Supplementary Information for Hallquist et al. RS-fcMRI manuscript

The Nuisance of Nuisance Regression: Spectral Misspecification in a Common Approach to Resting-State fMRI Preprocessing Reintroduces Noise and Obscures Functional Connectivity

S1. Detailed description of simultaneous bandpass filtering and nuisance regression (Simult)

The notion of spectral filtering using sinusoidal regressors extends from harmonic analysis, which seeks to estimate the strength of periodic components in the data when the periodicities are unknown (Bloomfield, 2000). Any time series can be decomposed into sinusoidal frequency components at the Fourier frequencies, f_j:

$$x_t = a_0 + \sum_{j=1}^{(n-1)/2} a_j \cos(2\pi f_j t) + b_j \sin(2\pi f_j t)$$

where $t = \Delta_t * (1, 2, \ldots n)$; $f_j = \frac{j}{n\Delta_t}$; n is the (odd) number of observations; $j = 1, 2, \ldots n/2$; and Δ_t is the sampling period (in fMRI, the repetition time [TR]) of the series (Shumway and Stoffer, 2010). The intercept of the regression equation is a_0, and the strength of the signal fluctuations at each Fourier frequency, f_j, are captured by the coefficients a_j and b_j, which correspond to the sine and cosine functions, respectively, at that frequency. Thus, a time series of length n may be perfectly reconstructed by n total sine and cosine functions, which form an orthogonal basis set that provides a unique mapping between the frequency and time domain representations of the series.

Using multiple regression, a bandpass filter can be applied by constructing a pair of sine and cosine regressors at each of the nuisance frequencies, f_j, regressing the input signal against these components, and retaining the residuals as the filtered series. Further, the bandpass sinusoids can be regressed alongside nuisance signals straightforwardly:

$$x_t = a_0 + \sum_{j=1}^{r} \left[a_j \cos\left(2\pi f_j t\right) + b_j \sin\left(2\pi f_j t\right) \right] + \sum_{k=1}^{p} [c_k z_{tk}] + \varepsilon_t$$

where there are r nuisance frequencies, $F = \{f_1, f_2, \dots, f_r\}$, and p nuisance signals. The coefficients a_j and b_j capture the strength of signal fluctuation at each nuisance frequency, f_j, whereas the coefficients c_k represent the association between each nuisance signal, z_k, and the outcome, x. This approach defines a single linear transformation of a time series that removes both nuisance signals and frequencies, avoiding the problem of which filtering step to perform first.

Moreover, in multiple regression the contribution of each predictor is net of the effects of all other predictors (i.e., the effect of one predictor, controlling for all others in the model; Cohen et al., 2002). Thus, by including bandpass filter sinusoids among the predictors, a given nuisance regressor (e.g., the WM signal), can only uniquely predict variance at frequencies not represented by the bandpass sinusoids in the regression. Consequently, all of the nuisance regressors are effectively bandpass filtered — that is, they only help to remove nuisance variance in the fMRI signal within the passband (e.g., .009 – .08 Hz). A practical ramification of this fact is that the same estimates for nuisance regressors will be obtained when the fMRI data and nuisance regressors are bandpass filtered prior to computing the regression. Said differently, whether using a truly simultaneous approach (i.e., one that includes both nuisance frequencies and nuisance regressors among the signals to be filtered at once) or whether one bandpass filters all regressors and fMRI data prior to nuisance regression, the effect is that the bandpass filter removes all fluctuations in the fMRI signal in the *stopband* (e.g., $f < .009$ Hz and $f > .08$ Hz) and the regression removes any nuisance regressor-related variability within the *passband* (e.g., $.009 \text{ Hz} < f < .08$ Hz).

Although the simultaneous filtering approach using multiple regression is conceptually elegant because it explicates how nuisance frequencies and noise sources can be treated identically in the same statistical model, creating a set of cosine and sine time series at the frequencies to be suppressed and incorporating these into a regression for each region of interest or voxel may be both laborious and computationally expensive. By contrast, modern spectral filtering methods based on the fast Fourier transform (FFT) are considerably faster (Bloomfield, 2000) and can be applied to voxelwise fMRI data reasonably quickly using existing software. Thus, for computational tractability, we suggest that simultaneous bandpass filtering and nuisance regression should be accomplished by bandpass filtering the regressors and fMRI time series, then conducting the nuisance regression on the filtered signals. This approach will yield identical results.

S2. Results for parcellating brain regions into functional brain networks using BpReg and Simult

In order to characterize differences between BpReg and Simult in the parcellation of brain regions into functionally related networks, we applied a community detection algorithm (Blondel et al., 2008) to the averaged and thresholded adjacency matrix for each approach, where connection weights represented the mean across subjects' 244 x 244 adjacency matrices. Average adjacency matrices were then thresholded at 15% edge density, which eliminated weak connections and negative weights. Because the threshold chosen for defining functional connections can influence the definition of networks using community detection methods, we also explored a weight-conserving algorithm that uses the connectivity strength to define functional networks (Rubinov and Sporns, 2011), rather than defining connections as present or absent at some connectivity threshold. To measure the similarity of community assignments between BpReg and Simult, normalized mutual information (NMI; see Newman, 2010) was calculated — NMI of 1.0 indicates that network compositions are identical between approaches, whereas NMI of 0.0 indicates no overlap in community assignments.

For the UPitt data, the algorithms classified 244 ROIs into six functional networks for both BpReg and Simult. The composition of these networks was consistent with previous literature (e.g., Fair et al., 2007), and we assigned the following labels: task control/default mode network, frontal-parietal attention control network, cingulo-opercular network, visual network, sensory-motor network, and cerebellar network (Figure S1).

For the UPitt data, community assignments derived from the 15% threshold algorithm largely overlapped between BpReg and Simult according to the normalized mutual information criterion, *NMI* = .87. Notably, however, we identified important differences between approaches in the composition of the task control/default mode, frontal-parietal attention control, and cingulo-opercular networks (highlighted by black circles in Figure S1). Using the weight-conserving modularity optimization algorithm, we found that NMI = .82 for BpReg versus Simult, corroborating a moderate to high level of correspondence between approaches.

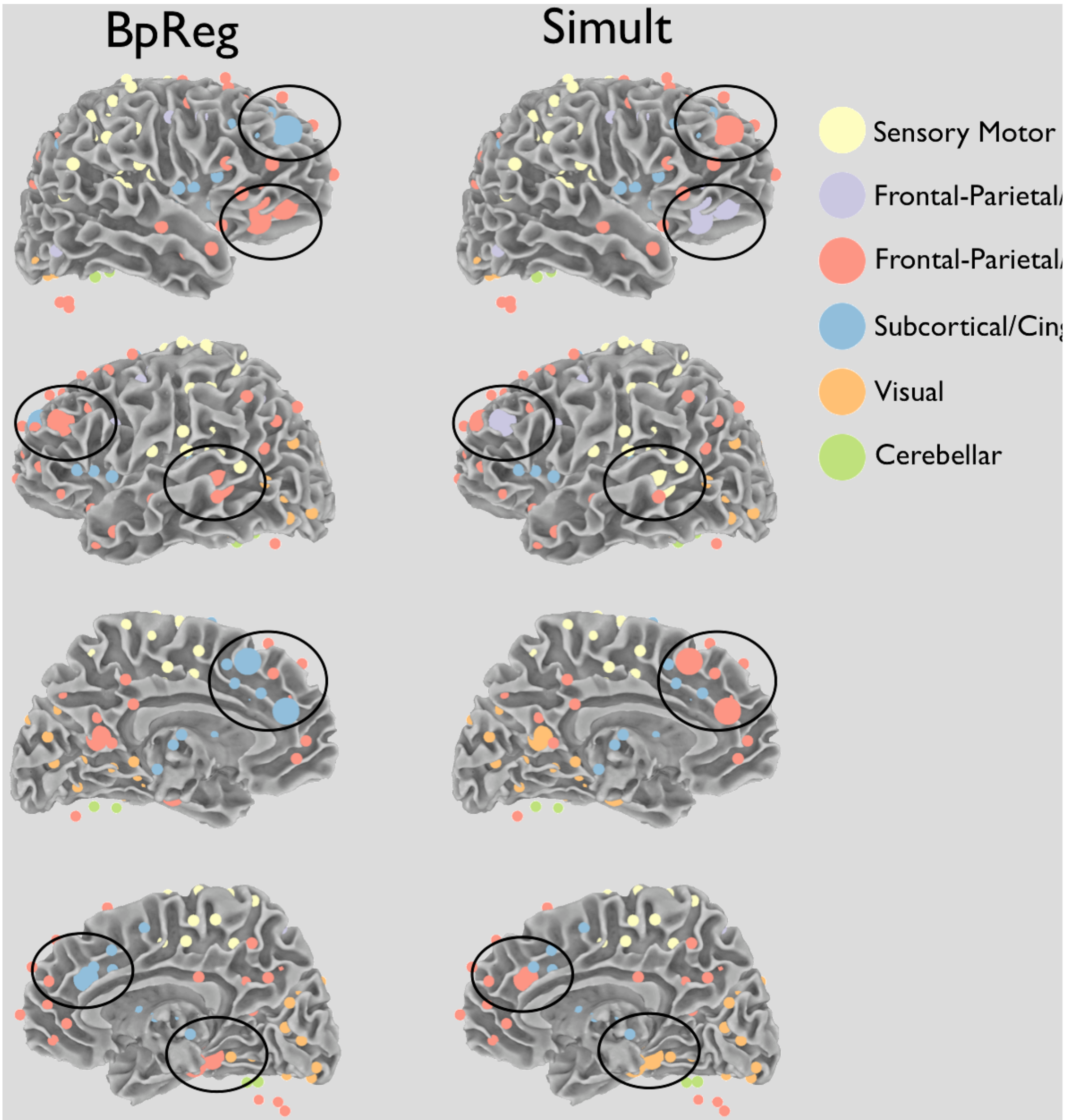
For the WashU data, using the 15% threshold approach, correspondence between BpReg and Simult network definitions was NMI = 0.66. For the weighted modularity method, we found a similarity of NMI = 0.77 between BpReg and Simult.

S3. Comparing the effects of motion scrubbing, nuisance regression, and bandpass filtering on functional connectivity estimates.

Figure 7 suggests the possibility that motion scrubbing and the nuisance regression and bandpass filtering approach may have relatively independent effects on functional connectivity estimates. Although we demonstrated that the Simult approach was significantly better at controlling the motion-connectivity association in both datasets, we did not address specifically whether changes in connectivity estimates due to motion scrubbing were related to changes attributable to the Simult approach. To quantify the whether scrubbing-related connectivity changes were associated with processing approach, for each subject, we computed connectivity changes due to 1) the difference between BpReg and Simult for unscrubbed data; 2) the difference between BpReg and Simult for scrubbed data; 3) scrubbing the Simult data; and 4) scrubbing the BpReg data. For each pairwise connection, we averaged scrubbing or approach-related differences across subjects from the above conditions to form four 264 x 264 group matrices where each cell represented the average connectivity change due to either scrubbing or use of Simult over BpReg. Finally, we correlated average connectivity changes across pairwise connections for the four matrices above to quantify the relationships among scrubbing and preprocessing approach effects. As detailed in Table S2, for both the UPitt and WashU cohorts, connectivity differences between Simult and BpReg were essentially comparable for scrubbed and unscrubbed data, suggesting that scrubbing had relatively little effect on preprocessing-related changes in connectivity. Conversely, the effects of motion scrubbing on connectivity estimates varied somewhat between BpReg and Simult (even though the same volumes were censored from the time series), as indicated by only moderate correlations between scrubbing-related connectivity changes for BpReg and Simult data. Notably, however, we found modest associations between approach-related and scrubbing-related connectivity changes, suggesting that these two procedures are largely independent. There was a small negative correlation between approach-related connectivity changes (BpReg – Simult) for the unscrubbed data and scrubbing-related effects for BpReg, which suggests that for individuals with greater motion-related inflation of connectivity estimates using the BpReg approach, scrubbing produces more pronounced reductions in functional connectivity estimates.

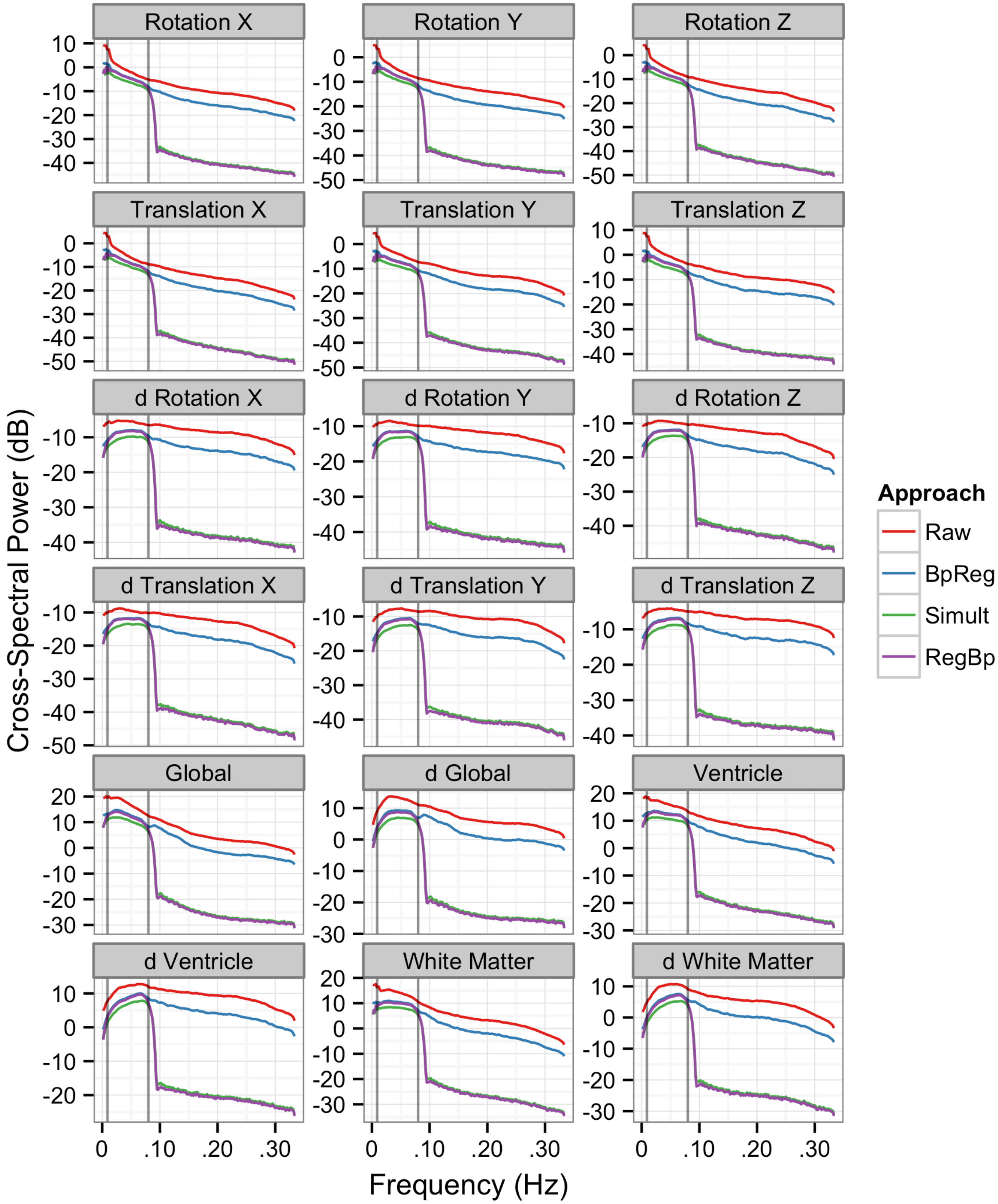
Our interpretation that scrubbing- and approach-related connectivity changes are relatively independent was bolstered by a comparison of Figures S4 and S5, which show that distance-dependent changes in connectivity estimates due to preprocessing approach are not qualitatively different after scrubbing, nor are distance-dependent changes due to scrubbing qualitatively different between BpReg and Simult (consistent with Figure 7). That said, our analyses indicate that the use of the Simult approach produced more pronounced reductions of the motion-connectivity association than scrubbing, and this may be related to the greater magnitude of connectivity changes related to use of the Simult approach (Figure S4) relative to scrubbing effects (Figure S5).

Figure S1. The parcellation of brain regions into functional networks differs between BpReg and Simult approaches.



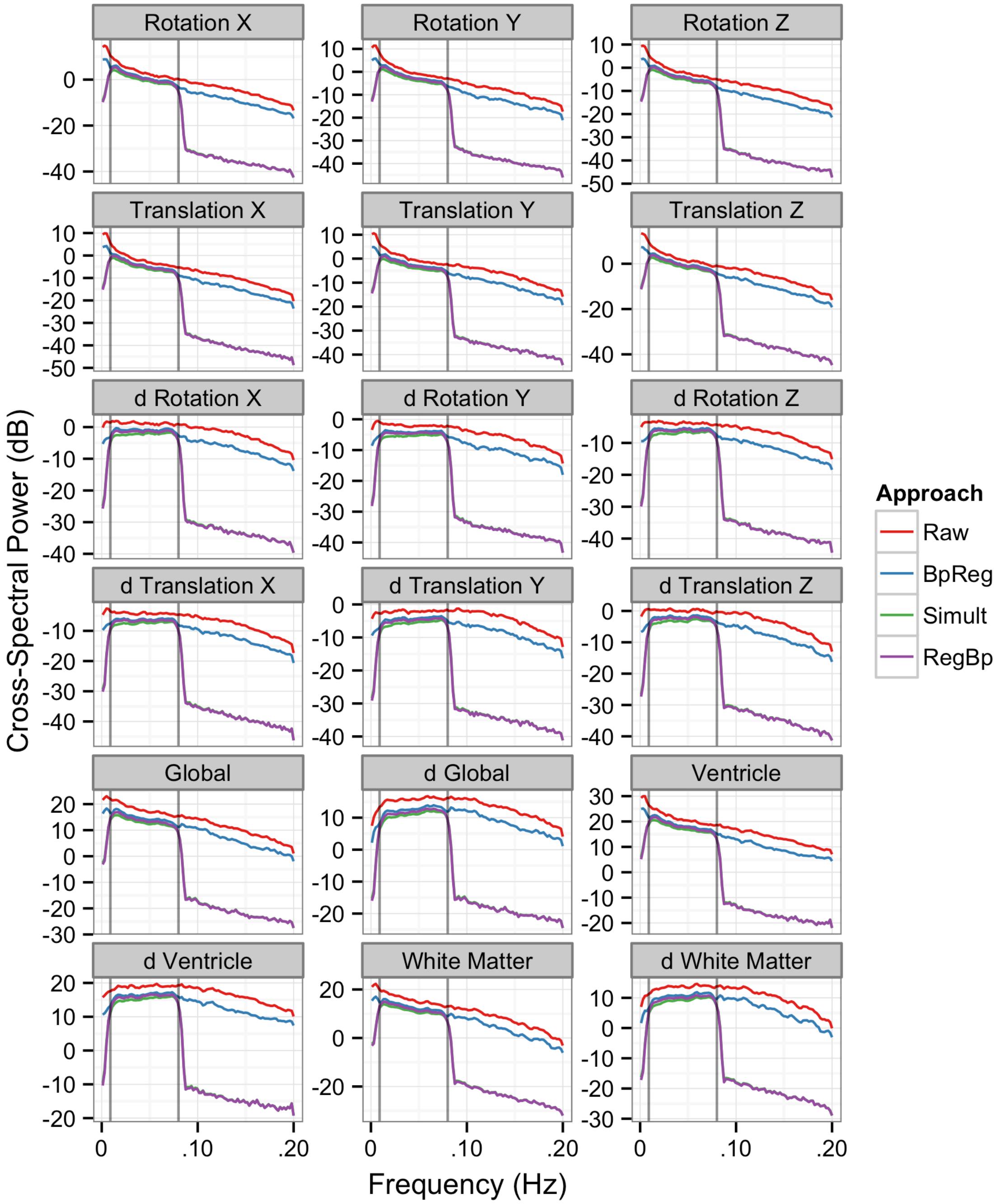
Note. Spheres are colored according to their community/network assignments from the modularity optimization algorithm. Black ovals highlight brain regions that differ in their community assignments between BpReg and Simult approaches. Results are displayed for the 15% threshold community detection algorithm applied to the UPitt data.

Figure S2. Mean cross-spectral power between RS-fcMRI time series and 18 nuisance regressors for the UPitt cohort.



Note. Each panel depicts the cross-spectral power between a nuisance regressor and RS-fMRI time series under different processing pipelines. *Raw* = fMRI time series prior to bandpass filtering and nuisance regression; *BpReg* = bandpass filtering, then nuisance regression; *Simult* = simultaneous bandpass filtering and nuisance regression*; RegBp* = nuisance regression, then bandpass filtering.

Figure S3. Mean cross-spectral power between RS-fcMRI time series and 18 nuisance regressors for the WashU cohort.



Note. Each panel depicts the cross-spectral power between a nuisance regressor and RS-fMRI time series under different processing pipelines. *Raw* = fMRI time series prior to bandpass filtering and nuisance regression; *BpReg* = bandpass filtering, then nuisance regression; *Simult* = simultaneous bandpass filtering and nuisance regression*; RegBp* = nuisance regression, then bandpass filtering.

Figure S4. Mean differences in functional connectivity estimates between the BpReg and Simult approaches as a function of interregional distance and motion scrubbing.

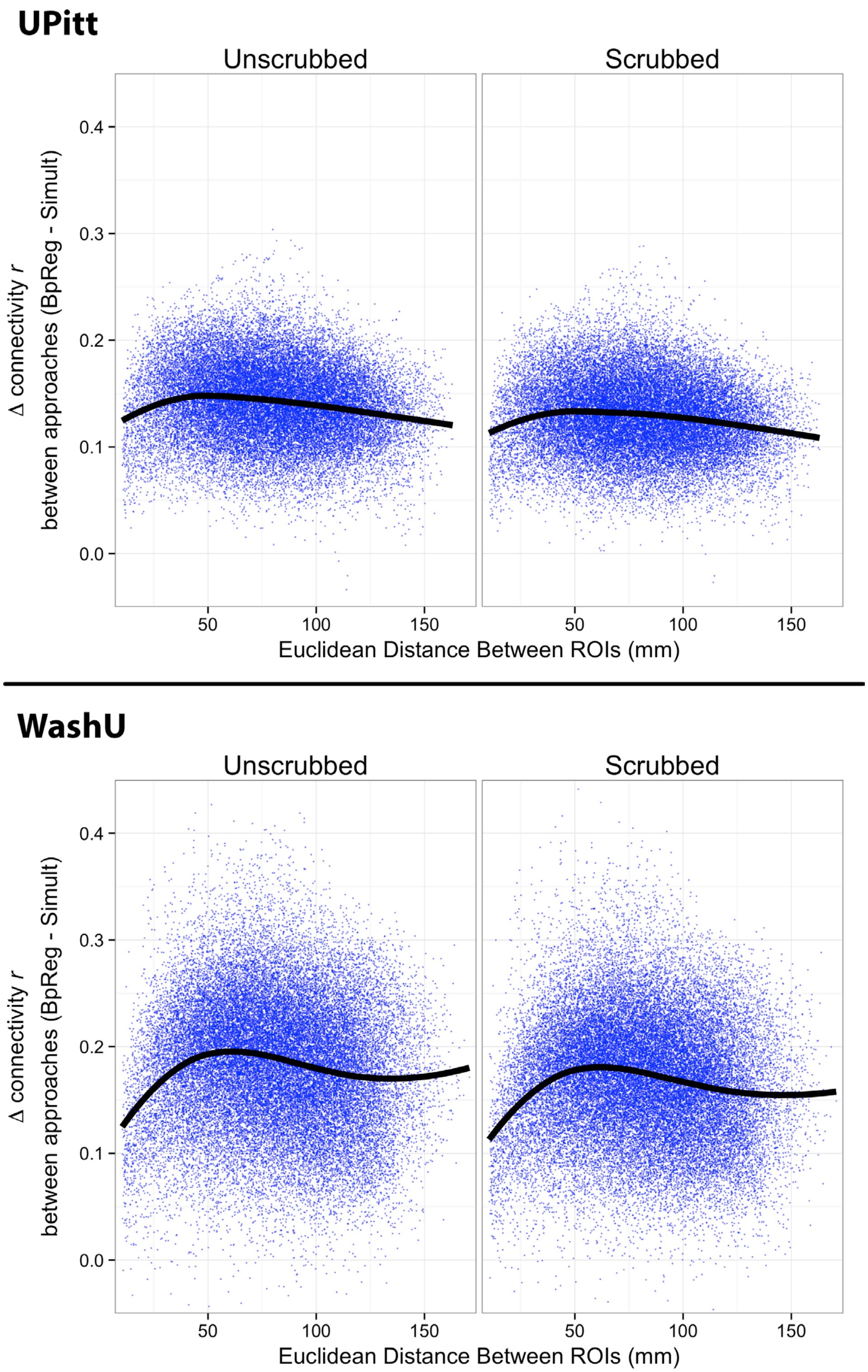


Figure S5. Mean differences in functional connectivity estimates due to motion scrubbing as a function of interregional distance and nuisance regression and bandpass filtering approcach.

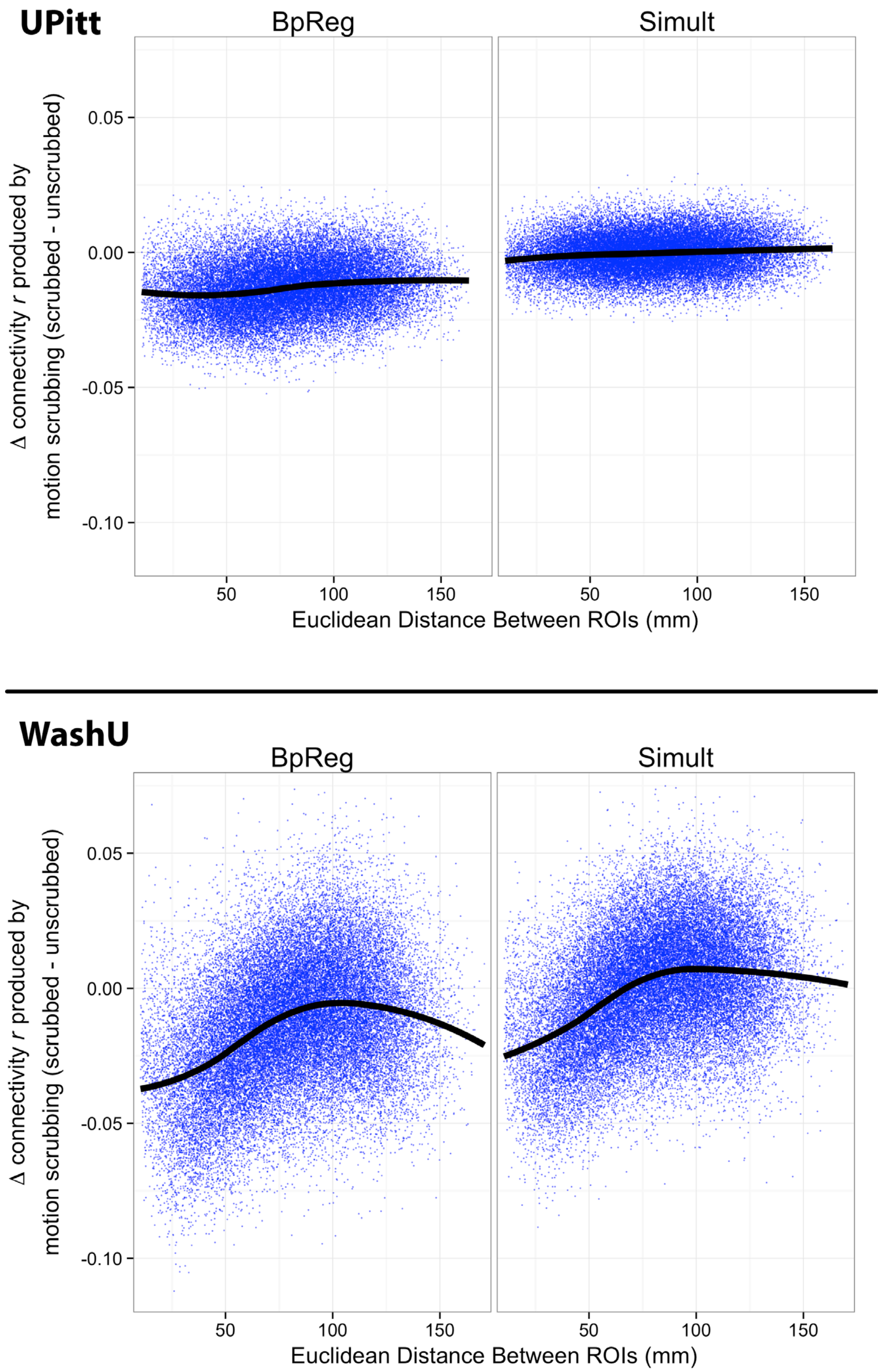


Table S1. *MNI coordinates of the 264 brain regions used throughout this paper, based on Power et al. (2012a) and Power et al. (2011).*

x	y	z	ROI
24	32	-18	3
-21	-22	-20	6
17	-28	-17	7
34	38	-12	12
-7	-52	61	13
-14	-18	40	14
0	-15	47	15
10	-2	45	16
-7	-21	65	17
-7	-33	72	18
13	-33	75	19
-54	-23	43	20
29	-17	71	21
10	-46	73	22
-23	-30	72	23
-40	-19	54	24
29	-39	59	25
50	-20	42	26
-38	-27	69	27
20	-29	60	28
44	-8	57	29
-29	-43	61	30
10	-17	74	31
22	-42	69	32
-45	-32	47	33
-21	-31	61	34
-13	-17	75	35
42	-20	55	36
-38	-15	69	37
-16	-46	73	38
2	-28	60	39
3	-17	58	40
38	-17	45	41
-49	-11	35	42
36	-9	14	43
51	-6	32	44
-53	-10	24	45
-3	2	53	47
54	-28	34	48

19	-8	64	49
-16	-5	71	50
-10	-2	42	51
37	1	-4	52
13	-1	70	53
7	8	51	54
-45	0	9	55
49	8	-1	56
-34	3	4	57
-51	8	-2	58
-5	18	34	59
36	10	1	60
32	-26	13	61
65	-33	20	62
58	-16	7	63
-38	-33	17	64
-60	-25	14	65
-49	-26	5	66
43	-23	20	67
-50	-34	26	68
-53	-22	23	69
-55	-9	12	70
56	-5	13	71
59	-17	29	72
-30	-27	12	73
-41	-75	26	74
6	67	-4	75
8	48	-15	76
-13	-40	1	77
-18	63	-9	78
-46	-61	21	79
43	-72	28	80
-44	12	-34	81
46	16	-30	82
27	16	-17	85
-44	-65	35	86
-39	-75	44	87
-7	-55	27	88
6	-59	35	89
-11	-56	16	90
-3	-49	13	91
8	-48	31	92
15	-63	26	93

-2	-37	44	94
11	-54	17	95
52	-59	36	96
23	33	48	97
-10	39	52	98
-16	29	53	99
-35	20	51	100
22	39	39	101
13	55	38	102
-10	55	39	103
-20	45	39	104
6	54	16	105
6	64	22	106
-7	51	-1	107
9	54	3	108
-3	44	-9	109
8	42	-5	110
-11	45	8	111
-2	38	36	112
-3	42	16	113
-20	64	19	114
-8	48	23	115
-56	-13	-10	117
-58	-30	-4	118
-68	-41	-5	120
13	30	59	121
12	36	20	122
52	-2	-16	123
-26	-40	-8	124
27	-37	-13	125
-34	-38	-16	126
28	-77	-32	127
52	7	-30	128
-53	3	-27	129
47	-50	29	130
-49	-42	1	131
-31	19	-19	132
-2	-35	31	133
-7	-71	42	134
11	-66	42	135
4	-48	51	136
-46	31	-13	137
-10	11	67	138

49	35	-12	139
8	-91	-7	140
17	-91	-14	141
18	-47	-10	143
40	-72	14	144
8	-72	11	145
-8	-81	7	146
-28	-79	19	147
20	-66	2	148
-24	-91	19	149
27	-59	-9	150
-15	-72	-8	151
-18	-68	5	152
43	-78	-12	153
-47	-76	-10	154
-14	-91	31	155
15	-87	37	156
29	-77	25	157
20	-86	-2	158
15	-77	31	159
-16	-52	-1	160
42	-66	-8	161
24	-87	24	162
6	-72	24	163
-42	-74	0	164
26	-79	-16	165
-16	-77	34	166
-3	-81	21	167
-40	-88	-6	168
37	-84	13	169
6	-81	6	170
-26	-90	3	171
-33	-79	-13	172
37	-81	1	173
-44	2	46	174
48	25	27	175
-47	11	23	176
-53	-49	43	177
-23	11	64	178
58	-53	-14	179
24	45	-15	180
-18	-76	-24	183
17	-80	-34	184

35	-67	-34	185
47	10	33	186
-41	6	33	187
-42	38	21	188
38	43	15	189
49	-42	45	190
-28	-58	48	191
44	-53	47	192
32	14	56	193
37	-65	40	194
-42	-55	45	195
40	18	40	196
-34	55	4	197
-42	45	-2	198
33	-53	44	199
43	49	-2	200
-42	25	30	201
-3	26	44	202
11	-39	50	203
55	-45	37	204
42	0	47	205
31	33	26	206
48	22	10	207
-35	20	0	208
36	22	3	209
37	32	-2	210
34	16	-8	211
-11	26	25	212
-1	15	44	213
-28	52	21	214
0	30	27	215
5	23	37	216
10	22	27	217
31	56	14	218
26	50	27	219
-39	51	17	220
2	-24	30	221
6	-24	0	222
-2	-13	12	223
-10	-18	7	224
12	-17	8	225
-5	-28	-4	226
-22	7	-5	227

-15	4	8	228
31	-14	2	229
23	10	1	230
29	1	4	231
-31	-11	0	232
15	5	7	233
9	-4	6	234
54	-43	22	235
-56	-50	10	236
-55	-40	14	237
52	-33	8	238
51	-29	-4	239
56	-46	11	240
53	33	1	241
-49	25	-1	242
-16	-65	-20	243
-32	-55	-25	244
22	-58	-23	245
1	-62	-18	246
10	-62	61	251
-52	-63	5	252
-47	-51	-21	253
46	-47	-17	254
47	-30	49	255
22	-65	48	256
46	-59	4	257
25	-58	60	258
-33	-46	47	259
-27	-71	37	260
-32	-1	54	261
-42	-60	-9	262
-17	-59	64	263
29	-5	54	264

Note. ROI numbers refer to the corresponding regions from Power and colleagues' (2011; 2012a) papers. Twenty of the 264 ROIs from these papers were not acquired by the UPitt EPI sequence and were omitted from UPitt analyses.

Table S2. Correlations among functional connectivity changes due to motion scrubbing versus the use of the Simult or BpReg approach.

Cohort	Connectivity Difference	1	2	3	4
UPitt	1: Δ*r* BpReg – Simult (unscrubbed)	1.0			
	2: Δ*r* BpReg – Simult (scrubbed)	.97	1.0		
	3: Δ*r* Scrubbed – Unscrubbed (BpReg)	-.29	-.11	1.0	
	4: Δ*r* Scrubbed – Unscrubbed (Simult)	-.04	-.10	.50	1.0
WashU	1: Δ*r* BpReg – Simult (unscrubbed)	1.0			
	2: Δ*r* BpReg – Simult (scrubbed)	.96	1.0		
	3: Δ*r* Scrubbed – Unscrubbed (BpReg)	-.24	-.11	1.0	
	4: Δ*r* Scrubbed – Unscrubbed (Simult)	.01	-.06	.72	1.0

References

Bloomfield, P. (2000). Fourier analysis of time series: An introduction (Wiley-Interscience).

Cohen, J., Cohen, P., West, S.G., and Aiken, L.S. (2002). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (Routledge Academic).

Fair, D.A., Dosenbach, N.U.F., Church, J.A., Cohen, A.L., Brahmbhatt, S., Miezin, F.M., Barch, D.M., Raichle, M.E., Petersen, S.E., and Schlaggar, B.L. (2007). Development of distinct control networks through segregation and integration. Proc. Natl. Acad. Sci. U.S.A. *104*, 13507–13512.

Newman, M. (2010). Networks: An Introduction (Oxford University Press, USA).

Rubinov, M., and Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. Neuroimage *56*, 2068–2079.

Shumway, R.H., and Stoffer, D.S. (2010). Time Series Analysis and Its Applications: With R Examples (Springer).