

1 **Author response for submission # 3770.** We thank the reviewers for their valuable feedback.

2 §R1. *limitations to using Gaussian-based methods . . .* BOLD time series were not binned into counts, but rather directly
3 analyzed as a continuous signal. GPFA was applied without conventional pre-processing steps for spikes (binning,
4 square-root transform, kernel smoothing). We also verified that a majority of the fMRI time series data satisfied
5 Gaussianity assumptions. For the HCP dataset, 78% of regional time series passed the Lilliefors test for normality. For
6 the MCI dataset, 81% passed the test for normality. These are now reported in the revision.

7 §R1. *Given the variances of BOLD counts is known to increase with the mean . . . Does a Poisson . . . correction need to*
8 *be considered?* To avoid the confound of variance scaling (or changing) with the mean, observed BOLD time series
9 were variance normalized by z-scoring (mean subtracted and divided by standard deviation) before GPFA. Future work
10 can incorporate a Poisson correction, so that the method can be directly applied to the observed BOLD data.

11 §R2. *the authors were able to decode the tasks so well - it does not allow for . . . insights into . . . how the brain works.*
12 Examining matrix C in the GPFA model (SI section 1, line 6), provides some insights into brain networks that contribute
13 to dynamics at specific timescales (e.g. map in Fig. 1D, top). Classification analysis with MCI patients, provides
14 insights into which networks, and at what timescales, may be implicated in cognitive decline (e.g. Fig 4C-D).

15 §R2. *the authors slightly miss-represent the current level of understanding . . .* We have revised the Introduction based
16 on the reviewer's suggestions to provide an up-to-date picture.

17 §R2. *how these dynamics are qualitatively different from the dynamics the authors study.* Dynamics in our study are
18 unlikely to be qualitatively different from those previously studied. Our approach permits quantifying the specific
19 timescale associated with each latent dimension, which enables linking dynamics at specific timescales with behavior.

20 §R2. *how one would be able to interpret any dynamics faster than . . . Nyquist frequency [of] 0.25 Hz.* The faster
21 timescales arise because HCP data is sampled at 0.72s (1.4 Hz). Interpreting timescales faster than 0.7 Hz is indeed
22 challenging. GPFA latents at these timescales likely reflect fits to the residual noise after accounting for slow latents.

23 §R2. *recommend framing it rather as a new method to study functional connectivity.* We have now emphasized the
24 novelty of the method for studying fMRI functional connectivity, and reduced the emphasis on slow dynamics.

25 §R2. *predicting MCIC, . . . requires a more detailed description . . .* We have now elaborated the description.

26 §R2. *unclear to me how the authors accounted for the class-imbalance . . . MCICs group was over-sampled?* The MCIC
27 group (the minority class) was over-sampled. To address class imbalance, all 22 MCIC subjects, and 22 MCIs subjects
28 sampled from the pool of 72, were used in each of 100 runs of the SVM. Each MCIs subject was sampled at least once
29 every 4 runs. For LOO cross-validation one pair of (1 MCIC and 1 MCIs) subjects was left out for testing, with the
30 remaining subjects being used to train the model. CV accuracies were averaged across the 100 runs.

31 §R2. *(CDR-SOB prediction) . . . Would a different CV strategy (5-fold for example) not be more appropriate?* We now
32 perform k-fold CV analysis with k=2,4,23,92 (92 subjects with CDR scores). Predicted vs observed score correlation
33 increased steadily with more folds, perhaps because of the small sample size and bias in the training data with few folds.

34 §R2. *. . . helpful to assess 'potential artifact' maps.* We have added an SI subsection that discusses these maps in detail.

35 §R2. *motion artifacts may be synchronized across subjects. . . ?* We acknowledge this in the revised Discussion.

36 §R2. *Would it be more accurate to state the GPFA method was at-least as good as conventional methods?* We
37 have revised the claim. Yet, only with GPFA did both spectra *and* time series provide high classification accuracies.
38 Significance was assessed with Wilcoxon signed rank test (SI section 2, lines 94-113).

39 §R2. *inform . . . the selection criteria of the behavioral measures* Behavioral scores were selected from Alertness,
40 Cognition, Motor categories (HCP Data Dictionary). Only scores that carried independent information were selected
41 (e.g. sensitivity and specificity, but not true/false positive rates). Scores were selected "blind" to the prediction analyses.

42 §R2. *correlation of .2 is not a very good prediction.* We have rephrased this claim more carefully.

43 §R2. *Improvements.* Addressed above (frequencies studied in fMRI; literature overview; details about MCI prediction).

44 §R3. *1. . . no causal effect to demonstrate that this improvement is related to fixing the slow hemodynamic issue!* We
45 do not propose a solution that overcomes slow hemodynamics. Rather, we estimate infra-slow dynamics by applying
46 GPFA to fMRI data (BOLD timeseries) directly, and demonstrate its relevance for cognitive state prediction.

47 §R3. *2. . . How will the performance . . . be changed if . . . fMRI datasets are collected from different machines or*
48 *locations?* Our results already incorporate this diversity. HCP data was acquired at Wash. Univ. in St. Louis. ADNI
49 data was collected at multiple centers in the US. Scanners were of different makes including, Siemens, Philips and GE.

50 §R3. *3. The proposed method must be clearly formulated.* Main paper (section 2) and SI (section 1) are revised.

51 §R3. *4. The limitations and applications of the proposed method . . .* Limitations include assumptions on Gaussianity,
52 linking hemodynamics to neural dynamics, and scalability for voxel-level estimation. Applications include predicting
53 behavioral scores, classifying pathologies and identifying signatures of cognitive decline. Discussed in the revision.

54 §R3. *5. . . what the effect of noise on the performance of the proposed method is?* GPFA reduces dimensionality by
55 identifying latent dimensions with maximal covariance among ROIs. With an explicit noise model, GPFA estimates
56 noise that is independent across ROIs (SI, lines 6-7). Thus, adding independent noise across ROIs generally does not
57 affect parameter estimates. Adding correlated noise enables GPFA to factor these into a separate latent dimension.

58 §R3. *6. Table S1 . . . is not informative-rich.* Publicly available data regarding subjects (gender and age-group) have now
59 been added in the revised SI (e.g. of 1000 subjects, 529 females, age range: 22-35 years, 10 subjects of age >36 years).