 aspects of the disease. Indeed, we found that EEG features showed very weak predictive power to key
288 symptoms of schizophrenia, suggesting that there is little information about individual differences in
289 psychopathology. Our results are an invitation to rethink the current approach in schizophrenia research
290 and suggest that new study designs conflating multiple features from the same and different paradigms
291 might be more adequate. In the next steps, it will be important to find the right set of features, which
292 may stem from EEG recordings but also potentially immunological markers, of which each may
293 contribute with a variety of features. Previous research has shown that combining features improved
294 classification and predictive performance in psychosis studies (Mothi et al., 2019; Yang et al., 2010).

295 Our study has several limitations. There are demographic differences between patients and controls,
296 which might affect our group comparisons. However, we attempted to minimize these demographic
297 effects by using education as a covariate and gender as factor in the analyses. Similarly, we cannot

298 exclude effects of medication in our results. Nonetheless, we find similar patterns of correlations
299 between EEG features, i.e., weak associations, in both patients and controls, suggesting that if there
300 is an effect of medication, it is small. Further, our sample size is relatively small for achieving reliable
301 estimates of predictive power (Poldrack et al., 2020; Schnack and Kahn, 2016; Varoquaux, 2018).

302 Our results may explain a deep mystery in schizophrenia research. Schizophrenia has an estimated
303 heritability of 70 to 85% (Burmeister et al., 2008). For example, the chance to also suffer from
304 schizophrenia for monozygotic twins is about 33% when the partner twin has the disease (Hilker et al.,
305 2018). Furthermore, about 0.25 to 0.75% people of a population suffer from schizophrenia and related
306 psychotic disorders (Kessler et al., 2005; Moreno-Küstner et al., 2018; Saha et al., 2005). These values
307 are rather stable across cultures (Simeone et al., 2015). Given that schizophrenia patients have less
308 offspring (Avila et al., 2001; Bassett et al., 1996; Keller and Miller, 2006; MacCabe et al., 2009), this
309 provokes the question why schizophrenia has not been extinguished during the course of evolution
310 (Keller and Miller, 2006; Liu et al., 2019). In the above-mentioned combinatorial scenario with many
311 redundant functions this may simply happen because evolution operates on the individual single-
312 nucleotide polymorphism (SNP) level and not on the combinatorial one. As long as most of the
313 population shows average functioning, there will be no change of the allele distributions. In the additive
314 scenario, evolution may extinct harmful alleles, of which each constitutes only a little risk, very slowly
315 and these may be replaced by harmful de novo mutations (Keller and Miller, 2006). To what extent
316 such considerations hold true will be shown by shallow rooting approaches using a plethora of
317 paradigms and a multiverse of analysis methods. In a nutshell, deep rooting will help to understand the
318 different aspects of the disorder, while shallow rooting will help to better diagnose schizophrenia by
319 finding subpopulations, leading to more personalized treatment.

Acknowledgements

We would like to thank Marc Reknow for his comments and Ben Lönnqvist for proofreading the manuscript. M.H.H., E.C., A.B., and M.R. designed the research; M.R. and E.C. performed the research; J.R.C., D.G., W.H.L., and O.F. analyzed data; J.R.C., D.G., O.F. A.B., P.F., and M.H.H. wrote the paper. The code that support the findings of this study are available upon request. This work was partially funded by the Fundação para a Ciência e a Tecnologia under grant FCT PD/BD/105785/2014 and the National Centre of Competence in Research (NCCR) Synapsy financed by the Swiss National Science Foundation under grant 51NF40-185897. The authors declare no competing interests.

Appendix A. Supplementary Information File

References

- Abdi, H., Williams, L.J., 2013. Partial Least Squares Methods: Partial Least Squares Correlation and Partial Least Square Regression, in: Reisfeld, B., Mayeno, A.N. (Eds.), *Computational Toxicology, Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp. 549–579. https://doi.org/10.1007/978-1-62703-059-5_23
- Andreasen, N.C., Paradiso, S., O’Leary, D.S., 1998. “Cognitive Dysmetria” as an Integrative Theory of Schizophrenia: A Dysfunction in Cortical-Subcortical-Cerebellar Circuitry? *Schizophr. Bull.* 24, 203–218. <https://doi.org/10.1093/oxfordjournals.schbul.a033321>
- Andreou, C., Leicht, G., Nolte, G., Polomac, N., Moritz, S., Karow, A., Hanganu-Opatz, I.L., Engel, A.K., Mulert, C., 2015. Resting-state theta-band connectivity and verbal memory in schizophrenia and in the high-risk state. *Schizophr. Res.* 161, 299–307. <https://doi.org/10.1016/j.schres.2014.12.018>

- 341 Avila, M., Thaker, G., Adami, H., 2001. Genetic epidemiology and schizophrenia: a study of
342 reproductive fitness. *Schizophr. Res.* 47, 233–241. [https://doi.org/10.1016/S0920-](https://doi.org/10.1016/S0920-9964(00)00062-1)
343 [9964\(00\)00062-1](https://doi.org/10.1016/S0920-9964(00)00062-1)
- 344 Bassett, A.S., Bury, A., Hodgkinson, K.A., Honer, W.G., 1996. Reproductive fitness in familial
345 schizophrenia. *Schizophr. Res.* 21, 151–160. [https://doi.org/10.1016/0920-9964\(96\)00018-7](https://doi.org/10.1016/0920-9964(96)00018-7)
- 346 Boutros, N.N., Arfken, C., Galderisi, S., Warrick, J., Pratt, G., Iacono, W., 2008. The status of spectral
347 EEG abnormality as a diagnostic test for schizophrenia. *Schizophr. Res.* 99, 225–237.
348 <https://doi.org/10.1016/j.schres.2007.11.020>
- 349 Braff, D.L., Freedman, R., Schork, N.J., Gottesman, I.I., 2007a. Deconstructing Schizophrenia: An
350 Overview of the Use of Endophenotypes in Order to Understand a Complex Disorder. *Schizophr.*
351 *Bull.* 33, 21–32. <https://doi.org/10.1093/schbul/sbl049>
- 352 Braff, D.L., Light, G.A., Swerdlow, N.R., 2007b. Prepulse Inhibition and P50 Suppression Are Both
353 Deficient but not Correlated in Schizophrenia Patients. *Biol. Psychiatry* 61, 1204–1207.
354 <https://doi.org/10.1016/j.biopsych.2006.08.015>
- 355 Burmeister, M., McInnis, M.G., Zöllner, S., 2008. Psychiatric genetics: progress amid controversy. *Nat.*
356 *Rev. Genet.* 9, 527–540. <https://doi.org/10.1038/nrg2381>
- 357 Burns, J., 2006. The social brain hypothesis of schizophrenia. *World Psychiatry Off. J. World Psychiatr.*
358 *Assoc. WPA* 5, 77–81.
- 359 da Cruz, J.R., Chicherov, V., Herzog, M.H., Figueiredo, P., 2018. An automatic pre-processing pipeline
360 for EEG analysis (APP) based on robust statistics. *Clin. Neurophysiol.* 129, 1427–1437.
361 <https://doi.org/10.1016/j.clinph.2018.04.600>
- 362 da Cruz, J.R., Favrod, O., Roinishvili, M., Chkonia, E., Brand, A., Mohr, C., Figueiredo, P., Herzog,
363 M.H., 2020a. EEG microstates are a candidate endophenotype for schizophrenia. *Nat. Commun.*
364 11, 3089. <https://doi.org/10.1038/s41467-020-16914-1>

- 365 da Cruz, J.R., Shaqiri, A., Roinishvili, M., Favrod, O., Chkonia, E., Brand, A., Figueiredo, P., Herzog,
366 M.H., 2020b. Neural Compensation Mechanisms of Siblings of Schizophrenia Patients as
367 Revealed by High-Density EEG. *Schizophr. Bull.* 46, 1009–1018.
368 <https://doi.org/10.1093/schbul/sbz133>
- 369 Di Lorenzo, G., Daverio, A., Ferrentino, F., Santarnecchi, E., Ciabattini, F., Monaco, L., Lisi, G., Barone,
370 Y., Di Lorenzo, C., Niolu, C., Seri, S., Siracusano, A., 2015. Altered resting-state EEG source
371 functional connectivity in schizophrenia: the effect of illness duration. *Front. Hum. Neurosci.* 9.
372 <https://doi.org/10.3389/fnhum.2015.00234>
- 373 Dickinson, D., Goldberg, T.E., Gold, J.M., Elvevag, B., Weinberger, D.R., 2011. Cognitive Factor
374 Structure and Invariance in People With Schizophrenia, Their Unaffected Siblings, and Controls.
375 *Schizophr. Bull.* 37, 1157–1167. <https://doi.org/10.1093/schbul/sbq018>
- 376 Favrod, O., Roinishvili, M., da Cruz, J.R., Brand, A., Okruashvili, M., Gamkrelidze, T., Figueiredo, P.,
377 Herzog, M.H., Chkonia, E., Shaqiri, A., 2018. Electrophysiological correlates of visual backward
378 masking in patients with first episode psychosis. *Psychiatry Res. Neuroimaging* 282, 64–72.
379 <https://doi.org/10.1016/j.psychresns.2018.10.008>
- 380 Friston, K., Brown, H.R., Siemerikus, J., Stephan, K.E., 2016. The dysconnection hypothesis (2016).
381 *Schizophr. Res.* 176, 83–94. <https://doi.org/10.1016/j.schres.2016.07.014>
- 382 Garobbio, S., Roinishvili, M., Favrod, O., da Cruz, J.R., Chkonia, E., Brand, A., Herzog, M.H., 2021.
383 Electrophysiological correlates of visual backward masking in patients with bipolar disorder.
384 *Psychiatry Res. Neuroimaging* 307, 111206. <https://doi.org/10.1016/j.psychresns.2020.111206>
- 385 Hilker, R., Helenius, D., Fagerlund, B., Skytthe, A., Christensen, K., Werge, T.M., Nordentoft, M.,
386 Glenthøj, B., 2018. Heritability of Schizophrenia and Schizophrenia Spectrum Based on the
387 Nationwide Danish Twin Register. *Biol. Psychiatry* 83, 492–498.
388 <https://doi.org/10.1016/j.biopsych.2017.08.017>

- 389 Howes, O.D., Kapur, S., 2009. The Dopamine Hypothesis of Schizophrenia: Version III--The Final
390 Common Pathway. *Schizophr. Bull.* 35, 549–562. <https://doi.org/10.1093/schbul/sbp006>
- 391 Hu, W., MacDonald, M.L., Elswick, D.E., Sweet, R.A., 2015. The glutamate hypothesis of
392 schizophrenia: evidence from human brain tissue studies. *Ann. N. Y. Acad. Sci.* 1338, 38–57.
393 <https://doi.org/10.1111/nyas.12547>
- 394 Keller, M.C., Miller, G., 2006. Resolving the paradox of common, harmful, heritable mental disorders:
395 Which evolutionary genetic models work best? *Behav. Brain Sci.* 29, 385–404.
396 <https://doi.org/10.1017/S0140525X06009095>
- 397 Kessler, R.C., Birnbaum, H., Demler, O., Falloon, I.R.H., Gagnon, E., Guyer, M., Howes, M.J., Kendler,
398 K.S., Shi, L., Walters, E., Wu, E.Q., 2005. The prevalence and correlates of nonaffective
399 psychosis in the National Comorbidity Survey Replication (NCS-R). *Biol. Psychiatry* 58, 668–
400 676. <https://doi.org/10.1016/j.biopsych.2005.04.034>
- 401 Kim, D.-J., Jeong, J., Chae, J.-H., Park, S., Yong Kim, S., Jin Go, H., Paik, I.-H., Kim, K.-S., Choi, B.,
402 2000. An estimation of the first positive Lyapunov exponent of the EEG in patients with
403 schizophrenia. *Psychiatry Res. Neuroimaging* 98, 177–189. [https://doi.org/10.1016/S0925-
404 4927\(00\)00052-4](https://doi.org/10.1016/S0925-4927(00)00052-4)
- 405 Krishnan, A., Williams, L.J., McIntosh, A.R., Abdi, H., 2011. Partial Least Squares (PLS) methods for
406 neuroimaging: A tutorial and review. *NeuroImage* 56, 455–475.
407 <https://doi.org/10.1016/j.neuroimage.2010.07.034>
- 408 Liu, C., Everall, I., Pantelis, C., Bousman, C., 2019. Interrogating the Evolutionary Paradox of
409 Schizophrenia: A Novel Framework and Evidence Supporting Recent Negative Selection of
410 Schizophrenia Risk Alleles. *Front. Genet.* 10. <https://doi.org/10.3389/fgene.2019.00389>
- 411 MacCabe, J.H., Koupil, I., Leon, D.A., 2009. Lifetime reproductive output over two generations in
412 patients with psychosis and their unaffected siblings: the Uppsala 1915–1929 Birth Cohort
413 Multigenerational Study. *Psychol. Med.* 39, 1667. <https://doi.org/10.1017/S0033291709005431>

- 414 McIntosh, A.R., Bookstein, F.L., Haxby, J.V., Grady, C.L., 1996. Spatial Pattern Analysis of Functional
415 Brain Images Using Partial Least Squares. *NeuroImage* 3, 143–157.
416 <https://doi.org/10.1006/nimg.1996.0016>
- 417 McIntosh, A.R., Chau, W.K., Protzner, A.B., 2004. Spatiotemporal analysis of event-related fMRI data
418 using partial least squares. *NeuroImage* 23, 764–775.
419 <https://doi.org/10.1016/j.neuroimage.2004.05.018>
- 420 Messamore, E., 2012. Niacin subsensitivity is associated with functional impairment in schizophrenia.
421 *Schizophr. Res.* 137, 180–184. <https://doi.org/10.1016/j.schres.2012.03.001>
- 422 Meyer-Lindenberg, A., Poline, J.-B., Kohn, P.D., Holt, J.L., Egan, M.F., Weinberger, D.R., Berman,
423 K.F., 2001. Evidence for Abnormal Cortical Functional Connectivity During Working Memory in
424 Schizophrenia. *Am. J. Psychiatry* 158, 1809–1817. <https://doi.org/10.1176/appi.ajp.158.11.1809>
- 425 Moreno-Küstner, B., Martín, C., Pastor, L., 2018. Prevalence of psychotic disorders and its association
426 with methodological issues. A systematic review and meta-analyses. *PLOS ONE* 13, e0195687.
427 <https://doi.org/10.1371/journal.pone.0195687>
- 428 Mothi, S.S., Sudarshan, M., Tandon, N., Tamminga, C., Pearlson, G., Sweeney, J., Clementz, B.,
429 Keshavan, M.S., 2019. Machine learning improved classification of psychoses using clinical and
430 biological stratification: Update from the bipolar-schizophrenia network for intermediate
431 phenotypes (B-SNIP). *Schizophr. Res.* 214, 60–69. <https://doi.org/10.1016/j.schres.2018.04.037>
- 432 Nikulin, V.V., Jönsson, E.G., Brismar, T., 2012. Attenuation of long-range temporal correlations in the
433 amplitude dynamics of alpha and beta neuronal oscillations in patients with schizophrenia.
434 *NeuroImage* 61, 162–169. <https://doi.org/10.1016/j.neuroimage.2012.03.008>
- 435 Poldrack, R.A., Huckins, G., Varoquaux, G., 2020. Establishment of Best Practices for Evidence for
436 Prediction: A Review. *JAMA Psychiatry* 77, 534.
437 <https://doi.org/10.1001/jamapsychiatry.2019.3671>

- 438 Price, G.W., Michie, P.T., Johnston, J., Innes-Brown, H., Kent, A., Clissa, P., Jablensky, A.V., 2006. A
439 Multivariate Electrophysiological Endophenotype, from a Unitary Cohort, Shows Greater
440 Research Utility than Any Single Feature in the Western Australian Family Study of
441 Schizophrenia. *Biol. Psychiatry* 60, 1–10. <https://doi.org/10.1016/j.biopsych.2005.09.010>
- 442 Rieger, K., Diaz Hernandez, L., Baenninger, A., Koenig, T., 2016. 15 Years of Microstate Research in
443 Schizophrenia – Where Are We? A Meta-Analysis. *Front. Psychiatry* 7.
444 <https://doi.org/10.3389/fpsy.2016.00022>
- 445 Rybakowski, J., Weterle, R., 1991. Niacin test in schizophrenia and affective illness. *Biol. Psychiatry*
446 29, 834–836. [https://doi.org/10.1016/0006-3223\(91\)90202-W](https://doi.org/10.1016/0006-3223(91)90202-W)
- 447 Saha, S., Chant, D., Welham, J., McGrath, J., 2005. A systematic review of the prevalence of
448 schizophrenia. *PLoS Med.* 2, e141. <https://doi.org/10.1371/journal.pmed.0020141>
- 449 Schnack, H.G., Kahn, R.S., 2016. Detecting Neuroimaging Biomarkers for Psychiatric Disorders:
450 Sample Size Matters. *Front. Psychiatry* 7. <https://doi.org/10.3389/fpsy.2016.00050>
- 451 Seidman, L.J., Helleman, G., Nuechterlein, K.H., Greenwood, T.A., Braff, D.L., Cadenhead, K.S.,
452 Calkins, M.E., Freedman, R., Gur, R.E., Gur, R.C., Lazzeroni, L.C., Light, G.A., Olincy, A.,
453 Radant, A.D., Siever, L.J., Silverman, J.M., Sprock, J., Stone, W.S., Sugar, C., Swerdlow, N.R.,
454 Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Green, M.F., 2015. Factor structure and heritability
455 of endophenotypes in schizophrenia: Findings from the Consortium on the Genetics of
456 Schizophrenia (COGS-1). *Schizophr. Res., Endophenotypes in Schizophrenia* 163, 73–79.
457 <https://doi.org/10.1016/j.schres.2015.01.027>
- 458 Simeone, J.C., Ward, A.J., Rotella, P., Collins, J., Windisch, R., 2015. An evaluation of variation in
459 published estimates of schizophrenia prevalence from 1990–2013: a systematic literature
460 review. *BMC Psychiatry* 15. <https://doi.org/10.1186/s12888-015-0578-7>

- 461 Sun, J., Tang, Y., Lim, K.O., Wang, J., Tong, S., Li, H., He, B., 2014. Abnormal Dynamics of EEG
462 Oscillations in Schizophrenia Patients on Multiple Time Scales. *IEEE Trans. Biomed. Eng.* 61,
463 1756–1764. <https://doi.org/10.1109/TBME.2014.2306424>
- 464 Toomey, R., Faraone, S.V., Seidman, L.J., Kremen, W.S., Pepple, J.R., Tsuang, M.T., 1998.
465 Association of neuropsychological vulnerability markers in relatives of schizophrenic patients.
466 *Schizophr. Res.* 31, 89–98. [https://doi.org/10.1016/S0920-9964\(98\)00025-5](https://doi.org/10.1016/S0920-9964(98)00025-5)
- 467 Tucker, L.R., 1958. An inter-battery method of factor analysis. *Psychometrika* 23, 111–136.
468 <https://doi.org/10.1007/BF02289009>
- 469 Uhlhaas, P.J., Singer, W., 2010. Abnormal neural oscillations and synchrony in schizophrenia. *Nat.*
470 *Rev. Neurosci.* 11, 100–113. <https://doi.org/10.1038/nrn2774>
- 471 Varoquaux, G., 2018. Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage*
472 180, 68–77. <https://doi.org/10.1016/j.neuroimage.2017.06.061>
- 473 Venables, N.C., Bernat, E.M., Sponheim, S.R., 2009. Genetic and Disorder-Specific Aspects of Resting
474 State EEG Abnormalities in Schizophrenia. *Schizophr. Bull.* 35, 826–839.
475 <https://doi.org/10.1093/schbul/sbn021>
- 476 Yang, H., Liu, J., Sui, J., Pearlson, G., Calhoun, V.D., 2010. A Hybrid Machine Learning Method for
477 Fusing fMRI and Genetic Data: Combining both Improves Classification of Schizophrenia. *Front.*
478 *Hum. Neurosci.* 4. <https://doi.org/10.3389/fnhum.2010.00192>
- 479 Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser.*
480 *B Stat. Methodol.* 67, 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

481