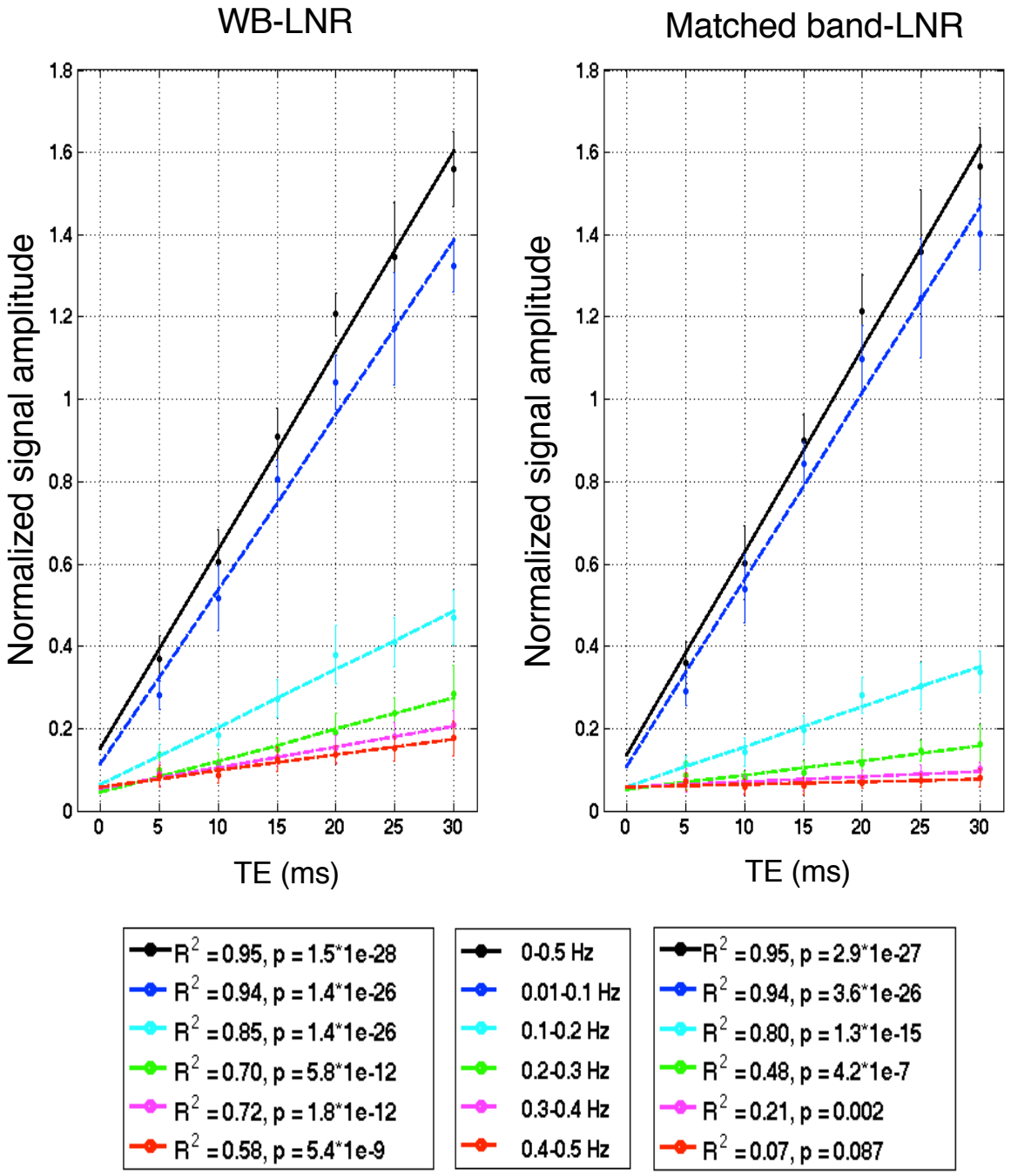

Fig. S1 Re-analyzing the TE dependence of functional signal amplitudes (Chen and Glover, 2015) using matched-band filtering ('mot6 + comb'). Functional connectivity signal amplitudes within the DMN vs. TE. Statistics shown here are standard linear model vs. Gaussian noise tests. To account for the inter-subject variability, each subject's correlated signal amplitudes were first normalized (divided by the mean correlated amplitude across different TEs of B0 band) before group fitting. Left: results estimated with WB-LNR ((Chen and Glover, 2015) Fig. 2, DMN); right: results estimated with matched-band filtering, demonstrating reduced HF response compared with original.

A. Dimensionality reduction using PCA

The principal components in temporal PCA are ordered by the amount of variance explained by each component. If the data is not pre-whitened by removing the temporal autocorrelation of each voxel, the *leading (not all)* PCs will be dominated by LF fluctuations, which contain the highest energy in the time series. As shown in A.1 Pitfalls of PCA, key parameters determining the *leading* PCs and the associated principal spaces do not deviate appreciably from results estimated solely from LF band signals. Therefore, HF components of these *leading* PCs and the data post dimensionality reduction (excluding PC dimensions with the least significance) would not preserve the intrinsic HF structures in the raw data but contain a distorted replication of information in the LF band if low dimensional leading PCs are preserved, in similar ways as LNR. For instance, Fig. S2 shows the DMN map of a typical subject post PCA. For both '*dummy*' and '*real*' dataset, spurious correlations start to occur in HB1 when PC dimensions are reduced, and the prominence of exhibited patterns scales inversely with the number of retained PCs.

A.1 Pitfalls of PCA

Let $S \in R^{n \times p}$ represent information from the raw time series, with *n* indicating the number of time frames and *p* indicating the number of voxels.

PCA decomposes *S* into orthogonal bases as:

$$T = SW$$

where columns of *T* refer to the derived principal components (PCs), and *W* is a *p*-by-*p* matrix whose columns are the eigenvectors of $S^T S$.

After reducing the data to *N* dimensions, i.e., projecting the data onto the space spanned by the leading *N* PCs:

$$S_N = T_N (T_N)^T S = (SW_N)(SW_N)^T S = S(W_N W_N^T S^T S)$$

T_N and W_N correspond to the first *N* columns of *T* and *W*, assuming they are ordered in descending orders of significance.

Let S_L and S_H correspond to the LF and HF fluctuations of S, since $S_L(S_H)^T, (S_H)(S_H)^T \ll S_L(S_L)^T$, leading eigenvectors of $S^T S$ (W_N) approach those derived from ${S_L}^T S_L$ (W_N^L), and

$$W_N W_N^T S^T S \approx W_N^L (W_N^L)^T S_L (S_L)^T$$

Thus,

$$S_N \approx S(W_N^L (W_N^L)^T S_L (S_L)^T)$$

Further filtering into LF and HF bands:

$$S_{N,L} \approx S_L (W_N^L (W_N^L)^T S_L (S_L)^T)$$
$$S_{N,H} \approx S_H (W_N^L (W_N^L)^T S_L (S_L)^T)$$

High frequency components of signals post PCA $S_{N,H}$ only reflect signal variances at low frequencies.

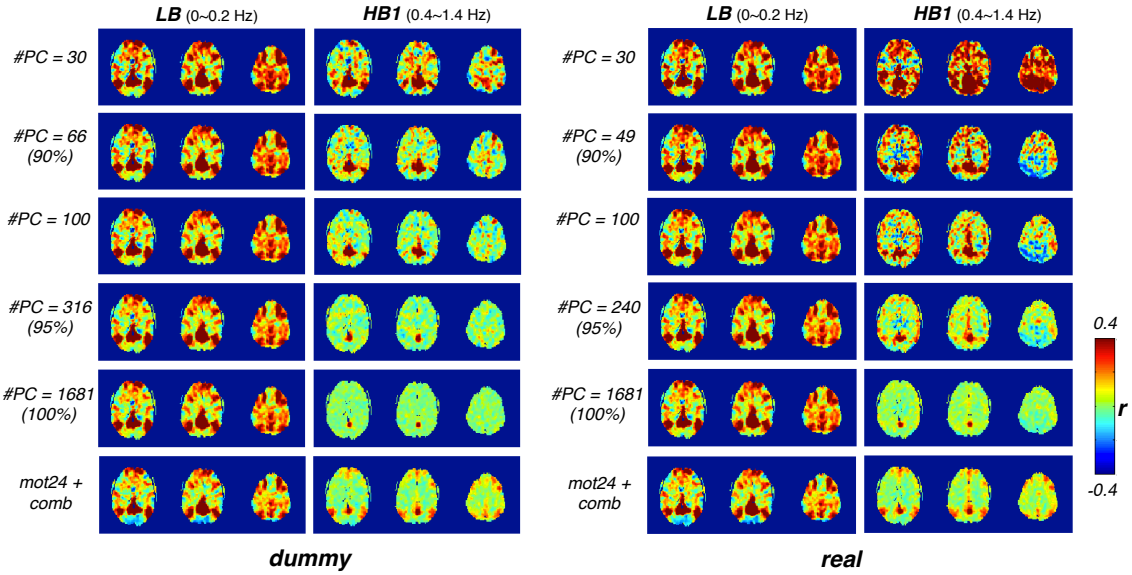

Figure S2 The DMN of '*dummy*' and '*real*' dataset post dimensionality reduction using temporal PCA, in a typical subject. Data are spatially smoothed with a Gaussian kernel (FWHM = 4mm) before PCA. Percentile values below the # of PCs indicates the percentage of variance explained by the retained PCs. The results post '*mot24+comb*' LNR are shown at the bottom for comparison.

B. Nuisance regression via Generalized least squares (GLS) instead of ordinary least squares (OLS)

The dominating role of LF fluctuations can be alleviated by approaches employing pre-whitening operations, for instance, generalized least squares (GLS, (Aitken, 1934)) based linear regression. Very briefly, GLS estimates the covariance of the residuals (Ω) by fitting observations to a theoretical model (AR(p) or low order ARMA, e.g., (Bullmore et al., 1996; Locascio et al., 1997; Purdon and Weisskoff, 1998)) or nonparametric approaches (e.g., (Woolrich et al., 2001)), and removes it out of the data via linear transformation ('pre-whitening') before OLS.

$$y = X\beta + \varepsilon \qquad \text{eqn. (1)}$$

OLS: $$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y \qquad \text{eqn. (2)}$$

GLS: $$\hat{\beta}_{GLS} = (X^T \Omega^{-1} X)^{-1} X^T \Omega^{-1} y \qquad \text{eqn. (3)}$$

In GLS, the whitened data favors no particular frequency band, and the issue of LF-RSFC information leaking into HF bands is hence diminished. Figure S3 shows the results of '*dummy*' dataset post GLS based nuisance regression, spurious HF FC is still present but significantly attenuated compared to OLS (Fig. 4, main text).

Since GLS accounts for the effective degree of temporal autocorrelation in the residuals, it leads to the best linear unbiased estimator for the scaling parameter β and has been commonly employed in task-based studies to generate statistically efficient inferences on task effects. Nonetheless, GLS is less conventional for nuisance regression than OLS in RS studies, Partly, instead of making inferences on the relation between nuisance factors and actual observation, LNR simply aims to project any fluctuations synchronized (linearly correlated) with nuisance sources out of observations, in which sense, OLS yields the most consistent and robust results. Secondly, the majority of existing RS literatures are still based on analyses of the auto-correlated, low-frequency driven residuals, regardless of whether OLS or GLS is employed in LNR. Thus, GLS is less optimum compared to OLS since it favors no specific frequency bands while OLS de-noises the low-frequency signals (major fluctuations) better. Finally, GLS relies on a robust estimator of the residual correlation model to avoid biased estimates.

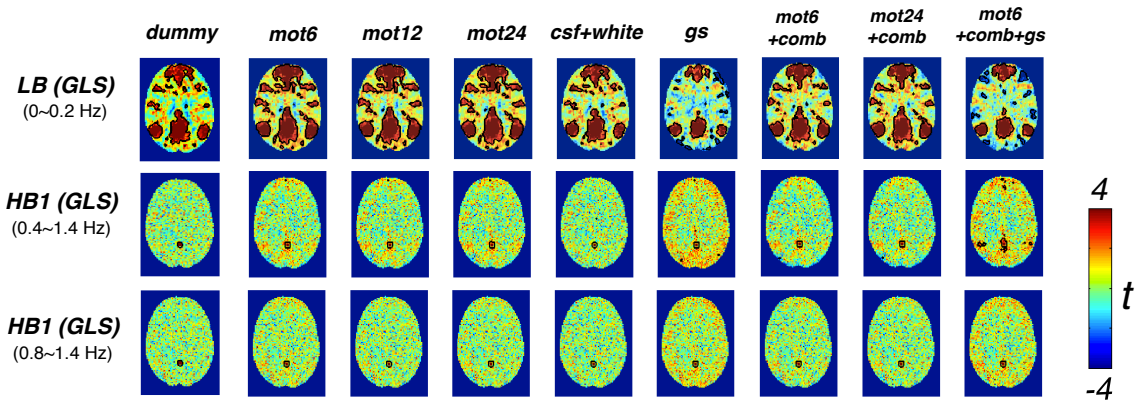

Figure S3 The group-level t-map of DMN within different frequency bands post generalized least square fitting based WB-LNR (the *'cleaned dummy'* dataset, across 10 subjects). Nuisance regression was performed using AFNI's 3DREMLfit, which adopts the AR1MA1 correlation model for each voxel's time series.

References

Aitken, A.C., 1934. On Least-squares and Linear Combinations of Observations. pp. 42-48.

Bullmore, E., Brammer, M., Williams, S., Rabehesketh, S., Janot, N., David, A., Mellers, J., Howard, R., Sham, P., 1996. Statistical methods of estimation and inference for functional MR image analysis. Magnetic Resonance in Medicine 35, 261-277.

Locascio, J., Jennings, P., Moore, C., Corkin, S., 1997. Time series analysis in the time domain and resampling methods for studies of functional magnetic resonance brain imaging. Human Brain Mapping 5, 168-193.

Purdon, P., Weisskoff, R., 1998. Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-level false-positive rates in fMRI. Human Brain Mapping 6, 239-249.

Woolrich, M., Ripley, B., Brady, M., Smith, S., 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage 14, 1370-1386.