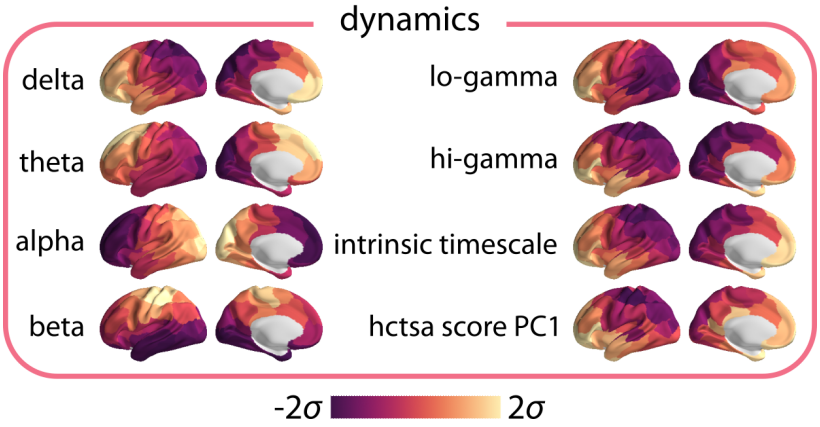
Supplementary Information: Neurophysiological signatures of cortical micro-architecture

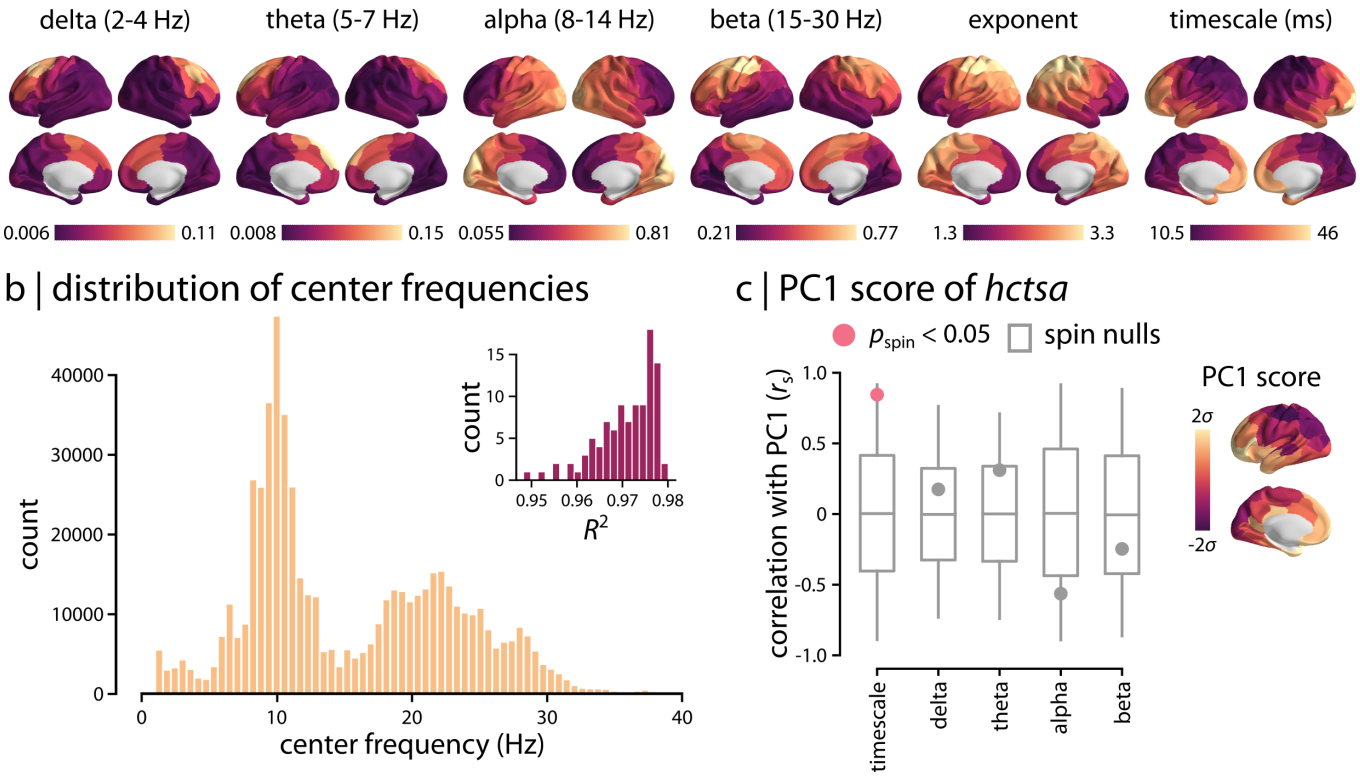
Golia Shafiei[1,2], Ben D. Fulcher[3], Bradley Voytek[4], Theodore D. Satterthwaite[2], Sylvain Baillet[1], Bratislav Misic[1,*]

[1]McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Montréal, Canada; [2]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; [3]School of Physics, The University of Sydney, NSW 2006, Australia; [4]Department of Cognitive Science, Halıcıoğlu Data Science Institute, University of California, San Diego, La Jolla, CA, USA
*Corresponding Author: Bratislav Misic. bratislav.misic@mcgill.ca

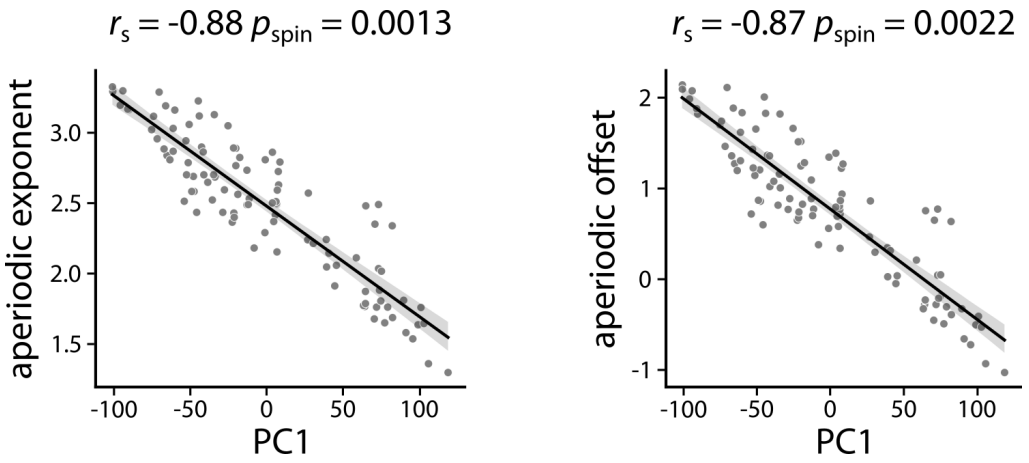
Supplementary Figures



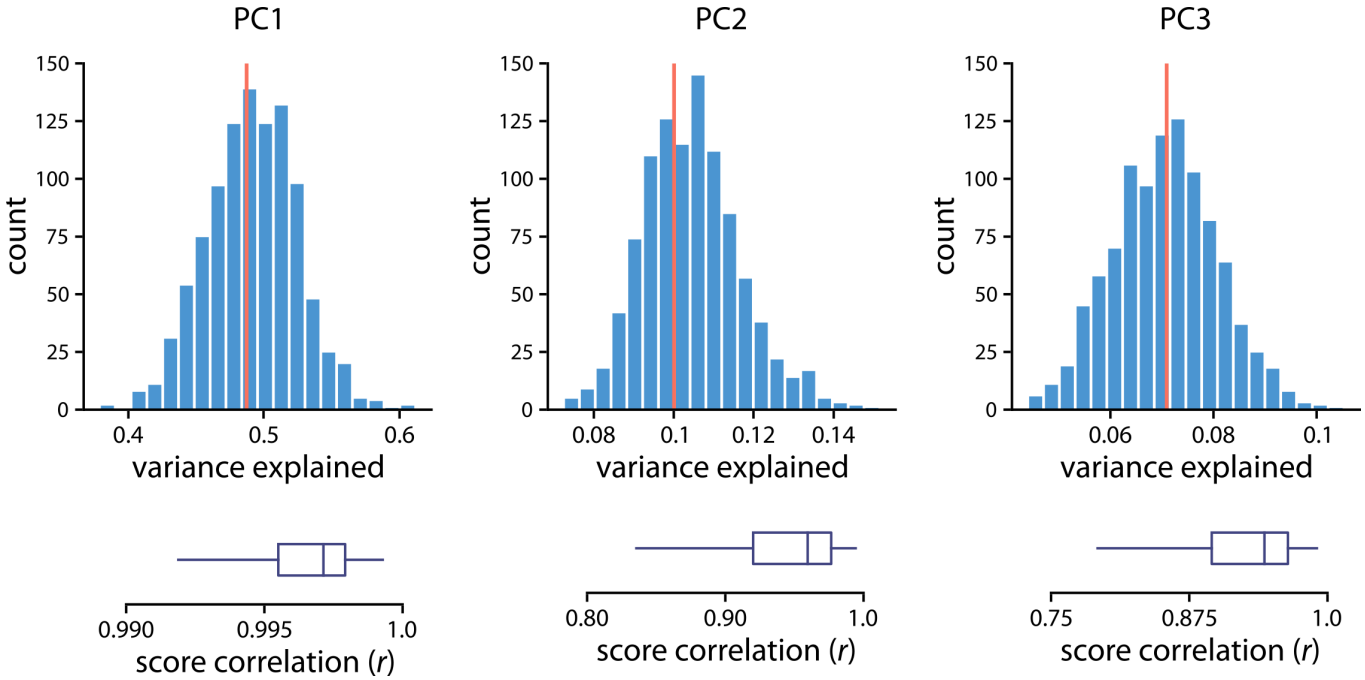
Supplementary Figure 1 . **Dynamical brain maps** | `Brainstorm` software was used to pre-process the resting-state MEG data and obtain power maps at six canonical electrophysiological bands (i.e., delta (δ: 2-4 Hz), theta (θ: 5-7 Hz), alpha (α: 8-12 Hz), beta (β: 15-29 Hz), low gamma (lo-γ: 30-59 Hz), and high gamma (hi-γ: 60-90Hz)) [1] (see *Methods* for more details). `FOOOF` algorithm was used to parametrize power spectral density and estimate the intrinsic timescale [2, 3] (see *Methods* for more details). Note that log-10 transformed intrinsic timescale map is shown here. Principal component analysis was used to estimate the principal component of the neurophysiological time-series features obtained from the `hctsa` toolbox (see Fig. 3). All obtained brains maps are depicted across the cortex at 95% confidence interval (Schaefer-100 atlas [4]).



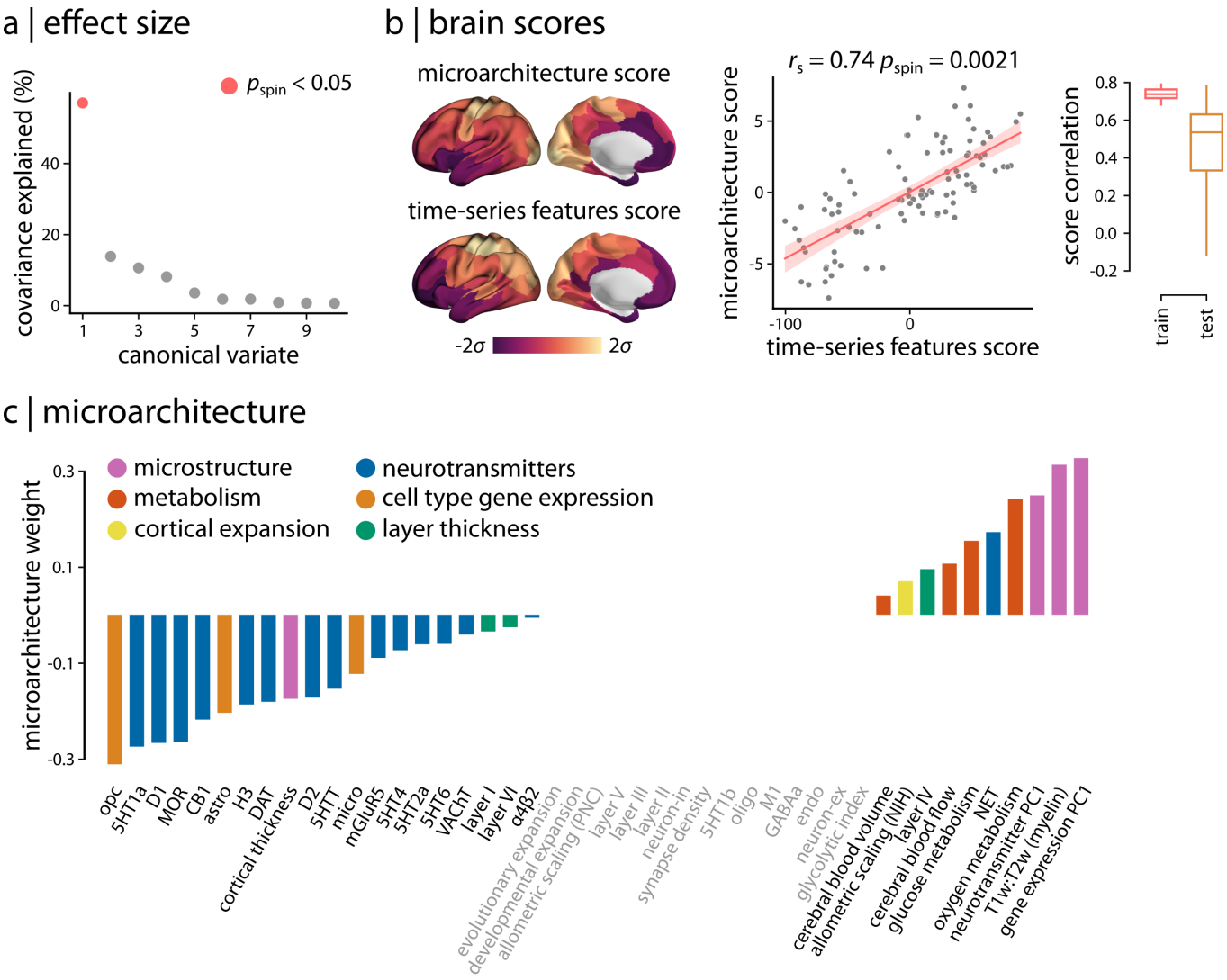
Supplementary Figure 2 . **Spectral parameterization of neurophysiological activity** | Spectral parameterization `FOOOF` toolbox was used to extract periodic and aperiodic components of the MEG power spectrum [2]. (a) The identified oscillatory peaks of the periodic component were used to estimate band-limited "oscillation score" maps at delta (2-4 Hz), theta (5-7 Hz), alpha (8-14 Hz), and beta (15-30 Hz) frequency bands. The oscillation scores reflect the average power at each region and frequency band, weighted by the probability of observing an oscillation peak at that region and band [2]. The aperiodic exponent and "knee" parameter (controls for the bend in the aperiodic component) were used to estimate intrinsic timescale (see *Methods* for details). Note that log-10 transformed intrinsic timescale map is depicted here for a more clear visualization. (b) Distribution of center frequencies of the identified periodic peaks are depicted across all vertices and participants. Visual inspection of the distribution shows clusters of peaks around the frequency bands shown in panel (a). Inset depicts the distribution of goodness-of-fit (R^2) for the `FOOOF` algorithm. (c) PC1 score map of `hctsa` time-series features was compared with the aperiodic-adjusted power maps and intrinsic timescale. Consistent with the results obtained with the total power maps at the canonical frequencies (Fig. 3), PC1 is significantly associated with the intrinsic timescale (FDR-corrected; 10 000 autocorrelation-preserving spin nulls). r_s denotes the Spearman's rank correlation coefficient.



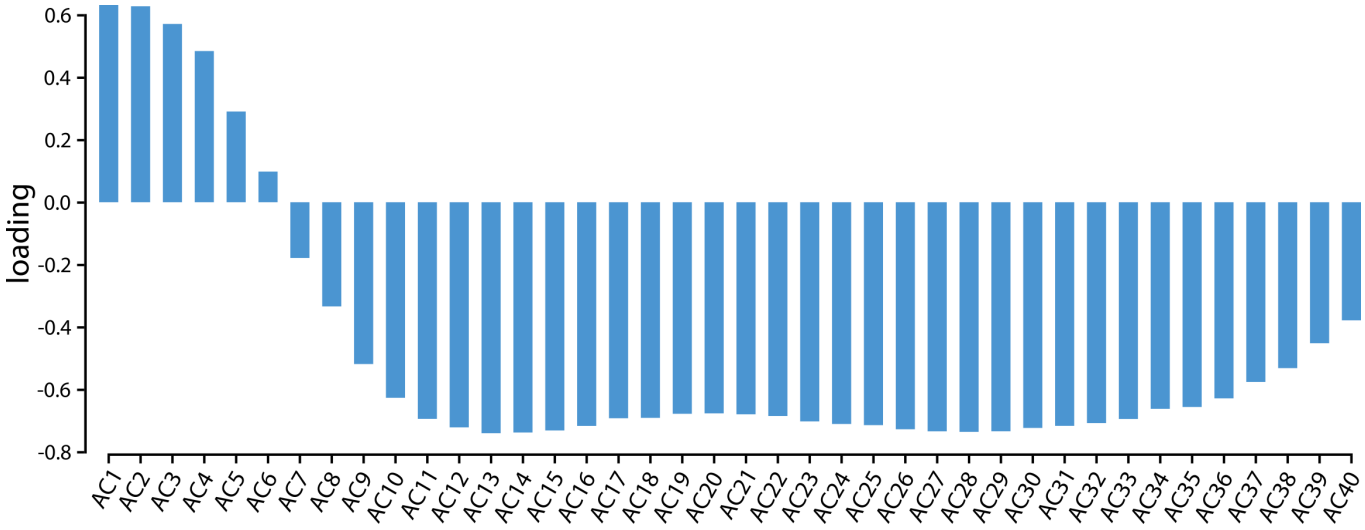
Supplementary Figure 3 . **Parameters of aperiodic component of power spectrum |** PC1 was correlated with the exponent and offset of the aperiodic component of the power spectrum. The exponent describes the "curve" or the overall "line" or the slope of the aperiodic component and the offset describes the overall vertical shift (up and down translation) of the whole power spectrum [2]. r_s denotes the Spearman's rank correlation coefficient.



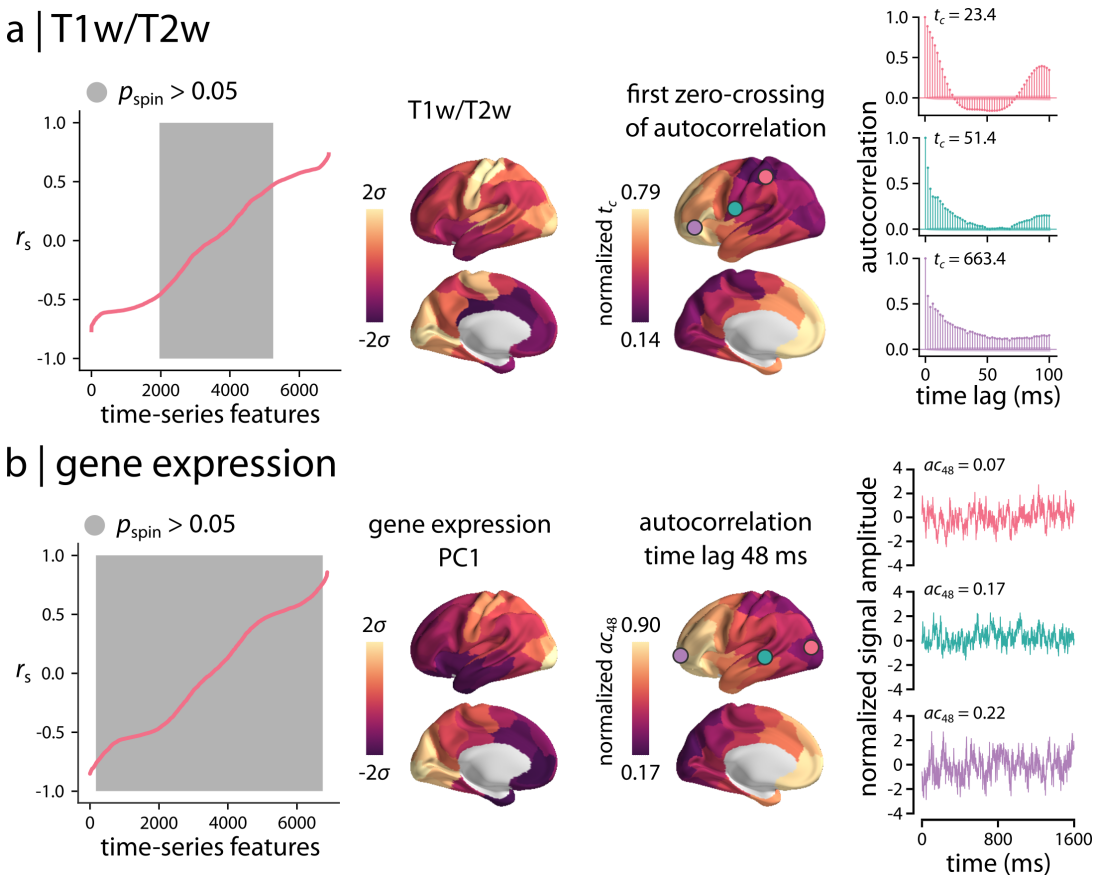
Supplementary Figure 4 . **PCA with reduced set of time-series features |** PCA was repeated 1 000 times with reduced sets of time-series features, where 100 time-series features were randomly selected from the original list of 6 880 features at each repetition. Variance explained by PC1, PC2, and PC3 as well as the Pearson correlation coefficients between the PCs from randomly selected feature matrices and the ones obtained from the original analysis are depicted. For each PC, variance explained obtained from the original analysis is shown with a red vertical line and the distribution of variance explained obtained from PCA on reduced feature sets is shown by blue histogram. Box plots depict the correlation coefficients between PC1 scores from the original PCA and the ones from reduced feature matrices.



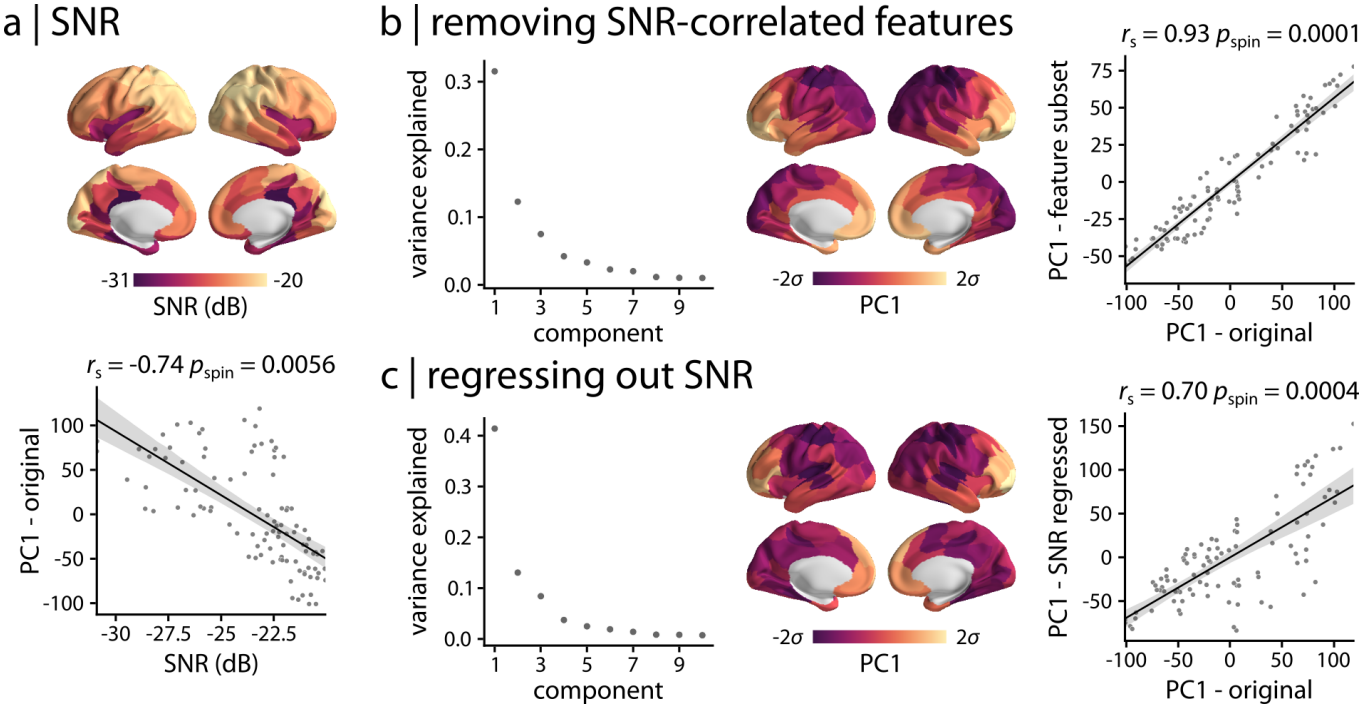
Supplementary Figure 5 . **Sparse canonical correlation analysis (sCCA)** | Sparse canonical correlation analysis (sCCA; [5]) was used as an alternative multivariate analysis approach to investigate the link between neurophysiological dynamics and micro-architectural features. The results are consistent with the findings from the original PLS analysis (Fig. 4). (a) First canonical variate explained 56.9% of the covariance, and (b) captured similar brain score patterns to PLS analysis. A distance-dependent cross-validation analysis was used to assess the out-of-sample relationship between sCCA-derived brain scores, identifying consistent relationship between sCCA scores in training and test sets (see *Methods*). (c) sCCA-derived weights for micro-architectural features are shown for each set of brain maps. Features labeled with grey color had zero weights due to the sparsity constraint imposed by sCCA.



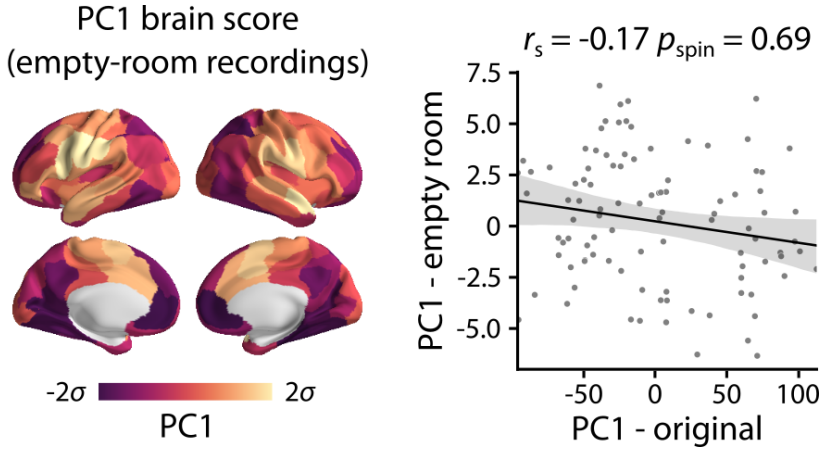
Supplementary Figure 6 . **PLS loadings for linear autocorrelation at various time lags |** The `hctsa` toolbox generates time-series features capturing the time-series linear autocorrelation at a range of time lags between 1 and 40 time steps (here corresponding to 2 ms and 80 ms, respectively). PLS loadings for all time lags are depicted as bar charts. ACX denotes autocorrelation at lag X, where X is the number of time steps (each corresponding to 2 ms). Note that autocorrelation at a lag of 48 ms (24 time steps) was shown as an example feature in Fig. 4c.



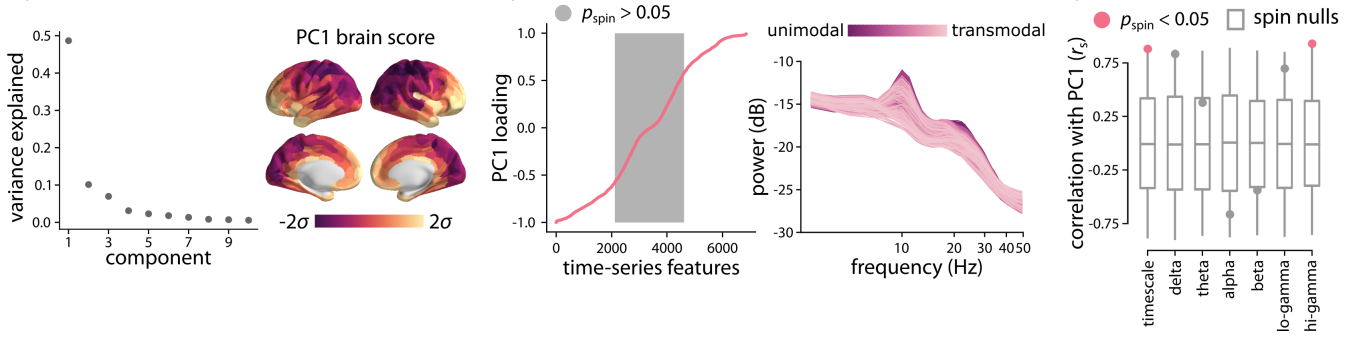
Supplementary Figure 7 . **Univariate analysis of neurophysiological time-series features** | Spearman's rank correlation coefficients (r_s) were used to investigate the univariate associations between `hctsa` time-series features of neurophysiological signal and two commonly-used micro-architectural maps: (a) T1w/Tw2 ratio as a proxy measure of intracortical myelin, and (b) principal component of gene expression. The resulting correlations were compared with null distributions of correlations obtained from 10 000 spatial autocorrelation-preserving nulls. Grey background indicates non-significant time-series features (FDR corrected). Examples of high loading time-series features are shown for each micro-architectural map. The group-average first zero-crossing time point of the autocorrelation function, t_c, is shown for T1w/T2w ratio. The group-average linear autocorrelation at a time lag of 48 ms, ac_{48}, is shown for principal component of gene expression. The autocorrelation function and short segments of raw time-series are also shown for a randomly selected participant at three different regions (circles on the brain surface: pink $\approx 5^{th}$ percentile, green $\approx 50^{th}$ percentile, purple $\approx 95^{th}$ percentile). Full lists of features, their correlation coefficients and p-values are available for T1w/T2w ratio and gene expression in the online Supplementary Datasets S5,6.



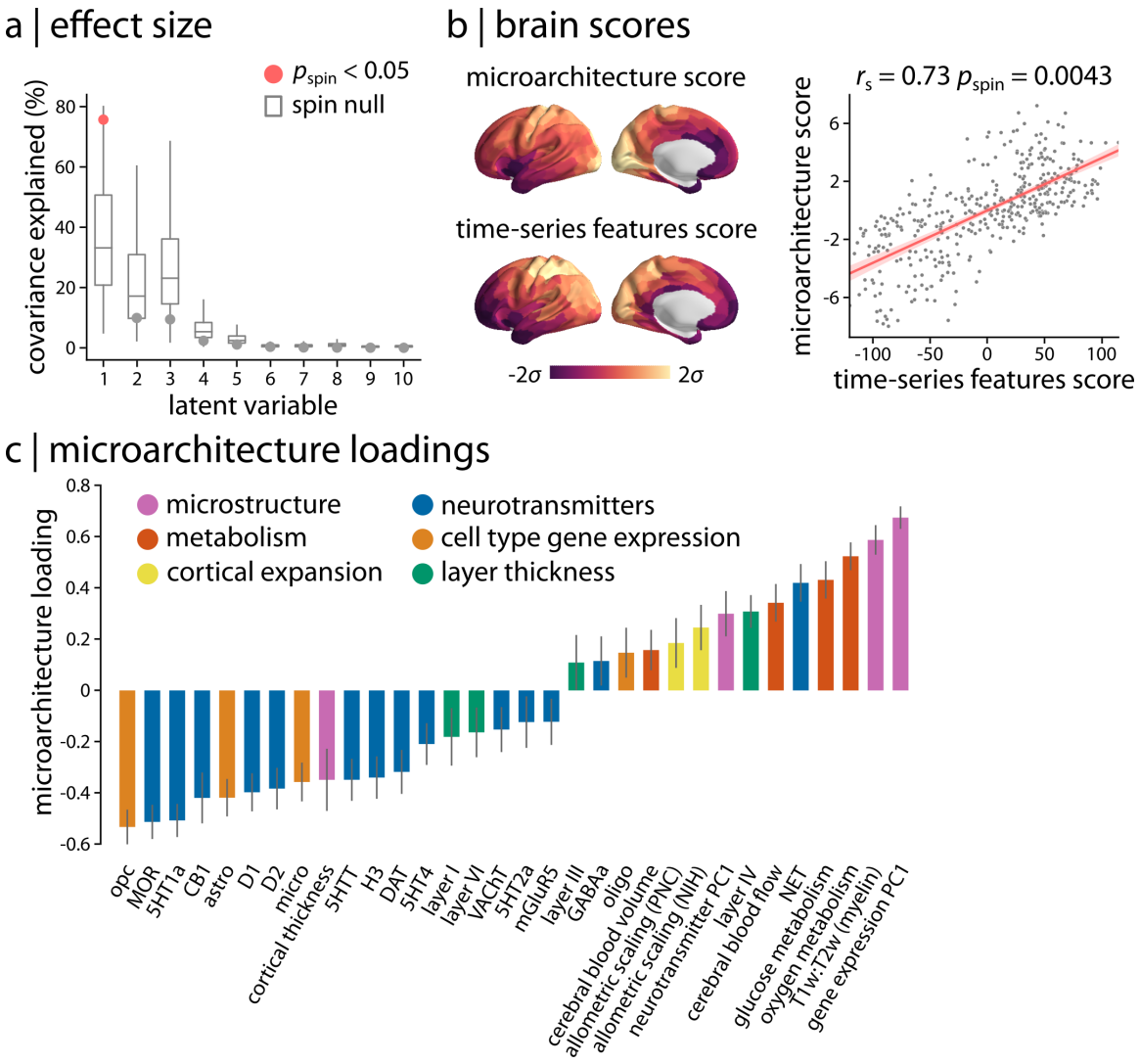
Supplementary Figure 8 . **Signal-to-noise ratio (SNR)** | (a) Source-level MEG signal-to-noise ratio (SNR) was estimated. Parcellated, group-average SNR map is depicted across the cortex. MEG SNR was compared with the principal component of neurophysiological dynamics (PC1 - original). (b) SNR was compared with full set of neurophysiological time-series features (i.e., 6 880 features) using univariate correlations. Features that were significantly correlated with SNR were removed without correcting for multiple comparisons ($p_{spin} < 0.05$, 10 000 spatial autocorrelation-preserving permutation tests) and PCA was repeated using the remaining 3 819 features. The principal component of the retained feature subset (PC1 - feature subset) explained 31.6% of the variance and was significantly correlated with the original PC1 from the full set of features. (c) SNR was linearly regressed out from the full set of time-series features. PCA was applied to the feature residuals. The principal component of SNR-regressed time-series features (PC1 - SNR regressed) explained 41.4% of the variance and was significantly correlated with the original PC1. r_s denotes the Spearman's rank correlation coefficient; linear regression lines are added to the scatter plots for visualization purposes only.



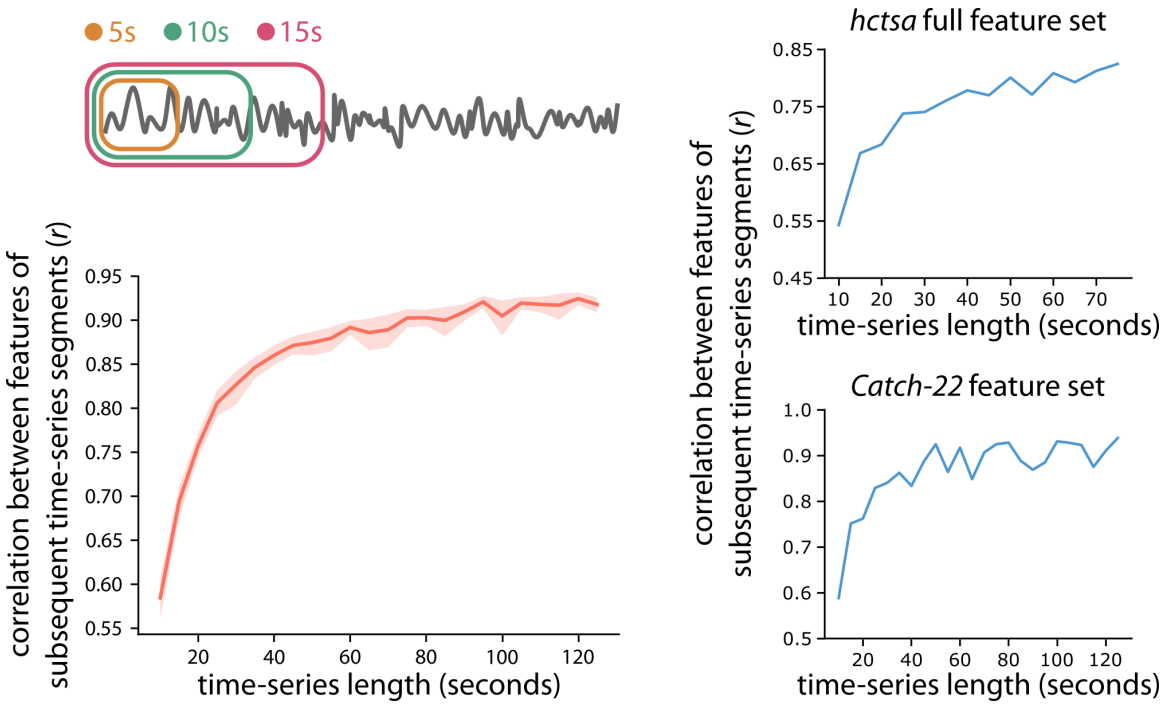
Supplementary Figure 9 . **Empty-room recordings** | PCA was applied to time-series features obtained from pre-processed empty-room MEG recordings. Following the Procrustes alignment of the resulting PCA weights with the PCA weights of the resting-state MEG recordings, the first principal components were compared between the two. The principal component of the empty-room time-series features (PC1 - empty room; variance explained = 41%) was not significantly correlated with the original PC1. r_s denotes the Spearman's rank correlation coefficient; linear regression line is added to the scatter plot for visualization purposes only.



Supplementary Figure 10 . **Topographic distribution of neurophysiological dynamics for Schaefer-400 |** (a) Principal component analysis (PCA) was applied to MEG time-series features at a higher resolution parcellation (i.e., Schaefer-400 atlas). The first principal component accounted for 48.6% of the variance. The spatial organization of features captured by PC1 is depicted across the cortex, displaying a consistent pattern with the original PCA results obtained for the Schaefer-100 atlas (Fig. 3). (b) Top loading features contributing to PC1 were identified using Pearson correlation coefficients between PC1 pattern and all time-series features. Grey background indicates non-significant features based on 10 000 spatial autocorrelation-preserving permutation tests (FDR corrected). Consistent with the original analysis, the top loading features were mainly related to power spectral density. Regional power spectral densities are depicted, where each line represents a brain region. Regions are coloured by their position in the putative unimodal–transmodal hierarchy [6]. The full list of features, their loadings and p-values are available in the online Supplementary Dataset S3. (c) PC1 pattern was directly correlated with MEG power maps at 6 canonical frequency bands and intrinsic timescales. Consistent with the original findings, PC1 score was significantly correlated with hi-gamma power and intrinsic timescale (FDR-corrected). r_s denotes the Spearman's rank correlation coefficient.



Supplementary Figure 11 . **Partial least squares (PLS) analysis for Schaefer-400** | PLS analysis was used to assess the multivariate relationship between micro-architectural and time-series features for the Schaefer-400 atlas. The results were consistent with the original findings for the Schaefer-100 atlas (Fig. 4). (a) PLS identified a single significant latent variable ($p_{\text{spin}} = 0.0083$, covariance explained $= 75.7\%$). (b) Spatial patterns of micro-architecture and time-series features scores are depicted for the first latent variable. The two brain score maps were significantly correlated, demonstrating similar patterns to the ones obtained for Schaefer-100 atlas (Fig. 4b). (c) Micro-architectural feature loadings were also consistent with the original findings (Fig. 4d). Full list of time-series feature loadings are included in the online Supplementary Dataset S4. Consistent with the original analysis, the top loading features were mainly related to the linear correlation structure of the signal. r_s denotes the Spearman's rank correlation coefficient.



Supplementary Figure 12 . **Stability of time-series features** | To identify the time-series length required to robustly estimate the time-series features, we calculated a subset of `hctsa` features using the `catch-22` toolbox [7] on subsequent segments of time-series with varying length for each participant. We extracted time-series features from short segments of data ranging from 5 to 125 seconds in increments of 5 seconds. To identify the optimal time-series length required to estimate robust and stable features, we calculated the Pearson correlation coefficient r between features of two subsequent segments (e.g., features estimated from 10 and 5 seconds of data). The group-average correlation coefficient between the estimated features started to stabilize at time-series segments of around 30 seconds, consistent with previous reports [8] (left). To compare the stability analysis of `catch-22` features with full `hctsa` features, the correlation coefficients between subsequent segments of time-series are shown for a randomly selected participant (right).

Supplementary References

[1] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, Computational intelligence and neuroscience **2011** (2011).

[2] T. Donoghue, M. Haller, E. J. Peterson, P. Varma, P. Sebastian, R. Gao, T. Noto, A. H. Lara, J. D. Wallis, R. T. Knight, *et al.*, Nature neuroscience **23**, 1655 (2020).

[3] R. Gao, R. L. van den Brink, T. Pfeffer, and B. Voytek, Elife **9**, e61277 (2020).

[4] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, and B. T. Yeo, Cerebral cortex **28**, 3095 (2018).

[5] D. M. Witten, R. Tibshirani, and T. Hastie, Biostatistics **10**, 515 (2009).

[6] D. S. Margulies, S. S. Ghosh, A. Goulas, M. Falkiewicz, J. M. Huntenburg, G. Langs, G. Bezgin, S. B. Eickhoff, F. X. Castellanos, M. Petrides, *et al.*, Proceedings of the National Academy of Sciences **113**, 12574 (2016).

[7] C. H. Lubba, S. S. Sethi, P. Knaute, S. R. Schultz, B. D. Fulcher, and N. S. Jones, Data Mining and Knowledge Discovery **33**, 1821 (2019).

[8] A. I. Wiesman, J. da Silva Castanheira, and S. Baillet, Neuroimage **247**, 118823 (2022).