1 **Title**

2 Low-Dimensional Spatio-Temporal Dynamics Underlie Cortex-Wide Neural Activity

3 **Authors**

4 Camden J. MacDowell[1,2,3] and Timothy J. Buschman[1,4,*]

5 1. Princeton Neuroscience Institute, Princeton University, Princeton, NJ, 08540, USA

6 2. Department of Molecular Biology, Princeton University, Princeton, NJ, 08540, USA

7 3. Rutgers Robert Wood Johnson Medical School, Piscataway, NJ, 08854, USA

8 4. Department of Psychology, Princeton University, Princeton, NJ, 08540, USA

9

10 * lead contact and corresponding author

11

12 **Abstract**

13 Cognition arises from the dynamic flow of neural activity through the brain. To capture these

14 dynamics, we used mesoscale calcium imaging to record neural activity across the dorsal cortex

15 of awake mice. We found that the large majority of variance in cortex-wide activity (~75%) could

16 be explained by a limited set of ~14 'motifs' of neural activity. Each motif captured a unique spatio-

17 temporal pattern of neural activity across the cortex. These motifs generalized across animals

18 and were seen in multiple behavioral environments. Motif expression differed across behavioral

19 states and specific motifs were engaged by sensory processing, suggesting the motifs reflect core

20 cortical computations. Together, our results show that cortex-wide neural activity is highly

21 dynamic, but that these dynamics are restricted to a low-dimensional set of motifs, potentially to

22 allow for efficient control of behavior.

23 **Introduction**

24 The brain is a complex, interconnected network of neurons. Neural activity flows through this

25 network, carrying and transforming information to support behavior. Previous work has associated

26 particular computations with specific spatio-temporal patterns of neural activity across the brain

27 (Buschman and Miller, 2007; Stringer et al., 2019a; Hutchison et al., 2013). For example,

28 sequential activation of primary sensory and then higher-order cortical regions underlies

29 perceptual decision making in both mice (Guo et al., 2014) and monkeys (Romo and de Lafuente,

30 2013; Siegel et al., 2011). Similarly, specific spatio-temporal patterns of cortical regions are

31 engaged during goal-directed behaviors (Allen et al., 2017), motor planning (Chen et al., 2017),

32 evidence accumulation (Pinto et al., 2019), motor learning (Makino et al., 2017), and sensory

33 processing (Mohajerani et al., 2013). Previous work has begun to codify these dynamics, either

34 in the synchronous activation of brain regions (Fries, 2015; Hutchison et al., 2013) or in the

35 propagation of waves of neural activity within and across cortical regions (Muller et al., 2018;

36 Zanos et al., 2015). Together, this work suggests cortical activity is highly dynamic, evolving over

37 both time and space, and that these dynamics play a computational role in cognition (Buonomano

38 and Maass, 2009; Miller and Wilson, 2008).

39 However, despite this work, the nature of cortical dynamics is still not well understood. Previous

40 work has been restricted to specific regions and/or specific behavioral states and so, we do not

41 yet know how neural activity evolves across the entire cortex, whether dynamics are similar across

42 individuals, or how dynamics relate to behavior. This is due, in part, to the difficulty of quantifying

43 the spatio-temporal dynamics of neural activity across the brain.

44 To address this, we used mesoscale imaging to measure neural activity across the dorsal cortical

45 surface of the mouse brain (Silasi et al., 2016). Then, using a convolutional factorization

46 approach, we identified dynamic 'motifs' of cortex-wide neural activity. Each motif captured a

47 unique spatio-temporal pattern of neural activity as it evolved across the cortex. Importantly,

48   because motifs captured the dynamic flow of neural activity across regions, they explained cortex-

49   wide neural activity better than 'functional connectivity' network measures.

50   Surprisingly, the motifs clustered into a limited set of ~14 different spatio-temporal 'basis' motifs

51   that were consistent across all animals. The basis motifs captured the majority of the variance in

52   neural activity in different behavioral states and in multiple sensory and social environments.

53   Specific motifs were selectively engaged by each environment and by sensory stimuli, suggesting

54   the motifs reflect core cortical computations, such as visual or tactile processing. Together, our

55   results suggest cortex-wide neural activity is highly dynamic but that these dynamics are low-

56   dimensional: they are constrained to a small set of possible spatio-temporal patterns.

57   **Results**

58   <u>Discovery of spatio-temporal motifs of cortical activity in awake, head-fixed mice</u>

59   We performed widefield 'mesoscale' calcium imaging of the dorsal cerebral cortex of awake,

60   head-fixed mice expressing the fluorescent calcium indicator GCaMP6f in cortical pyramidal

61   neurons (Fig. 1A; see Methods for details, Chen et al., 2013). A translucent-skull prep provided

62   optical access to dorsal cortex, allowing us to track the dynamic evolution of neural activity across

63   multiple brain regions, including visual, somatosensory, retrosplenial, parietal, and motor cortex

64   (Fig. 1A, inset and Fig. 1 supplement 1, Silasi et al., 2016). We initially characterized the dynamics

65   of 'spontaneous' neural activity when mice were not explicitly performing a specific behavior (Fig.

66   1B, N=48 sessions across 9 mice, 5-6 sessions per mouse, each session lasted 12 minutes,

67   yielding a total of 9.6 hours of imaging). These recordings revealed rich dynamics in the spatio-

68   temporal patterns of neural activity across the cortex (Supplemental Movie 1, as also seen by

69   Cramer et al., 2019; Shimaoka et al., 2019; Murphy et al., 2016; Mohajerani et al., 2013)

70   Our goal was to capture, quantify, and characterize the dynamic patterns of activity in an unbiased

71   manner. To do so, we used convolutional non-negative matrix factorization (CNMF, Mackevicius

72   et al., 2019) to discover repeated spatio-temporal patterns in the neural activity, in an

73   unsupervised way (Fig. 1C; see Methods and Fig. 1 supplement 2 for details, including robustness

74   to parameters). CNMF identified 'motifs' of neural activity; these are dynamic patterns of neural

75   activity that extend over space and time (Fig. 1C, bottom left, shows an example motif and the

76   corresponding original data). Once identified, the algorithm uses these motifs to reconstruct the

77   original dataset by temporally weighting them across the entire recording session (Fig. 1C;

78   transients in temporal weightings indicate motif expression, see Supplemental Movie 2 for

79   comparison of reconstruction to original data). Importantly, overlapping temporal weightings

80   between motifs are penalized, which biases factorization towards only one motif being active at a

81   given point in time. This allows us to capture the spatio-temporal dynamics of neural activity as a

82   whole, rather than decomposing activity into separate spatial and temporal parts (see Methods

83   for details).

84   Figure 1D shows three example motifs identified by CNMF from a single 2-minute recording

85   epoch. Many of the identified motifs show dynamic neural activity that involves the sequential

86   activation of multiple regions of cortex (top two rows in Fig. 1D). For example, example motif 1

87   starts in somatosensory/motor regions and, over the course of a few hundred milliseconds,

88   propagates posteriorly before ending in the parietal and visual cortices (Fig. 1D, top row). To aid

89   in visualizing these dynamics, the arrows overlaid on Figure 1D show the direction and magnitude

90   of activity propagation of the top 50% most active pixels between subsequent timepoints (as in

91   Afrashteh et al., 2017; see Methods for details). Other motifs were more spatially restricted,

92   engaging either one (or more) brain regions simultaneously (e.g. third row in Fig. 1D). In total, we

93   identified 2622 motifs across 144 different 2-minute epochs of imaging (3 independent epochs

94   from each of 48, 12-minute recording sessions; Fig. 1B, light blue 'discovery epochs').

95   Motifs captured the flow of activity across the brain during a brief time period (~1 second). By

96   tiling different motifs across time, the entire 2-minute recording epoch could be reconstructed. On

97   average, each 2-minute recording epoch could be reconstructed by combining ~19 motifs (Fig

98   2A. median; 95% Confidence Interval (CI): 18-20). This captured 89.05% of the total variance of

99   neural activity on average (Fig. 2B; 89.05% median explained variance, CI: 87.78-89.68%; N=144

100   discovery epochs; see Methods for details). To achieve this, individual motifs occurred repeatedly

101   during a recording session. Over half of the motifs occurred at least 3 times during a 2-minute

102   epoch, with motifs occurring 2.48 times per minute on average (Fig. 2C; CI: 2.38-2.59). The broad

103   distribution of the frequency of motifs suggests all motifs are required to explain cortex-wide

104   neural activity. Indeed, the cumulative percent explained variance (PEV) in neural activity

105   captured by individual motifs shows a relatively gradual incline (Fig. 2D, see Methods for details).

106   On average, no single motif captured more than 20% of the variance of the recording epoch, and

107   14 motifs were needed to capture over 90% of the relative PEV. Importantly, the number of

108   discovered motifs and their explained variance was robust to the changes in the regularization

109   hyperparameter of the CNMF algorithm, suggesting it is a true estimate of the number of motifs

110   needed and not a consequence of our analytical approach (Fig. 1 supplement 2, see Methods for

111   details).

112   <u>Motifs capture the dynamic flow of neural activity across the cortex</u>

113   Next, we tested whether motifs simply reflected the co-activation of brain regions or if they

114   captured the dynamic flow of neural activity between regions. Previous work has found neural

115   activity can be explained by the simultaneous activation of a coherent network of brain regions

116   (i.e. zero-lag, first-order correlations, as seen in functional connectivity analyses; Hutchison et al.,

117   2013). The CNMF approach used here is a generalization of such approaches; it can capture

118   spatio-temporal dynamics in the motifs, but it is not required to do so if dynamics are not

119   necessary to capture variance in neural activity. Therefore, to test whether neural activity is

120   dynamic, we tested whether dynamics were a necessary component of the motifs.

121 First, we determined whether motifs simply reflected the static engagement of a network of

122 regions. To this end, we measured the autocorrelation of neural activity during the timecourse of

123 each motif. Consistent with dynamic motifs, the correlation of activity patterns within a motif

124 quickly decayed with time (Fig. 3A; mean half-life τ across all motifs was 113ms +/- 2ms,

125 bootstrap; when fit to individual motifs, 25%-50%-75% of τ was 66ms - 117ms - 210ms; see

126 Methods for details). While activity patterns at adjacent motif timepoints (75ms apart) were

127 spatially correlated (Pearson's r=0.39 CI: 0.38-0.40, p<10^-16, Wilcoxon Signed-Rank Test versus

128 r=0, right-tailed; N=2622 Motifs), this similarity quickly declined when time points were farther

129 apart (Pearson's r=0.098 CI: 0.095-0.10 at 375ms and r=0.043, CI: 0.039-0.047 at 600ms; a

130 decrease of 0.29 and 0.35, both p<10^-16, Wilcoxon Signed-Rank Test). Similarly, the mean spatial

131 pattern of activity of a given motif, averaged across the timecourse of the motif, was dissimilar

132 from individual timepoints within the motif (Fig. 3B; median dissimilarity 0.58, CI: 0.57-0.58 across

133 motifs).

134 Second, we tested whether dynamics were necessary to fit neural data. To do this, we compared

135 the fit of CNMF-derived motifs to alternative decomposition approaches that do not consider

136 temporal dynamics. We used two 'static' decomposition techniques that are standards in the field:

137 spatial Principal Components Analysis (sPCA) and spatial Non-Negative Matrix Factorization

138 (sNMF; see Methods for details). Both approaches required >3 times more dimensions to capture

139 the same amount of variance as the motifs (Fig. 3E; on average, 64.5 and 93 dimensions for

140 sPCA and sNMF, respectively). If restricted to 19 dimensions, sPCA and sNMF explained

141 significantly less variance in neural activity than motifs (Fig. 3E; sPCA: 79.87% CI: 78.72-81.46%

142 a difference of 9.18%, p<10^-16; sNMF 77.95% CI: 76.63-79.42% a difference of 11.10%, p<10^-16,

143 Wilcoxon Signed-Rank Test). Although there are differences in the number of terms in each

144 dimension, the fact that dynamic motifs can capture significantly more variance than temporally

6

145   constrained approaches suggests brain activity has complex spatial-temporal dynamics that are

146   not captured by traditional decomposition methods but can be captured by our motifs.

147   <u>Motifs generalize to withheld data and across animals</u>

148   If the neural dynamics captured by CNMF reflect true, repeated, motifs of neural activity, then the

149   motifs identified in one recording session should generalize to other recording sessions. To test

150   if motifs generalized, we refit the motifs identified during a recording 'discovery' epoch to withheld

151   data (Fig. 1B, purple, N=144 'withheld epochs'). Motifs were fit to new epochs by only optimizing

152   the motif weightings over time (i.e. not changing the motifs themselves, see Methods for details).

153   Indeed, the motifs generalized; the same motifs could explain 74.82% of the variance in neural

154   activity in withheld data from the same recording session (Fig. 3D, purple, CI: 73.92-76.05%; see

155   Fig. 3 supplement 1A for robustness to sparsity parameter; see Methods for details). This was

156   not just due to fitting activity on average: motifs captured neural activity at each timepoint during

157   a recording epoch, explaining the majority of the variance in the spatial distribution of neural

158   activity in any given frame (60-80%, Fig. 3 supplement 1B).

159   Dynamics were important for the ability to generalize. To show this, we created 'static networks'

160   by averaging neural activity across the timecourse of each motif. This maintained the overall

161   spatial pattern of activity, ensuring the same network of brain regions was activated, but removed

162   any temporal dynamics within a motif (Fig. 3E; see Methods for details). When the static networks

163   were fit to withheld data, they captured significantly less variance in neural activity compared to

164   the dynamic motifs (Fig. 3D, gray; static networks captured 55.50%, CI: 53.74-57.09%; a 19.32%

165   reduction, $p<10^{-16}$, Wilcoxon Signed-Rank Test).

166   Similarly, motifs generalized across animals: motifs identified in one animal cross-generalized to

167   capture 68.19% of the variance in neural activity in other animals (Fig; 4A, green; CI: 66.74-

168   69.35%, a decrease of 6.63% compared to generalizing within animals, purple, $p<10^{-16}$, Wilcoxon

7

169   Signed-Rank Test; N=144 withheld epochs; see Methods for details). Together, our results show

170   motifs can generalize across recording session and across animals, suggesting they reflect

171   repeated spatio-temporal dynamics in neural activity.

172   <u>Motifs cluster into a low-dimensional set of basis motifs</u>

173   The ability of motifs to generalize across time and animals suggests there may be a set of 'basis

174   motifs' that capture canonical patterns of spatio-temporal dynamics. To identify these basis motifs,

175   we used an unsupervised clustering algorithm to cluster all 2622 motifs that were identified across

176   all discovery epochs (clustering done with the Phenograph algorithm using peak of cross-

177   correlation as the distance metric between motifs, Nicosia et al., 2009; Levine et al., 2015, see

178   Methods for details). Motifs clustered into a set of 14 unique groups (Fig. 4B). For each cluster,

179   we defined the basis motif as the mean of the motifs within the 'core-community' of each cluster

180   (taken as those motifs with the top 10% most within-cluster nearest neighbors, see Methods for

181   details).

182   Similar to the motifs discovered within a single session, the basis motifs captured the dynamic

183   engagement of one or more brain regions (Fig. 4C; all basis motifs shown in Supplemental Movie

184   3). While some basis motifs engaged a single brain region (e.g. motif 3, Fig. 4C), most of the

185   basis motifs captured the propagation of activity across cortex. For example, motif 4 captures the

186   posterior-lateral flow of activity from retrosplenial to visual cortex. Similarly, motif 11 captures a

187   cortex-wide anterior-to-posterior wave of activity that has been previously studied (Greenberg et

188   al., 2018; Matsui et al., 2016; Mitra et al., 2018). As expected, these dynamics were similar to

189   those found in individual recording sessions (e.g. basis motif 9 matches example motif 2 in Fig.

190   1D).

191   At the same time, the same brain region, or network of regions, can be engaged in multiple basis

192   motifs. For instance, parietal cortex is engaged in motifs 3, 5 and 8 (Fig. 4C). In motif 3, neural

8

193   activity remains local to parietal cortex for the duration of the motif. However, in motif 5 parietal

194   activity is prefaced by a burst of activity in rostrolateral cortex. In motif 8, activity starts in parietal

195   areas before spreading across the entire dorsal cortex. Similarly, several motifs (6, 8, 11, and 14)

196   involve coactivation of a network of anterolateral somatosensory and primary motor cortices; a

197   coupling observed in previous mesoscale imaging studies (Musall et al., 2019; Silasi et al., 2016;

198   Vanni et al., 2017). Thus, basis motifs reflect the ordered engagement of multiple brain regions,

199   likely reflecting a specific flow of information through the brain.

200   Basis motifs explained the large majority of the variance in neural activity across animals (73.91%

201   CI: 73.14-75.19%, Fig. 4A, orange; N=144 withheld epochs). This is about the same amount of

202   variance explained by motifs defined within the same animal (Fig. 4A, purple vs. orange; a 0.91%

203   reduction, p=0.074; Wilcoxon Signed-Rank Test, N=144 withheld epochs). It is significantly more

204   variance explained than when using motifs defined in another animal (Fig. 4A, orange vs. green

205   plots; a 5.72% increase in explained variance; $p<10^{-16}$, Wilcoxon Signed-Rank Test). This

206   improvement is likely because basis motifs are averaged across many instances, removing the

207   spurious noise that exists in individual motifs and resulting in a better estimate of the underlying

208   'true' motif that exist across animals.

209   As before, dynamics were important for basis motifs; when spatial-temporal dynamics were

210   removed, the variance explained dropped significantly (Fig. 4A, gray vs orange plots; static

211   networks captured 48.99% CI: 47.15-51.30% of variance, 24.92% less than dynamic motifs, $p<10^{-16}$

212   , N=144 withheld epochs, Wilcoxon Signed-Rank Test). Furthermore, all basis motifs were

213   necessary to explain neural activity; the cumulative PEV of motifs followed a gradual rise and no

214   basis motif contributed less than 2% of relative PEV on average (Fig. 4 supplement 1A).

215   The high explanatory power of the 14 basis motifs suggests they provide a low-dimensional basis

216   for capturing the dynamics of neural activity in the cortex. This is consistent with the number of

217   motifs (~19) identified in each recording session (the slightly lower number of basis motifs could

218   reflect spurious noise in individual sessions). Importantly, the number of discovered basis motifs

219   was robust to CNMF parameters (Fig. 4 supplement 1B) and potential hemodynamic contributions

220   to basis motifs were minimal (Fig. 4 supplement 2, see Methods for details). In addition, the low

221   number of basis motifs was not due to the resolution of our approach. We estimated the functional

222   resolution of our imaging approach by correlating pixels across time (Fig. 4 supplement 3). This

223   revealed ~18 separate functional regions in dorsal cortex (Fig. 4 supplement 3). Individual motifs

224   engaged multiple of these regions over time (Fig. 4C and supplement 3), consistent with the idea

225   that motifs were not constrained by our imaging approach. Indeed, the number of motifs observed

226   was substantially less than the possible number of motifs; even if motifs engaged only 1-2 of these

227   regions, there are still 18^2=324 different potential motifs, much higher than the 14 we observed.

228   Finally, low dimensionality of basis motifs was not due to compositionality of motifs across time,

229   as this was penalized in the discovery algorithm and the temporal dependency between motifs

230   was weak (Fig. 4 supplement 4; see Methods for details).

231   <u>Basis motifs generalize across behaviors</u>

232   So far, we have only described the motifs of neural activity in animals 'at rest'. To test whether

233   these same motifs can explain neural activity in multiple behavioral states, we imaged dorsal

234   cortex while animals were engaged in a variety of behaviors.

235   To begin, we measured the expression of basis motifs in different spontaneous behavioral states.

236   Two mice, which were not used to define the original basis motifs, were imaged for 1 hour while

237   head-fixed on a transparent treadmill (Fig. 5A). As with the original mice, basis motifs captured

238   the majority of variance in neural activity in both animals (Mouse 1: 64.42%, Mouse 2: 66.66%).

239   The ability of basis motifs to generalize outside the set of animals in which they were discovered

240   provides further support for the idea that basis motifs capture core, repeated, spatio-temporal

241   dynamics in neural activity.

10

242   Next, we were interested in whether basis motifs were specific to different behavioral states in the

243   freely behaving mice. Building from recent work, we used infrared cameras to track three different

244   behavior measures: limb speed, whisker pad motion energy, and nose motion energy (Mathis et

245   al., 2018; Musall et al., 2019). To classify behavioral state at each point in time, we fit a gaussian

246   mixed model to the distribution of these three measures (Fig. 5B; see Methods for details).

247   Behavior fell into two broad categories: an 'active' state (high whisker pad energy, low nose

248   motion energy, high limb speed) and an 'inactive' state (low whisker pad energy, high nose motion

249   energy, low limb speed; Fig 5A). Behavioral states typically lasted for several seconds (median

250   duration for active states: 2.65s and 2.46s; inactive states: 7.88s and 11.54s in mouse 1 and 2

251   respectively), a longer timescale than the motifs (which are all less than 1 second).

252   All motifs occurred in both behavioral states (Fig. 5B-C). Follow-up, detailed analyses did not find

253   specific motifs were associated with any specific behaviors, at least as captured by our three

254   tracked measures of behavior (i.e. no motif was associated with grooming, onset of walking,

255   stopping walking, paw repositioning, sniffing, whisking, etc.; data not shown). However, the

256   activity of motifs did vary across the two broad behavioral states (Fig. 5B-C; the activity of 9/14

257   and 10/14 motifs were different across behavioral states at p<0.05 in Mouse 1 and 2, respectively,

258   Mann-Whitney U-test; this is significantly more than chance, $p=10^{-11}$ and $p=10^{-12}$, binomial test).

259   Similarly, the distribution of motif responses could be used to classify which behavioral state the

260   animal was in (mean classification AUC was 63.42% and 66.51% on withheld data, 99/100 cross-

261   validations were above chance for both animals, 50%; see Methods for details). Therefore, while

262   motifs were not specific to an individual behavioral state, how often a motif was expressed differed

263   across states.

264   Similar results were seen when animals were engaged in social behaviors. Using a novel paired-

265   imaging paradigm, two mice were simultaneously imaged under the same widefield macroscope

266   (Fig. 5D; see Methods for details). Mice were head-fixed near one another (~5mm snout-to-snout),

11

267   enabling sharing of social cues (e.g. whisking, sight, vocalizations, olfaction). To add richness to

268   the sensory environment, mice were intermittently exposed to 'auditory movies' that consisted of

269   male mouse vocalizations and synthetic tones (see Methods for details). In this way, the social

270   environment provided a rich, unstructured sensory environment that is fundamentally different

271   from the solo, low-sensory environment used to define the basis motifs. Even in this vastly

272   different environment, basis motifs defined in the original environment captured 73.41% (CI:

273   71.85-75.23%) of the variance in neural activity (Fig. 5E, right orange plot). This was similar to

274   the variance explained in the solo environment (Fig. 5E, left orange plot; 73.91% CI: 73.14-

275   75.19%; difference between the solo and social environments=0.50%, p=0.49, N=144 solo

276   epochs, N=123 social epochs; Mann-Whitney U-test). Similarly, the opposite was generally true:

277   11 basis motifs were identified in the social recordings, which captured slightly less of the neural

278   activity in the solo environment (Fig. 5E blue plots; solo=70.83% CI: 69.77-72.09%,

279   social=73.72% CI: 71.54-75.05%; difference=2.89%, p=0.0047, Mann-Whitney U-test).

280   As before, many basis motifs changed their expression when the behavioral state of the animal

281   changed. Half of the basis motifs had a significant change in their relative explained variance in

282   the social environment compared to baseline (Fig. 5F; 7/14 were different at $p_{Bonferroni}$<0.05, Mann-

283   Whitney U-test; significantly more than chance, p=10^{-16}, binomial test). Given the nature of social

284   interactions in mice, one would expect tactile- and visual-associated motifs to be increased.

285   Consistent with this prediction, the motifs elevated in the social environment (motifs 1, 9, and 10)

286   involved somatosensory and/or visual regions (Fig. 5G and 4C).

287   Specific motifs capture visual and tactile sensory processing

288   To begin to understand the computational role of specific motifs, we measured the response of

289   the motifs to tactile and visual stimuli (Fig. 6A, see Methods for details). Solo, awake mice (N=9)

290   were imaged while presented with either moving gratings on a screen or airpuffs to their whiskers

291   (see Methods for details). As before, the distribution of motifs differed between both sensory

292　environments and the original, resting environment; a large proportion of motifs showed a

293　significant change in their relative PEV (Fig. 6B-C: visual: 9/14, tactile: 11/14, significantly different

294　at $p_{Bonferroni}$<0.05, Mann-Whitney U-test; N=1109 visual and N=1110 tactile presentations, N=144

295　original epochs). Most of these changes were reductions in explained variance, with only a few

296　motifs increasing their expression to the stimuli (Fig. 6B-C).

297　Specific motifs were evoked in response to visual or tactile stimulation. Motif 10 was selectively

298　induced by visual stimulation (Fig. 6D), while motif 1 was selectively induced by tactile stimulation

299　(Fig. 6E). Consistent with a role in sensory processing, both motifs had increased expression in

300　their respective environments (Fig. 6B-C). Furthermore, the spatio-temporal pattern of activity in

301　each motif matched the trial-averaged response to the associated stimulus; motif 10 captured

302　activity in visual cortex, while motif 1 captured activity in motor, somatosensory and parietal

303　cortices (Fig. 6F). Reflecting their overlap, both motifs were significantly correlated with the trial-

304　averaged response (Fig. 6F, and Pearson's R=0.81, $p<10^{-16}$, for correlation between the average

305　visual response and motif 10, and R=0.90, $p<10^{-16}$, for correlation between the average tactile

306　response and motif 1, all taken during the 13 timepoints post stimulation onset).

307　While certain basis motifs captured the sensory-evoked neural activity, other motifs were able to

308　capture trial-by-trial 'noise' in neural activity. For both stimuli, basis motifs, identified at rest,

309　captured the majority of the variance in neural activity (Fig. 6G; visual: 58.13% CI: 57.11-59.18%;

310　tactile: 62.37% CI: 61.42-63.62%, N=1110 tactile and 1109 visual stimulus presentations; see

311　Methods for details). This was significantly more variance than could be explained by the mean

312　response to each sensory stimulus alone (Fig. 6G; Mean response fits: visual: 17.25% CI: 16.61-

313　17.72%; tactile: 34.42% CI: 33.34-35.40%; difference between motif and mean fits: visual:

314　40.88%, $p<10^{-16}$; tactile: 27.95%, $p<10^{-16}$; Wilcoxon Signed-Rank Test). This highlights the high

315　variability in responses to a sensory stimulus across trials. Typically, such variability would be

316　discarded as 'noise' unrelated to sensory processing. Instead, our results suggest this variability

317   has structure: it is due to the engagement of other motifs that are, presumably, related to other,

318   ongoing, computations.

319   Finally, we sought to determine whether motifs reflect general stimulus processing or specific

320   stimulus features. To this end, we compared motif expression in response to two visual stimuli

321   (Fig. 7; gratings moving medial to lateral or lateral to medial; see Methods for details and Fig. 7

322   supplement 1 for similar analysis of tactile stimuli). Unlike comparisons across sensory modalities,

323   the relative percent explained variance in neural activity captured by each basis motif was the

324   same for both visual stimuli (p>0.24 for all 14 motifs; Mann-Whitney U-Test; N=554 and 555

325   stimulus presentations for medial to lateral and lateral to medial grating respectively). For

326   example, the visually responsive motif 10 responded equally to both stimuli (Fig. 7A). This was

327   not due to limits in spatial resolution of our imaging approach or analytical smoothing. Figure 7B

328   shows pixel-wise classification of the same data can decode stimulus identify (p=0.022, N=9

329   animals, one-sample t-test; see Methods for details). However, the same classification analysis

330   on data reconstructed from motif activity failed to distinguish between stimuli (p=0.87, N=9

331   animals, one-sample t-test; difference between classification on original and reconstructed data

332   was significant, p=0.016, paired sample t-test). Thus, the specifics of visual stimuli were encoded

333   in the residuals after fitting the motifs. However, these details contributed minimally to the overall

334   neural activity during the stimulus. The stimulus-specific residuals captured only 3.85% +/- 0.70%

335   SEM of the explainable variance (Fig. 7C). In contrast, motifs captured the vast majority of

336   explainable variance (19.23% +/- 3.23% SEM for stimulus-specific motif 10; 76.92% +/- 3.86%

337   SEM, for remaining motifs; Fig. 7C). Taken together, our results show that motifs capture large-

338   scale patterns of neural activity but are generally agnostic to the finer-grain local activity that

339   captured specifics of stimuli. This is consistent with the idea that motifs capture the broader flow

340   of information across cortical regions.

341   **Discussion**

14

342   <u>Spatio-temporal dynamics of cortex-wide activity</u>

343   Our results show that neural activity is highly dynamic, evolving in both time and space.

344   Leveraging mesoscale calcium imaging in mice, we tracked the spatio-temporal dynamics of

345   neural activity across the dorsal surface of the cortex. Using a convolutional factoring analysis,

346   we identified 'motifs' in neural activity. Each motif reflected a different spatio-temporal pattern of

347   activity, with many motifs capturing the sequential activation of multiple, functionally diverse,

348   cortical regions (Figs. 1-3). Together, these motifs explained the large majority of variance in

349   neural activity across different animals (Fig. 4) and in novel behavioral situations (Fig. 5-6).

350   A couple of the motifs captured patterns of activity observed in previous work, supporting the

351   validity of the CNMF approach. For example, previous work has studied spatio-temporal waves

352   of activity that propagate anterior-to-posteriorly across the cortex at different temporal scales

353   (Greenberg et al., 2018; Matsui et al., 2016; Mitra et al., 2018). Our motifs (1 and 11) recapitulate

354   these waves, along with their temporal diversity (motif 1 = fast, motif 11 = slow). In addition to

355   these motifs, we also discovered several additional, spatio-temporally distinct, anterior-to-

356   posterior propagating waves (motifs 2, 4, and 9).

357   Similarly, brain regions that were often co-activated in motifs were aligned with previously

358   reported spatial patterns of co-activation in the mouse cortex (Musall et al., 2019; Silasi et al.,

359   2016; Vanni et al., 2017). For example, motifs 6, 8, 11, and 14 include coactivation of anterolateral

360   somatosensory and motor regions. This pattern is observed often and reflects the close functional

361   relationship between motor activity and somatosensory processing. Here, we extend this work by

362   showing neural activity can flow within and between these networks in different ways.

363   Relatedly, previous work using mesoscale imaging demonstrated that the mouse cortex exhibits

364   repeating patterns of activity (Mohajerani et al., 2013). However, this work relied on identifying

365   average patterns evoked by sensory stimuli (visual, tactile, auditory) and correlating the spatially

15

366   and temporally static templates of those patterns to activity in resting animals. As we demonstrate,

367   stimulus-evoked patterns capture considerably less variance in neural activity than our approach,

368   even in response to sensory stimuli themselves (average stimuli responses: ~15-35% of activity

369   versus motifs: ~60%). In addition, previous work has used zero-lag correlations to show the brain

370   transitions through different functional network states over time (Ashourvan et al., 2017;

371   Hutchison et al., 2013; Preti et al., 2017; Vidaurre et al., 2017). Here, we show that these

372   functional network states themselves have rich dynamics, reflecting specific sequential patterns

373   of activity across the network. By encapsulating these dynamics, motifs are able to capture

374   significantly more of the variance in neural activity compared to static networks.

375   <u>Motifs of neural activity may reflect cognitive computations</u>

376   Each motif captured a different spatio-temporal pattern of neural activity. As neural activity passes

377   through the neural network of a brain region, it is thought to be functionally transformed in a

378   behaviorally-relevant manner (e.g. visual processing in visual cortex or decision-making in parietal

379   cortex). Therefore, the dynamic activation of multiple regions in a motif, could reflect a specific,

380   multi-step, ordered transformation of information. In this way, the basis motifs would reflect a set

381   of 'core computations' carried out by the brain.

382   Consistent with this hypothesis, the distribution of motifs differed across behavioral states (Fig. 5)

383   and in response to social and sensory stimuli (Figs. 6-7). Specific motifs were also associated

384   with specific cognitive processes, such as tactile and visual processing (Fig. 7). This follows

385   previous work showing the engagement of brain networks is specific to the current behavior

386   (Mattar et al., 2015; Telesford et al., 2016) and that disrupting these networks underlies numerous

387   pathologies (Badhwar et al., 2017; Braun et al., 2016; Harlalka et al., 2019).

388   In this way, our results are consistent with a hierarchical relationship between behavior and

389   cortical dynamics. Behavior engages a set of core cognitive computations, reflected in the

16

390   activation of multiple motifs in each behavioral state (Fig. 5). Multiple behavioral states engage

391   the same motifs, but they engage them to a different extent. This suggests different behavioral

392   states do not require independent patterns of neural activities, but rather reflect a re-weighting of

393   the expression of core computations.

394   <u>Low dimensional structure of cortex-wide dynamics may facilitate cognitive control of neural</u>

395   <u>activity</u>

396   Our results show that the dynamics of the cortex are low dimensional. Motifs identified in different

397   animals and recording sessions clustered into a limited set of 14 unique 'basis' motifs. This limited

398   number of basis motifs captured the large majority of variance in neural activity (~75%) across

399   animals, across behavioral states, and in different behavioral environments. Such a low-

400   dimensional repertoire of cortical activity is consistent with previous work using zero-lag

401   correlations of neural activity to measure functional connectivity between brain regions (e.g. 17

402   functional networks in the human cortex using fMRI, Yeo et al., 2011).

403   Why might cortex-wide neural activity be low dimensional? Given the multitude of sensory inputs,

404   internal states, and motor actions, one might expect neural activity to be extremely high

405   dimensional. Indeed, recent large-scale electrophysiology and imaging studies have found high

406   dimensional (>100) representations within a brain region, with neural activity often representing

407   small aspects of behavior (e.g. facial twitches, limb position, etc; Stringer et al., 2019a, 2019b;

408   Lieber and Bensmaia, 2019). However, while this high dimensionality of neural representations is

409   great for capturing the fullness of an experience, its complexity presents a problem for cognitive

410   control.

411   In most situations, we do not *a priori* know the exact set of computations to engage. Instead, we

412   must learn which computations are optimal for the current situation. One way to do this is to

413   search through the space of possible computations until the 'best' computation is found. This

414   could explain the diversity of motifs observed in our animals at rest; the expression of different

415   motifs reflects the 'sampling' of different computations as the animals search for the one that fits

416   their current, novel, situation. Similarly, the greater consistency in the motifs expressed during

417   social or sensory environments may reflect a reduction in the uncertainty about what

418   computations to engage in.

419   However, in this framework, a high dimensionality is costly. As the number of possible

420   computations (and their neural representations) increases, it becomes harder to find the optimum

421   (the classic 'curse of dimensionality' problem; Bellman, 2003). Therefore, limiting the number of

422   motifs may make it easier to find a contextually appropriate computation. The caveat to a low

423   dimensional repertoire of motifs is that it necessitates a coarser sampling of computational space,

424   which may limit how well the best computation approximates the true optimal computation. In this

425   way, the dimensionality of dynamics may reflect a trade-off between the time it takes to identify

426   an appropriate computation and the optimality of that computation.

427   More broadly, our results are consistent with a hierarchical control of information processing in

428   the brain (Mearns et al., 2019; Deco and Kringelbach, 2017; Ashourvan et al., 2019; Park and

429   Friston, 2013; Botvinick, 2008). Behavioral state changes slowly and reflects the animal's broad

430   behavioral goal (Wiltschko et al., 2015). On a shorter timescale, control mechanisms coordinate

431   broad dynamics across cortical regions (motifs). This allows the animal to engage in a

432   behaviorally relevant category of computations (e.g. tactile versus visual processing). High

433   dimensional representations in local circuits then produce more nuanced processing (e.g. the

434   specific identity of a visual stimulus). In this way, by activating specific motifs, control processes

435   could direct the broader flow of information across brain regions (a low-dimensional problem)

436   while avoiding the difficulty of directing detailed single-neuron processing (a high-dimensional

437   problem).

438   Future Directions

18

439   Our approach has several limitations that motivate future research. First, although we are able to

440   capture a large fraction of the cortex, we are limited to the dorsal cortex and so miss out on lateral

441   auditory cortex, cortical regions deep along the midline, and all sub-cortical regions. Second, while

442   we probed activity in several different environments, imaging was always restricted to head-fixed

443   animals. Third, the number and identity of motifs, as well as the relative contributions of spatial

444   and temporal dynamics to variance in neural activity, is likely influenced by the nature of our

445   approach. The relatively slow timecourse of GCaMP6f (Chen et al., 2013), and biases in the

446   neural activity underlying the calcium signal (Allen et al., 2017; Makino et al., 2017) may have

447   decreased the spatial resolution and slowed the temporal dynamics. However, it is important to

448   note that the spatial resolution used in our imaging approach ($\sim$136 μm^2/pixel) is higher than the

449   broad activation of brain regions observed in the motifs and was high enough to capture pixel-

450   specific information about stimuli beyond the motifs. Furthermore, our functional resolution was

451   able to distinguish 18 distinct regions. Motifs engaged multiple regions, suggesting motifs were

452   broader than the functional resolution (and therefore not limited by our approach). Finally, even

453   given these constraints, the number of potential spatio-temporal patterns is far higher than the 14

454   basis motifs found.

455   In addition to addressing these limitations, future work is needed to understand the computations

456   associated with each motif. Here we've used simple behavioral paradigms to associate a few

457   motifs with sensory processing. However, the computation underlying many of the motifs remains

458   unknown – by cataloging motif expression across experiments and behaviors, we can begin to

459   understand the function of each motif and gain a more holistic understanding of how and why

460   neural activity evolves across the brain in support of behavior.

461   **Acknowledgments**

19

464   Henrickson for their detailed feedback during the writing of this manuscript. We also thank

465   Stephan Thiberge for designing and constructing the widefield macroscope. We thank the

466   Princeton Laboratory Animal Resources staff for their support. This work was funded by NIH

467   DP2 EY025446.

468   **References**

469   Afrashteh, N., Inayat, S., Mohsenvand, M., and Mohajerani, M.H. (2017). Optical-flow analysis toolbox

470   for characterization of spatiotemporal dynamics in mesoscale optical imaging of brain activity.

471   Neuroimage *153*, 58–74.

472   Allen, W.E., Kauvar, I.V., Chen, M.Z., Richman, E.B., Yang, S.J., Chan, K., Gradinaru, V., Deverman, B.E.,

473   Luo, L., and Deisseroth, K. (2017). Global Representations of Goal-Directed Behavior in Distinct Cell

474   Types of Mouse Neocortex. Neuron *94*, 891-907 e6.

475   Ashourvan, A., Gu, S., Mattar, M.G., Vettel, J.M., and Bassett, D.S. (2017). The energy landscape

476   underpinning module dynamics in the human brain connectome. Neuroimage *157*, 364–380.

477   Ashourvan, A., Telesford, Q.K., Verstynen, T., Vettel, J.M., and Bassett, D.S. (2019). Multi-scale detection

478   of hierarchical community architecture in structural and functional brain networks. PLOS ONE *14*,

479   e0215520.

480   Badhwar, A., Tam, A., Dansereau, C., Orban, P., Hoffstaedter, F., and Bellec, P. (2017). Resting-state

481   network dysfunction in Alzheimer's disease: A systematic review and meta-analysis. Alzheimers Dement

482   (Amst) *8*, 73–85.

483   Bellman, R.E. (2003). Dynamic Programming (Courier Corporation).

484   Botvinick, M.M. (2008). Hierarchical models of behavior and prefrontal function. Trends Cogn Sci *12*,

485   201–208.

486   Brainard, D. (1997). The Psychophysics Toolbox. Spatial Vision *10*.

487   Braun, U., Schafer, A., Bassett, D.S., Rausch, F., Schweiger, J.I., Bilek, E., Erk, S., Romanczuk-Seiferth, N.,

488   Grimm, O., Geiger, L.S., et al. (2016). Dynamic brain network reconfiguration as a potential

489   schizophrenia genetic risk mechanism modulated by NMDA receptor function. Proc Natl Acad Sci U S A

490   *113*, 12568–12573.

491   Buonomano, D.V., and Maass, W. (2009). State-dependent computations: spatiotemporal processing in

492   cortical networks. Nature Reviews Neuroscience *10*, 113–125.

493   Buschman, T.J., and Miller, E.K. (2007). Top-down versus bottom-up control of attention in the

494   prefrontal and posterior parietal cortices. Science *315*, 1860–1862.

495   Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A.,

496   Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal

497   activity. Nature *499*, 295–300.

498   Chen, T.-W., Li, N., Daie, K., and Svoboda, K. (2017). A Map of Anticipatory Activity in Mouse Motor

499   Cortex. Neuron *94*, 866-879.e4.

500   Cramer, J.V., Gesierich, B., Roth, S., Dichgans, M., Düring, M., and Liesz, A. (2019). In vivo widefield

501   calcium imaging of the mouse cortex for analysis of network connectivity in health and brain disease.

502   NeuroImage *199*, 570–584.

503   Dana, H., Chen, T.-W., Hu, A., Shields, B.C., Guo, C., Looger, L.L., Kim, D.S., and Svoboda, K. (2014). Thy1-

504   GCaMP6 Transgenic Mice for Neuronal Population Imaging In Vivo. PLOS ONE *9*, e108697.

505   Deco, G., and Kringelbach, M.L. (2017). Hierarchy of Information Processing in the Brain: A Novel

506   'Intrinsic Ignition' Framework. Neuron *94*, 961–968.

507   Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. Neuron *88*, 220–235.

508   Friston, K.J. (2004). Chapter 49 - Functional Connectivity: Eigenimages and Multivariate Analyses. In

509   Human Brain Function (Second Edition), R.S.J. Frackowiak, C.D. Frith, R.J. Dolan, C.J. Price, S. Zeki, J.T.

510   Ashburner, and W.D. Penny, eds. (Burlington: Academic Press), pp. 999–1018.

511   Greenberg, A., Abadchi, J.K., Dickson, C.T., and Mohajerani, M.H. (2018). New waves: Rhythmic electrical

512   field stimulation systematically alters spontaneous slow dynamics across mouse neocortex. NeuroImage

513   *174*, 328–339.

514   Guo, Z.V., Li, N., Huber, D., Ophir, E., Gutnisky, D., Ting, J.T., Feng, G., and Svoboda, K. (2014). Flow of

515   cortical activity underlying a tactile decision in mice. Neuron *81*, 179–194.

516   Harlalka, V., Bapi, R.S., Vinod, P.K., and Roy, D. (2019). Atypical Flexibility in Dynamic Functional

517   Connectivity Quantifies the Severity in Autism Spectrum Disorder. Front Hum Neurosci *13*, 6.

518   Horn, B.K.P., and Schunck, B.G. (1981). Determining Optical Flow. 19.

519   Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Della Penna,

520   S., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., et al. (2013). Dynamic functional connectivity: promise,

521   issues, and interpretations. Neuroimage *80*, 360–378.

522   Leonardi, N., Richiardi, J., Gschwind, M., Simioni, S., Annoni, J.-M., Schluep, M., Vuilleumier, P., and Van

523   De Ville, D. (2013). Principal components of functional connectivity: a new approach to study dynamic

524   brain connectivity during rest. Neuroimage *83*, 937–950.

525   Lerner, T.N., Shilyansky, C., Davidson, T.J., Evans, K.E., Beier, K.T., Zalocusky, K.A., Crow, A.K., Malenka,

526   R.C., Luo, L., Tomer, R., et al. (2015). Intact-Brain Analyses Reveal Distinct Information Carried by SNc

527   Dopamine Subcircuits. Cell *162*, 635–647.

528   Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir el, A.D., Tadmor, M.D., Litvin, O., Fienberg,

529   H.G., Jager, A., Zunder, E.R., et al. (2015). Data-Driven Phenotypic Dissection of AML Reveals Progenitor-

530   like Cells that Correlate with Prognosis. Cell *162*, 184–197.

531   Lieber, J.D., and Bensmaia, S.J. (2019). High-dimensional representation of texture in somatosensory

532   cortex of primates. Proc Natl Acad Sci U S A *116*, 3268–3277.

533   Ma, Y., Shaik, M.A., Kozberg, M.G., Kim, S.H., Portes, J.P., Timerman, D., and Hillman, E.M.C. (2016a).

534   Resting-state hemodynamics are spatiotemporally coupled to synchronized and symmetric neural

535   activity in excitatory neurons. PNAS *113*, E8463–E8471.

536   Ma, Y., Shaik, M.A., Kim, S.H., Kozberg, M.G., Thibodeaux, D.N., Zhao, H.T., Yu, H., and Hillman, E.M.C.

537   (2016b). Wide-field optical mapping of neural activity and brain haemodynamics: considerations and

538   novel approaches. Philos Trans R Soc Lond B Biol Sci *371*.

539   Mackevicius, E.L., Bahle, A.H., Williams, A.H., Gu, S., Denisenko, N.I., Goldman, M.S., and Fee, M.S.

540   (2019). Unsupervised discovery of temporal sequences in high-dimensional datasets, with applications

541   to neuroscience. Elife *8*.

542   Makino, H., Ren, C., Liu, H., Kim, A.N., Kondapaneni, N., Liu, X., Kuzum, D., and Komiyama, T. (2017).

543   Transformation of Cortex-wide Emergent Properties during Motor Learning. Neuron *94*, 880-890 e8.

544   Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018).

545   DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nature

546   Neuroscience *21*, 1281–1289.

547   Matsui, T., Murakami, T., and Ohki, K. (2016). Transient neuronal coactivations embedded in globally

548   propagating waves underlie resting-state functional connectivity. PNAS *113*, 6556–6561.

549   Mattar, M.G., Cole, M.W., Thompson-Schill, S.L., and Bassett, D.S. (2015). A Functional Cartography of

550   Cognitive Systems. PLoS Comput Biol *11*, e1004533.

551   Mearns, D.S., Donovan, J.C., Fernandes, A.M., Semmelhack, J.L., and Baier, H. (2019). Deconstructing

552   Hunting Behavior Reveals a Tightly Coupled Stimulus-Response Loop. Current Biology *0*.

553   Miller, E.K., and Wilson, M.A. (2008). All My Circuits: Using Multiple Electrodes to Understand

554   Functioning Neural Networks. Neuron *60*, 483–488.

555   Mitra, A., Kraft, A., Wright, P., Acland, B., Snyder, A.Z., Rosenthal, Z., Czerniewski, L., Bauer, A., Snyder,

556   L., Culver, J., et al. (2018). Spontaneous Infra-slow Brain Activity Has Unique Spatiotemporal Dynamics

557   and Laminar Structure. Neuron *98*, 297-305.e6.

558   Mohajerani, M.H., Chan, A.W., Mohsenvand, M., LeDue, J., Liu, R., McVea, D.A., Boyd, J.D., Wang, Y.T.,

559   Reimers, M., and Murphy, T.H. (2013). Spontaneous cortical activity alternates between motifs defined

560   by regional axonal projections. Nat Neurosci *16*, 1426–1435.

561   Muller, L., Chavane, F., Reynolds, J., and Sejnowski, T.J. (2018). Cortical travelling waves: mechanisms

562   and computational principles. Nat Rev Neurosci *19*, 255–268.

563   Murphy, T.H., Boyd, J.D., Bolaños, F., Vanni, M.P., Silasi, G., Haupt, D., and LeDue, J.M. (2016). High-

564   throughput automated home-cage mesoscopic functional imaging of mouse cortex. Nature

565   Communications *7*, 11611.

566   Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural

567   dynamics are dominated by richly varied movements. Nature Neuroscience *22*, 1677–1686.

568   Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M., and Mathis, M.W. (2019). Using DeepLabCut for 3D

569   markerless pose estimation across species and behaviors. Nature Protocols *14*, 2152–2176.

570   Nicosia, V., Mangioni, G., Carchiolo, V., and Malgeri, M. (2009). Extending the definition of modularity to

571   directed graphs with overlapping communities. J. Stat. Mech. *2009*, P03024.

572   Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., et

573   al. (2014). A mesoscale connectome of the mouse brain. Nature *508*, 207–214.

574   Park, H.-J., and Friston, K. (2013). Structural and functional brain networks: from connections to

575   cognition. Science *342*, 1238411.

576   Pereira, F., Mitchell, T., and Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial

577   overview. Neuroimage *45*, S199-209.

578   Pinto, L., Rajan, K., DePasquale, B., Thiberge, S.Y., Tank, D.W., and Brody, C.D. (2019). Task-Dependent

579   Changes in the Large-Scale Dynamics and Necessity of Cortical Regions. Neuron *104*, 810-824.e9.

580   Preti, M.G., and Ville, D.V.D. (2017). Dynamics of functional connectivity at high spatial resolution reveal

581   long-range interactions and fine-scale organization. Scientific Reports *7*, 12773.

582   Preti, M.G., Bolton, T.A., and Van De Ville, D. (2017). The dynamic functional connectome: State-of-the-

583   art and perspectives. Neuroimage *160*, 41–54.

584   Romo, R., and de Lafuente, V. (2013). Conversion of sensory signals into perceptual decisions. Prog

585   Neurobiol *103*, 41–75.

586   Schmeisser, M.J., Ey, E., Wegener, S., Bockmann, J., Stempel, a V., Kuebler, A., Janssen, A.-L., Udvardi,

587   P.T., Shiban, E., Spilker, C., et al. (2012). Autistic-like behaviours and hyperactivity in mice lacking

588   ProSAP1/Shank2. Nature *486*, 256–260.

589   Seabrook, T.A., Burbridge, T.J., Crair, M.C., and Huberman, A.D. (2017). Architecture, Function, and

590   Assembly of the Mouse Visual System. Annual Review of Neuroscience *40*, 499–538.

591   Shimaoka, D., Steinmetz, N.A., Harris, K.D., and Carandini, M. (2019). The impact of bilateral ongoing

592   activity on evoked responses in mouse cortex.

593   Siegel, M., Engel, A.K., and Donner, T.H. (2011). Cortical network dynamics of perceptual decision-

594   making in the human brain. Front Hum Neurosci *5*, 21.

595   Silasi, G., Xiao, D., Vanni, M.P., Chen, A.C.N., and Murphy, T.H. (2016). Intact skull chronic windows for

596   mesoscopic wide-field imaging in awake mice. J Neurosci Methods *267*, 141–149.

597   Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019a).

598   Spontaneous behaviors drive multidimensional, brainwide activity. Science *364*, 255.

599   Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., and Harris, K.D. (2019b). High-dimensional

600   geometry of population responses in visual cortex. Nature 1.

26

601   Telesford, Q.K., Lynall, M.E., Vettel, J., Miller, M.B., Grafton, S.T., and Bassett, D.S. (2016). Detection of

602   functional brain network reconfiguration during task-driven cognitive states. Neuroimage *142*, 198–210.

603   Torquet, N., de Chaumont, F., Faure, P., Bourgeron, T., and Ey, E. (2016). mouseTube – a database to

604   collaboratively unravel mouse ultrasonic communication. F1000Research *5*, 2332.

605   Van Segbroeck, M., Knoll, A.T., Levitt, P., and Narayanan, S. (2017). MUPET-Mouse Ultrasonic Profile

606   ExTraction: A Signal Processing Tool for Rapid and Unsupervised Analysis of Ultrasonic Vocalizations.

607   Neuron *94*, 465-485.e5.

608   Vanni, M.P., and Murphy, T.H. (2014). Mesoscale Transcranial Spontaneous Activity Mapping in GCaMP3

609   Transgenic Mice Reveals Extensive Reciprocal Connections between Areas of Somatomotor Cortex. J

610   Neurosci *34*, 15931–15946.

611   Vanni, M.P., Chan, A.W., Balbi, M., Silasi, G., and Murphy, T.H. (2017). Mesoscale Mapping of Mouse

612   Cortex Reveals Frequency-Dependent Cycling between Distinct Macroscale Functional Modules. J

613   Neurosci *37*, 7513–7533.

614   Vidaurre, D., Smith, S.M., and Woolrich, M.W. (2017). Brain network dynamics are hierarchically

615   organized in time. Proc Natl Acad Sci U S A *114*, 12827–12832.

616   Wiltschko, A.B., Johnson, M.J., Iurilli, G., Peterson, R.E., Katon, J.M., Pashkovski, S.L., Abraira, V.E.,

617   Adams, R.P., and Datta, S.R. (2015). Mapping Sub-Second Structure in Mouse Behavior. Neuron *88*,

618   1121–1135.

619   Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller,

620   J.W., Zollei, L., Polimeni, J.R., et al. (2011). The organization of the human cerebral cortex estimated by

621   intrinsic functional connectivity. J Neurophysiol *106*, 1125–1165.

622   Zanos, T.P., Mineault, P.J., Nasiotis, K.T., Guitton, D., and Pack, C.C. (2015). A sensorimotor role for

623   traveling waves in primate visual cortex. Neuron *85*, 615–627.

624

A Imaging Paradigm

B Schematic of Data Partitioning

Data: 48 sessions, 144 Discovery Epochs, 144 Withheld Epochs, 9 Mice

C Unsupervised Discovery Of Spatio-Temporal Motifs In Neural Data

Convolutional Non-Negative Matrix Factorization With Regularization Identifies **'Motifs'**

Data Reconstruction Using Motifs
(3 out of 16 motifs shown)

Motif Temporal Weights
(Spike = Motif Occurs)

Reconstruction Reproduces Original Data

D Example Motifs From Single Epoch

625

626 **Figure 1. Discovery of spatio-temporal patterns in cortical activity of awake, head-fixed**

627 **mice.**

628 **(A)** Schematic of imaging paradigm. Mice expressing GCaMP6f in cortical pyramidal neurons

629 underwent a translucent skull prep to allow mesoscale imaging of neural activity across the

630 majority of dorsal cortex. Red dot denotes bregma. Cortical parcellation follows Allen Brain Atlas.

631 General anatomical parcels are labeled. Motor, motor cortex; SS, somatosensory cortex; Parietal,

632 parietal cortex; RSP, retrosplenial cortex; Visual: visual cortex (See Fig. 1 supplement 1 for

633 complete parcellation of 24 regions; 12 per hemisphere).

634 **(B)** Schematic of data partitioning. 9 mice were imaged for 12 minutes a day for 5-6 consecutive

635 days. Recording sessions (N=48) were divided into 2-minute epochs (N=144). Alternating epochs

636 were used for discovering spatio-temporal motifs in neural activity or were withheld for testing

637 generalization of motifs.

638 **(C)** Schematic of unsupervised discovery of spatio-temporal motifs from a single epoch.

639 Mesoscale calcium imaging captured patterns of neural activity, measured by change in

640 fluorescence (ΔF/F), across the dorsal cortex (top-left). Movies were vectorized (top-middle,

641 black=activity) and then decomposed into dynamic, spatio-temporal motifs (bottom-right; 3

642 example motifs shown along left, temporal weightings along top). Convolving motifs with temporal

643 weightings reconstructed the original movie (bottom-left; snapshot of data is highlighted in green

644 throughout, corresponding to activity in motif 3, which is shown in image format in bottom row of

645 **D**). Note: only 3 out of 16 example motifs and their corresponding temporal weightings are shown;

646 data reconstruction in bottom right used all 16 motifs.

647 **(D)** Timecourses of the 3 example motifs in panel **C** showing spatio-temporal patterns of neural

648 activity across dorsal cortex. Arrows indicate direction of flow of activity across subsequent

649 timepoints (see Methods for details). Blue dot denotes center of mass of the most active pixels in

30

650   each hemisphere (>=95% intensity). Red dot denotes bregma. Dotted white lines outline

651   anatomical parcels as in **A**. Only every other timepoint is shown. For visualization, motifs were

652   filtered with 3D gaussian (across space and time), and intensity scale is normalized for each motif.

653   Intensity value is arbitrary as responses are convolved with independently scaled temporal

654   weightings to reconstruct the normalized ΔF/F fluorescence trace (see Methods for details).

655

656

657

658

659

660

661

662

663

664

665

666

Characterizing Motifs in Cortical Neural Activity

667
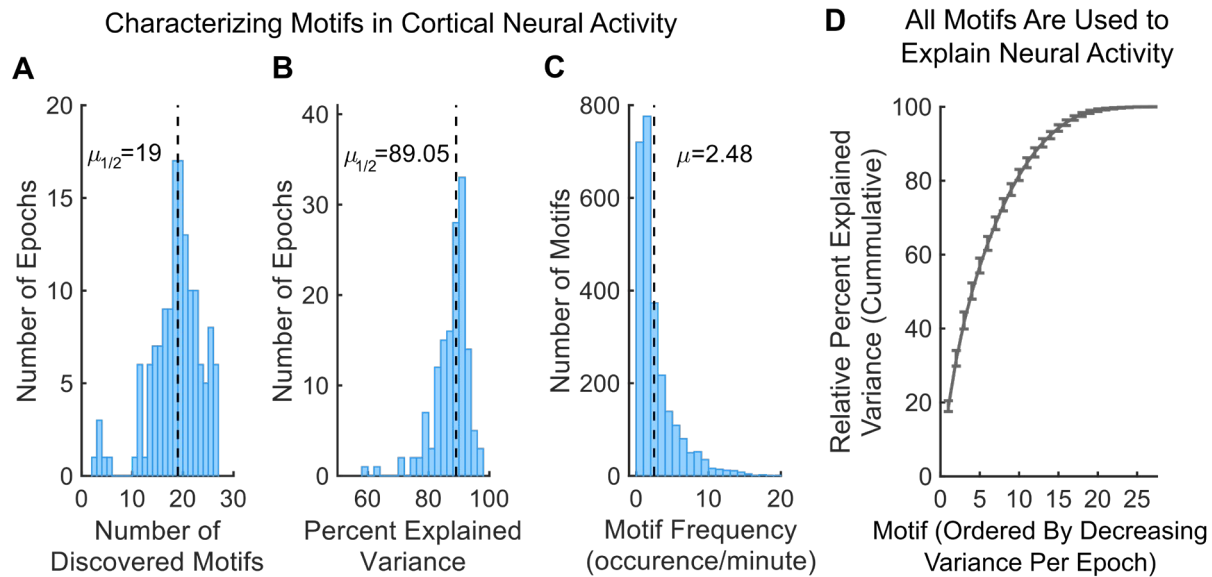
Figure 2. Motifs capture majority of variance in neural activity

(A) Distribution of number of discovered motifs per discovery epoch (N=144). Dotted line indicates median. 2622 motifs were discovered in total.

(B) Distribution of the total percent of variance in neural activity explained by motifs per discovery epoch. Dotted line indicates median.

(C) Distribution of how often motifs occurred during discovery epochs. A motif was considered active when its temporal weighting was 1 standard deviation above its mean (e.g. transients in Fig. 1C in occurrences per minute; see Methods for details). Dotted line indicates mean.

(D) Cumulative sum of relative percent explained variance (PEV) of each motif in withheld epochs. Relative PEV was calculated as the PEV of each motif divided by the sum of all motif PEVs in an epoch. For each epoch, motifs are ordered by their relative PEV (i.e. the first motif is the most common motif, which is not necessarily the same motif for all epochs). Line and error bars indicate mean and 95% CI, respectively.

All p-values estimated with Wilcoxon Signed-Rank tests.
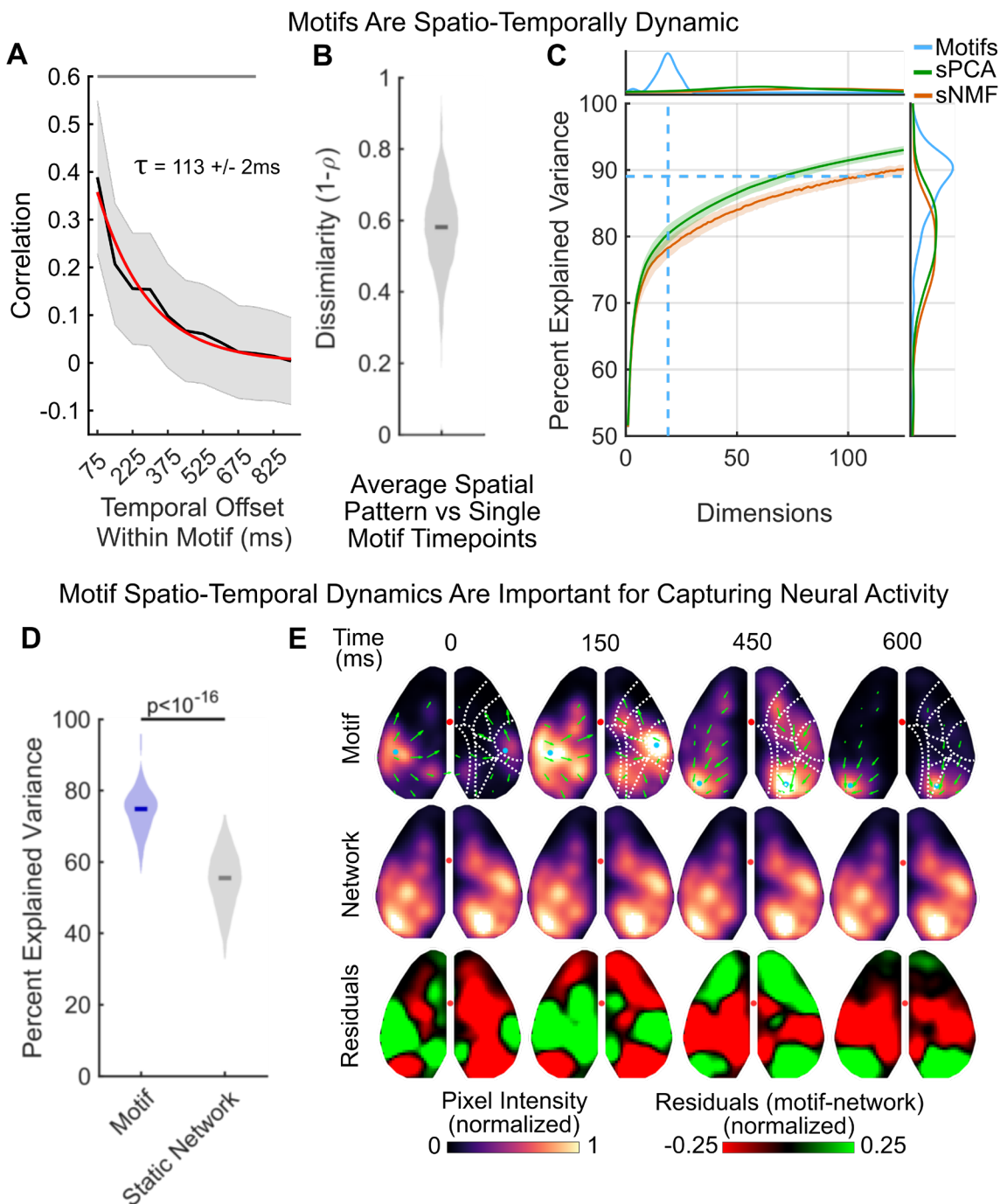
Figure 3. Motifs capture the flow of neural activity across the cortex

(A) Cross-temporal autocorrelation of motifs (N=2622). Average spatial correlation of activity (y-axis) was calculated for different temporal offsets (x-axis) within a motif. For example, an offset of 75ms indicates the correlation between timepoint N and timepoints N-1 and N+1 (given

687   sampling frequency of 13.33 Hz). Black line and gray shading denote mean and standard

688   deviation, respectively, across all motifs. Red line shows exponential fit to autocorrelation decay.

689   Mean half-life of autocorrelation decay (τ) across all motifs was 113ms +/- 2ms SEM.

690   **(B)** Average dissimilarity between each timepoint of a dynamic motif and the mean spatial activity

691   of that motif; averaged across frames of the motifs. Full distribution shown, depicts average

692   dissimilarity per motif (N=2622 motifs). Dark line indicates median.

693   **(C)** Comparison of reconstruction of neural data by CNMF motifs (blue), spatial principal

694   components analysis (sPCA, green) and spatial non-negative matrix factorization (sNMF,

695   orange). Central plot shows percent of variance in neural activity explained (y-axis) as a function

696   of number of dimensions included (x-axis). Lines show median variance explained, shaded

697   regions show 95% confidence interval. Dashed light blue lines show median number of motifs

698   discovered (vertical) and the median percent explained variance (horizontal) captured by CNMF

699   motifs across discovery epochs (N=144). **(top)** Plot shows the probability density function (PDF)

700   of number of motifs discovered per discovery epoch (blue), as well as the PDF of the minimum

701   number of dimensions needed to capture the same amount of variance using sPCA (orange)

702   and sNMF (green). **(right)** PDF of percent explained variance by motif reconstructions (blue)

703   across epochs, as well as the PDF of percent of variance explained by sPCA (orange) and

704   sNMF (green) when the number of dimensions is restricted to match the number of discovered

705   motifs in each epoch. For visualization, x-axis is cropped to 125 dimensions.

706   **(D)** Percent of variance in neural activity explained by dynamic motifs (blue) and static networks

707   (grey), defined as the average activity across the motif. Both static networks and motifs are fit to

708   the data in the same manner (see Methods for details). Full distribution shown; dark lines indicate

709   median. Analyses performed on withheld epochs (N=144).

710   **(E)** An example motif (top row; example motif 2 from Fig. 1D) and its corresponding static

711   network (middle row). Bottom row shows normalized residuals between dynamic motif and static

712   network. For calculation of residuals, motif and networks were scaled to the same mean pixel

713   value per timepoint. Display follows Figure 1D.

714   All p-values estimated with Wilcoxon Signed-Rank tests.

715
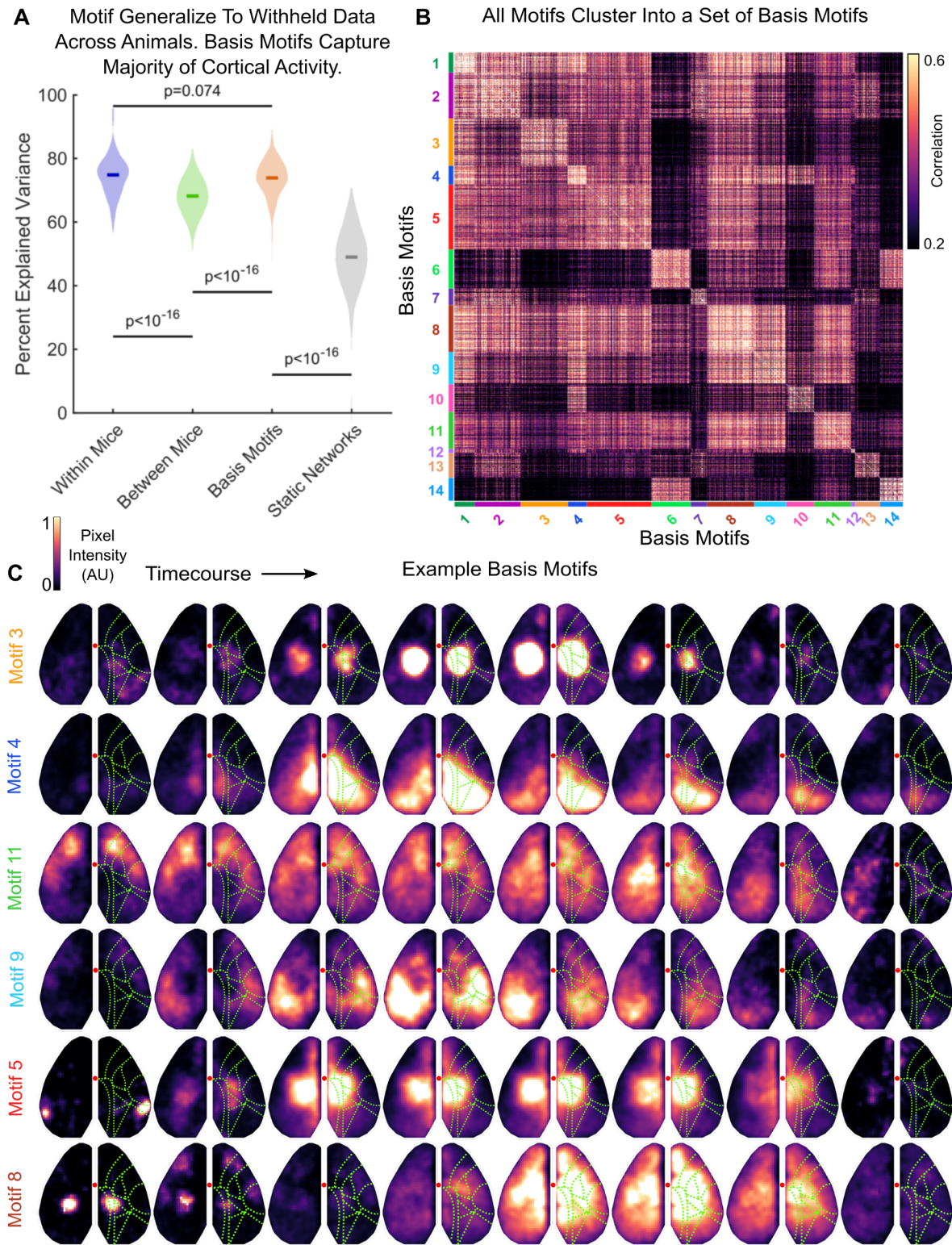
Figure 4. Motifs cluster into a low-dimensional set of basis motifs.

718   **(A)** Comparison of the percent of variance in neural activity explained by motifs from the same

719   mouse (within; purple), by motifs from other mice (between; green), by basis motifs (orange), and

720   by static network versions of basis motifs (gray). Static networks for each basis motifs were

721   derived as in Figure 3 (see Methods for details). All show fit to withheld data (N=144). Full

722   distribution shown; dark lines indicate median. Horizontal lines indicate pairwise comparisons. All

723   p-values estimated with Mann-Whitney U-test.

724   **(B)** Pairwise peak cross-correlation between all 2622 discovered motifs. Motifs are grouped by

725   their membership in basis motif clusters. Basis motif identity is indicated with color code along

726   axes. Group numbering (and thus the sorting of the correlation matrix) is determined by relative

727   variance explained by each basis motif (see Fig. 4 supplement 1A).

728   **(C)** Representative timepoints from example basis motifs. Display follows Fig. 1D, except without

729   direction of flow arrows.
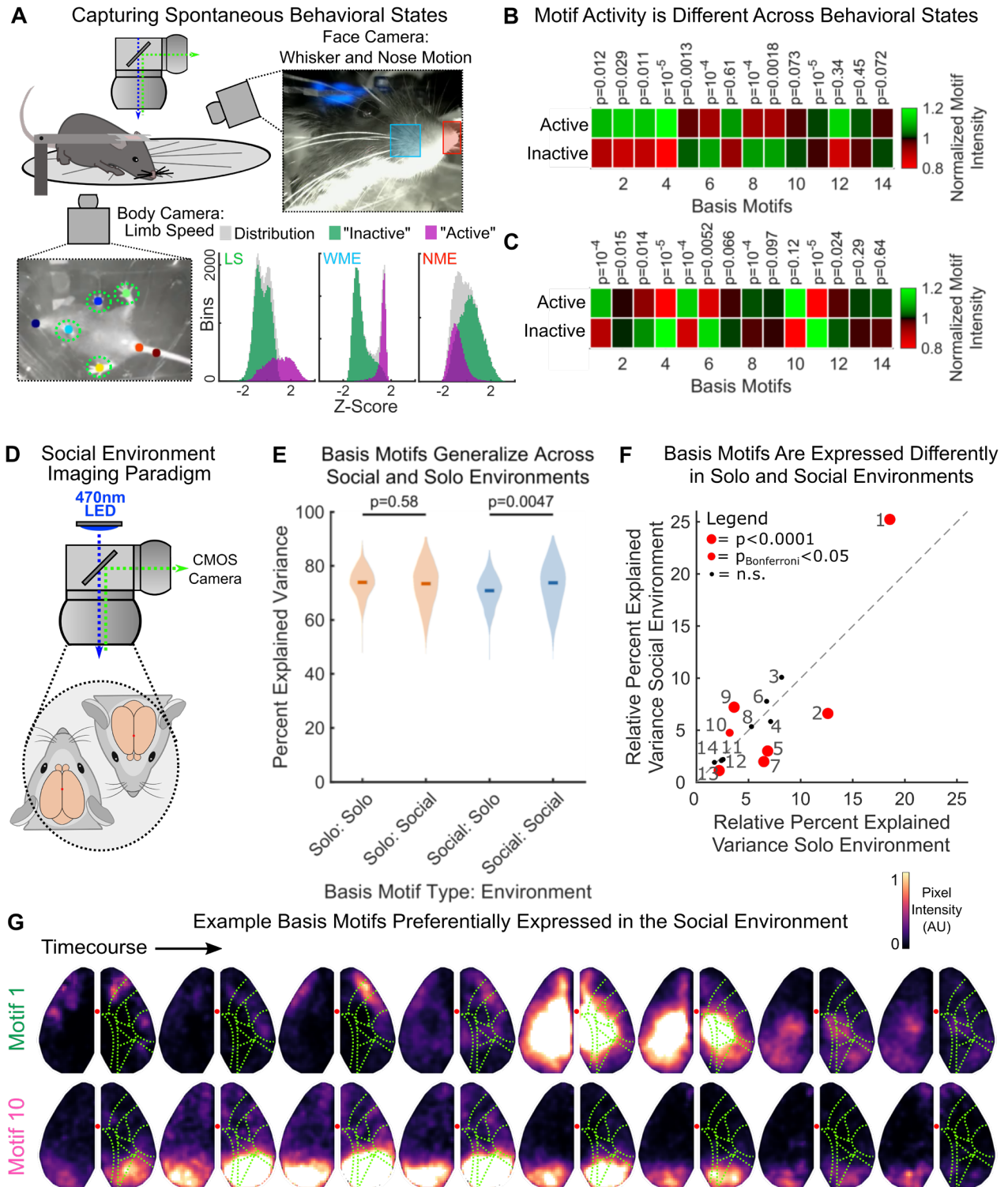
Figure 5. Basis motifs generalize to new animals, across behavioral states, and to social environments.

733   **(A)** Schematic of simultaneous mesoscale imaging and video capture of spontaneous behavior

734   of head-fixed mice on a transparent treadmill. Two new mice, not from the original cohort were

735   used. Blue and red squares indicate area used to measure whisker pad motion energy (WME)

736   and nose motion energy (NME), respectively. Colored dots indicate tracked position of the

737   forelimbs, nose, and tail. These were used to estimate limb speed (LS). Distribution of behavioral

738   variables are shown in three histograms along bottom. Gaussian mixture models fit to the

739   distributions of LS, WME, and NME simultaneously. Two states were discovered: an "active" and

740   "inactive" state (inset; purple and green respectively, see Methods for details).

741   **(B-C)** Motif activity evaluated across different behavioral states in mouse 1 (**B**) and mouse 2 (**C**).

742   Heatmaps show the mean activity of each motif (x-axis) during a behavioral state (y-axis)

743   normalized to the mean activity for that motif across all states. Motif activity was estimated as the

744   mean pixel value of a motif reconstruction over time (see Methods for details). P-values compare

745   motif activity across the two behavioral states. Note: behavioral states across mice are not directly

746   comparable since "active" and "inactive" states reflect a diversity of possible behaviors, as

747   confirmed by manual inspection of videos.

748   **(D)** Schematic of social environment imaging paradigm.

749   **(E)** Comparison of the percent of variance in neural activity that could be explained when animal

750   was alone (at rest, 'solo') or when paired with another animal ('social'). Basis motifs were

751   estimated in each setting and then fit to withheld data in both settings (as in Figure 4C; N=144

752   and N=123 withheld epochs for solo and social settings, respectively). Labels indicate

753   identification environment: fitting environment (e.g. 'Solo: Social' indicates basis motifs from solo

754   environment fit to withheld social data). Full distribution shown; dark lines indicate median.

755   **(F)**. Scatter plot of the relative PEV for each basis motif in the solo environment (x-axis; N=144

756   epochs) versus the social environment (y-axis; N=123 epochs). Motif labels are indicated with
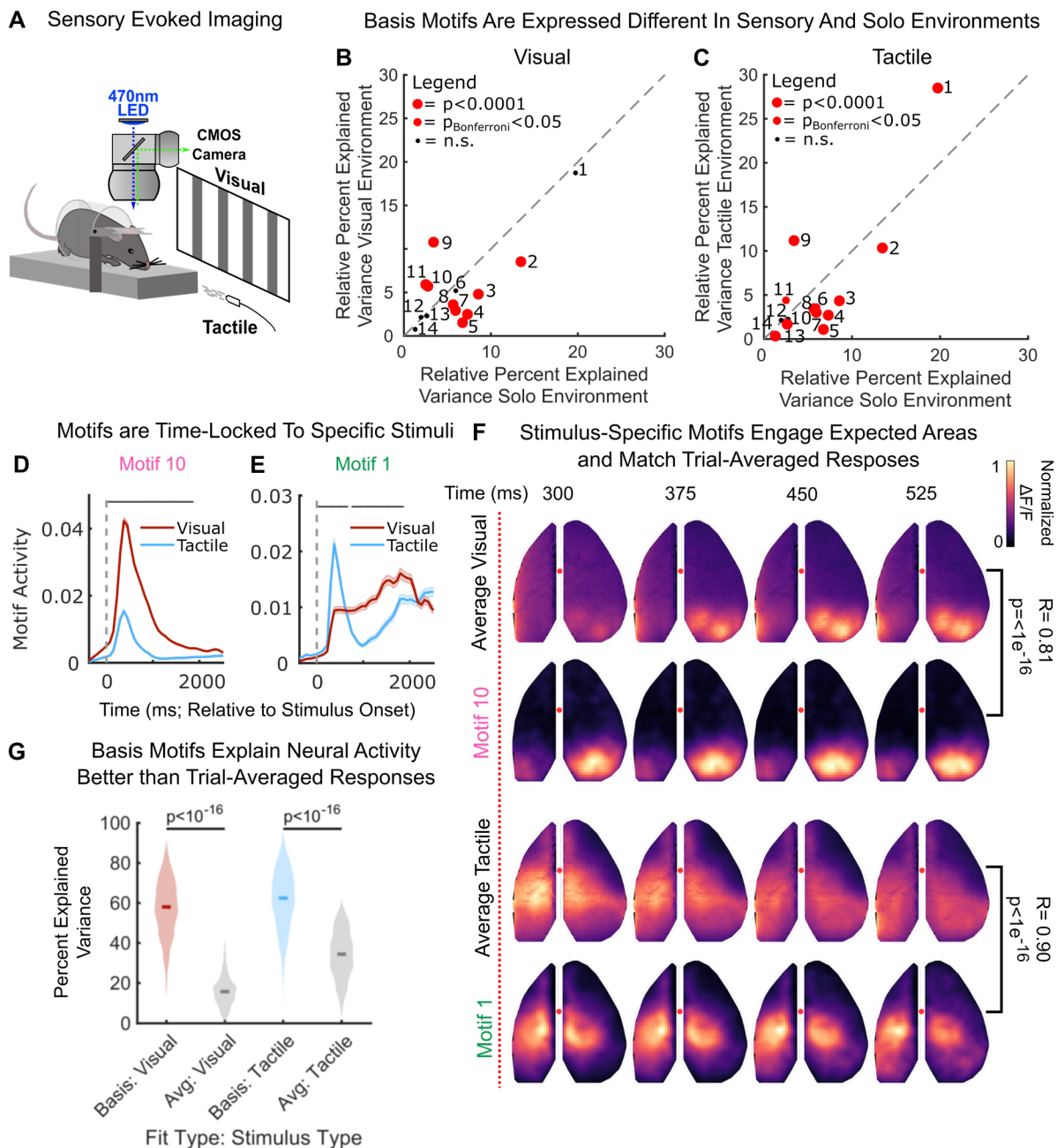
757   numbers, red markers indicate significant differences in expression rate between environments.

758   Identity line shown along diagonal.

759   **(G)** Example basis motifs preferentially expressed in the social environment. Display and motif

760   labels follow Figure 4.

761   All p-values estimated with Mann-Whitney U-test.

Figure 6. Specific basis motifs reflect processing of specific stimulus modalities

(A) Schematic of sensory stimulation paradigm. All stimuli were delivered on animals' left side.

(B-C) Scatter plot of the relative PEV for each basis motif in the solo environment (x-axis; N=144) versus the **(B)** visual or **(C)** tactile environment (y-axis; N=1109 and N=1110 visual and tactile

767   samples, respectively). Display follows Figure 5F, showing activation in solo environment at rest

768   (x-axis) versus activation in response to sensory stimuli. Significance computed with Mann-

769   Whitney U-test.

770   **(D-E)** Timecourse of **(D)** motif 10 and **(E)** motif 1 activity relative to stimulus onset (vertical

771   dotted line). Lines and shaded regions indicate mean +/- SEM motif activity in response to

772   visual (red; N=1109) and tactile (blue; N=1110) stimulation. Motif activity calculated as the mean

773   pixel value of a motif reconstruction over time (see Methods for details). Horizontal grey bar

774   indicates significant difference in motif activity between visual and tactile stimuli ($p_{Bonferroni} < 0.05$,

775   two-sample t-test; see Methods for details).

776   **(F)** Comparison of the trial-averaged stimulus-evoked response and the stimulus-evoked

777   response of the selective basis motifs in **B-C**. First two rows show responses for visual stimuli;

778   third and fourth rows for tactile stimuli. Correlation between average response and basis motif is

779   indicated along right side (pixelwise correlation, Pearson's ρ). As amplitude of response is

780   arbitrary, pixel intensities were normalized from 0-to-1 before correlation.

781   **(G)** Basis motifs explain more of the variance in neural activity than the average stimulus

782   response. Distributions show the percent of variance in neural activity during the 5 s after stimulus

783   explained by basis motifs (colors) or average responses (gray). Full distribution shown; dark lines

784   indicate median. Significance computed with Wilcoxon Signed-Rank test.

785
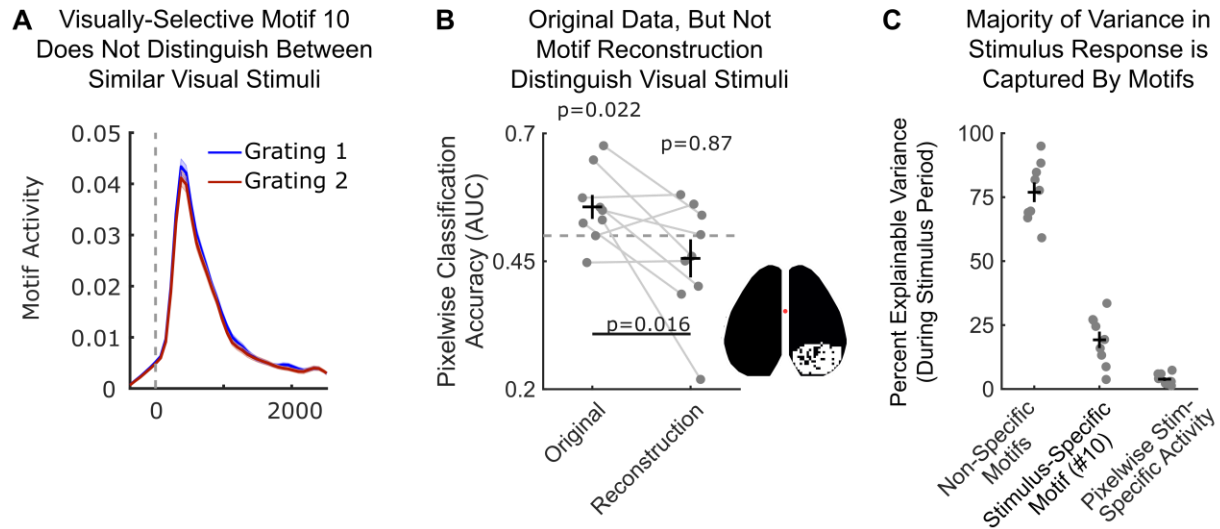
Figure 7. Basis motif 10 reflects general visual stimulus processing.

(A) Timecourse of motif 10 intensity relative to onset (vertical dotted line) of visual grating 1 (blue; N=554) or visual grating 2 (red; N=555). Display follows Figure 6D-E. No significant differences were observed between stimuli at any timepoint. p>0.11 for all timepoints; two-sample t-test.

(B) Visual stimuli can be decoded from neural response but not motif response. Classification was done using a support vector machine (SVM) classifier (see Methods for details) and accuracy is shown for withheld validation trials. Markers indicate classifier performance (measured with AUC) for each animal (N=9). Classifiers were trained on either raw pixel values (left column) or reconstructed motif response (right column). Inset shows pixels used for classification (see Methods for details). Dotted line denotes chance (AUC=0.5). Significance computed with one-sample t-test.

(C) Percent explainable variance in neural activity in response to visual stimuli captured by motifs and stimulus-specific residuals. Stimulus specific residuals are the trial-averaged residuals of motif reconstructions to each visual stimulus type. Data points correspond to mice (N=9). Black bars denote mean (horizontal) and SEM (vertical).

43

802   **Material and Methods**

803   **Key Resources Table**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Mouse: Thy1-GCaMP6f: *C57BL/6J-Tg(Thy1-GCaMP6f)GP5.3Dkim/J*	The Jackson Laboratory	#028280
Software and Algorithms		
MATLAB 2017a-2018b	Mathworks	N/A
seqNMF MATLAB Toolbox	Mackevicius et al., 2019	N/A
MUPET MATLAB Toolbox	Van Segbroeck et al., 2017	N/A
OFAMM MATLAB Toolbox	Afrashteh et al., 2017	N/A
Psychtoolbox-3 MATLAB Toolbox	http://psychtoolbox.org/ (Brainard, 1997)	N/A
Python version 3.6.4	Python Software Foundation	N/A
DeepLabCut Python Library	Mathis et al., 2018	N/A
Signal Processing and Analysis code for Widefield Imaging data	This paper	N/A

804

805   **Lead Contact and Material Availability**

806   Further information and request for resources and reagents should be addressed to Lead Contact,

807   Timothy J. Buschman (tbuschma@princeton.edu)

808   **Data and Software Availability**

809   Preprocessed data is available on Dryad data repository as image stacks (saved in Matlab file

810   format; DOI: 10.5061/dryad.kkwh70s1v; url: https://datadryad.org/stash/share/-q39l6jEbN--

811   voeSe2-5Y3Z1pfbeYEFLO-Kf_f-cjpE). The data has been preprocessed as described below

812   (spatially binned, masked, filtered, and then thresholded). Due to file size constraints, the full

813   raw data is not available on the Dryad repository but is available upon request. Example data

814   and figure generation code will be available on GitHub (https://github.com/buschman-lab) upon

815   acceptance.

816   **Experimental Model**

817 All experiments and procedures were carried out in accordance with the standards of the Animal

818 Care and Use Committee (IACUC) of Princeton University and the National Institutes of Health.

819 All mice were ~6-8 weeks of age at the start of experiments. Mice (N=11) were group housed

820 prior to surgery and single housed post-surgery on a reverse 12-hr light cycle. All experiments

821 were performed during the dark period, typically between 12:00 and 18:00. Animals received

822 standard rodent diets and water *ad libitum*. Both female (N=5) and male (N=6) mice were used.

823 All mice were C57BL/6J-Tg(Thy1-GCaMP6f)GP5.3Dkim/J (The Jackson Laboratory; Dana et al.,

824 2014). 9 mice were used for solo (rest) and sensory environment widefield imaging experiments.

825 These mice were control animals from a larger study. In that context, these animals were the

826 offspring of female mice that received a single intraperitoneal injection (0.6-0.66mL, depending

827 on animal weight) of sterile saline while pregnant. A subset of these mice (N=7) were used for

828 social imaging experiments. Separate mice (N=2) were used for spontaneous behavioral state

829 and hemodynamic correction experiments (these mice did not receive *in utero* exposure to saline).

830 **Surgical Procedures**

831 Surgical procedures closely followed Guo et al. (2014). Mice were anesthetized with isoflurane

832 (induction ~2.5%; maintenance ~1%). Buprenorphine (0.1mg/kg), Meloxicam (1mg/kg), and

833 sterile saline (0.01mL/g) were administered at the start of surgery. Anesthesia depth was

834 confirmed by toe pinch. Hair was removed from the dorsal scalp (Wahl, Series 8655 Hair

835 Trimmer), the area was disinfected with 3 repeat applications of betadine and 70% isopropanol,

836 and the skin removed. Periosteum was removed and the skull was dried. A thin, even layer of

837 clear dental acrylic was applied to the exposed bone and let dry for ~15 minutes (C&B Metabond

838 Quick Cement System). Acrylic was polished until even and translucent using a rotary tool

839 (Dremel, Series 7700) with rubber acrylic polishing tip (Shofu, part #0321). A custom titanium

840 headplate with a 11mm trapezoidal window was cemented to the skull with dental acrylic (C&B

841 Metabond). After the cement was fixed (~15 minutes), a thin layer of clear nail polish (Electron

45

842   Microscopy Sciences, part #72180) was applied to the translucent skull window and allowed to

843   dry (~10 minutes). A custom acrylic cover screwed to the headplate protected the translucent

844   skull after surgery and between imaging sessions. After surgery, mice were placed in a clean

845   home cage to recover. Mice were administered Meloxicam (1mg/kg) 24 hours post-surgery and

846   single housed for the duration of the study.

847   **Widefield Imaging**

848   Imaging took place in a quiet, dark, dedicated imaging room. For all experiments except

849   spontaneous behavioral monitoring (detailed below), mice were head-fixed in a 1.5 inch diameter

850   x 4 inch long polycarbonate tube (Fig. 1A) and placed under a custom-built fluorescence

851   macroscope consisting of back-to-back 50 mm objective lens (Leica, 0.63x and 1x magnification),

852   separated by a 495nm dichroic mirror (Semrock Inc, FF495-Di03-50x70). Excitation light (470nm,

853    $0.4mW/mm^2$) was delivered through the objective lens from an LED (Luxeon, 470nm Rebel LED,

854   part #SP-03-B4) with a 470/22 clean-up bandpass filter (Semrock, FF01-470/22-25).

855   Fluorescence was captured in 75ms exposures (FPS = 13.3Hz) by an Optimos CMOS Camera

856   (Photometrics). Prior to imaging, the macroscope was focused ~500um below the dorsal cranium,

857   below surface blood vessels. Fluorescence activity was captured at 980x540 resolution

858   (~34um/pixel) when the animal was imaged alone.

859   Images were captured using Micro-Manager software (version 1.4, Edelstein et al., 2014) on a

860   dedicated imaging computer (Microsoft, Windows 7). Image capture was triggered by an analog

861   voltage signal from a separate timing acquisition computer. Custom MATLAB (Mathworks) code

862   controlled stimulus delivery, recorded gross animal movement via a piezo sensor (SparkFun, part

863   #09197) attached to the animal holding tube, and captured camera exposure timing through a

864   DAQ card (National Instruments, PCIe-6323 X Series, part #7481045-01). Timing of all camera

865   exposures, triggers, behavioral measures, and stimulus delivery were captured for post-hoc

866   timing validation. No frames were dropped across any imaging experiments. A camera allowed

867   remote animal monitoring for signs of distress.

868   For recordings of spontaneous cortical activity, mice were head-fixed in the imaging rig and

869   habituated for 5 minutes. After habituation, cortical activity was recorded for 12 consecutive

870   minutes and stored as 3, 4-minute stacks of TIFF images. Qualitative real-time assessment of

871   behavioral videos and post-hoc analysis of activity (captured by piezo sensor) revealed minimal

872   episodes of extensive motor activity (e.g. struggling) during imaging. As our goal is to capture all

873   behavioral states, we did not exclude these moments from our analysis. Instead, motifs captured

874   these events alongside other cortical events.

875   **Widefield Imaging: Spontaneous Behavioral State Monitoring**

876   Widefield imaging was performed as above with minor modifications. Mice were head-fixed on a

877   custom transparent acrylic treadmill and illuminated with infrared light (Univivi 850nm IR

878   Illuminator). During imaging, behavioral measures were captured using two cameras: a PS3 EYE

879   webcam (640x480 pixel resolution) focused on the animal's whole body and a GearHead webcam

880   (320x240 pixel resolution) focused on the animal's face. Custom python (v3.6.4) scripts

881   synchronized the frame exposure of the behavioral cameras at 60Hz.

882   **Widefield Imaging: Paired Social Environment**

883   The macroscope objectives from the above widefield imaging paradigm were replaced with 0.63x

884   and 1.6x magnification back-to-back objectives, permitting an ~30x20mm field of view (lens order:

885   mouse, 0.63x, 1.6x, CMOS camera). Images were acquired at 1960 x 1080 resolution

886   (~34um/pixel). Animals were precisely positioned to be the same distance from the objective.

887   Mice faced one another, approximately eye-to-eye in the anterior-posterior axis. Their snouts

888   were separated along the medial-lateral axis by a 5-7 mm gap; close enough to permit whisking

889   and social contact but prevent adverse physical interactions. A 1mm plexiglass divider at snout

47

890   level ensured no paw/limb contact. Mice were positioned in individual plexiglass tubes. Pairs were

891   imaged together for 12 consecutive minutes, once each recording day. Some pairings included

892   mice outside this study cohort. 76 recordings from the experimental cohort were collected. After

893   each recording, the imaging apparatus was thoroughly cleaned with ethanol and dried before

894   imaging of the next pair (removing olfactory cues).

895   Animal pairs were provided with sensory stimuli consisting of playback of pre-recorded,

896   naturalistic ultrasonic vocalizations (USVs) between adult mice, synthetic USVs, or 'background'

897   noise. Naturalistic USV stimuli were obtained from the mouseTube database (Torquet et al.,

898   2016). In particular, we used four recordings of 3 min interactions between male and estrus female

899   wildtype C57BL/6J mice (files S2-4-4, S2-4-105, S2-4-123, S2-4-138). Details on the methods

900   used to record these interactions are described in the original study by Schmeisser et al. (2012).

901   To produce more salient stimuli, we reduced these 3 min recordings into 1 min recordings by

902   using Praat software (version 6.0.23) to shorten the silent periods between USV bouts. We

903   bandpass filtered these recordings to the 40-100 kHz range (Hann filter with 100 Hz smoothing)

904   to reduce extraneous background noise, and down sampled the recordings to 200 kHz.

905   Synthetic USV stimuli were generated using a customized MATLAB script that created artificial

906   sine wave tones matching the spectro-temporal properties of naturalistic stimuli. Specifically,

907   synthetic stimuli had the same rate (calls per minute), average duration, and mean frequency as

908   naturalistic USVs (we used MUPET to characterize USV properties; Van Segbroeck et al., 2017).

909   Tones were evenly spaced throughout the synthetic stimulus. Background noise was generated

910   from the silent periods of the 3-minute vocalization recordings. Each recording session contained

911   1 epoch of naturalistic USVs, 1 epoch of synthetic USVs, and 1 epoch with background noise. All

912   acoustic stimuli were presented at ~70dB through a MF1-S speaker (Tucker Davis Technologies)

913   placed 10cm away from both subjects.

914   **Widefield imaging: Structured Sensory Environments**

915   Widefield imaging was performed as in the original (solo) condition. All stimuli were provided to

916   the animals' left side. Recordings were 15-minutes long, divided into 90 trials of 8000ms duration.

917   Trials were structured with a 3000ms baseline period, 2000ms stimulus period, and 3000ms post-

918   stimulus period. Trials were separated by an inter-trial interval randomly drawn between 1000 and

919   1750ms.

920   Air puffs (10psi) were gated by solenoids (NResearch, Solenoid valve, part #161K011) and were

921   directed at the whisker pad in either the anterior-to-posterior or posterior-to-anterior direction.

922   Visual stimuli were gratings of 2.1cm bar width, 100% contrast, delivered on a 10inch monitor

923   (Eyoyo, 10-inch 1920x1200 IPS LED), positioned 14cm away from animals' left eye. Gratings

924   drifted from medial-to-lateral or lateral-to-medial at 8 cycles per second. Visual stimuli were

925   presented for 2000ms. During these recordings, mice also received trials of auditory stimuli (e.g.

926   2 tones). These data were not analyzed since auditory cortex was not imaged and so no evoked

927   response was observed. Each recording captured 30 trials of each stimulus modality. 3329 trials

928   were captured in total across 9 animals: resulting in 1110 tactile, 1109 visual (and 1110 auditory)

929   trials. 1 visual trial was lost due to timing issues.

930   Due to light-artifact of visual stimuli leaking through the ipsilateral cortical bone in a subset of

931   recordings, a more conservative mask on the ipsilateral hemisphere was used for all sensory

932   environment analyses (as shown in Fig. 6F). Accordingly, this mask was used for all analyses in

933   Figures 6-7, including quantification of motifs in the original (solo) environment.

934   **Statistical Analysis:**

935   All analyses were performed in MATLAB (Mathworks). Number of mice used was based on

936   previously published studies (Makino et al., 2017; Allen et al., 2017; Mohajerani et al., 2013). As

937   described throughout methods and main text, analyses were performed on 11 separate animals,

938   across multiple recording sessions, and 4 behavioral environments (e.g. biological replicates).

939   Analyses were validated across a range of processing parameters (e.g. technical replicates). All

940   statistical tests, significance values, and associated statistics are denoted in the main text. P-

941   values below machine precision are reported as $p<10^{-16}$. All 95% Confidence intervals were

942   computed using MATLAB *bootci* function (1000 bootstrap samples).

943   **Widefield Imaging Preprocessing**

944   Image stacks were cropped to a 540x540 pixel outline of the cortical window. Images were aligned

945   within and across recordings using user-drawn fiducials denoting the sagittal sinus midline and

946   bregma for each recording. For anatomical reference (Figs. 1A and supplement 1), recordings

947   were aligned to a 2D projection of the Allen Brain Atlas, version CCFv3 using bregma coordinates

948   (Oh et al., 2014; ABA API interfacing with MATLAB adapted from *https://github.com/Sainsbury*

949   *WellcomeCentre/AllenBrainAPI*). The complete 2D projection is shown in Fig. 1 Supplement 1.

950   As they are only intended to be local references, the parcel outlines overlaid in Figures 1, 4, 5

951   were created by manually tracing this 2D projection (Inkscape Vector Graphics Software).

952   Changes in fluorescence due to hemodynamic fluctuations may confound the neural activity

953   captured by widefield imaging (Allen et al., 2017; Ma et al., 2016a, 2016b). However, previous

954   work has found hemodynamic contributions to fluorescent signal using similar widefield imaging

955   approaches are minimal (Cramer et al., 2019; Murphy et al., 2016; Vanni and Murphy, 2014) ,

956   and can be mitigated by removing pixels corresponding to vasculature (Makino et al., 2017). To

957   mitigate impact of hemodynamic contributions, we masked pixels corresponding to vasculature.

958   To identify vasculature, the middle image of each recording was smoothed with a 2D median filter

959   (neighborhood 125 pixels^2) and subtracted from the raw image. As vasculature pixels are much

960   darker than pixels of neural tissue, we created a vasculature mask by thresholding the reference

961   image to pixels intensities >= 2.5 standard deviations below the mean. To remove noise, the mask

962   was morphologically closed with a 2-pixel disk structuring element. A vasculature mask was

963　created for each recording. Supplemental experiments, outlined below, demonstrated that

964　vascular masks successfully mitigated the contribution of hemodynamics to our signal.

965　Vasculature masks were combined with a manually drawn outline of the optically accessible

966　cortical surface and applied to each recording to conservatively mask non-neural pixels. Masks

967　removed the sagittal sinus, mitigating vascular dilation artifacts. Additionally, masks removed

968　peripheral lateral regions, such as dorsal auditory cortex, where fluorescence contributions across

969　animals may be differentially influenced by individual skull curvature. After alignment and

970　registration, recordings were spatially binned to 135x135 pixels (~68µm^2/pixel). Masked pixels

971　were ignored for spatial binning. Normalized activity was computed as change in fluorescence,

972　e.g. ΔF/F over time according to $\frac{F-F_0}{F_0}$. Baseline fluorescence, F_0 was computed using a 9750ms

973　(130 timepoints) rolling mean. To remove slow fluctuation in signal (e.g. due to change in

974　excitation intensity), pixels traces were detrended using linear least squares fit.

975　Recordings were bandpass filtered at 0.1 to 4Hz (10th order Butterworth filter). Pixel traces were

976　thresholded at 2 standard deviations per pixel in order to remove noise in spontaneous activity.

977　Values below this threshold were set to zero. Thresholding had minimal impact on subsequent

978　analyses and conclusions. For example, similar numbers of basis motifs (~14) were discovered

979　in non-thresholded data. After filtering and thresholding, recordings were spatially binned again

980　to a final size of 68x68 pixels (~136µm^2/pixel). Pixels with zero variance during an epoch (e.g.

981　masked pixels) were ignored for all subsequent analyses. Subsequent factorizations require non-

982　negative pixel values so recordings where normalized to range of 0 to 1 using the maximum and

983　minimum pixel values per recording. The 12-minute solo and social recordings were divided into

984　six, 2-minute epochs; alternating epochs were used for motif discovery or withheld for testing (Fig.

985　1B). For solo recordings this resulted in 144 'discovery' and 144 'withheld' epochs.

986   For social recordings, cortices of individual animals were cropped to 540x540 pixels and

987   preprocessing followed as above. Again, recordings were divided into 2-minute epochs, resulting

988   in a total of 228 'discovery' and 'withheld' epochs. Given the proximity of the animals, whiskers

989   from one animal sometimes entered the imaging field of view of the paired animal, creating

990   artifacts easily detected upon manual inspection. All epochs were manually inspected and epochs

991   with any whisker artifacts (N=105) were removed, resulting in 123 'discovery' and 'withheld'

992   epochs (8.2 hours in total).

993   For sensory trials, which were 8 seconds in length, each trial's ΔF/F was calculated using the

994   mean of the first 26 timepoints (~2s) as baseline fluorescence. Burst in activity were discovered

995   by thresholding traces at 1 standard deviation per pixel. All other preprocessing steps were

996   followed as above.

997   **Multiwavelength Hemodynamic Correction**

998   Additional experiments using multiwavelength hemodynamic correction were performed to

999   confirm that vasculature masking mitigated hemodynamic contributions to motifs (Fig. 4

1000   supplement 2). Hemodynamic correction followed (Musall et al., 2019). In brief, widefield imaging

1001   was performed while strobing between illumination with a blue LED (470nm, $0.4mW/mm^2$) and

1002   violet LED (410nm, LuxDrive LED, part #A008-UV400-65 with a 405/10 clean-up bandpass filter;

1003   Edmund Optics part #65-678). Each exposure was 35.5ms and light from both LEDs were

1004   collimated and coupled to the same excitation path using a 425nm dichroic (Thorlabs part

1005   #DMLP425). Illumination wavelengths alternated each frame. Strobing was controlled using an

1006   Arduino Due with custom MOSFET circuits coupled to frame exposure of the macroscope (as in

1007   Pinto et al., 2019). After vasculature masking and spatial-binning to 135x135 pixels, violet-

1008   exposed frames (e.g. non-calcium dependent GCaMP6f fluorescence, Lerner et al., 2015) were

1009   rescaled to match the intensity of blue-exposed frames. ΔF/F was then computed as $\Delta F/F_{blue}$ -

1010 $\Delta F/F_{violet}$. Remaining preprocessing steps followed original (solo) experiments.

Motif Discovery

1012 We used the *seqNMF* algorithm (MATLAB toolbox from Mackevicius et al., 2019) to discover

1013 spatio-temporal sequences in widefield imaging data. This method employs convolutional non-

1014 negative matrix factorization (CNMF) with a penalty term to facilitate discovery of repeating

1015 sequences. All equations below are reproduced from the main text and Tables 1 and 2 of

1016 Mackevicius et al (2019). For interpretability, we maintained the nomenclature of the original

1017 paper where possible.

1018 We consider a given image as a *P x 1* vector of pixel values and a recording image sequence (i.e.

1019 recording epoch) as a *P x T* matrix, where *T* is the number of timepoints in the recording. This

1020 matrix can be factorized into a set of *K* smaller matrices of size *P x L* representing short sequences

1021 of events (e.g. *motifs)*. Collectively this set of motifs is termed *W* (a *P x K x L* tensor).

1022 Each pattern is expressed over time according to a *K x T* temporal weighting matrix termed *H.*

1023 Thus, the original data matrix can be approximated as the sum of K convolutions between the

1024 motifs in *W* and their corresponding temporal weightings in *H*:

1026
$$X_{pt} \approx \tilde{X}_{pt} = \sum_{k=1}^{K} \sum_{\ell=0}^{L-1} W_{pk\ell} H_{k(t-\ell)} \equiv (W \circledast H)_{pt} \quad equation\ 1$$

1025

1027 Here, $\circledast$ indicates the convolution operator. The values of W and H were found iteratively using

1028 a multiplicative update algorithm. The number of timepoints for each motif was set to 13 frames

1029 (975ms). This *L* value was chosen because it is well above the duration of GCaMP6f event

1030 kinetics and qualitative assessment of imaging recordings suggested most spontaneous events

1031 were < 1000ms in duration, agreeing with previous literature (Stringer et al., 2019a). Active

1032 timepoints in motifs may be shorter than *L*; in which case unused timepoints are zero-padded.

1033 Preliminary tests at *L > 13* produced similar results to *L*=13, and explained variance captured by

53

1034   motifs reconstructions plateaued at *L* values >=13 (Fig. 1 supplement 2C). As increasing *L*

1035   increased computation time, *L*=13 was used for practical purposes.

1036

1037   The maximum number of possible motifs (*K*) per recording was chosen by iteratively sweeping a

1038   range of *K* values and evaluating the number of basis motif clusters at each value (see below

1039   regarding basis motif clustering). Figure 4 supplement 1B shows the number of basis motifs

1040   asymptotes at *K* >= *12*. K = 28 was used for all motif discovery experiments. *K* = 28 was chosen

1041   as it was well above the elbow of the asymptote and produced a total number of basis motifs on

1042   the upper range of the test *K* values (thus not analytically limiting the observe dimensionality). As

1043   with *L,* if fewer than *K* motifs were found, the remaining motifs would be populated with zeros.

1044   Again, we constrained *K* because computational demands when adding extra, unnecessary

1045   (blank) motifs.

1046   The *seqNMF* algorithm improves upon typical CNMF by including a spatio-temporal penalty term

1047   into the cost function of the multiplicative update algorithm. In brief, this reduces redundancy

1048   between motifs: 1) multiple motifs do not describe the same sequence of activity; 2) a single motif

1049   is not temporally split into separate motifs; and 3) motifs are encouraged to be non-overlapping

1050   in time. This penalty termed is implemented as follows:

1051
$$R = \lambda \left\| \left(W \overset{T}{\circledast} X \right) SH^T \right\|_{1, i \neq j}$$

1052   Here, temporal overlap (correlation) between motifs is captured by SH^T. *S* is a *T x T* temporal

1053   smoothing matrix where $S_{ij} = 1$ when $|i - j| < L$; otherwise $S_{ij} = 0$. Thus, each temporal

1054   weighting in *H* is smoothed by a square window of length 2*L*-1, increasing the product of motifs

1055   that temporally overlap within that window.

1056   Competition between spatio-temporal structure of motifs is achieved by calculating the overlap of

1057   motifs in *W* with the original data as follows

54

1058
$$(W \overset{T}{\circledast} X)_{kt} = \sum_{\ell} \sum_{p} W_{pk\ell} X_{p(t+\ell)} \quad equation\ 3$$

1059 Motifs containing similar patterns will overlap with the original data matrix at the same times. The

1060 notation $||\cdot||_{1,i \neq j}$ ignores penalizing terms along the diagonal such that spatial and temporal

1061 autocorrelation of motifs is not penalized. λ is a tunable parameter that controls the magnitude of

1062 the penalty term. The result of this penalty, when implemented as a cost function in the

1063 multiplicative update algorithm for fitting H and W, is to bias factorization such that only one motif

1064 is active at a given timepoint.

1065 We followed the approach used by Mackevicius et al., to determine the magnitude of the λ penalty

1066 term. λ was swept across 6 orders of magnitude from 10^{-6} to 1 (Fig. 1 supplement 2A). The impact

1067 on reconstruction cost, motif spatio-temporal correlation, explained variance, and number of

1068 factors was evaluated. A value slightly above the cross-over point in reconstruction cost and

1069 explained variance was chosen for subsequent experiments (Fig. 1 supplement 2A). Importantly,

1070 a similar number of motifs were found for all values of λ, suggesting our results did not depend

1071 on the exact value.

1072 Additionally, the *seqNMF* algorithm contains optional orthogonality constraints to bias

1073 factorizations towards parts-based and events-based factorizations. In widefield imaging data,

1074 parts-based would preferentially detect spatially independent motifs. Events-based factorizations,

1075 used in this study, would preferentially discover temporally independent motifs that correspond to

1076 specific instantiations of sequential activity patterns. This was achieved with an additional

1077 smoothed orthogonality cost term penalizing overlap in motif temporal weightings:

1078
$$R_{Hortho} = \frac{\lambda_{Hortho}}{2} ||HSH^{T}||_{1,i \neq j} \quad equation\ 4$$

1079 The magnitude of λ_{orthoH} was set to 1 based on preliminary experiments and remained

1080 unchanged for all motif discovery experiments, including the parameter sweeps of λ and K

55

1081   described above. All fitting processes were run for 300 iterations; at which point the cost function

1082   leveled out (Fig. 1 supplement 2B).

1083   Supplemental Table 1 contains a complete list of all adjustable parameter values of the seqNMF

1084   algorithm toolbox used for each experiment. Descriptions for parameters not discussed above

1085   can be found in original work. To discover motifs in paired social recordings, the same K and L

1086   values (28 and 13) were used. λ was refit following the procedure described above.

1087   Note, one change was made to the *seqNMF* algorithm toolbox. For convenience, the smoothing

1088   matrix S was multiplied by 0.01. Specifically, line 98 of original seqNMF code: *smoothkernel =*

1089   *zeros(1,(2*L)-1)* replaced with *smoothkernel = 0.01*zeros(1,(2*L)-1).* This allowed λ values to be

1090   100x larger and, therefore, easier to read.

Visualizing Motif Activity Flow

1092   To aid in visualization of motif patterns (Figs. 1 and 3), we used the Horn-Schunck optical flow

1093   method to calculate the velocity vector fields between subsequent timepoints of motifs (Horn and

1094   Schunck, 1981; implemented using *HS* function from OFAMM MATLAB toolbox, Afrashteh et al.,

1095   2017). For basis motif images, optical flow was only performed for timepoints with non-zero

1096   variance across pixels. For all plots, arrows depict the direction and velocity of flow for the top

1097   50% intensity pixels of each timepoint. The number of arrows were downsampled by a factor of

1098   three for clarity in visualization.

Comparing Motifs to sPCA and sNMF

1100   Data preprocessing for Motif, sPCA, and sNMF discovery was identical. For sPCA, pixels were

1101   treated as variables and timepoints as observations. sNMF was performed by setting L to 1

1102   (effectively reducing *equation 1* to matrix multiplication) and removing all temporal and spatial

1103   sparsity terms.

56

Refitting Motifs to Withheld Data

To refit motifs to withheld data, we used the same *seqNMF* algorithm, but *W* was fixed to the previously discovered motifs (or basis motifs) during fitting. Thus, the only updatable features of the factorization were motif temporal weightings, *H*. See Supplemental Table 1 for a list of all parameters used during fitting. Importantly, since only temporal weightings were refit, λ was set to zero. Thus, unlike motif discovery where motifs were highly spatially and temporally independent, refit motifs did not have this constraint and thus some combinatorial motif activation was observed. However, increasing λ_{Hortho} had little impact on the percent of variance explained when refitting the motifs, suggesting that any compositionality had minimal impact on results (Fig. 3 supplement 1). Importantly, all motifs, static networks, and time-varying networks were refit to withheld data in the same way and therefore their percent explained variance can be directly compared.

For solo and social recordings, the part-based factorization orthogonality bias was maintained as in initial discovery (see Supplemental Table 1). Given the short duration of sensory trials, the orthogonality bias was not used. The same process was followed for refitting basis motifs and social basis motifs to withheld data. For all refitting processes, motifs and withheld data were first spatially smoothed with a 2D gaussian filter ($\sigma = [1,1]$). Only pixels with non-zero variance in both motifs and withheld data were used.

Generating Basis Motifs

To identify basis motifs, we used an unsupervised clustering algorithm (Phenograph, Nicosia et al., 2009; Levine et al., 2015). For clustering, motifs were renormalized to 0-to-1 and spatiotemporally smoothed with a 3D gaussian filter ($\sigma = [1,1,0.1]$). The Phenograph algorithm generates a directed graph, where each node is connected to its *k* nearest neighbors. Louvain community detection is then performed on this graph to cluster nodes into groups. For finding

57

1128   neighbors, distances between motifs were computed as the peak in their temporal cross-

1129   correlation. The only tunable parameter of Phenograph is the number of nearest neighbors (*k*).

1130   *k*=15 was chosen based on initial experiments. Similar number of clusters (10-20) were observed

1131   for *k*=10 and *k*=20.

1132   Basis motifs were generated by taking the mean of the core community of motifs in each cluster.

1133   The core community was defined as the top 10% of motifs in each cluster with the most within-

1134   cluster nearest neighbors. Prior to averaging, motifs were aligned to a 'template' motif. The

1135   template motif shared the most zero-lag peak temporal cross-correlations with all other motifs. If

1136   there were multiple templates, one was chosen at random. All motifs were zero-padded to a

1137   length of *3L* (39 timepoints) and aligned to these templates by their maximal cross-correlation

1138   lag and then basis motifs calculated from the core communities. Basis motifs were then aligned

1139   to one another by shifting the center of mass of activity to the middle timepoint. Timepoints with

1140   no variance across all basis motifs were removed, resulting in basis motifs that were 26

1141   timepoints (~2s) long.

1142   **Percent Explained Variance Calculations**

1143   For all experiments, the percent explained variance (PEV) in neural activity was defined as

1144 $$PEV \ = \ 100 * \left(1 - \frac{(\sigma_X{}^2 - \sigma_{\tilde{X}}{}^2)}{\sigma_X{}^2} \right)$$

1145   Where $\sigma_X{}^2$ and $\sigma_{\tilde{X}}{}^2$ denote the spatio-temporal variance of the original data and reconstructed

1146   data respectively.

1147   The PEV of individual motifs was calculated by convolving the motif with its temporal weighting

1148   and computing PEV as above (see *helper.reconstruct* function from *seqNMF* toolbox). Thus, the

1149   PEV of individual motifs reflected both the frequency of motif occurrence and the quality of fit to

1150   the data. Relative PEVs of individual motifs were calculated by dividing the PEV of each motif by

1151   the total PEV across all motifs for that epoch.

1152   Timepoint-wise PEV (Fig. 3 supplement 1B) used the same calculation as above but was

1153   performed separately on each timepoint of an epoch. Thus, these analyses reflect solely the

1154   spatial variance in activity captured by reconstructed data for each individual timepoint.

1155   **Cross-Temporal Autocorrelation Analysis**

1156   The cross-temporal autocorrelation of each motif was calculated by computing the spatial

1157   correlation between frames of the motif at vary temporal lags. The resulting autocorrelation was

1158   then fit with an exponential to estimate the half-life of decay in the autocorrelation (τ). Pixels with

1159   no variance across motif timepoints were ignored when calculating correlation.

1160   **Static Networks**

1161   Static networks were generated for each motif by replacing all of the active timepoints of a motif

1162   with the mean activation of that motif (see Fig. 3E for example). Active timepoints were defined

1163   as any timepoints with variance across pixels greater than zero. Thus, static networks represent

1164   a constant 'state' of activation of the same brain regions for the duration of that motif. The temporal

1165   weightings of these static networks were refit to the withheld data the same way as dynamic motifs

1166   (i.e. the activation of these states could vary throughout an epoch; described above).

1167   **Estimating the Working Resolution of Widefield Imaging**

1168   One concern is that the spatial resolution of our approach may have artificially limited the

1169   dimensionality of the observed motifs. Therefore, we sought to estimate the 'working resolution'

1170   of our approach by grouping pixels into functional clusters, defined as contiguous groups of pixels

1171   with correlated activity (Fig. 4 supplement 3).

1172 To identify functional clusters, we divided each recording epoch into 1 second time periods

1173 (17,280 total 1-second periods, each with 13 frames). A pixelwise correlation matrix was

1174 computed for each 1-second period. Next, PCA was applied to each correlation matrix, producing

1175 a (pixel x component) matrix of 'eigenconnectivities'. These eigenconnectivities reflected the

1176 dominant correlation patterns across pixels during that 1-second period (similar to approaches by

1177 Leonardi et al., 2013; Preti and Ville, 2017)

1178 The first eigenconnectivity from each 1-second time period were concatenated together, creating

1179 a (pixel x 17,280) matrix that captured the wide variety of different possible spatial patterns across

1180 the cortex. We then used Phenograph (k=15) to group pixels according to their correlation across

1181 these spatial patterns. This produced 37 'functional clusters' of highly correlated pixels (Fig. 4

1182 supplement 3; 18 in left and 19 in right hemisphere. Similar results (~20 clusters per hemisphere)

1183 were observed by performing PCA on each 1-second time period and then clustering as above

1184 (i.e. without first creating a correlation matrix; this creates an 'eigenimage' instead of an

1185 'eigenconnectivity' as in Friston, 2004)

1186 Several lines of evidence suggest our motifs were not constrained by the functional resolution of

1187 our approach. First, as shown in Figures 1, 3, and 4, most motifs are dynamic – different regions

1188 are activated over time. This suggests motifs are not limited by the spatial resolution of our

1189 approach, as the motifs capture the flow of activity across distinct spatial regions. Second, our

1190 imaging approach can still capture pixel-wise activity (Fig. 7B), suggesting the functional

1191 clustering is a lower-bound on our resolution. Even using this lower bound, the dimensionality of

1192 our motifs is far less than the possible dimensionality of our approach. For example, even if motifs

1193 engaged only 1-2 functional clusters, there are still 18^2=324 possible patterns.

1194 **Quantifying Spontaneous Behavioral States**

1195   Behavioral state was quantified using the binarized intensity of nose motion energy, whisker pad

1196   motion energy, and total limb speed. Nose and whisker pad motion energy was quantified as

1197   the mean absolute temporal derivative of pixels within a manually selected ROIs (shown on Fig.

1198   5A). The whisking speed of the mouse was faster than the frame rate of the behavioral videos.

1199   This caused the whiskers to be blurred when the animal was intensely whisking, resulting in a

1200   low value for the whisker energy measure. Therefore, whisking motion energy was inverted for

1201   all analyses (e.g. $M_{whisk} = \max(M_{whisk\ original}) - M_{whisk\ original}$) so that the axes were

1202   consistent across the three behavioral variables (i.e. higher values indicate higher energy).

1203   Markerless tracking of paw position (DeepLabCut; Mathis et al., 2018) was used to measure

1204   limb speed. Training and validation of DeepLabCut neural network followed Nath et al.'s, 2019

1205   published protocol, with the network trained to identify paw position as the center point between

1206   the 1st and 5th digit of each paw. To improve accuracy of limb position estimates, nose, tail base,

1207   and tail root were also tracked. 360 frames from 3 animals were used for network training. One

1208   refinement iteration was performed. The network was trained until loss plateaued (Fig. 5

1209   supplement 1C, 120000 iterations). Total limb speed was calculated as the summed absolute

1210   temporal derivatives of the x and y position of all four paws.

1211   To categorize behavioral state, a gaussian mixture model (GMM) was fit to the distributions of

1212   these three behavioral variables (MATLAB; *fitgmdist* function; 2 components). Timepoints were

1213   assigned to one of two behavioral states according to the 2-second running product of the

1214   posterior probability of being in each state. Any timepoints with a likelihood of being in either

1215   state that was less than 0.5 were excluded from analysis (this was less than 0.01% of all

1216   timepoints in both animals).

1217   To decode behavioral state based on motif activity, we used a support vector machine classifier

1218   with radial basis function kernels (MATLAB; *fitsvm* function; Sequential Minimal Optimization

1219   solver). Prior to training a classifier, 20% of each trial type was held out as a validation test data

1220   set. For each classifier, two hyperparameters, "box-constraint" (a regularization parameter to

1221   prevent overfitting) and "kernel scale" were tuned using cross validation within the training set

1222   data (5 folds, balanced trial types: using the MATLAB functions *fitsvm* and *cvpartition*).

1223   Hyperparameters were optimized using Bayesian optimization (MATLAB, *bayesopt* function).

1224   Tuned classifiers were then tested on withheld validation data. Classifier accuracy was

1225   quantified using the area under the curve (AUC) of the receiver operator characteristic function

1226   (MATLAB, *perfcurve* function). Results of this classification analysis are reported in the main

1227   text.

Fitting Average Traces to Sensory Trials

1228

1229   Average stimulus responses were calculated by taking the mean of all trials across animals for

1230   visual and tactile stimuli. Different stimuli within a modality were combined to generate the

1231   average trace (e.g. visual grating 1 and 2 were averaged together). Temporal weightings of

1232   average traces were refit to trials using the same seqNMF algorithm and parameters as when

1233   refitting basis motifs, except now the average stimulus response was used in place of the motifs.

1234   Additionally, prior to fitting, trials were zero-padded which allowed the algorithm to flexibility shift

1235   the average trace timing to best match the evoked response timing of each trial. This allowed us

1236   to directly compare the PEV of the motifs and the average stimulus response.

Comparing Stimulus Evoked Motif Responses

1237

1238   We sought to compare the motif responses evoked by different stimuli (Fig. 6D-E). Motif activity

1239   was estimated by convolving each motif with its temporal weight on each trial (as described above;

1240   see *seqNMF Toolbox helper.reconstruct* function). Stimulus selectivity of a motif was estimated

1241   by comparing the evoked responses to each stimulus (two-sample t-test). This was done over

1242   time, using 300ms windows, stepped every 150ms to give a timecourse of similarity. Resulting

1243   p-values were Bonferroni corrected for multiple comparisons across time and conditions (i.e. 24

1244   comparisons in Fig. 6D-E, 12 windows x 2 modalities).

1245   Pixelwise Classifier Analysis Comparing Visual Stimuli

1246   As described above, mice were presented with two different visual stimuli: visual stimulus 1

1247   (grating drifting from medial-to-lateral) and visual stimulus 2 (grating drifting from lateral-to-

1248   medial). Mice received 15 trials of each stimulus on 4-5 consecutive recording days (60-75 total

1249   of each type per mouse). As described in the main text, the response of the 'visual' motif (#10)

1250   did not differ between the two stimuli. To test whether this was due to a limitation in our imaging

1251   approach, we tested whether we could classify neural responses as belonging to the two stimuli

1252   (Fig. 7B)

1253   To classify stimulation response, we used a support vector machine classifier (see detailed

1254   description above). Classification was performed on pixels restricted to the right hemisphere

1255   (contralateral to stimulus) and in the top 95% intensity percentile of motif 10 (visually-evoked

1256   motif). This resulted in a total of 6006 possible feature pixels (231 pixels across the 26 timepoints

1257   of visual stimulus presentation). For each classifier, we performed a cross-validated ANOVA

1258   feature selection to select a subset of these pixels as classification features (Pereira et al., 2009).

1259   The p-value for a one-way ANOVA comparing responses to each stimulus type was computed

1260   for each pixel in the training data. The 50 pixels with the highest -log10(p-values) were used for

1261   classification. After determining the feature vector, the classification hyperparameters were tuned

1262   within the training data, as described above.

1263   For each mouse, pixelwise classification was performed on two datasets (Fig. 7B). First it was

1264   performed on the original data (after the preprocessing steps above). Second it was performed

1265   on the reconstructed activity of the motifs fitted to this original data. Thus, for both classification

1266   procedures, the data were the same spatial resolution. The only difference between the data sets

1267   were the residuals of the motif fitting procedure. Therefore, a total of 18 classifiers were trained

1268   (2 per animal).

Evaluating Motif Expression During Two Tactile Stimuli

1270   Animals were also presented with two different tactile stimuli; airpuffs that either traveled from

1271   anterior to posterior or from posterior to anterior. Motifs captured the vast majority of explainable

1272   variance in the response to tactile stimuli. In particular, Motif 1 (tactile-specific; Fig. 6E) captured

1273   a large portion of response to both stimuli (Fig. 6 supplement 1; 61.08% +/-1.97% SEM for non-

1274   stimulus specific motifs; 36.90% +/- 1.90% SEM for stimulus-specific motif 1).

1275   However, it is important to note that, unlike the visual stimuli, these two tactile stimuli were likely

1276   significantly different in their impact on behavior. First, the behavioral connotation of an anterior-

1277   approaching stimulus is likely different than that of a posterior stimulus. Second, although both

1278   airpuffs were directed at the whisker pad, differences in the solenoids used to control the airflow

1279   impacted flow rate and air pressure. Thus, one may expect these stimuli to evoke different

1280   intensity responses (and thus different motif temporal weightings). Consistent with this

1281   hypothesis, the relative percent variance in neural activity captured by motifs differed in response

1282   to the two airpuffs (Fig. 7 supplement 1A). Because of these large differences at the level of

1283   motifs, there was no need for to classify the tactile stimuli. However, stimulus-specific activity in

1284   the residuals still only captured a small part of the total explainable variance (2.02% +/- 0.38%

1285   SEM; Fig. 7 supplement 1B). Thus, as we saw for visual stimuli, most cortex-wide neural activity

1286   can be explained by the expression of motifs.

64

1290

1291

1292

1293

1294

Supplemental Information:

Supplemental Movie 1. Example Widefield Imaging of   Cortical Activity. Shows an example of widefield imaging (left) alongside a video of animal's behavior (right).

Supplemental Movie 2. Example Fit of Data Reconstructed from Discovered Motifs. (A) Temporal weightings of discovered motifs. **(B)** Original data. Color indicates fluorescence intensity of original data (i.e. $\hat{F}_O$: normalized between 0 to 1, using the 99.9th percentile of all pixels). **(C)** Reconstructed data. Created by convolving motifs with their temporal weightings (see Method for details). Color indicates intensity of reconstructed data ($\hat{F}_R$); color scale follows **B**. **(D)** Residual of fit (i.e. $\hat{F}_O$ - $\hat{F}_R$). Color scale indicates the quality of fit of reconstructed to original data. Red dots indicate bregma.

Supplemental Movie 3. All Basis Motifs. Basis motif display follows Figure 4C. For video visualization, all motifs were convolved with a 450ms gaussian temporal weighting vector to mimic a 'spike' in activity.

1308

1309

1310

1311

65

1312

1313

1314

1315

1316

1317

1318

1319

1320

1321

1322

1323

1324

1325

	Discovery: Solo	Fit: Solo	Fit: Paired Social	Fit: Sensory	Discovery: Paired Social	Fit: Paired Social
K	28	14 (# basis motifs)	14 (# basis motifs)	14 (# basis motifs)	28	11 (# social basis motifs)
L	13	13	13	13	13	13
λ	0.0005	0	0	0	0.003	0
W_{init}	Random	Basis Motifs	Basis Motifs	Basis Motifs	Random	Complex Basis Motifs
H_{init}	Random	Random	Random	Random	Random	Random
λ_{orthoH}	1	1	1	0	1	1
λ_{orthoW}	0	0	0	0	0	0
Iterations	300	100	100	100	300	100
Tolerance	0	0	0	0	0	0
Shift	0	0	0	0	0	0
$W_{\lambda L1}$	0	0	0	0	0	0
$H_{\lambda L1}$	1	0	0	0	1	0
W_{fixed}	0	1	1	1	0	1
SortFactors	0	0	0	0	0	0
W_{update}	1	0	0	0	1	0

1326

1327 **Supplemental Table 1. CNMF parameters used in each experiment.** See methods for

1328 descriptions of parameter choices and fitting procedures. For complete details refer to seqNMF

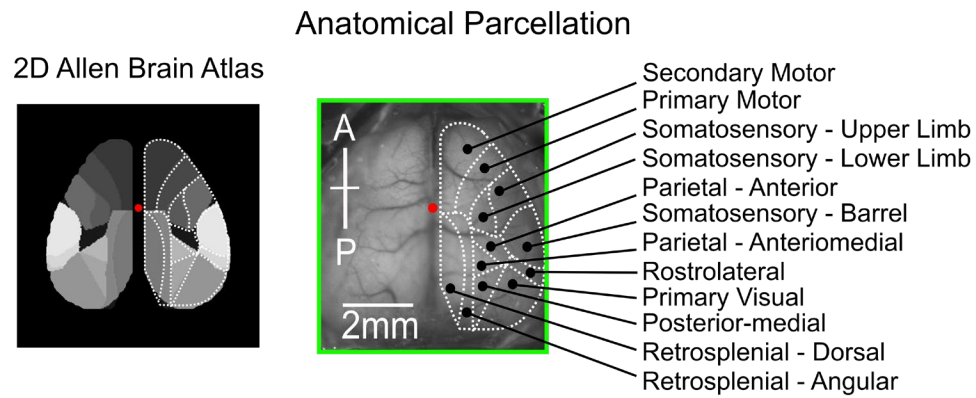1329 MATLAB Toolbox (Mackevicius et al., 2019).

1330

1331

1332

1333

1334

Anatomical Parcellation

2D Allen Brain Atlas

Secondary Motor
Primary Motor
Somatosensory - Upper Limb
Somatosensory - Lower Limb
Parietal - Anterior
Somatosensory - Barrel
Parietal - Anteriomedial
Rostrolateral
Primary Visual
Posterior-medial
Retrosplenial - Dorsal
Retrosplenial - Angular

1335

1336 **Figure 1, Supplement 1. Detailed anatomical parcellation. Left:** 2D projection of Allen Brain

1337 Atlas anatomical parcelation. **Right:** Allen Brain Atlas anatomical region labels overlaid on

1338 example mouse brain. Dotted white lines indicate manually drawn region outlines overlayed on

1339 motifs in main figure text.

1340
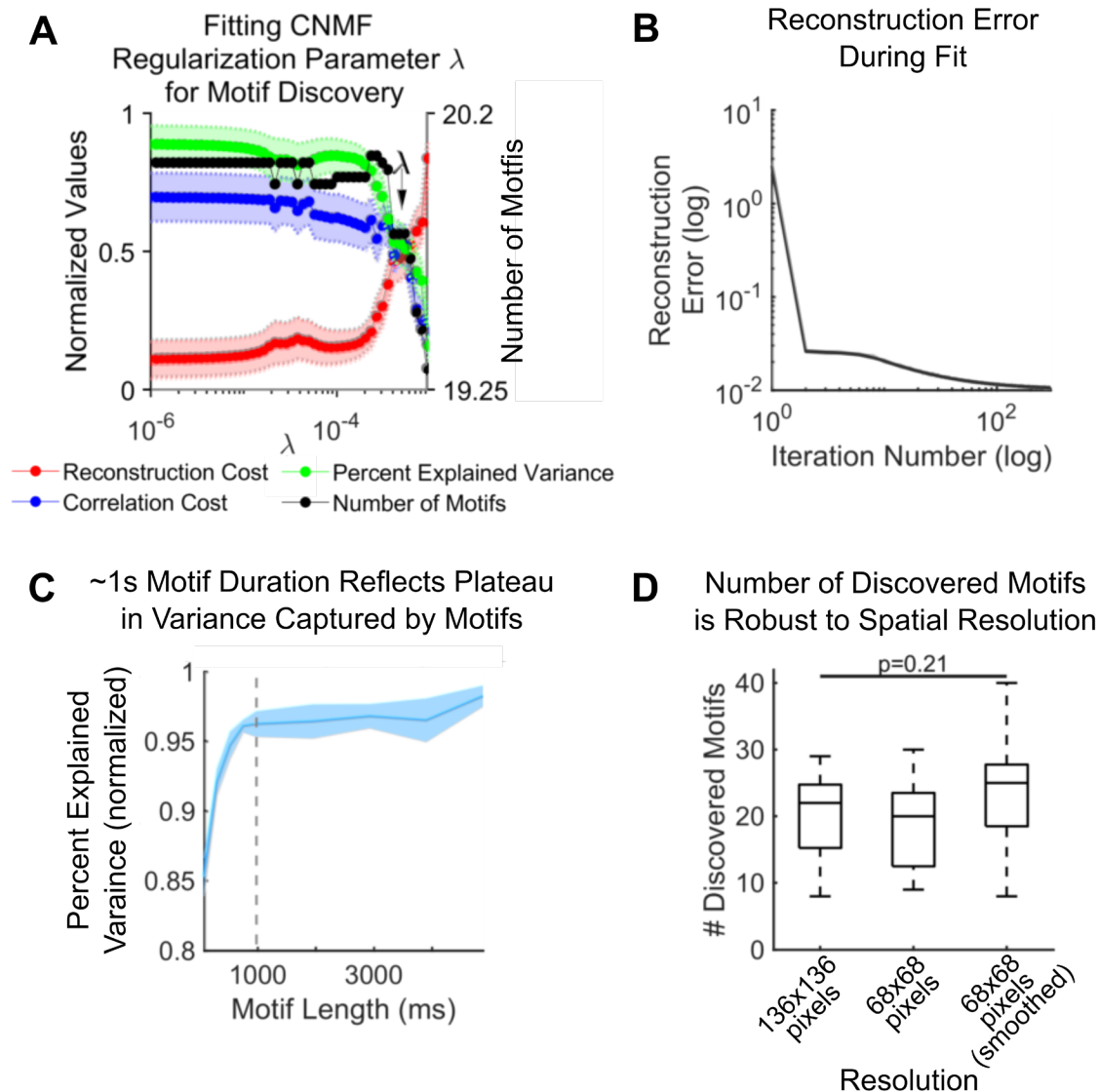
1341

Figure 1, Supplement 2. Testing impact of CNMF parameters on motif discovery. (A) The effect of changing the spatio-temporal regulation paramater (λ) in the CNMF algorithm on reconstruction cost (red), correlation cost (blue), explained variance (green) and number of identified motifs (black). Each data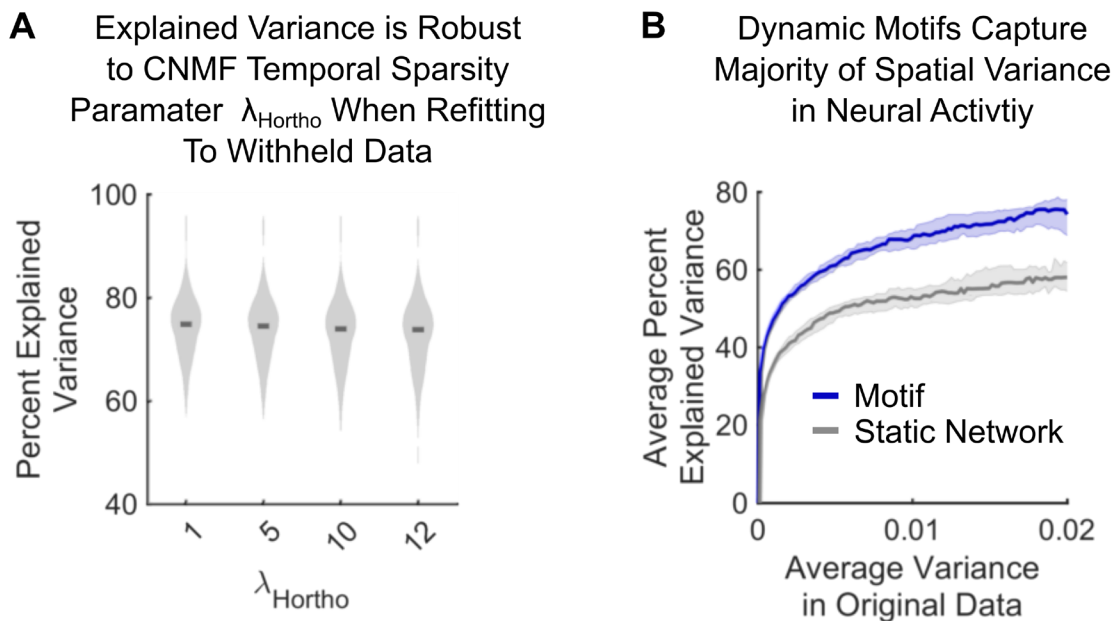 point indicates the mean value from 20 fit epochs (randomly selected; no replacement); shaded regions indicate SEM. Y-axis units are arbitrary; values were normalized between 0 and 1 across λ values for each of the 20 fits. Chosen lamda is indicated by arrow. **(B)** Reconstruction error as a function of iteration number of CNMF algorithm. All motif discovery factorizations were run for 300 iterations, at which point there was minimal improvement

69

1350   in reconstruction error. **(C)** Post-hoc validation of choice of motif duration. Motifs were discovered

1351   using motif durations up to 5 seconds. 1 second motifs captured comparable variance to longer

1352   motif durations. Shaded regions and dark line show mean and SEM, respectively. **(D)** Spatial

1353   resolution did not change number of discovered motifs. Spatial resolution indicated along x-axis.

1354   When smoothed, a 2D gaussian filter ($\sigma = [1,1]$ pixels; see Methods for details) was convolved

1355   across each frame. Line, box, and whiskers denote median, 25th-75th percentile, and range

1356   repectively. Significance estimated with one-way ANOVA.

A Explained Variance is Robust to CNMF Temporal Sparsity Paramater λ_{Hortho} When Refitting To Withheld Data

B Dynamic Motifs Capture Majority of Spatial Variance in Neural Activtiy

1357

1358   **Figure 3, Supplement 1 (A)** Percent of variance in neural activity explained by motif

1359   reconstructions as a function of temporal sparsity parameter λ_{Hortho}. Full distribution shown. Dark

1360   lines indicate median. **(B)** Percent of variance in neural activity explained by motif reconstructions

1361   (purple) and static networks (gray) of withheld epochs per timepoint. The explained variance was

1362   separately calculated for each timepoint of each epoch (1560 timepoints per epoch, 144 epochs;

1363   see Methods for details). Timepoints were then binned according to the variance across the image

1364   in the original data (100 equal bins). The explained variance captured by the reconstruction was

1365   averaged per bin per epoch. Dark lines indicate median of 144 epochs. Shading indicates 95%

1366   confidence interval. Analyses performed on withheld data within the same animals (as in Fig. 3D;

1367   purple).

1368
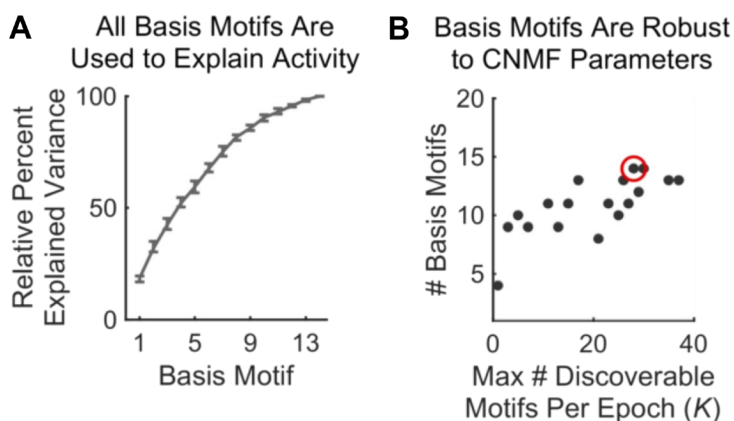
1369

Figure 4, Supplement 1. Additional characterizations of basis motifs (A) Cumulative sum of relative percent explained variance of basis motifs. Relative PEV was defined as the fraction of total PEV (of all motifs) captured by each motif per epoch. Basis motifs are in descending order by their relative PEV; these labels are used for basis motifs throughout manuscript. Line and error bars denote mean and 95% CI, respectively. **(B)** Number of basis motifs discovered (y-axis) as a function of CNMF hyperparameter K, the maximum number of discoverable motifs allowed in a single epoch (x-axis). Motif discovery and clustering was repeated for each K value (see Methods for details). Regardless of parameters, 10-14 basis motifs were identified. Red circle denotes K value (28) used for all experiments in the main text, conservatively chosen to maximize the number of basis motifs discovered.

1380

Hemodynamics Contributed Minimally to Motifs
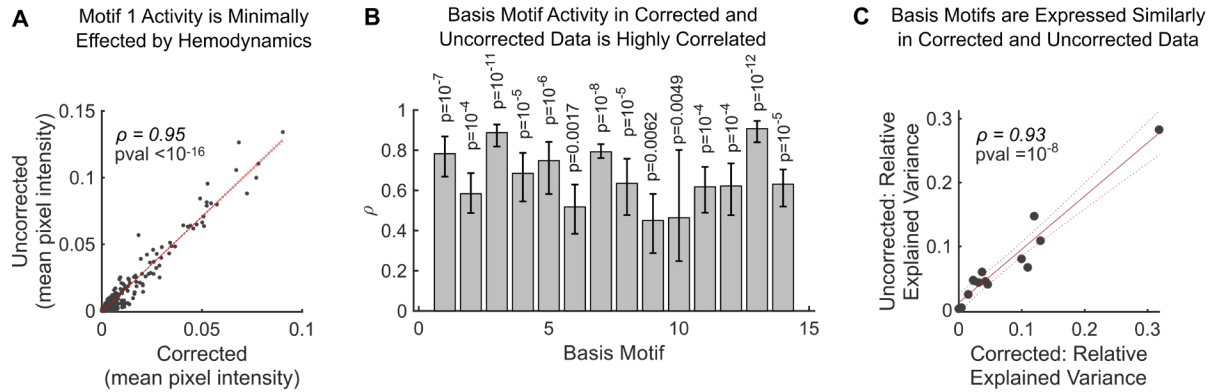
Figure 4, Supplement 2. Multiwavelength hemodynamic correction. (A) Example correlation between average pixel intensity of motif 1 reconstruction before and after multiwavelength hemodynamic correction (see Methods for details). Gray markers represent mean pixel intensity per timepoint. Solid and dotted red lines show linear least squares fit and 95% confidence bounds respectively. **(B)** The correlation in activity between corrected and uncorrected data was high for all motifs. Correlation is shown for N=30 2-min epochs across 2 animals. Mean and confidence inte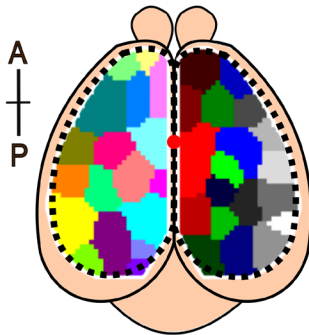rvals calculated on fisher z-transformed data before reconverting to Pearsons correlation coefficient. **(C)** Average relative variance explained by motifs in corrected and uncorrected epochs. Data points show mean of N=30 2-min epochs. Display follows **A**.

Functional Parcellation
of Mouse Cortex

1392

Figure 4, Supplement 3. Estimating the 'working resolution' of widefield imaging approach.

Parcellation of mouse cortex into functional clusters (N = 18 and 19 for left and right hemisphere,

respectively). Functional clusters grouped pixels that were correlated over time (see Methods for

details). Each color denotes a separate functional cluster. Red dot indicates bregma.

1397
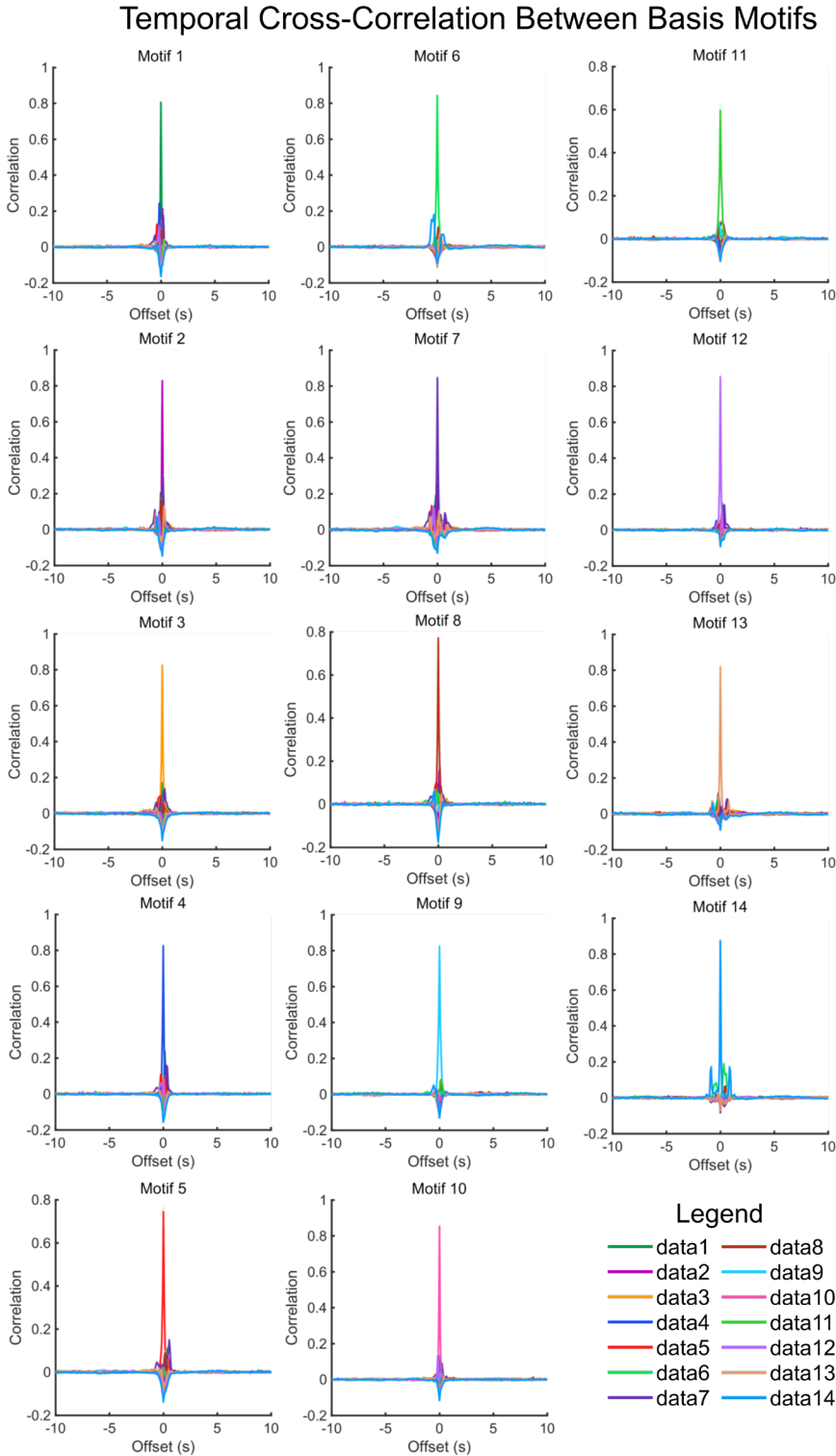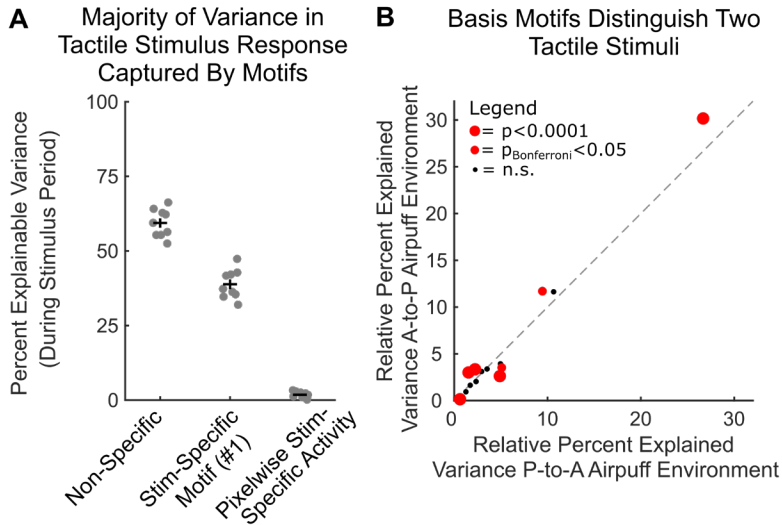
Figure 4, Supplement 4. Temporal cross correlation between motifs. Temporal cross-correlations (and autocorrelations) performed on the temporal weightings of basis motifs refit to

1401   N=144 withheld epochs. Line and shading reflect mean and SEM respectively. No obvious

1402   hierarchical structure was observed in the activation of different motifs.

1403

1404   **Figure 7, Supplement 1. Basis motif expression differs in response to two tactile stimuli.**

1405   **(A)** The majority of variance in neural activity could be explained by motif activity, not stimulus-

1406   specific activation. Plot shows the percent of explainable variance in the neural response to tactile

1407   stimuli that is captured by non-specific motifs (left column), the stimulus specific motif (motif 1;

1408   middle column) and stimulus-specific residuals (right column). Follows Figure 7C. Data points

1409   correspond to mice (N=9). Black horizontal bars indicate mean and vertical bars indicate SEM.

1410   **(B)** Plot shows the relative percent explained variance for each basis motif in response to anterior-

1411   to-posterior airpuffs and a posterior-to-anterior airpuffs directed at the whisker pad. Presentation

1412   follows Figure 6B.

1413