# Results

### Static Functional Network Connectivity

As a pre-analysis sanity test, we compared static functional network connectivity (sFNC) in the same fashion as (Du et al. 2020) to confirm that the data were not compromised. Figure (a) displays the mean sFNC of the FBIRN dataset, (b) the *t*-statistic returned by two-sample *t*-test comparisons between controls and patients, and (c) the connections from (b) which survive Bonferroni multiple comparison correction. No corresponding matrices are displayed in (Du et al. 2020) for comparison; however, the extant sFNC matrices display the diagonal block pattern expected of such matrices, which suggests that the data have been correctly reconstructed.

### Dynamic Functional Network Connectivity

Data were provided as sliding-window correlation matrices with 137 time points. The tapered sliding window generated by convolving a rectangular window of 40 seconds with a normal distribution of zero mean and a standard deviation of (Du et al. 2020). Vectorization of each dFNC matrix produced a data array of correlation values for time point , thus producing a array of dFNC values for each subject. Applying the Marčenko-Pasteur theorem to this array, after concatenation over all subjects, reliably detected 171 temporal independent components (tICs), which captures just over 80% of variance in the dFNC matrix. Selecting a tIC number within the proposed window of 16 to 22 tICs, on the other hand, captures less than 44% of data variance. Within this window, silhouette and Calinski-Harabasz scores are universally poor, with the least bad occurring at 17 tICs. By contrast, Davies-Bouldin scores propose the use of 22 tICs. These contradictory results may reflect the very low numbers (16-22) contained within this window. The chart of the cumulative sum of variance captured only begins to form an appreciable “elbow” in the vicinity of 100 tICs (~70% of variance), suggesting that the true dimensionality of dFNC may be substantially higher than previously reported. It is also possible that the conflicting results of the silhouette, Calinski-Harabasz, and Davies-Bouldin scores are due to the fact that these metrics were designed for clustering rather than blind source separation algorithms. As such, it is very likely that they do not accurately capture the ideal number of tICs necessary to represent dFNC dynamics.

### Joint Entropy

As each tIC is statistically independent from any other, each IC’s entropy and entropy rate can be estimated independently of the others. Estimation is necessary because we do not have access to the data’s complete probability distribution, which the formal definition of entropy requires. As such, we employed a variation on the Kozachenko-Leonenko estimator (Kozachenko and Leonenko 1987; Delattre and Fournier 2017) based on *k*-nearest-neighbor distances between sample points (Singh et al. 2003; Goria et al. 2005). Running this estimation algorithm on each subject’s tIC time courses produced a array of values, where is the number of ICs (171) and the number of subjects (311). To compute subject-level joint entropy production rate, we took the sum of this matrix along its first dimension to produce a array of joint entropy rate values.

A Kolmogorov-Smirnov two-sample test of the 171 tICs identified by the Marcěnko-Pasteur method showed that patient and subject joint entropy rate values differ (Figure []), with patients displaying a small but statistically significant elevation compared to controls , Table 1). A two-sample *t*-test and a difference-of-means permutation test (Krol 2021) qualitatively confirmed this result. Repeating the analysis with 22 tICs produces a similar result , Table 1). Schizophrenia patients thus appear to consistently generate entropy at a higher rate than healthy controls.

### Component Entropy

Having confirmed that subject joint entropy rate differs between populations, the next obvious question was whether these differences are evenly distributed amongst tICs or concentrate within a subset of them. To answer this question, we compared the group-level entropy rates of each tIC using the Kolmogorov-Smirnov two-sample test, with the false discovery rate (Benjamini and Hochberg 1995) correcting for multiple comparison. The Jarque-Bera test suggested that component-level subject entropies do not follow a normal distribution, so the *t*-test was dispensed with. Of the 171 tICs detected by the Marcěnko-Pasteur method, twelve (12) survive multiple comparison correction by the false discovery rate, while two (2) of the 22 Davies-Bouldin tICs survive. In all cases, patients displayed elevated entropy rate compared to controls.

### Entropy and Clinical Scores

A long-term goal of neuroimaging research is to identify links between identifiable neuroimaging data features and observable clinical or behavioral traits. To that end, we correlated positive and negative PANSS scores against both joint and tIC entropy rate distributions. Of the 171 tICs highlighted by the Marcěnko-Pasteur distribution, twenty (20) display significant correlations with clinical variables. Specifically, a diagnosis of schizophrenia correlates positively with joint entropy rate and entropy rate of components 31, 32, 58, 66, 69, 79, 103, and 116. Processing speed displays negative correlations with joint entropy and components 22, 32, and 58; verbal learning, with joint entropy and components 31 and 159; CMINDS composite score with joint entropy and components 31, 49, 58, and 159; age with components 37, 73, and 93; attention and vigilance, with components 40; working memory, with components 40 and 58; CPZ with component 43; visual learning, with components 49 and 58; reasoning and problem solving, with components 49, 109, 130, 132. Of the Davies-Bouldin tIC scores, on the other hand, only joint entropy rate and the entropy rate of component 13 survive multiple-comparison correction; both display significant positive correlations with a diagnosis of schizophrenia.