#### **SUPPLEMENTAL MATERIAL**

#### **Network Density Across Pipelines**

The proportion of negative edges across pipelines is presented in Supplemental Figure 1. The proportion of negative edges out of the full 766,656 (264 x 264 nodes x 11) sliding windows) was computed for each subject for each pipeline. In line with previous observations (e.g., Murphy et al., 2009), pipelines making use of GSR (e.g., 9P, 36P) tended to have a greater proportion of negative edges relative to pipelines that did not use GSR (e.g., 2P).



*Figure S1*. Proportion negative edges across all pipelines.

Network density across pipelines is presented in Supplemental Figure 2. The proportion of non-zero, positive edges out of the full 766,656 (264 x 264 nodes x 11 sliding windows) was computed for each subject for each pipeline. As expected based on the number of negative edges across pipelines, pipelines making use of GSR (e.g.,9P) tended to have lower density networks relative to pipelines that did not use GSR (e.g., 2P).



*Figure S2*. Proportion non-zero edges across all pipelines.

The average positive edge weight of networks across pipelines is presented in Supplemental Figure 3. The average of all non-zero, positive edges was computed for each subject for each pipeline.



*Figure S3.* Mean positive edge weight across all pipelines.

#### **Results with Negative Edge Weights Included**

This section reports results parallel to those in the manuscript using functional connectivity estimates that retained negative edge weights. The dispersion measure used in the main manuscript was not appropriate for use with data including negative edge weights as a negative ratio of variance to the mean would be uninterpretable. As such, the dispersion measures used in the case of networks including negative edge weights was the standard deviation of the edge weights.

#### *Edge Dispersion-Motion Association is Minimal Across Pipelines*

Distributions of *dispersion-motion* correlations are presented in Figure S4. Pairedsample *t*-tests indicated significant differences in the mean *dispersion-motion* correlation across all pipelines (Table S1). The median absolute *dispersion-motion* correlations ranged between 0.04 and 0.10, indicating small associations between *dispersion* and participant motion following the application of the de-noising pipelines.

## **Table S1**

Paired samples *t*-test comparisons of the mean dispersion-motion correlations across all pipelines



*Notes*: Elements of matrix represent *t*-values. Degrees of freedom for all *t*-tests = 34715. Columns represent reference level (i.e., positive t-values indicate higher mean dispersion-motion correlations for the pipeline indicated by the column versus the row heading). All *p*-values were <0.001.



*Figure S4*. Distributions of all edgewise dispersion-motion correlations after de-noising for each de-noising pipeline. The median absolute value of the correlation and the standard deviation of the correlation is displayed in the top right of each panel.

While the range of *dispersion-motion* correlations was small across pipelines, heterogeneity in pipeline performance was most notable in the percent of edges whose dispersion was related to motion (Figure S5). The percent of edges whose dispersion was associated with motion was generally small (i.e.,  $\lt 1\%$ ). However, the ICA-AROMA pipeline emerged as a clear outlier with over 30% of edges having dispersion values that remained associated with subject motion following de-noising. The performance of ICA-AROMA in the case in which negative edge weights were retained, paralleled its performance in the case in which negative edge weights were not retained (Figure 4 in main manuscript).



*Figure S5.* Percent of edges significantly related to motion after de-noising for the dispersion index. More effective de-noising pipelines reduced the relationship between dispersion and motion. Bars are ordered such that the least effective de-noising pipelines are on the left and the most effective are on the right. Panel A illustrates the results after correcting for multiple comparisons. Panel B depicts the results without controlling for multiple comparisons.

#### *Dispersion Distance-Dependent Motion Artifact Varies Across Pipelines*

Correlations between the Euclidean distance separating the nodes and the magnitude of the dispersion-motion correlation for the edge connecting those nodes ranged between -0.09 and 0.02 (Figure S6). Tests of equality of correlations revealed significant differences in the *distance-dependent effects of motion* across many pipelines (Table S2). Particularly poorly-performing pipelines included the 6P, 24P, wmLOCAL, and tCOMPCOR pipelines. These pipelines were also poorly-performing pipelines in the context in which negative edge weights were not retained (Figure 5 in main manuscript).

## **Table S2**

Tests of equality of correlation coefficients representing the association between the Euclidean distance separating nodes (mm) and the magnitude of the *dispersion-motion* correlation between the edges connecting the nodes.



*Notes*:  $n = 34716$ . Elements of matrix represent *z*-values. Columns represent reference level (i.e., positive t-values indicate higher distance-motion correlations for the pipeline indicated by the column versus the row heading). Shaded rows indicate non-significant differences between pipelines.



*Figure S6*. Hexbin plots of the association between the Euclidean distance separating nodes (mm) and the magnitude of the *dispersion-motion* correlation between the edge connecting the nodes (y-axis). A trend line for each pipeline is indicated in blue and the magnitude of the correlation is presented in the bottom right of each panel. *Note*: \*\*\**p*<0.001.

#### *Marked Heterogeneity in Sub-Network Identification Across Pipelines*

There were marked differences across de-noising pipelines in the extent to which sub-networks were identifiable, as operationalized by the *network modularity quality (Q)*. Paired sample *t*-tests revealed significant differences across the *Q* values of most pipelines (Table S3). Neither the difference between the *Q* values from the 2P and the 9P pipelines nor between the tCOMPCOR and wmMEAN pipelines were significant. The 24P pipeline was the least effective at allowing the identification of sub-networks, exhibiting a mean *Q* value of 13.17 (see Figure S7). AROMA+GSR emerged as the most effective pipeline, with a mean *Q* value 19.88. The relatively poor performances of the 24P and 6P pipelines and the relatively good performance of the AROMA+GSR pipeline on sub-network identification parallels the performance of these pipelines in the context in which negative edge weights were not retained (Figure 6 in main manuscript).

### **Table S3.**

**Paired samples** *t***-test comparisons of the mean** *Q* **across all pipelines.**



*Notes*: Elements of matrix represent *t-*values. Degrees of freedom for all *t*-tests = 392. Columns represent reference level (i.e., positive t-values indicate higher mean *Q* values for the pipeline indicated by the column versus the row heading). Shaded rows indicate nonsignificant differences between pipelines.



*Figure S7*. Scatterplot of the association between subject motion (x-axis) and modularity quality (*Q*, y-axis) with trend line. The partial correlation between subject motion and *Q* controlling for age and sex is presented in the top right of each panel. The mean *Q* value for each pipeline is presented in the top left of each panel. \*\*\**p*<0.001; \*\**p*<0.01; \**p*<0.05.

#### *Q-Motion Associations Reduced by Effective Pipelines*

Correlations between *subject motion and Q* ranged between -0.42 and 0.18 (Figure S7). Pipelines that were the least effective at allowing the identification of subnetworks tended to be the pipelines that were the least effective at mitigating subject motion artifacts,  $r(10)=0.80$ ,  $p=0.002$ . For example, pipelines 6P and 24P performed poorly across both indices while AROMA+GSR and 36P+DESPIKE performed well. Results mirror those in the context in which negative edge weights were not retained (Figure 6 in main manuscript).

#### *Node Flexibility-Motion Correlations Generally Small Across Pipelines*

Distributions of *node flexibility-motion* correlations for each pipeline are presented in Figure S8. Paired sample *t*-tests comparing the mean node flexibility-motion correlations cross pipelines are presented in Table S4. Absolute median correlations were small in magnitude across pipelines and ranged between 0.04 and 0.08.

## **Table S4.**

Paired samples *t*-test comparisons of the mean *node flexibility-motion* correlations across all pipelines.



*Notes*: Elements of matrix represent *t-*values. Degrees of freedom for all *t*-tests = 263. Columns represent reference level (i.e., positive *t*-values indicate higher mean flexibility-motion correlations for the pipeline indicated by the column versus the row heading). Shaded rows indicate non-significant differences between pipelines.



*Figure S8*. Distributions of all *node flexibility-motion* correlations after de-noising for each pipeline. The median absolute value and the standard deviation of the correlation is displayed in the top right of each panel.

When examining the *percentage of nodes related to subject motion* following denoising (Figure S9), we found that there was greater variability across pipelines in the extent to which subject motion artifacts were reduced. wmLOCAL and wmMEAN emerged as the least successful pipelines, with over 10% of nodes displaying flexibility values that were significantly associated with motion after de-noising. No nodes displayed flexibility values that were significantly associated with subject motion for the aCOMPCOR pipeline. The locations of the nodes significantly associated with subject

motion are presented in Figure S10. The good performance of aCOMPCOR paralleled its performance in the context in which negative edge weights were not retained (Figure 9 in main manuscript). ICA-AROMA was the most poorly performing pipeline in the context in which negative edge weights were not retained, while in this context it performed better than the wmMEAN and wmLOCAL pipelines.



*Figure S9*. Percent of nodes significantly related to motion after de-noising for node flexibility. More effective de-noising pipelines reduced the relationship between flexibility and motion. Bars are ordered such that the least effective de-noising pipelines are on the left and the most effective are on the right. Panel A illustrates the results after correcting for multiple comparisons. Panel B depicts the results without controlling for multiple comparisons.



*Figure S10.* The proportion of nodes within each subnetwork of the Power *et al.* (2011) parcellation with flexibility values significantly related to subject motion, without controlling for multiple comparisons. AUD=auditory; CB=cerebellum; CON=cingulateopercular network; DAN=dorsal attention network; DMN=default mode network; FPN=frontoparietal network; MEM=memory network; SN=salience; SUB=subcortical; VAN=ventral attention network; VIS=visual network.

When examining the association between *global flexibility and subject motion* (Fig. S11), no significant associations between subject movement and global flexibility emerged for 3 (aCOMPCOR, ICA-AROMA, AROMA+GSR) of the 12 pipelines. wmLOCAL emerged as the least effective pipeline for mitigating the association between subject motion and global flexibility. The good performance of aCOMPCOR was consistent with its performance in the context in which negative edge weights were not retained (Figure 11 in main manuscript).



*Figure S11*. Scatterplot of the association between subject motion (x-axis) and global flexibility (y-axis) with trend line. The partial correlation between subject motion and global flexibility controlling for age and sex is presented in the bottom right of each panel. \*\*\**p*<0.001; \*\**p*<0.01; \**p*<0.05.

#### *Node Promiscuity-Motion Correlations Are Generally Small Across Pipelines*

Distributions of *node promiscuity-motion* correlations are presented in Figure S12. Paired sample *t*-tests comparing the mean node promiscuity-motion correlations cross pipelines are presented in Table S5. Absolute median correlations were generally small. The values across pipelines ranged between 0.03 and 0.06.

### MOTION MITIGATION IN DYNAMIC CONNECTIVITY S21

### **Table S5.**

Paired samples *t*-test comparisons of the mean *node promiscuity-motion* correlations across all pipelines.



*Notes*: degrees of freedom for all *t*-tests = 263. Columns represent reference level (i.e., positive *t*-values indicate higher mean promiscuity-motion correlations for the pipeline indicated by the column versus the row heading). Shaded rows indicate nonsignificant differences between pipelines.



*Figure S12*. Distributions of all *node promiscuity-motion* correlations after de-noising for each pipeline. The median absolute correlation is displayed in the top right of each panel.

When examining the percentage of nodes significantly associated with subject motion following de-noising, we found that there was marked variability in the effectiveness of reducing motion artifact across pipelines (Figure S13). The ICA-AROMA pipeline had the highest number of nodes significantly associated with motion, although with only 3% of nodes associated with motion after de-noising this performance was still good. The performance of the ICA-AROMA pipeline paralleled its performance in the context in which negative edge weights were not retained (Figure 13 in main manuscript). All other pipelines performed well, with many showing no nodes associated

with motion. There were many more pipelines in the current case with no significant associations with motion than in the case in which negative edge weights were not retained. The locations of the nodes significantly associated with subject motion are presented in Figure S14.



*Figure S13*. Percent of nodes significantly related to motion after de-noising for node promiscuity. More effective de-noising pipelines reduced the relationship between promiscuity and motion. Bars are ordered such that the least effective de-noising pipelines are on the left and the most effective are on the right. Panel A illustrates the results after correcting for multiple comparisons. Panel B depicts the results without controlling for multiple comparisons.



*Figure S14.* The proportion of nodes within each subnetwork of the Power *et al.* (2011) parcellation with promiscuity values significantly related to subject motion, without controlling for multiple comparisons) AUD=auditory; CB=cerebellum; CON=cingulateopercular network; DAN=dorsal attention network; DMN=default mode network; FPN=frontoparietal network; MEM=memory network; SN=salience; SUB=subcortical; VAN=ventral attention network; VIS=visual network.

When examining the association between *global promiscuity and subject motion*

(Figure S15), no significant associations with motion emerged.



*Figure S15.* Scatterplot of the association between subject motion (x-axis) and global promiscuity (y-axis) with trend line. The partial correlation between subject motion and global promiscuity controlling for age and sex is presented in the bottom right of each panel.

### MOTION MITIGATION IN DYNAMIC CONNECTIVITY S26

# **Table S6.**

Paired samples *t*-tests comparisons of the mean dispersion-motion correlations across all pipelines.



*Notes*: degrees of freedom for all *t*-tests = 34715. Columns represent reference level (i.e., positive *t*-values indicate higher mean dispersion-motion correlations for the pipeline indicated by the column versus the row heading). All *p*-values were  $\leq 0.001$ .

## **Table S7**

Tests of equality of correlation coefficients representing the association between the Euclidean distance separating nodes (mm) and the magnitude of the *dispersion-motion* correlation between the edges connecting the nodes.



*Notes*: *n* = 34716. Shaded rows indicate non-significant differences between pipelines. Columns represent reference level (i.e., positive *z*-values indicate higher mean distance-motion correlations for the pipeline indicated by the column versus the row heading).

# **Table S8.**

Paired samples *t*-test comparisons of the mean *Q* across all pipelines .



*Notes*: degrees of freedom for all *t*-tests = 392. Columns represent reference level (i.e., positive *t*-values indicate higher mean *Q* values for the pipeline indicated by the column versus the row heading). Shaded rows indicate non-significant differences between pipelines.

### **Table S9**

Paired samples *t*-test comparisons of the mean *node flexibility-motion* correlations across all pipelines.



*Notes*: Degrees of freedom for all *t*-tests = 263. Columns represent reference level (i.e., positive *t*-values indicate higher mean flexibility-motion correlations for the pipeline indicated by the column versus the row heading). Shaded rows indicate non-significant differences between pipelines.

### MOTION MITIGATION IN DYNAMIC CONNECTIVITY S30

### **Table S10**

Paired samples *t*-test comparisons of the mean *node promiscuity-motion* correlations across all pipelines.



*Notes*: degrees of freedom for all *t*-tests = 263. Columns represent reference level (i.e., positive *t*-values indicate higher mean promiscuity-motion correlations for the pipeline indicated by the column versus the row heading). Shaded rows indicate nonsignificant differences between pipelines.

#### **Mean** *Q* **Across Pipelines Controlling for Graph Density**

To examine whether differences in the average *Q* value across pipelines was driven by differences in network density across pipelines, we used a multilevel model where subjects' Q values across the 12 pipelines were the dependent variables and both pipeline (dummy coded with pipelines with no GSR as the reference category), network density, age, and sex as independent variables. Results (Table S11) revealed that pipelines with GSR had significantly higher *Q* values relative to pipelines without GSR, suggesting that differences in mean *Q* values were not simply a product of differences in network density across pipelines.



# **Table S11**

*Notes*: GSR Pipelines was dummy coded such that the reference category included pipelines without global signal regression. \*\*\**p*<0.001, \**p*<0.05. AIC = Akaike information criterion;  $BIC =$  Bayesian information criterion.

## **References**

Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., & Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anticorrelated networks introduced? *NeuroImage*, *44*(3), 893-905.