TREBLE: a generalizable framework for high-throughput behavioral analysis.

Ryan A. York[1*], Lisa M. Giocomo[1], and Thomas R. Clandinin[1*]

Affiliations

[1]Department of Neurobiology, Stanford University, Stanford, CA 94305

[*]Correspondence: ryanyork@stanford.edu; trc@stanford.edu

Abstract

Understanding the relationships between neural activity and behavior represents a critical challenge, one that requires generalizable statistical tools that can capture complex structures within large datasets. We developed Time-REsolved BehavioraL Embedding (TREBLE), a flexible method for analyzing behavioral data from freely moving animals. Using data from synthetic trajectories, fruit flies, and mice we show how TREBLE can capture both continuous and discrete behavioral dynamics, can uncover behavioral variation across individuals, and can detect the effects of optogenetic perturbation in an unbiased fashion. By applying TREBLE to the freely moving mouse, and medial entorhinal cortex (MEC) recordings, we show that nearly all MEC neurons encode information relevant to specific movement patterns, expanding our understanding of how navigation is related to the execution of locomotion. Thus, TREBLE provides a flexible framework for describing the structure of complex behaviors and their relationships to neural activity.

Introduction

Accurate descriptions of animal behavior are essential to understanding brain function. However, naturalistic behavior in freely moving animals is often continuous, has structure across multiple timescales, and can vary broadly between individuals and contexts. Thus, statistical tools that can capture the temporal structural of behavior, can operate on large data sets derived from many individuals, and are generalizable across systems and experiments are of central interest.

Advances in tracking technology have enabled measurements of animal movement from a wide range of species (Calhoun & Murthy 2017; Brown & de Bivort 2017; Datta et al. 2019). A variety of methods have begun to be used to parse such behavioral structure and to measure behavioral changes caused by ever more powerful experimental perturbations (Krakauer et al. 2017; Datta et al. 2019). Many of these methods train algorithms to identify discrete movement motifs and to order these motifs into probabilistic descriptions of behavior (Brown et al. 2013; Wiltschko et al. 2015; Robie et al. 2017; Marques et al. 2018; Calhoun et al. 2019; Markowitz et al. 2018). In these approaches, experimental perturbations can change either individual movements, the probabilities with which they emerge, or both. At the same time, animal movements can be continuous, and sensitive to both experimental context and manipulation (Berman et al. 2014; Katsov et al. 2017; Cande et al. 2018). Generalizable methods that can capture both the discrete structure of behavior as well as its continuity - and can operate rapidly across organisms and contexts - would be of widespread interest. Here, we present Time-REsolved BehavioraL Embedding (TREBLE), an easy-to-use method

that achieves these goals, and that can be used to assess the effects of experimental perturbations across the entire space of behavioral measurements.

We reasoned that individual movements, by analogy with some genomic analyses, could be treated akin to conserved nucleotide sequence blocks of variable length. In genomics, quantitative and qualitative differences in sequence can be efficiently revealed by shotgunning long sequences into smaller, overlapping blocks that preserve local structure (Staden 1979; Wang et al. 2009). These blocks are then assembled into libraries that densely sample the structure, frequency and surrounding sequence of each block, while smoothly reconstructing the entire sequence. Quantitative variation in these features can then be assessed using standardized computational methods often including dimensionality reduction followed by statistical testing. Building on this conceptual parallel, we developed TREBLE as a method for extracting all behavioral sequences from a dataset (analogous to shotgun sequencing), assembling 'libraries', and creating a shared 'behavior space' using dimensionality reduction.

In the TREBLE framework, behavior is first quantitatively measured (Figure 1A). Relevant measurements such as centroid velocities, or changes in body or limb position, are computed from video data, and segmented into highly overlapped temporal windows, the size of which is constrained by the temporal structure of behavior. These windows are then collected into a large library encompassing all individuals and experimental perturbations and assembled in a low-dimensional behavior space. Depending on window size, this space can flexibly capture a range of temporal dynamics, ranging from unique trajectories to recurring patterns of movement. The

resulting space can be leveraged to decode complex patterns of movement in a time-resolved fashion, facilitating analyses such as the rapid comparison of individuals or identification of neural perturbations and stimulus effects.

Here, we use TREBLE to analyze a variety of data sources (synthetic trajectories, fruit flies, mice) and applications (behavioral repertoire description, experimental perturbation, intersection of behavior with multielectrode neuron recordings). We show that TREBLE's parameters can be rationally chosen using empirical analyses and that the resulting behavior spaces are regularly structured and display recurrent dynamics. Individual movement bouts can be easily decoded and classified, the representation of which can be tailored to a user's need and allowing for detailed dissections of behavior into its constituent components. TREBLE can handle data from many individuals and millions of data points simultaneously, making possible the detection of otherwise invisible behavioral variation. Finally, we demonstrate the ability of TREBLE to rapidly uncover neurobehavioral relationships during optogenetic perturbations in flies and multielectrode recordings in freely moving mice.

Results

Calibrating TREBLE

To calibrate this approach, we first used a large, synthetic dataset comprised of correlated random walks to explore the relationships between window size and the emergent structure of behavior space (Bovet 1988)(Figure 1B; Figure S1A). Individual random walks were completely described by a combination of velocity features (translational velocity, angular velocity, and side slip; Figure 1C) with characteristic

temporal dependencies determined by the correlated random walk generator (and highlighted by their autocorrelations; Figures 1D-E). We reasoned that window sizes smaller than the underlying correlations within and between velocities would artificially granularize features, while windows larger than the underlying correlations would combine uncorrelated features. Given this, an 'ideal' window size should show less variance across trials and display recurrent dynamics, paths through the space that repeat independently in the dataset since stereotyped movements that repeat would necessarily lead to recurrent paths in such a space.

To explore this, we assessed the effect of window size on feature representation, sweeping sizes from 10ms to 2 seconds (see Methods). Windows extracted from all replicate walk trajectories were then embedded into behavior space using the UMAP algorithm, a computationally efficient, non-linear dimensionality reduction approach (McInnes et al. 2018). The structure of behavior space varied broadly as a function of window size; topologies ranged between disordered (10-30ms), recurrent (40~350ms), and unique paths (450ms ~ 2 seconds) (Figure 1F; Figure S1B). These observations confirmed our initial intuition that window size could have a range of effects on the structure of behavior space.

We next sought to quantify this topological variation. We employed two statistics, one targeting local structural differences and the other focused on global variation. To assess local variation, we measured the average Euclidean distance between temporally adjacent points, reasoning that spaces with smoother paths would display smaller distances with less variance. Mean values of Euclidean distance displayed a trough spanning window sizes of between 450 to 1250ms (Figure 1G). The coefficient of

variation (CV) highlighted a broad overlapping region of similarity, spanning window sizes of between 10 to 950ms (Figure S1C). To assess global variation, we measured Procrustes distance (McInnes et al. 2018; Dryden & Mardia 1998). This metric compares the difference between configurations of points, used here to quantify the distance between replicate behavior spaces of a given window size. This measure displayed a pattern similar to Euclidean distance, with a large trough between 450 and 1250ms, and a consistent CV up to window sizes of 950ms (Figure 1H; Figure S1D). Thus, these metrics allow quantification of the effects of window size, in this case revealing stability in topologies of behavior space across a range of windows.

We then explored how the choice of window size affects the temporal progression of trajectories through behavior space by measuring the frequency with which portions of a trajectory were repeated. We delineated a neighborhood around each point in behavior space (see Methods) and determined the amount of time between visits to each neighborhood. This was then represented as the proportion of all neighborhoods that were revisited (Bruno et al. 2017). The distribution of these values was consistent across replicates *within* a given window size but varied across sizes (Figure 1I). Window sizes of 150 to 650ms displayed a dominant return time of approximately 250ms, with a secondary peak at 1 second (Figure 1I). Windows larger than 650ms displayed increasingly noisy return time distributions (Figure 1I). Reflecting this, plots of mean recurrence times revealed a structured tuning curve in which the mean and variance were minimized at a window size of 150ms (Figure 1J). Combining these observations, the broad tuning curves in both the spatial and temporal properties

of behavior space indicate that while window size can have appreciable effects, this representation of behavior can be robust across multiple parameter values.

Analysis of fruit fly locomotion

How does TREBLE perform when applied to biological data? First, we applied the TREBLE to fruit fly locomotor behavior collected from individual animals walking on an air-cushioned ball while exploring a virtual world (Haberkern et al. 2019). We extracted rotational, translational and slip velocity components from these data and applied the empirical window analysis outlined above. Structural variation analysis of behavior space revealed a minima in the Procrustes distance distribution at 180ms window size (Figure 2A; Figure S2A) while mean recurrence time was minimized at 140ms (Figure 2B; Figures S2D-F). Given that these metrics displayed only a small amount of variation between these two window sizes, we focused further analyses on a behavior space constructed using a window size of 160ms.

We constructed a behavior space from the movements of 14 individual walking fruit flies (Figure 2C; 399,869 windows). Individual points in this space corresponded to unique windows with specific combinations of velocity features (Figure 2C). Connecting these points based on their temporal order revealed stereotyped paths in which individual flies repeated the same patterns of movement (Figure 2D; Figure S3A). As a result, the mean vector field produced by these was highly structured (Figure 2E). These pathways traversed regions of space defined by the input velocity features, meaning that position within the space could be used to infer the underlying pattern of movement (Figures 2F-H). Analyzing continuous movement through these regions

183 demonstrated that individual behavioral bouts and sequences could be identified
184 (Figures 2I-L). Finally, consistent with the notion that these sequences were recurrent
185 (Figure 2B), plotting the positions of individual fly trajectories in behavior space over
186 time reveals periodicity (Figure 2M, and data not shown). Taken together, these
187 observations show how TREBLE can be used to identify repeated, intuitive, and
188 interpretable patterns of behavior over time.

189 The capacity of TREBLE to co-embed many trials or individuals in the same
190 space facilitates direct measurements of trial-to-trial or individual-to-individual variation.
191 For example, we found that each of the 14 flies moved in grossly similar ways, as
192 indicated by the fact that 83% of the behavior space was explored by all individuals,
193 while less than 1% of the behavior spaces was explored by only one individual (Figures
194 S3B). At the same time, individual trajectories varied greatly in how often they traversed
195 different parts of the space. To measure this, we computed a probability density map
196 for each individual (Figure S3D), and identified small, specific regions of behavior space
197 that had the greatest variance between individuals (Figure 2N; Figures S3C-D).
198 Hierarchical clustering of these maps revealed a variety of behavioral repertoires across
199 individuals (Figure 2O). Thus, in this case, behavior space was composed of movement
200 types common to all individuals and could therefore act as template to compare the
201 unique statistics of each.

202

203 **Detecting behavioral changes due to optogenetic manipulation**

204 To test the ability of TREBLE to detect behavioral changes arising from an
205 experimental manipulation, we analyzed locomotor behavior during optogenetic

206 activation of gustatory sensory neurons expressing Gr5a, a sugar receptor (Haberkern
207 et al. 2019) (Figure 3A; see Methods). To induce local search behavior, a 200 ms
208 optogenetic stimulation was delivered every time a fly reached a pre-specified area and
209 was repeated for every return visit. We processed nineteen optogenetic trials
210 (1,110,025 windows) and co-embedded these windows with those from the control flies
211 (described above) in the same behavior space. The movements contained within
212 optogenetically stimulated trials overlapped extensively with those in control flies, with
213 >97% of the space occupied by both datasets (Figures S3E-F). However, on average,
214 the frequencies with which specific patterns emerged in the two datasets diverged
215 dramatically (Figures 3B-C) and were unevenly distributed across behavior space (per
216 bin Kruskal-Wallis test, see Methods; Figure 3D). These observations suggest that,
217 while the structure of locomotion is conserved overall, these two groups display quite
218 different temporal patterns of behavior.
219 To further explore this difference, we leveraged the continuous, time-resolved
220 nature of TREBLE to map the evolution of average responses to neuronal activation
221 over time (Figure 3E). Compared to pre-stimulus epochs, overall behavioral responses
222 displayed an initial period of slowing (spanning from 0 to 120ms after stimulus offset),
223 followed by increased turning (120ms to 280ms) (Figure 3E). These dynamics are
224 consistent with known local search behaviors (Haberkern et al. 2019; Corfas et al.
225 2019). Despite these population-level patterns, we noted a surprising amount of
226 behavioral heterogeneity amongst optogenetically activated flies, such that the behavior
227 of some optogenetically activated flies overlapped with that of controls (Figure 3F;

228 Figure S3G). Thus, TREBLE can assess the effects of a behavioral perturbation to
229 detect both population and individual level variation.
230
231 **Testing the generality of TREBLE using mouse locomotion**
232 We next sought to probe the utility of TREBLE in terms of its generalizability
233 across animals and contexts using mouse foraging as a model. As above, we extracted
234 time-varying velocity measures from positional data of individual mice as they were
235 foraging for randomly scattered food rewards in a 1 meter by 1 meter box (Figure 4A).
236 We then collected windows from these data using iterative selection procedure to
237 choose a window size of 400ms (Figure S4). As we previously observed with the
238 correlated random walk and *Drosophila* spaces, the resulting behavior space ordered
239 points by velocity features (Figure S5A) and connected them with continuous and
240 directionally recurrent pathways (Figures 4B-D). Distinct behavioral bouts – such as
241 running, turning, and stopping – could be easily decoded from the space and mapped
242 back onto real XY coordinates (Figure 4E). Given this, we conclude that TREBLE is
243 generalizable across animal taxa and behavioral contexts.
244
245 **Using TREBLE to characterize neural encoding of behavior**
246 A common, and often difficult, goal in neuroscience is to relate neural activity to
247 behavior. We therefore used TREBLE to identify behavioral coding in 794 medial
248 entorhinal cortical (MEC) neurons recorded during the foraging trials presented above
249 (n = 14 mice, 327 trials; Hardcastle et al. 2017) (Figure 4A). The MEC is hypothesized
250 to support navigation and contains a population of functionally-defined neurons that

251 encode a variety of behavioral variables such as an animal's position in external space
252 (e.g. grid and border cells), head direction, and running speed (Kropff et al. 2015;
253 Sargolini et al. 2006; Solstad et al. 2008).

254 We related MEC neuron activity to movement in the behavior space defined by
255 TREBLE. To do so, we first created a 2-dimensional behavioral tuning curve for each
256 neuron. Each cell's average activity was mapped as a function of the animal's position
257 in behavior space, revealing a wide variety of patterns (Figure S5G). These
258 relationships could be constrained to specific regions of behavior space as captured by
259 both negative and positive correlations (Figure 4F-H). Consistent with the MEC's role in
260 navigation, permutation tests revealed that these correlations were overwhelmingly non-
261 random, with only 63 of the 794 MEC neurons displaying non-significant relationships
262 with locomotor behavior (Figure 4I). Furthermore, clustering neural activity patterns
263 across behavior space revealed multiple distinct types of relationships between MEC
264 neurons and locomotor behavior (Figure 4J; Figure S5G). Most neurons were active in
265 relatively small and specific regions of behavior space, suggesting that they encode
266 information relevant to particular locomotor movements (Figure S5G). Other neurons
267 displayed distributed activity across behavior space, consistent with less selective
268 coding of locomotor information (Figure S5G). Thus, these various patterns can be
269 considered 2-d tuning curves, representing the relationship between neural activity and
270 behavior for each neuron.

271 To what extent do these patterns reflect previously described representations of
272 navigational coding in the MEC? To address this, we intersected TREBLE correlations
273 with other commonly calculated MEC coding variables (speed, head direction, spatial

stability, border proximity; see Methods). Unsurprisingly, we found that behavior space correlated with speed coding (Figure 5K) compared to the other measures (Figures 5L-N). However, a more precise measure of coding capacity requires a comparison between the amount of information jointly shared between position in behavior space, speed score, and neural activity. To do this, we computed the mutual information (MI) between TREBLE position and neural activity for each neuron and then related these measures to the commonly calculated MEC coding variable (Figures 5O-R). Higher MI values here indicate that spatial coordinates in behavior space contain more information about a given neuron's activity, reflecting stronger behavioral coding. Strikingly, MI was uncorrelated with speed score (Figure 5O), indicating that encoding of locomotor behavior was not solely related to variation in speed. This raised the possibility that these neurons may encode information about other locomotor variables including longer timescale and more complex maneuvers.

Using TREBLE to identify temporal variation in the neural coding of behavior

Neural activity can relate to continuous behaviors on a variety of timescales, the discovery of which is an area of active interest (Glaser & Kording 2016; Krakauer et al. 2017; Datta et al. 2019). We therefore wanted to test the ability of TREBLE to provide a path for relating neural activity with behavior and took advantage of the apparent rich capacity of MEC neurons to encode locomotor behavior over a variety of timescales.

To do this, we first computed the cross-correlation of TREBLE position and activity for each neuron (Figures 6A-C). This analysis identifies when a neuron's activity correlates with position in behavior space by calculating correlations between the two

variables using a sliding window centered around a temporal offset of zero (i.e. simultaneity). We found a variety of temporal offsets in the cross-correlation distributions for each neuron. Peak correlations could both precede (Figure 6A) and follow (Figure 6B) changes in behavior. Clustering the distribution of these cross-correlation coefficients revealed a variety of neural classes, including negative, positive, and even mixed negative and positive cross-correlation profiles (Figure 6C; Figure S6A). Notably, cross-correlation peaks could be of varying widths; some were tightly centered (e.g. Figures 6A-B) while others displayed correlations lasting seconds (Figure S6A). Therefore, in addition to their instantaneous relationships, the activity of MEC neurons can covary with behavior over a range of temporal scales and offsets.

Given this observation we hypothesized that individual MEC neurons may most reliably encode behavior over specific timescales. To test this, we calculated each neuron's activity over a range of temporal windows and measured their associations with behavior space via correlation, mutual information (MI), and entropy (i.e. how narrowly neural activity was distributed across behavior space) (Figures 6D-F). For most neurons, the strength of the correlation between activity and behavior as well as entropy plateaued (Figures 6D, 6F) while MI displayed a clear, and at times very specific, maximum value as a function of temporal window size (Figure 6F; Figure S6B). This pattern suggested that, given temporal binning that maximized MI, neural activity could be more specifically associated with distinct regions in behavior space. This was a general pattern: across all neurons the MI maximized window size encoded more information than instantaneous spikes (Figure 6F). The distribution of these MI maximized window sizes varied broadly, ranging from instantaneous spiking to window

sizes up to 2 seconds, reflecting the diversity of patterns observed in the cross-correlation analyses (Figure 6G; Figure 6C). MEC neurons thus possess specific and, in some cases, finely tuned temporal scales over which their activity most informatively encodes movement.

Finally, we asked whether these temporal relationships might be further refined by considering how the magnitude of each neuron's activity constrains its behavioral coding. Do extrema in neuronal firing rates reflect specific behaviors? To answer this question, we calculated percentile cutoffs for each neuron's activity across the trial using the MI maximized window sizes (Figure 7A). We then examined whether increasing activity rates were associated with consistent changes in movement through behavior space, indicating increasing behavioral specificity. To do so, we analyzed the distribution of movements within behavior space by computing the entropy of the distribution associated with neural activity bouts above each cutoff (Figures 7B-C). In this metric, lower entropy values correspond increased stereotypy in behavior space. For many neurons, as activity increased the corresponding behavior space entropy decreased, meaning that high levels of neuronal activity tended to be associated with specific behavioral patterns (Figure 7C; Figure S6C). Hierarchical clustering of these entropy distributions separated neurons into coherent groups with temporally similar profiles (Figures 7D-E). Notably, a number of clusters displayed entropy minima that were significantly offset from zero, meaning that peaks in neural activity could either precede or follow changes in behavior (Figure 7F; Figure S6C). Taken together, these results demonstrate that MEC neurons can encode locomotor information on temporally varying scales, both before and after changes in behaviors occur. Moreover, these

analyses demonstrate how TREBLE can be used to uncover the relationship between neuron activity and continuous behavior across a rich neural dataset from freely moving animals.

Discussion

Technological developments have led to substantial shifts in how researchers are attempting to unlock the development, expression, and evolution of animal behavior. Facilitated by high throughput data analysis, a number of ambitious conceptual projects have recently been proposed. These include considering behavior as a natural extension of physics (Brown & de Bivort 2017), attempting to capture many species' complete behavioral repertoires (Anderson & Perona 2014), or even reconciling the two dominant lineages of behavioral research, comparative psychology and ethology (Datta et al. 2019). Achieving these goals requires the study of individuals across a diversity of animal species, in both controlled and naturalistic contexts, and in conjunction with other data streams, critically including neural activity. Given this, behavioral analytical methods that are statistically robust, broadly applicable, and easily linked to the function of nervous systems are of general interest.

All such methods inherently make tradeoffs between computational ease, temporal resolution, and generalizability. TREBLE does not require extensive model fitting, is agnostic to feature type, and can accommodate many millions of data points. Moreover, we describe how the key parameters needed, such as window size, can be systematically explored and rationally chosen to capture behavioral structure of interest. By cutting behavior into defined temporal windows that are extremely densely sampled,

TREBLE simultaneously captures both long and short temporal structures in behavior. Furthermore, representation of these can be handled flexibly. For example, here we represent temporal structure in two ways: 1) as continuously varying time series and 2) as information theoretic summary measures of entire distributions over time (using entropy and mutual information). It is also possible for users to discretize TREBLE's output simply by segmenting movement through behavior space into sub regions. Finally, TREBLE is generalizable in three different ways. First, as we demonstrate, the framework can be applied to a variety of movement sources, including synthetic trajectories and multiple animal species. Second, the position of TREBLE in analysis pipelines is adjustable. While TREBLE can be implemented as a standalone method for behavioral analysis from beginning to end, it can also be used as a screening tool for identifying statistically significant behavioral variation to be followed by other approaches specifically tailored to the question of interest. Finally, as we show using optogenetic activation in flies and neural recordings in mouse MEC, TREBLE can be used to find meaningful links between nervous systems and behavior in a variety of contexts. This versatile nature of the TREBLE framework therefore opens up the possibility for customizable high-throughput behavioral analyses.

Uncovering behavioral variation

TREBLE can uncover otherwise opaque behavioral variation by directly comparing behavioral data across many individuals and conditions. Here, we show that free walking fruit flies possess substantial variation in locomotor dynamics and that this variation is associated with specific regions of behavior space. Surprisingly, we find that

this variation in free walking behavior overlaps with the diversity of responses observed after an optogenetic perturbation. A limited set of pre-determined behavioral measurements, as would have been used typically, may have missed this variation. Similarly, even unbiased statistical methods, without clear grounds for comparison across animals or experimental conditions, would have been unlikely to have detected these patterns. This highlights the benefit of using unsupervised statistical descriptions such as TREBLE for assessing the structure of behavior across diverse individuals. Furthermore, we envision that TREBLE might be employed on control data sets to measure variability across baseline individuals in order to facilitate statistical power calculations across the full range of behavioral dynamics that might be present. Such catalogs of variability could be thus be developed across species' wildtype behavioral repertoires, allowing researchers to better account for inborn variance or threshold effects in applications such as genetic mapping, behavioral manipulations, and neural perturbations (López-Alonso et al. 2014; Ayroles et al. 2015; Buchanan et al. 2015; Duistermars et al. 2018; Allen et al. 2015).

Linking neural activity to continuous behavior

Associating neural activity with continuous behavior presents two key challenges. First, identifying the specific aspects of behavior that activity corresponds to is difficult to determine in the context of a large behavior repertoire. Second, the temporal scales over which this relationship occurs can be hard to know a priori. Here, we show how TREBLE can be used to address these issues by characterizing the coding of locomotor behavior in mouse medial entorhinal cortex (MEC) neurons as a proof of concept.

We find that TREBLE captures known components of speed coding in MEC neurons while also uncovering extensive variation in temporal and rate-based coding. Previous work has found that MEC neurons prospectively code for changes in an animal's speed and position ~50-80ms in the future (Kropff et al. 2015). Our analyses add to these findings, suggesting that MEC neurons can encode changes both before *and* after behavioral events, doing so across a variety of timescales. These patterns may reflect the presence a multitude of temporal encoding strategies for behavior in the MEC, similar to what has been seen in basal ganglia movement selection (Jin & Costa 2015; Markowitz et al. 2018), orofacial rhythms (Moore et al. 2013), and whole-brain activity during *C. elegans* locomotion (Kaplan et al. 2020). We speculate that this may arise from some MEC neurons encoding short time-scale behavioral events (e.g. turning left or right) while others update based on longer-term navigational behaviors such as goal-oriented searching or foraging. These results demonstrate how TREBLE can be leveraged to identify specific aspects of behavior that are associated with the activity of individual neurons and to uncover the temporal structure of such relationships. More broadly, we anticipate that TREBLE may be useful in uncovering the statistical structure of behaviorally relevant activity across many diverse populations of neurons.

Future directions

There are a number of areas in which TREBLE may be further developed and employed. First, movement through behavior space may be classified on timescale that are longer than those presented here, affording descriptions of behavioral state over time or the emergence of behavior in development. Second, the ability of TREBLE to

co-embed multiple individuals with temporal resolution may make it particularly amenable to studying social and collective behavioral dynamics. Third, by readily capturing variability in behavior, TREBLE may be amenable to exploring differences across individuals arising from factors imbuing variance such as reaction norms, behavioral syndromes, and environmental or genetic variation. Finally, it is especially intriguing to consider how TREBLE may be leveraged to jointly infer neural coding principles by applying the framework to the structure of behavioral and neural dynamics in parallel.

Acknowledgements

We thank members of the Clandinin lab and Kiah Hardcastle for useful discussion and input. This work was supported by the NIH (U19NS104655; T.R.C.), the Simons Foundation (T.R.C., L.M.G.), and the Stanford School of Medicine Deans Office (R.A.Y.).

Author contributions

R.A.Y. and T.R.C. conceived the study. R.A.Y. wrote the TREBLE code and performed data analysis. R.A.Y., L.M.G. and T.R.C. prepared the manuscript.

Competing interests statement

The authors declare no competing interests.

481 **Methods**

482 ***Datasets***

483 Correlated random walks were produced using the `TrajGenerate` function in
484 the `trajr R` package[22]. Ten replicate walks were produced per parameter tested (e.g.
485 window size), sampled at a rate of 100 frames per second and consisting of 10,000
486 frames. All walks were generated using the same underlying angular and linear error
487 distributions (Normal distribution; Mean = trajectory length; Standard deviation = 0.5
488 (angular)/0.2 (linear)).

489 Details of the *Drosophila* walking dataset can be found in Haberkern *et al.* 2019.
490 Briefly, animals in the free-walking dataset ('WTB hybrid' genotype) were allowed to
491 explore a circular matt acrylic platform (radius 11.4 cm) surrounded by a siliconized
492 acrylic cylinder to prevent climbing. Videos were recorded from 120 cm above the
493 platform using a Flea3 camera at 12.3 Hz with a spatial resolution ~40 pixel/cm. Trials
494 lasted 10 minutes after a 1-2 minute acclimation period.

495 The local search data were produced from female flies possessing
496 optogenetically-accessible sugar receptor neurons (Gr64f-Gal4 > ChrimsonR). Animals
497 were allowed to explore a virtual landscape consisting of distributed 'cone forest' while
498 walking on a circular treadmill. Treadmill movement data (fly's position and velocity)
499 were collected at 360 Hz. Optogenetic stimulation was triggered whenever a fly crossed
500 within a 10mm radius of a given cone and persisted for 200ms.

501 Details of the mouse dataset can be found in Hardcastle *et al.* 2017. Behavior
502 and neural recordings were collected from two cohorts of adult wildtype mice: 5 male
503 and 2 female C57BL/6 mice (405 MEC cells recorded) and 7 male C57BL/6J:129SVEV

mice (389 MEC cells). Two polymide-coated platinum iridium 90%-10% tetrodes were implanted in each mouse prior to the experimental period. Behavioral data were collected approximately one week after surgery in large open environments with black walls containing chocolate flavored cereal to induce foraging (sizes varying; see Hardcastle *et al*. 2017 for details). The majority of recording sessions lasted between 30-35 minutes with a small number ranging in time between 12 and 122 minutes. During each session, position, head direction, and running speed were recorded every 20ms and single unite spikes were recorded at a 10kHz sampling rate.

Generation of correlated random walk behavior spaces

The trajectories and velocity distributions of replicate correlated random walks were extracted using the custom function `iterative_umap`. First, a novel walk trajectory was generated using `TrajGenerate` (as described above). Instantaneous velocity components were then calculated (translation, angular velocity, sideways velocity) and sampled using windows of a specified width (denoted here as `w`) and step size (denoted here as `s`). The sampling procedure was as follows.

First, given frame `i`, the velocity components were extracted for frames `i:i+w`. Angular and sideways velocity values were normalized to the first frame in the window so that each velocity vector to originate from zero. Since we weren't concerned with information pertaining to turn direction (i.e. left vs. right) the proceeding velocity values were then adjusted so that the second frame was always positive. The resulting velocity vectors were linearized and concatenated, resulting in a single vector of length `3w`. This procedure was then repeated iteratively every `s` frames for the length of the trial (`t`).

After each iteration the concatenated vectors were appended to a matrix with `3w` rows to ultimately create `t/s` columns.

This procedure yielded a library of densely sampled velocity fragments corresponding to behavioral dynamics for each moment in the trial. The `R` implementation of the `UMAP` algorithm[17] was then used to embed these fragments into a low-dimensional behavior space (each point in space corresponding to a window). The resulting space thus provided a 2-dimensional xy position for each point/window in the trajectory. For downstream analyses we computed simplified positional information for each space using the `bin_space` function. This function decomposes behavior space into a grid of desired size (*n* bins x *n* bins). For example, for a desired grid of 16 bins x 16 bins, a new set of xy coordinates will be calculated corresponding to 16 intervals spanning the minimum/maximum xy coordinates in the original space. The position of each point in the space will then be compared to the new coordinates and associated with the closest bin, in this example case generating a reduced set of 16x16, or 256, unique bins.

Analyzing correlated random walk parameter space

We assessed the relationship of parameter choice and behavior space structure by sweeping the two main window sampling parameters: window width and step size. We examined 24 different window widths, ranging from 10ms to 2 seconds (10-40ms sampled at 10ms steps; 50ms-2 seconds sampled at 100ms steps). Ten correlated random walk replicates were generated for each window size and then processed and embedded into behavior space using the `iterative_umap` function. We also

explored the effect of step size by producing replicate behavior spaces from trajectories sampled at step sizes varying from 10 ms (i.e. 1 frame) to 1 second, separated by 100 millisecond intervals. Again, for each step size 10 correlated random walk replicates were generated, processed, and embedded into behavior space using the `iterative_umap` function.

Intra-point Euclidean distance was calculated for each space by comparing the spatial position of temporally adjacent points. Specifically, the position of a given point (corresponding to a unique temporal window) in behavior space (x_t, y_t) was compared to the position of the next point/window in time (x_{t+1}, y_{t+1}). The Euclidean distance between these points was then calculated and stored into a vector, allowing for the distribution of distances to be compared across replicates and parameter conditions. The coefficient of variation was also calculated for these measures to control for substantial variation in the magnitude of effect; this was done by dividing the distribution's standard deviation by its mean.

Procrustes distance was used to measure the similarity between replicate behavior spaces of a certain parameter combination[18]. Briefly, this method treats spaces and their component points as a set of landmarks for pairwise comparison. To do so, pairs of behavior spaces are scaled to be similar sizes then shifted and rotated to possess the same position and orientation in space[18]. The distance between corresponding points in the two adjusted landmark sets is then calculated using Procrustes distance (reported here as Root Mean Square Error; RMSE). We performed this procedure and calculated Procrustes distance for all pairwise combinations of the

572 10 replicate behavior spaces within a given parameter set. Coefficient of variation for
573 this measure was also calculated as above.
574 To assess recurrence, we measured the average return time for all points in each
575 space. We first used the binning procedure outlined above to create a 32x32 grid so
576 that the resulting space was composed of 1024 unique bins. We then computed a
577 distance threshold to delineate a “neighborhood” around each bin (5% smallest intra-bin
578 distance). For a given bin, each time the trajectory passed through its threshold the
579 event was recorded and its duration stored. This distribution of return delays was
580 analyzed by computing the proportion of points that displayed a return for a given time
581 delay. For example, given a delay of say 150ms, all bins would be scanned and the
582 proportion possessing returns that occurred took between 150 and 160ms would be
583 calculated. We calculated proportions in this manner for delays between 0ms and 2
584 seconds at 10ms intervals (as plotted in Figure 1I). Mean recurrence time was
585 measured by collecting all observed return times for each replicate and calculating the
586 mean.
587
588 ***Generation of free-walking fruit fly behavior spaces***
589 The movement trajectories of 20 female WTB flies walking in a circular area were
590 collected from Haberkern *et al.* 2019. For each fly we computed velocity components as
591 above and interpolated the resulting vectors from 12.3 Hz to 50 Hz to increase
592 smoothness in the downstream behavior space. Initial analyses suggested the removal
593 of 6 trials due to missing data or lack of substantial movement, resulting in a final
594 dataset of 14 flies.

As above, we performed an empirical test to decide on the optimal window size for behavior space creation. To do so, we sampled the first 2000 frames from each trial (to increase computational efficiency) and swept through window sizes ranging from 20ms to 1 second (70 ms step size), creating a single behavior space for each of the 14 flies per window size. The resulting spaces were then directly compared using the Euclidean distance, Procrustes distance, and recurrence metrics previously described (Figure S2).

Analyzing free-walking fruit fly behavior spaces

After sweeping window parameters, we generated a behavior space composed of all 14 free-walking trials using a window size of 160ms (399,869 windows). We then used `bin_space` to calculate point coordinates in a 64x64 grid. Stereotypy in movement through the space was visualized using a vector field transformation (Figure 2D). All instances in which trajectories through the space passed into a given bin were collected and then used to calculate the mean x and y vectors of the trajectory leaving the bin. These mean values were then represented visually using arrows that originated out of the corresponding bin, the direction and magnitude of which were dictated by the mean x and y vectors. Velocity distributions across the space (Figures. 2F-H) were visualized by calculating the mean value for each velocity component per window. These values were then used as input to the color function determining the hue of each point in the space.

Intra-fly variation in behavior space was assessed using the reduced 64x64 grid. For each fly we calculated the number of times its trajectory passed through a given bin,

618 in addition to a binary measure of whether that bin was visited at all ('0' if not; '1' if
619 visited at least once). The continuous counts of bin visits were then used to calculate
620 bin-wise mean and variance across all flies, the latter of which is visualized in Figure 3F.
621 The binary measure was used to assess variation in overall space occupancy across
622 flies (Figure S3).

623 We used 2-dimensional histograms to compare the overall behavioral patterns of
624 individual flies. To do so, we created an individual map for each trial comparing the
625 binned distributions of x and y coordinates via 2-dimensional kernel density estimation
626 using the function `kde2d` (bandwidth = 2; 32 grid points in each direction) in the `MASS`
627 `R` package[23]. To facilitate comparisons across trials, the resulting density maps were
628 normalized to the max density value of each and then linearized so that each was
629 represented by a single vector of density values. The relationships between trials were
630 inferred by calculating a distance matrix of these density vectors which was then used
631 as input for hierarchical clustering (`hclust` function; Figure 3G).

632

633 ***Generation of optogenetic local search behavior spaces***

634 The movement trajectories of 19 Gr64f-Gal4 > ChrimsonR flies in virtual reality
635 was collected from Haberkern *et al.* 2019. We processed and calculated velocity from
636 these files as above, down sampling from 360 to 50 Hz to match the sampling rate of
637 the free-walking dataset (for downstream comparisons). Windows were extracted from
638 all 19 trials using the same width (160 ms) as above, yielding a library of 1,110,025
639 windows. We then used the `predict` function in the `UMAP R` package to embed all
640 windows, one fly at a time, using the free-walking fly behavior space as a template. This

procedure produced a combined behavior space composed of 33 trials (1,509,894 windows) that facilitated comparisons across experimental conditions and individuals.

Variation in space occupancy between free-walking and local search trials was assessed using 2-d density maps. As before, we used `bin_space` to get new coordinates for each trial, and then calculated density maps using `kde2d` (bandwidth = 1; 200 grid points in each direction), which were then normalized by dividing each density estimate by the maximum value in the map. As in Figure 3G these density maps were then linearized, combined in a matrix, and clustered using Euclidean distance and hierarchical clustering to produce the tree in Figure 4F.

Analyzing optogenetic local search behavior spaces

We used a bin-wise Kruskal-Wallis test to statistically analyze differences in space occupancy between the two groups. For a given fly/trial we calculated the percent occupancy at each bin (number of visits to bin divided by total number of windows). A Kruskal-Wallis test was then used to compare the percent occupancies of the individual trials between free-walking and local search flies from which the test statistic and accompanying *p*-values were collected. We used Bonferroni correction to adjust these *p*-values, controlling for the number of tests performed (4,096). The adjusted *p*-values were then used to visually assess regions of greater differentiation between the free walking and local search trials, plotted as the -log10 transformation of the *p*-values (as seen in Figure 3F).

Time-evolving responses to optogenetic stimulation were assessed using density maps. To do so, we sampled the behavior space trajectories of each trial before and

after all bouts of optogenetic stimulation. For each bout of stimulation, we extracted positions in behavior space for the second before stimulation in order to represent a baseline behavioral distribution. We then extracted positions immediately after stimulation using 1 second windows and step size of 100ms, from 0ms to 1 second afterward (Figure 4E). We combined positions across bouts and individuals within window to produce an aggregated response profile using kernel density estimation (bandwidth = 2; 64 grid points in each direction; common x and y limits across windows). If individuals displayed common behavioral responses to stimulation in a specific window, then the related density map should show structure in its distribution (i.e. concentrated red regions in Figure 4E).

Generation of mouse locomotor behavior space

Behavior data and neural recordings were previously published in from Hardcastle *et al*. 2017. For each cell, full trial spike event data (collected at 10kHz) and positional coordinates (collected at 50hz) were extracted. As above, velocity components over the course of the whole trial were computed from the positional coordinates. We then used the empirical window-size test to identify the optimal window size for behavior space creation. 30 trials were randomly sampled for the iterative window test. The first 3000 frames of each trial were used to sweep through window sizes ranging from 20ms to 1 second (40ms step size; Figure S4). The resulting spaces were then directly compared using the Euclidean distance, Procrustes distance, and recurrence metrics (Figure S4). After analyzing these metrics a window size of 400ms was chosen for all downstream analyses.

687 Given the large size of the dataset (41,850,995 windows) we opted to perform a
688 seed-embedding procedure to produce individual behavior spaces for each trial. To do
689 so, we extracted and combined together the first 5,000 windows from each trial and
690 created a seed behavior space using UMAP. We then individually embedded the
691 remaining windows for each trial using the `predict` function in the `R` implementation of
692 UMAP. Time points and XY coordinates corresponding to movement through these
693 individual behavior spaces were then combined and used jointly for overall annotation of
694 the mouse locomotion (As seen in Figures 5B-E). This all-trial behavior space was used
695 to calculate 64x64 bin coordinates, allowing all mouse trials to be directly compared
696 within the same behavior space architecture. Vector field representation (Figure 5D)
697 and physical XY space movement decoding (Figure 5E) were computed as previously
698 described.
699 The spike times for all cells in a given trial were then associated with the
700 corresponding behavioral timepoints. To do so, spike rates were calculated over 20ms
701 bins corresponding to the sampling rate of mouse positional data (50hz) and, thus, the
702 rate of movement through behavior space. The rate of activity was also calculated over
703 1 second (1hz) scales to explore the extent to which broad temporal differences might
704 be present between neural activity and behavior. For each cell/trial pair the Pearson
705 correlation between behavior space position and firing rate was calculated. The
706 significance of these correlations was measured using permutation tests. For each cell,
707 the spike rate data were shuffled 10,000 times and then correlated with position in
708 behavior space. These correlations were used to calculate *p*-values by comparing the
709 number of times shuffled correlations were greater than the observed value, divided by

710 the number of permutations (10,000; Figure 5I). 2-dimendionsal tuning curves (as in
711 Figures 5F-G and Figure S5) were computed for each cell by calculating the mean spike
712 rate in each bin using the 64x64 representation previously calculated. Mutual
713 information between the cells and behavior space was calculated (in bits) by comparing
714 the resulting 2-d tuning curves to the distribution of average occupancy time per bin for
715 each trial using the function `mi.plugin` from the `R` package `entropy`. The canonical
716 MEC coding variables (speed score, spatial stability, head direction, border score;
717 Figures 5K-R) used were previously computed following Hardcastle *et al.* 2017. Cross-
718 correlations between neural activity and behavior space position were calculated with
719 the function `ccf` in `R` (lag of 16 seconds as seen in Figure 6C).

720 Temporal variation in the association of neural activity and behavior space (as in
721 Figures 6D-H) was assessed by calculating the correlation coefficient, entropy, and
722 mutual information between the two across a range of windows (0-4 seconds, 20ms
723 steps). Correlations and mutual information were calculated for each window size as
724 above. Shannon entropy (in bits) was calculated using the distribution of mean firing
725 values from 2-d neural tuning curve for each window (using the `entropy` function from
726 the `entropy R` package). MI maximized window sizes were chosen for each cell using
727 the timescale at which the maximum mutual information between behavior space
728 position and neural activity occurred.

729 Rate-based differences in coding were assessed using the MI maximized window
730 sizes identifying above. For each cell, we calculated the activity rates corresponding to
731 the 75th, 80th, 90th, 95th, and 99th percentiles. We then identified the time points at
732 which each cell's activity occurred above the respective percentile. The corresponding

positions in behavior space were then extracted in 8 second windows surrounding each timepoint (from 4 seconds before to 4 seconds after; 20ms bins). 2-d tuning curves were then constructed from the distribution of points in behavior space for all events that occurred across the corresponding 20ms bins. For example, the behavior space positions occurring 100-80ms before all events above the 99th percentile would be compared, followed by the same calculation for all events occurring 80-60ms before, then 60-40ms before, etc. This procedure would thus produce 400 2-d tuning curves per percentile, corresponding to a sampling rate of 50hz over a course of 8 seconds. The Shannon entropy (in bits) of each 2-d tuning curve was then calculated and plotted as a curve over time to detected consistent changes in its distribution (as in Figures 7C and 7E). To compare these entropy distributions across all cells each was divided by its maximum value. The normalized entropy measures for the 75th and 99th percentile values were then used as input to hierarchically cluster all cells. Clusters were identified using a dynamic tree cutting algorithm via the `cutreeHybrid` function in the R package `dynamicTreeCut` (minimum cluster size of 10).

References

1. Allen, B.D., Singer, A.C., and Boyden, E.S. (2015). Principles of designing interpretable optogenetic behavior experiments. Learn Mem 22, 232-238.
2. Anderson, D.J., and Perona, P. (2014). Toward a science of computational ethology. Neuron 84, 18-31.
3. Ayroles, J.F., Buchanan, S.M., O'Leary, C., Skutt-Kakaria, K., Grenier, J.K., Clark, A.G., Hartl, D.L., and de Bivort, B.L. (2015). Behavioral idiosyncrasy

756 reveals genetic control of phenotypic variability. Proc Natl Acad Sci U S A 112,
757 6706-6711.
758 4. Berman, G.J., Choi, D.M., Bialek, W., and Shaevitz, J.W. (2014). Mapping the
759 stereotyped behaviour of freely moving fruit flies. J R Soc Interface 11.
760 5. Bovet, P.B., S. (1988). Spatial analysis of animal's movements using a correlated
761 random walk model. Journal of Theoretical Biology 131, 419-433.
762 6. Brown, A.E., Yemini, E.I., Grundy, L.J., Jucikas, T., and Schafer, W.R. (2013). A
763 dictionary of behavioral motifs reveals clusters of genes affecting Caenorhabditis
764 elegans locomotion. Proc Natl Acad Sci U S A 110, 791-796.
765 7. Brown, A.E.X., and de Bivort, B. (2018). Ethology as a physical science. Nature
766 Physics 14, 653-657.
767 8. Bruno, A.M., Frost, W.N., and Humphries, M.D. (2017). A spiral attractor network
768 drives rhythmic locomotion. Elife 6.
769 9. Buchanan, S.M., Kain, J.S., and de Bivort, B.L. (2015). Neuronal control of
770 locomotor handedness in Drosophila. Proc Natl Acad Sci U S A 112, 6700-6705.
771 10. Calhoun, A.J., and Murthy, M. (2017). Quantifying behavior to solve sensorimotor
772 transformations: advances from worms and flies. Curr Opin Neurobiol 46, 90-98.
773 11. Calhoun, A.J., Pillow, J.W., and Murthy, M. (2019). Unsupervised identification of
774 the internal states that shape natural behavior. Nat Neurosci 22, 2040-2049.
775 12. Cande, J., Namiki, S., Qiu, J., Korff, W., Card, G.M., Shaevitz, J.W., Stern, D.L.,
776 and Berman, G.J. (2018). Optogenetic dissection of descending behavioral
777 control in Drosophila. Elife 7.

13. Corfas, R.A., Sharma, T., and Dickinson, M.H. (2019). Diverse Food-Sensing Neurons Trigger Idiothetic Local Search in Drosophila. Curr Biol 29, 1660-1668 e1664.
14. Datta, S.R., Anderson, D.J., Branson, K., Perona, P., and Leifer, A. (2019). Computational Neuroethology: A Call to Action. Neuron 104, 11-24.
15. Dryden, I.L., and Mardia, K.V. (1998). Statistical shape analysis (Chichester: John Wiley and Sons).
16. Duistermars, B.J., Pfeiffer, B.D., Hoopfer, E.D., and Anderson, D.J. (2018). A Brain Module for Scalable Control of Complex, Multi-motor Threat Displays. Neuron 100, 1474-1490 e1474.
17. Glaser, J.I., and Kording, K.P. (2016). The Development and Analysis of Integrated Neuroscience Data. Front Comput Neurosci 10, 11.
18. Haberkern, H., Basnak, M.A., Ahanonu, B., Schauder, D., Cohen, J.D., Bolstad, M., Bruns, C., and Jayaraman, V. (2019). Visually Guided Behavior and Optogenetically Induced Learning in Head-Fixed Flies Exploring a Virtual Landscape. Curr Biol 29, 1647-1659 e1648.
19. Hafting, T., Fyhn, M., Molden, S., Moser, M.B., and Moser, E.I. (2005). Microstructure of a spatial map in the entorhinal cortex. Nature 436, 801-806.
20. Hardcastle, K., Maheswaranathan, N., Ganguli, S., and Giocomo, L.M. (2017). A Multiplexed, Heterogeneous, and Adaptive Code for Navigation in Medial Entorhinal Cortex. Neuron 94, 375-387 e377.
21. Jin, X., and Costa, R.M. (2015). Shaping action sequences in basal ganglia circuits. Curr Opin Neurobiol 33, 188-196.

801 22. Kaplan, H.S., Salazar Thula, O., Khoss, N., and Zimmer, M. (2020). Nested
802 Neuronal Dynamics Orchestrate a Behavioral Hierarchy across Timescales.
803 Neuron 105, 562-576 e569.
804 23. Katsov, A.Y., Freifeld, L., Horowitz, M., Kuehn, S., and Clandinin, T.R. (2017).
805 Dynamic structure of locomotor behavior in walking fruit flies. Elife 6.
806 24. Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., and Poeppel,
807 D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias.
808 Neuron 93, 480-490.
809 25. Kropff, E., Carmichael, J.E., Moser, M.B., and Moser, E.I. (2015). Speed cells in
810 the medial entorhinal cortex. Nature 523, 419-424.
811 26. Lopez-Alonso, V., Cheeran, B., Rio-Rodriguez, D., and Fernandez-Del-Olmo, M.
812 (2014). Inter-individual variability in response to non-invasive brain stimulation
813 paradigms. Brain Stimul 7, 372-380.
814 27. Markowitz, J.E., Gillis, W.F., Beron, C.C., Neufeld, S.Q., Robertson, K., Bhagat,
815 N.D., Peterson, R.E., Peterson, E., Hyun, M., Linderman, S.W., et al. (2018). The
816 Striatum Organizes 3D Behavior via Moment-to-Moment Action Selection. Cell
817 174, 44-58 e17.
818 28. Marques, J.C., Lackner, S., Felix, R., and Orger, M.B. (2018). Structure of the
819 Zebrafish Locomotor Repertoire Revealed with Unsupervised Behavioral
820 Clustering. Curr Biol 28, 181-195 e185.
821 29. McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold
822 Approximation and Projection for Dimension Reduction. In arXiv e-prints, pp.
823 arXiv:1802.03426.

824 30. McLean, D.J.S.V., M.A. (2018). trajr: An R package for characterisation of
825 animal trajectories. Ethology 00, 1-9.
826 31. Moore, J.D., Deschenes, M., Furuta, T., Huber, D., Smear, M.C., Demers, M.,
827 and Kleinfeld, D. (2013). Hierarchy of orofacial rhythms revealed through
828 whisking and breathing. Nature 497, 205-210.
829 32. Robie, A.A., Hirokawa, J., Edwards, A.W., Umayam, L.A., Lee, A., Phillips, M.L.,
830 Card, G.M., Korff, W., Rubin, G.M., Simpson, J.H., et al. (2017). Mapping the
831 Neural Substrates of Behavior. Cell 170, 393-406 e328.
832 33. Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B.L., Witter, M.P., Moser, M.B.,
833 and Moser, E.I. (2006). Conjunctive representation of position, direction, and
834 velocity in entorhinal cortex. Science 312, 758-762.
835 34. Solstad, T., Boccara, C.N., Kropff, E., Moser, M.B., and Moser, E.I. (2008).
836 Representation of geometric borders in the entorhinal cortex. Science 322, 1865-
837 1868.
838 35. Staden, R. (1979). A strategy of DNA sequencing employing computer programs.
839 Nucleic Acids Res 6, 2601-2610.
840 36. Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for
841 transcriptomics. Nat Rev Genet 10, 57-63.
842 37. Wiltschko, A.B., Johnson, M.J., Iurilli, G., Peterson, R.E., Katon, J.M.,
843 Pashkovski, S.L., Abraira, V.E., Adams, R.P., and Datta, S.R. (2015). Mapping
844 Sub-Second Structure in Mouse Behavior. Neuron 88, 1121-1135.

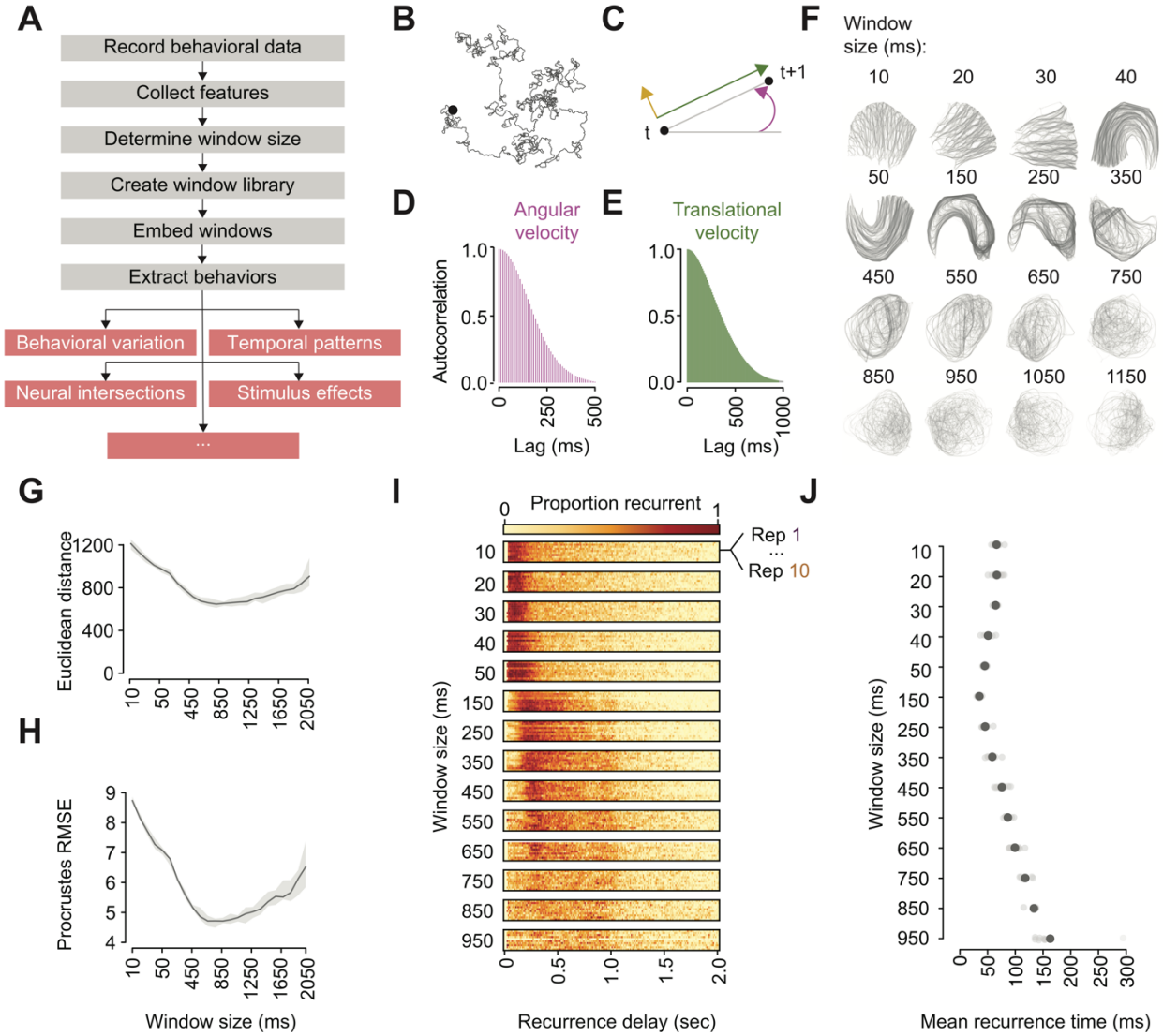


Figure 1 The TREBLE framework and its application to synthetic data

(A) Outline of the TREBLE framework. Behavioral data are recorded. Relevant features (such as velocities) are extracted these recordings following which empirical analyses are used to determine the optimal sampling window size. The chosen window size is then used to create a library spanning all time points in the data set. The resulting window library is embedded into a low-dimensional behavior space from which recurrent behaviors can be decoded and used for a number of analyses (examples labeled in red boxes).

(B) Example correlated random walk used for parameter tuning (see also Figure S1).

(C) Velocity components that can be calculated from movements in a 2-d plane. The black point at time *t* denotes the beginning of the trajectory which proceeds to *t+1*. The purple arrow corresponds to the angular velocity of this trajectory while the green and yellow arrows represent translational velocity and side slip, respectively.

(D) The observed autocorrelation distribution of angular velocity computed from all correlated random walks.

(E) The observed autocorrelation distribution of translational velocity.

(F) Example behavior spaces for a range of window sizes. Spaces are plotted here by connecting temporally adjacent points (corresponding to feature windows), gray lines. Darker lines reflect repeated visits along the same pathway.

(G) Mean intra-point Euclidean distances as a function of window size. Mean values (dark gray); Standard error of the mean (shading).

(H) Procrustes distance RMSE measures as a function of window size. Mean values (dark gray); Standard error of the mean (light gray shading).

(I) Recurrence plot of behavior spaces produced from the correlated random walk dataset. The proportion of recurrent points given a range of time delays spanning from 0 to 2 seconds, is indicated by the color of the corresponding bins (ranging from light yellow to dark red). Each window size contains the distributions for 10 replicate walks.

(J) Mean recurrence times as a function of window size. Population mean (large dark circle); Individual replicates (small light circle).

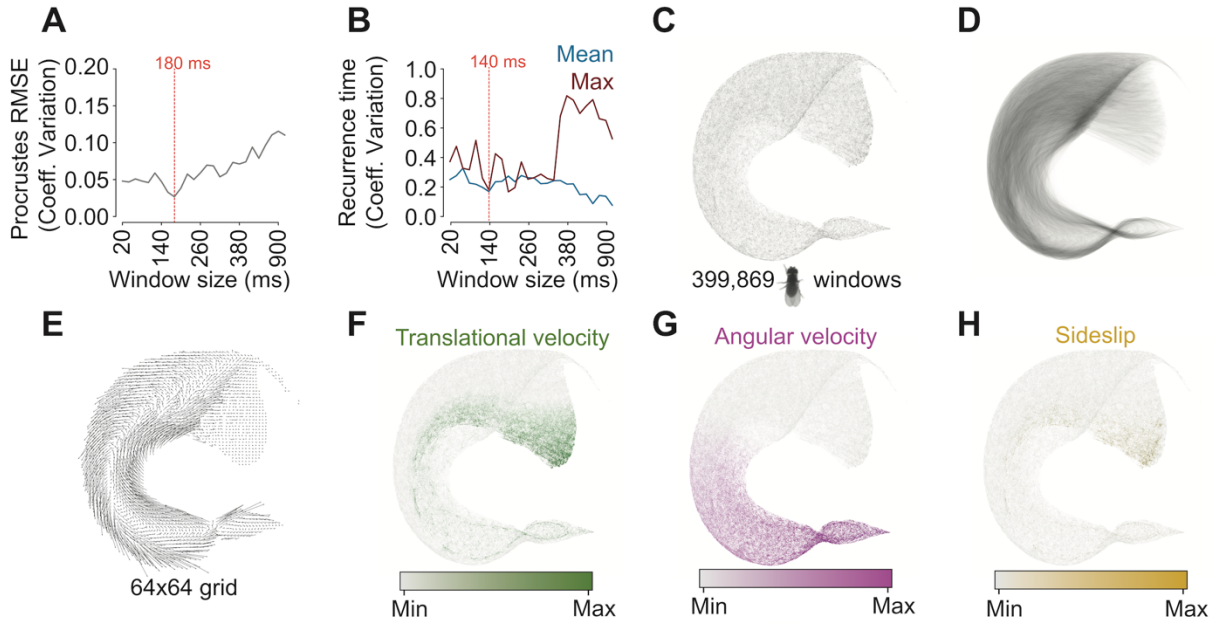


Figure 2 Analyzing fruit fly locomotion with TREBLE

(A) The coefficient of variation of Procrustes distance as a function of window size. The observed minimum at 180ms is denoted by the red dotted line.

(B) Coefficient of variation for the mean (blue line) and maximum (i.e. highest proportion of time bins displaying recurrence; dark red line) recurrence times as a function of window size. The optimal tradeoff between the mean and maximum times is denoted at 140ms with a red dotted line.

(C) Fly locomotor behavior space. Each point corresponds to a feature window (399,869 windows in total) as extracted from the 14 individual flies.

(D) Pathways through the locomotor behavior space, produced by connecting temporally adjacent windows with partially transparent lines (as in Figure 1F).

(E) Walking fruit fly behavior space represented as a vector field. Arrow direction and magnitude correspond to the angle and mean direction taken after visiting each bin.

(F) The distribution of translational velocity across behavior space. Darker green corresponds to larger values of translational velocity.

(G) The distribution of angular velocity across behavior space. Darker purple corresponds to larger values.

(H) The distribution of side slip across behavior space. Darker gold corresponds to larger values.

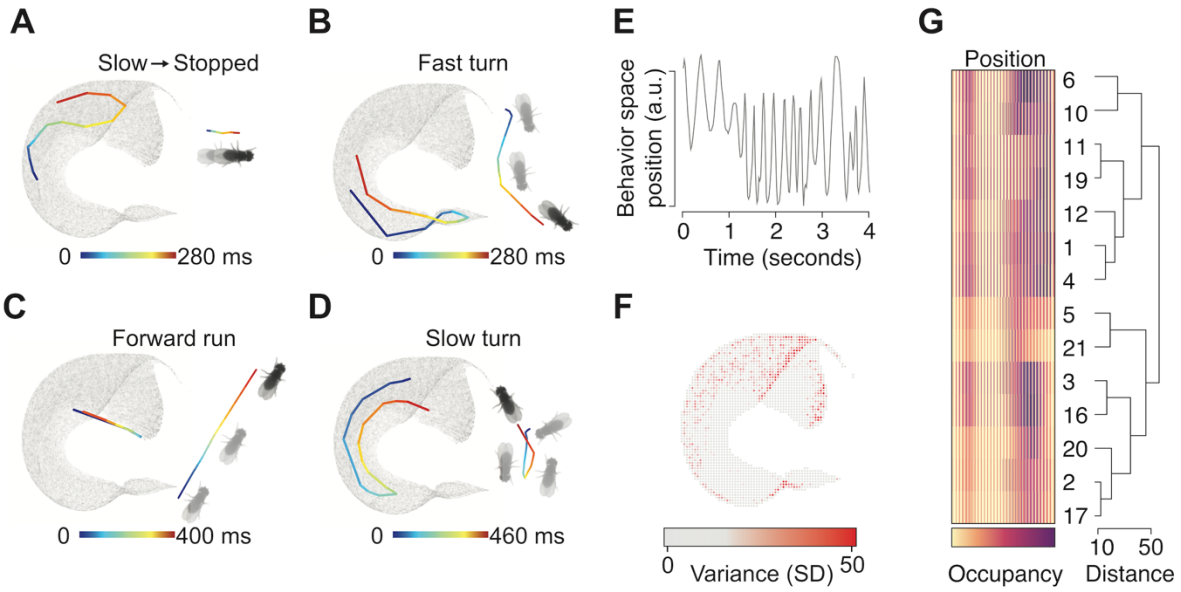


Figure 3 Decoding and comparing behavior across many individuals

(A) Example bout of stopping. The pathway through behavior space is represented on the left. XY coordinates of the actual bout are plotted on the right. Both representations are colored by elapsed time.

(**B-D)** Same as **(A)** but for fast turns, forward runs, and slow turns, respectively.

(E) Sample trajectory through behavior space as represented by a 1-dimensional coordinate value.

(F) Distribution of per-bin standard deviation across locomotor behavior space. Color corresponds to the variance of bin-wise visitation frequency across all 14 flies.

(G) Hierarchical clustering of individual fly density maps. The dendrogram on the right represents the relationships between all 14 flies (numbered at the tips of each branch). Individual fly density maps are presented on the left in 1-dimension by converting the density map matrix into a single vector (presented in 2-dimensions in Figure S7). Darker colors reflect more time spent ('occupancy') in a specific region.

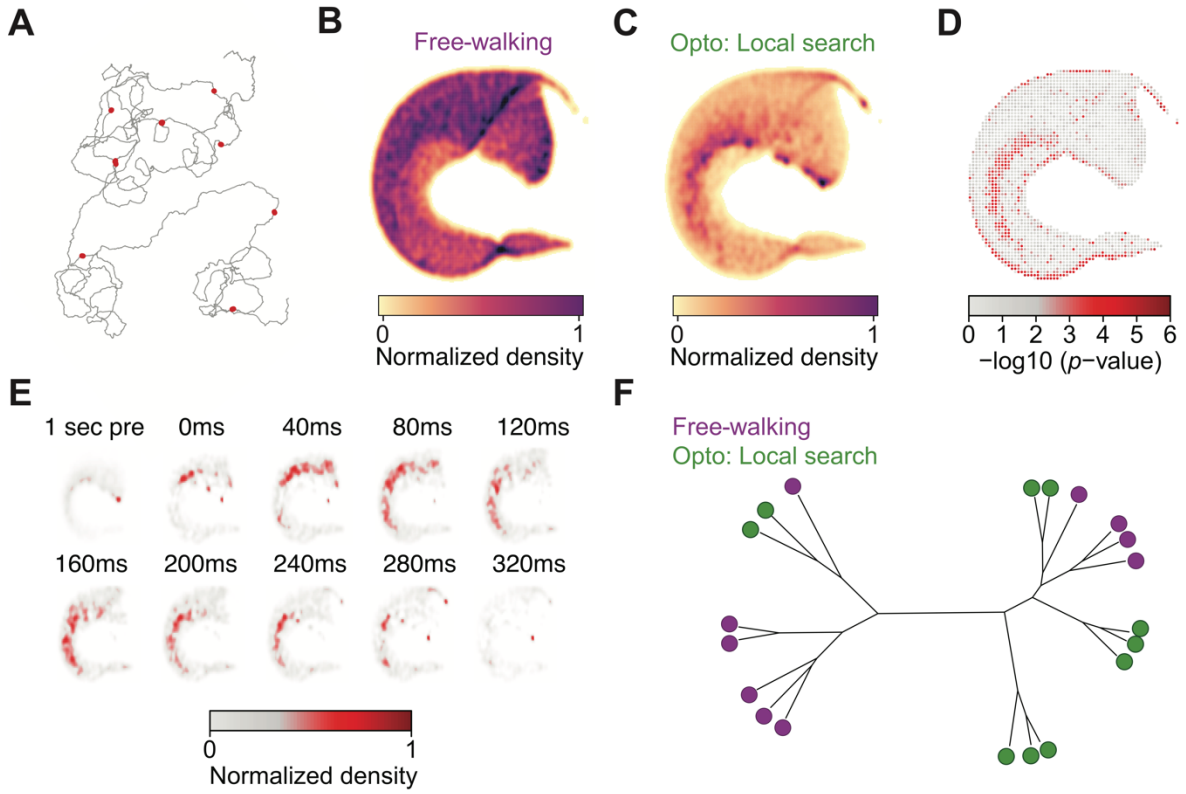


Figure 4 Identifying behavioral perturbations via optogenetic maniuplation

(A) Example trajectory for a single optogenetic trial. The fly's path through the virtual world is represented by the dark grey line. Red dots correspond to locations in which optogenetic stimulation was presented.

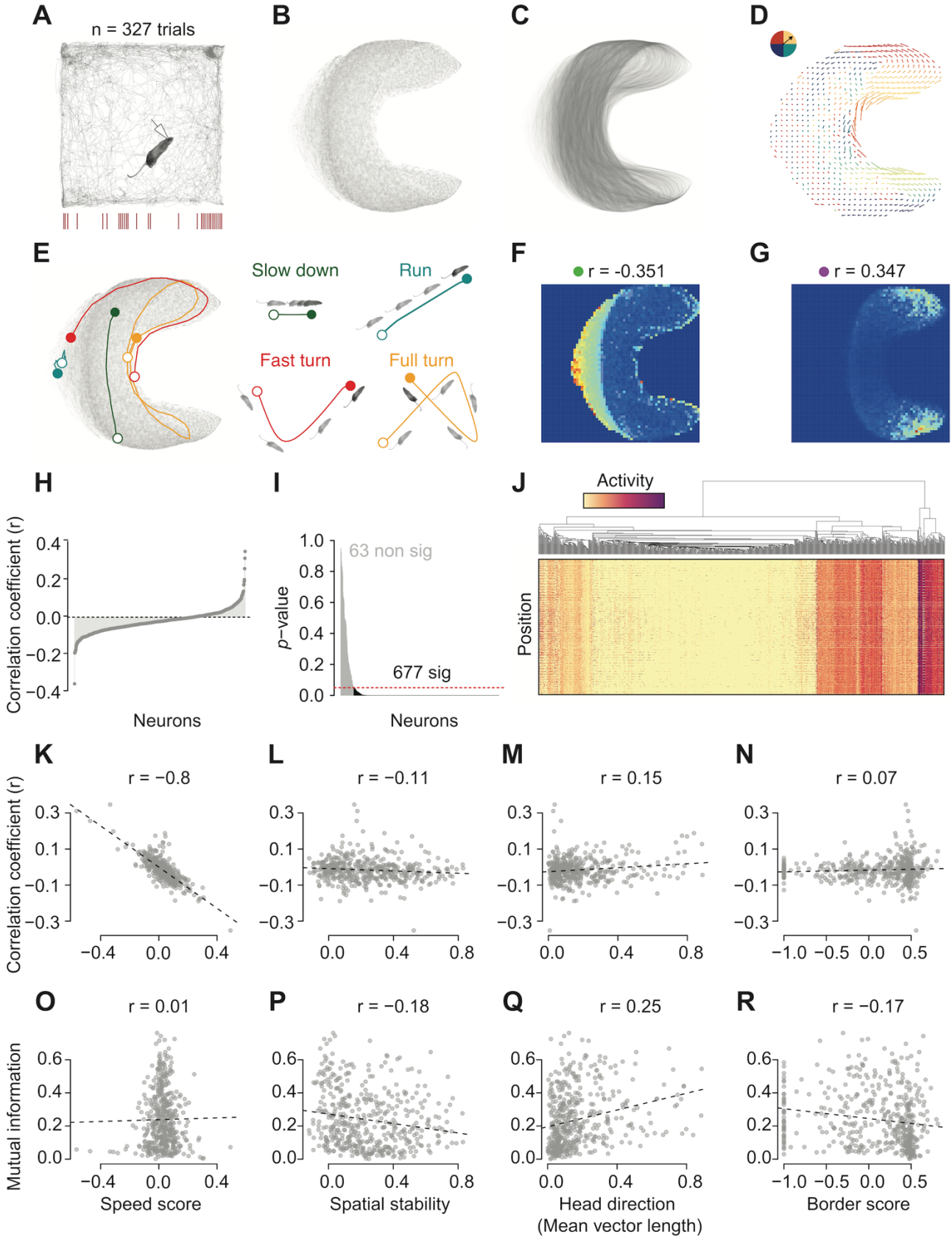
(B) Average density map computed from the 14 free-walking control flies (darker color corresponds to more time spent in a specific region).

(C) Average density map computed from the 19 optogenetically activated flies.

(D) Bin-wise differences between control and optogenetically activated flies. Color corresponds to the significance (-log10 transformation of the *p*-value; Kruskal-Wallis test) of the differences in density between the two groups. Darker red corresponds to increasingly significant differences.

(E) Time evolving responses to optogenetic stimulation in behavior space. Each space reflects a specific time window before (first space) and after stimulation (each subsequent space). Color corresponds to the average response to stimulation across all 19 flies (represented as normalized density; see methods). Windows span 1 second of time, beginning at the value represented above each space.

(F) The behavioral relationships between all free-walking control and optogenetically activated flies. The tree represents the results of hierarchical clustering on the density maps of each individual fly (as in Figure 2O). Each branch tip is associated with an individual fly, group is represented by color (Controls = purple; Optogenetic trials = dark green).


A
n = 327 trials
B
C
D
E
Slow down
Run
Fast turn
Full turn
F
r = -0.351
G
r = 0.347
H
Correlation coefficient (r)
Neurons
I
63 non sig
677 sig
p−value
Neurons
J
Activity
Position
K
r = −0.8
Correlation coefficient (r)
L
r = −0.11
M
r = 0.15
N
r = 0.07
O
r = 0.01
Mutual information
Speed score
P
r = −0.18
Spatial stability
Q
r = 0.25
Head direction
(Mean vector length)
R
r = −0.17
Border score

975

Figure 5 Mapping mouse locomotion and navigational coding

(A) Example trajectory for of a foraging mouse with paired spike recordings via an implanted electrode (drawn).

(B) Mouse locomotor behavior space, each point corresponds to a temporal window.

(C) Pathways through mouse locomotor behavior space, produced by connecting temporally adjacent windows with partially transparent lines.

(D) Mouse locomotor behavior space represented as a vector field. Arrow direction and magnitude correspond to the angle and mean direction taken after visiting each bin. Arrows are colored by the degree of the direction vector (corresponding to circle in upper left-hand corner).

(E) Example behavior bouts decoded from behavior space. Pathways through behavior space are plotted on the left. Bout starts are indicated by the open circle, ends by the closed circle. Positions in real XY space are plotted on the right.

(F) Example 2-d tuning curve of a neuron that is negatively correlated with behavior space (activity measured as spikes/second; labeled with a green dot for reference in later figures). Colors range from blue to red, corresponding to minimum activity (blue) to maximum activity (red). Pearson correlation (r) is denoted above the plot.

(G) Example 2-d tuning curve of a neuron that is positively correlated with behavior space (activity measured as spikes/second; labeled with a purple dot for reference in later figures). Colors range from blue to red, corresponding to minimum activity (blue) to maximum activity (red). Pearson correlation (r) is denoted above the plot.

(H) Distribution of correlation coefficients between neural activity and behavior space position for all MEC neurons (activity measured as spikes/second).

(I) Distribution of *p*-values (plotted as -log10 (*p*-value)) resulting from permutation tests (10,000 shuffles) of the correlations plotted in **(H)**. The red dotted line indicates $p = 0.05$.

(J) Clustered 2-d tuning curves for all neurons. The heatmap corresponds to a linearized version of the 2-d tuning curve, colored by average activity as a function of position in behavior space.

(K) Scatterplot of the relationship between the correlation coefficient of the relationship between behavior space and neural activity (y-axis) and speed score (x-axis). Pearson correlation (r) is denoted above the plot.

(L) Scatterplot of the relationship between the correlation coefficient and spatial stability.
(M) Scatterplot of the relationship between the correlation coefficient and head direction (HD) mean vector length.
(N) Scatterplot of the relationship between the correlation coefficient and border score.
(O) Scatterplot of the relationship between the mutual information (MI) of behavior space and neural activity (y-axis) and speed score (x-axis).
(P) Scatterplot of the relationship between MI and spatial stability.
(Q) Scatterplot of the relationship between MI and head direction (HD) mean vector length.
(R) Scatterplot of the relationship between MI and border score.

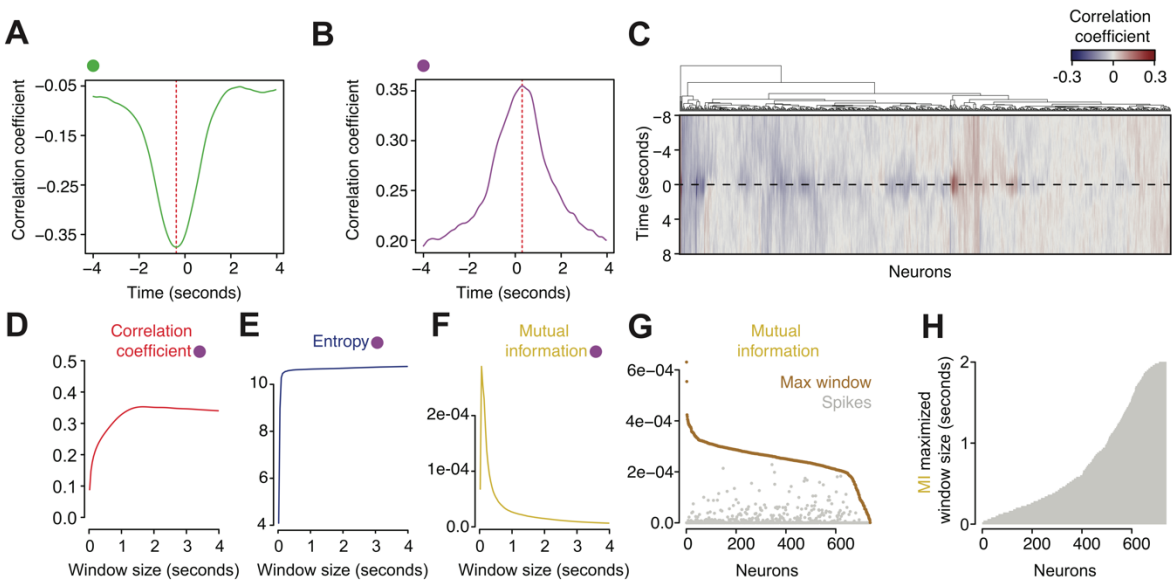


Figure 6 Temporal variation in mouse navigational coding

(A) Cross-correlation distribution of a neuron that is negatively correlated with behavior space (same as in **4F**)

(B) Cross-correlation distribution of a neuron that is positively correlated with behavior space (same as in **4G**).

(C) Hierarchical clustering of cross-correlation distributions across all neurons. Color corresponds to the correlation coefficient, blue indicating negative and red indicating positive correlations. The y-axis represents time before and after instantaneity, ranging from 8 seconds before to 8 seconds after. The x-axis corresponds to the relationships of all MEC cells based on their cross-correlations with behavior (which are shown via the colored heatmap). In essence, each column contains the same information as **Figs 6A-B** but here represents the distributions using variation in color over time, rather than a 1-d tuning curve.

(D) Correlation coefficient distribution as a function of varied window sizes for the neuron plotted in **4G** and **5B**.

(E-F) same as **(D)** but for entropy **(E)** and mutual information (MI).

(G) Distribution of MI for all neurons, comparing the maximally informative window size (in gold) and that measured from spikes (grey).

(H) Barplot of the distribution of MI maximized window sizes for all neurons.

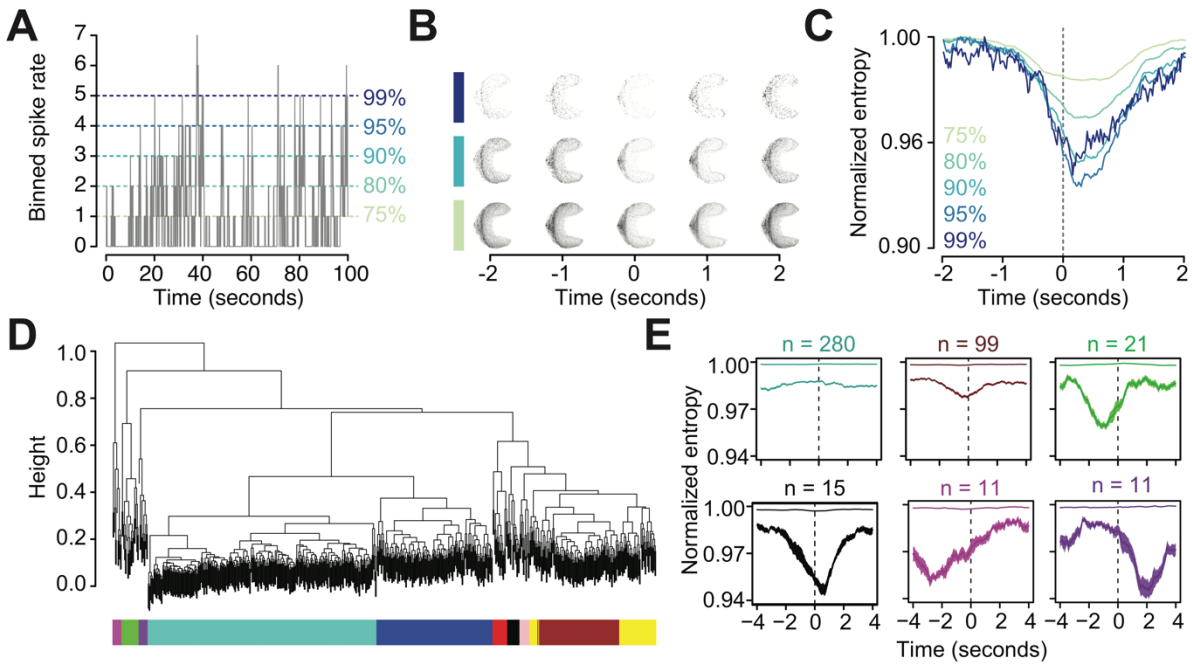


Figure 7 Rate-based coding of locomotion in medial entorhinal cortex

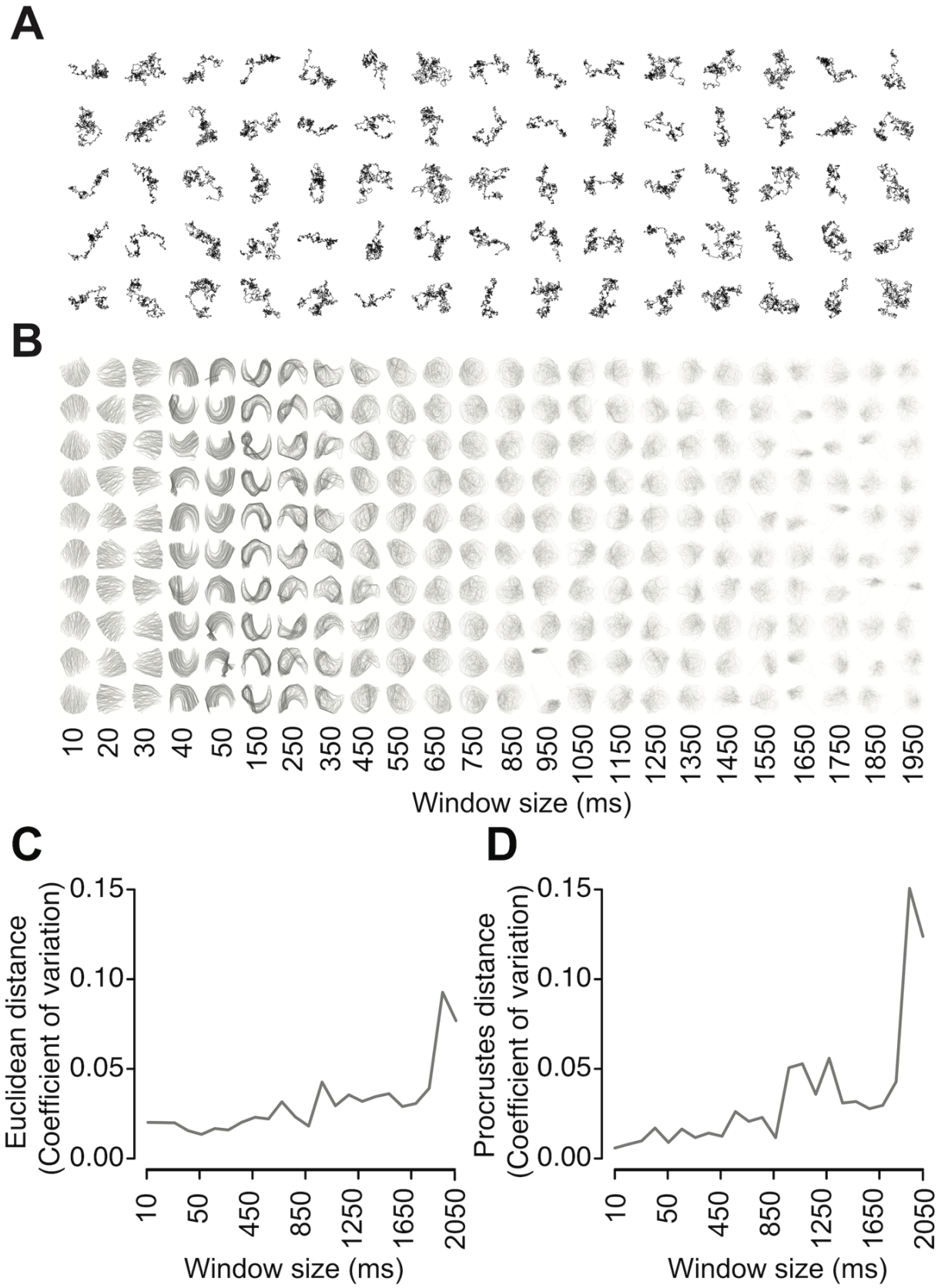
(A) Example neural activity trace (using MI maximized window size) and corresponding percentile cutoffs (at 75%, 80%, 90%, 95%, and 99%).

(B) Temporal distributions in behavior space (1 second windows) surrounding peaks in activity (75%, 90%, and 99% cutoffs).

(C) Entropy distributions of surrounding activity bouts corresponding to each percentile (labeled and colored on plot). Normalized entropy measures are shown to be able to compare across percentiles with different sample sizes.

(D) Hierarchical clustering of entropy distributions across all neurons. Colored bars indicate clusters as measured using dynamic tree trimming (see Methods).

(E) Example mean entropy distributions for four clusters identified using dynamic tree trimming. Plotted are the distributions for the 75% (top) and 99% (bottom) cutoffs. Mean and standard error are plotted as well as sample sizes (above).


A
B
10
20
30
40
50
150
250
350
450
550
650
750
850
950
1050
1150
1250
1350
1450
1550
1650
1750
1850
1950
Window size (ms)
C
Euclidean distance
(Coefficient of variation)
0.15
0.10
0.05
0.00
10
50
450
850
1250
1650
2050
Window size (ms)
D
Procrustes distance
(Coefficient of variation)
0.15
0.10
0.05
0.00
10
50
450
850
1250
1650
2050
Window size (ms)

1078

Figure S1 Correlated random walk iterative window tests

(A) A set of 100 example correlated random walks used for parameter tuning. Each is 10,000 frames long and sampled at 100 FPS.

(B) Trajectories through behavior space are plotted for each of the 10 correlated random walk replicates per window size tested. Replicates are arranged in rows while window sizes are arranged in columns. Trajectories were visualized by connecting temporally adjacent points in behavior space with lines. Recurrent dynamics are represented by highly overlapping lines, reflecting repeated excursions through those specific paths.

(C) Coefficient of variation of mean intra-point Euclidean distance as a function of window size.

(D) Coefficient of variation of Procrustes distance RMSE as a function of window size.

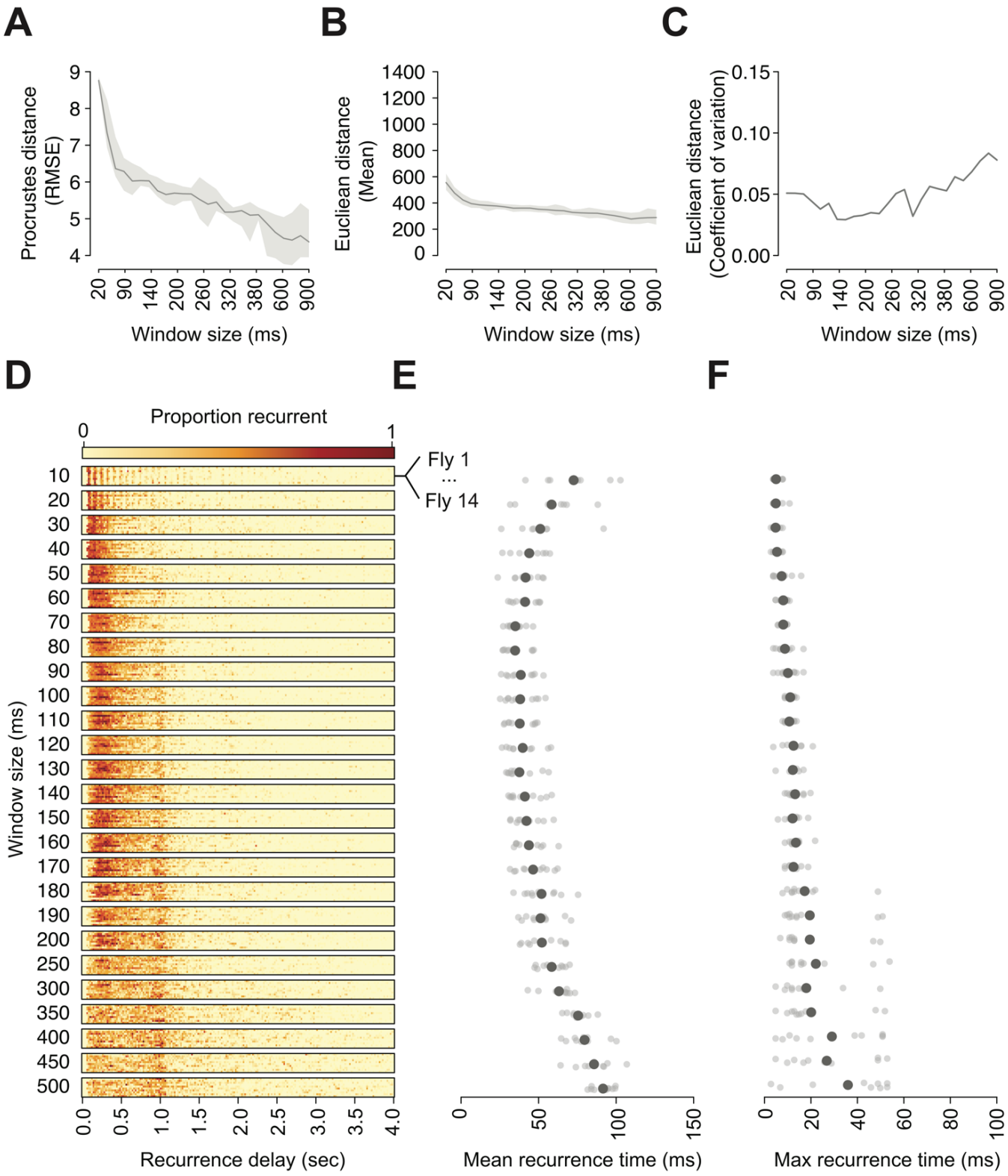

A
Procrustes distance (RMSE)
Window size (ms)
B
Euclean distance (Mean)
Window size (ms)
C
Euclean distance (Coefficient of variation)
Window size (ms)
D
Proportion recurrent
Fly 1
...
Fly 14
Window size (ms)
Recurrence delay (sec)
E
Mean recurrence time (ms)
F
Max recurrence time (ms)

Figure S2 Assessing structural and temporal components of the free-walking fly dataset

(A) Procrustes distance RMSE measures as a function of window size. The darker line corresponds to the mean value while the shading reflects standard error.

(B) Mean intra-point Euclidean distances. The darker line corresponds to the mean value while the shading reflects standard error.

(C) Coefficient of variation of mean intra-point Euclidean distance as a function of window size.

(D) Recurrence plot of behavior spaces produced from the free-walking dataset. The proportion of recurrent points given a range of time delays ranging from 0 to 4 seconds is indicated by the color of the corresponding bins (from light yellow to dark red). Each horizontal window size bar includes all 14 flies tested.

(E) Mean recurrence times for all 14 flies as a function of window size. The larger circle corresponds to the population mean while smaller circles correspond to each fly.

(F) Maximum recurrence times for all 14 flies as a function of window size. Each point reflects the bin in which the largest proportion of points displayed recurrence. The larger circle corresponds to the population mean while smaller circles correspond to each fly.

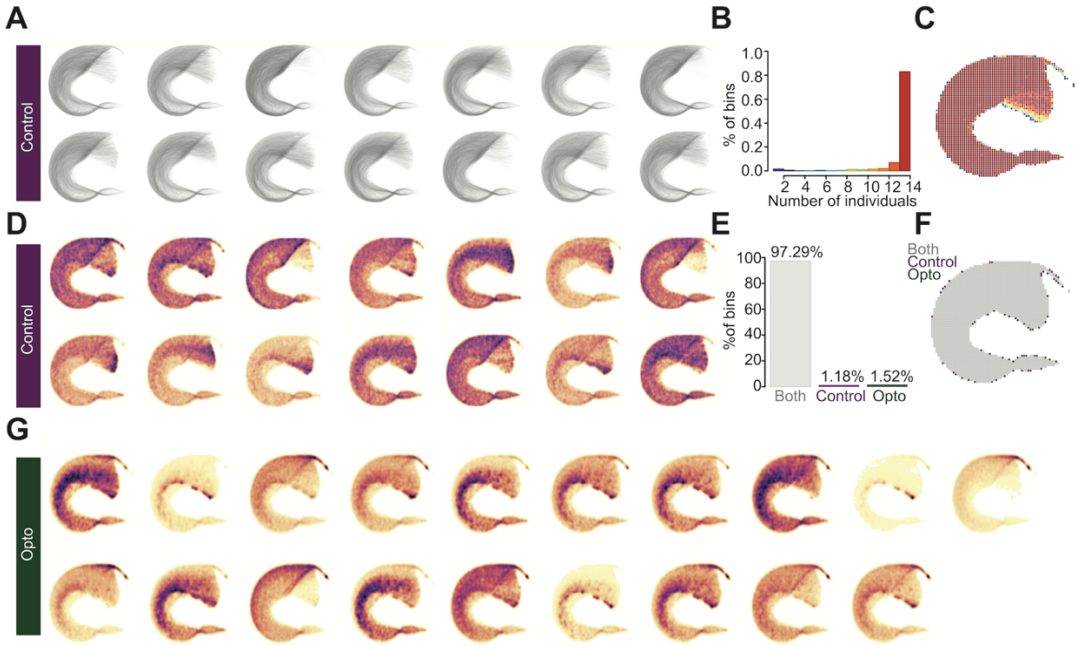


Figure S3 Comparing control and optogenetic fly behavior space distributions

(A) Pathways through the locomotor behavior space for all 14 control flies, produced by connecting temporally adjacent windows with partially transparent lines.

(b) Bar plot comparing the percent of behavior space bins that were visited by at least *n* individuals. Bars are color coded from blue to red to correspond to **(c)**.

(c) Locomotor behavior space binned to a 64x64 grid. Each bin is colored corresponding to the number of individuals that visited it.

(d) Density maps for all control flies. Each behavior space corresponds to an individual and is colored by the kernel density estimate generated for that fly's trajectory in behavior space. Darker color corresponds to a greater density in a given region.

(E) Bar plot comparing the percentage of bins in a 64x64 gridded behavior space that were visited by control flies (purple), optogenetically activated flies (green), or both (grey).

(F) Behavior space colored by the distribution of overlapping and unique bins. Each bin is colored by the designations in the bar plot.

(G) Density maps for optogenetically activated flies. Each behavior space corresponds to a unique fly (n = 19) and is colored by the kernel density estimate generated for that fly's trajectory through behavior space. Darker colors correspond to a greater density in a given region.

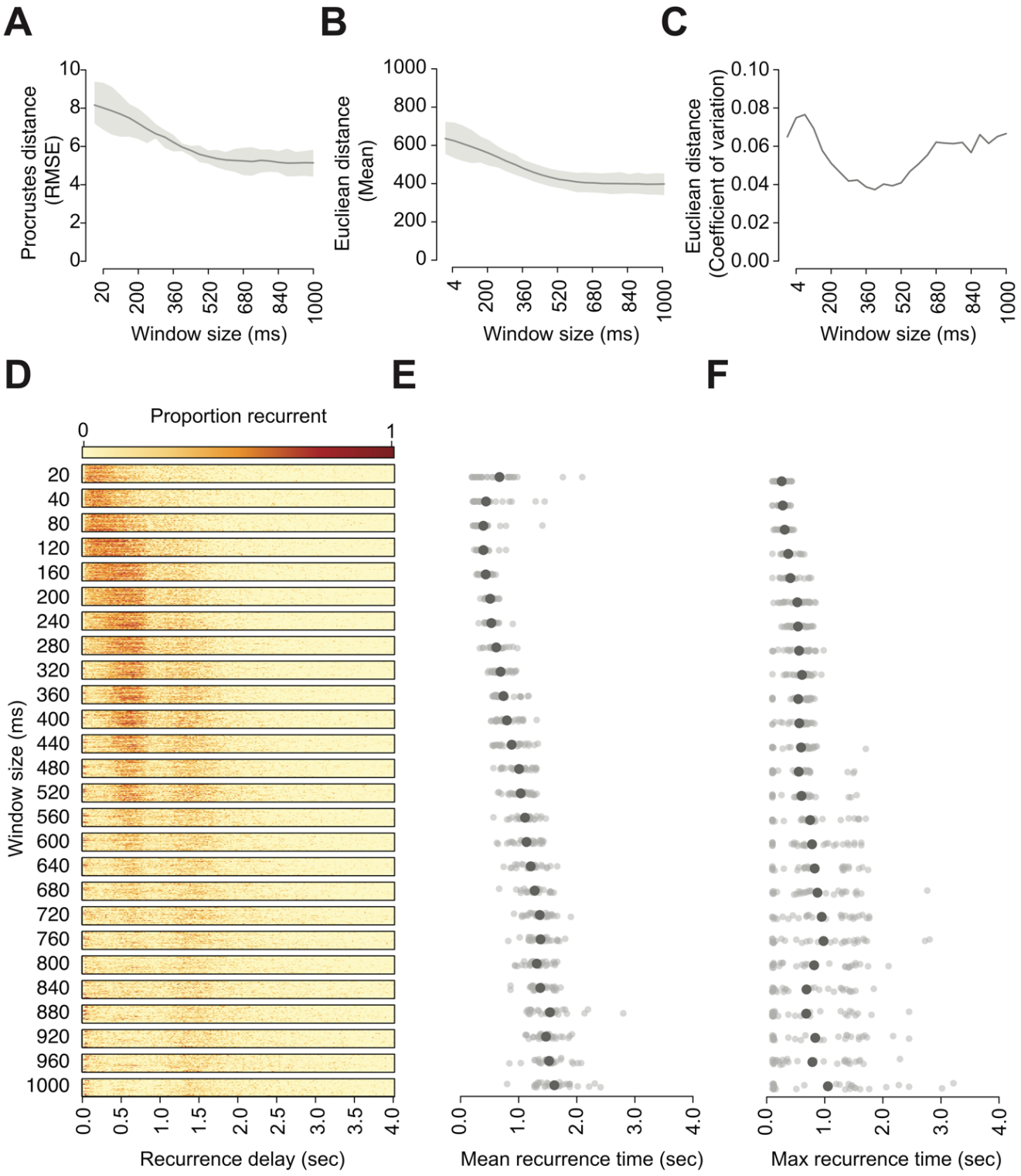

A
Procrustes distance (RMSE)
Window size (ms)
B
Euclidean distance (Mean)
Window size (ms)
C
Euclidean distance (Coefficient of variation)
Window size (ms)
D
Proportion recurrent
Window size (ms)
Recurrence delay (sec)
E
Mean recurrence time (sec)
F
Max recurrence time (sec)

Figure S4 Structural and temporal components of mouse locomotor space

(A) Procrustes distance RMSE measures as a function of window size. The darker line corresponds to the mean value while the shading reflects standard error.

(B) Mean intra-point Euclidean distances. The darker line corresponds to the mean value while the shading reflects standard error.

(C) Coefficient of variation of mean intra-point Euclidean distance as a function of window size.

(D) Recurrence plot of behavior spaces taken from 30 random mouse trials. The proportion of recurrent points given a range of time delays ranging from 0 to 4 seconds is indicated by the color of the corresponding bins (from light yellow to dark red). Each horizontal window size bar includes all 30 mice tested.

(E) Mean recurrence times as a function of window size. The larger circle corresponds to the population mean while smaller circles correspond to each mouse trial.

(F) Maximum recurrence times as a function of window size. Each point reflects the bin in which the largest proportion of points displayed recurrence. The larger circle corresponds to the population mean while smaller circles correspond to each mouse trial.

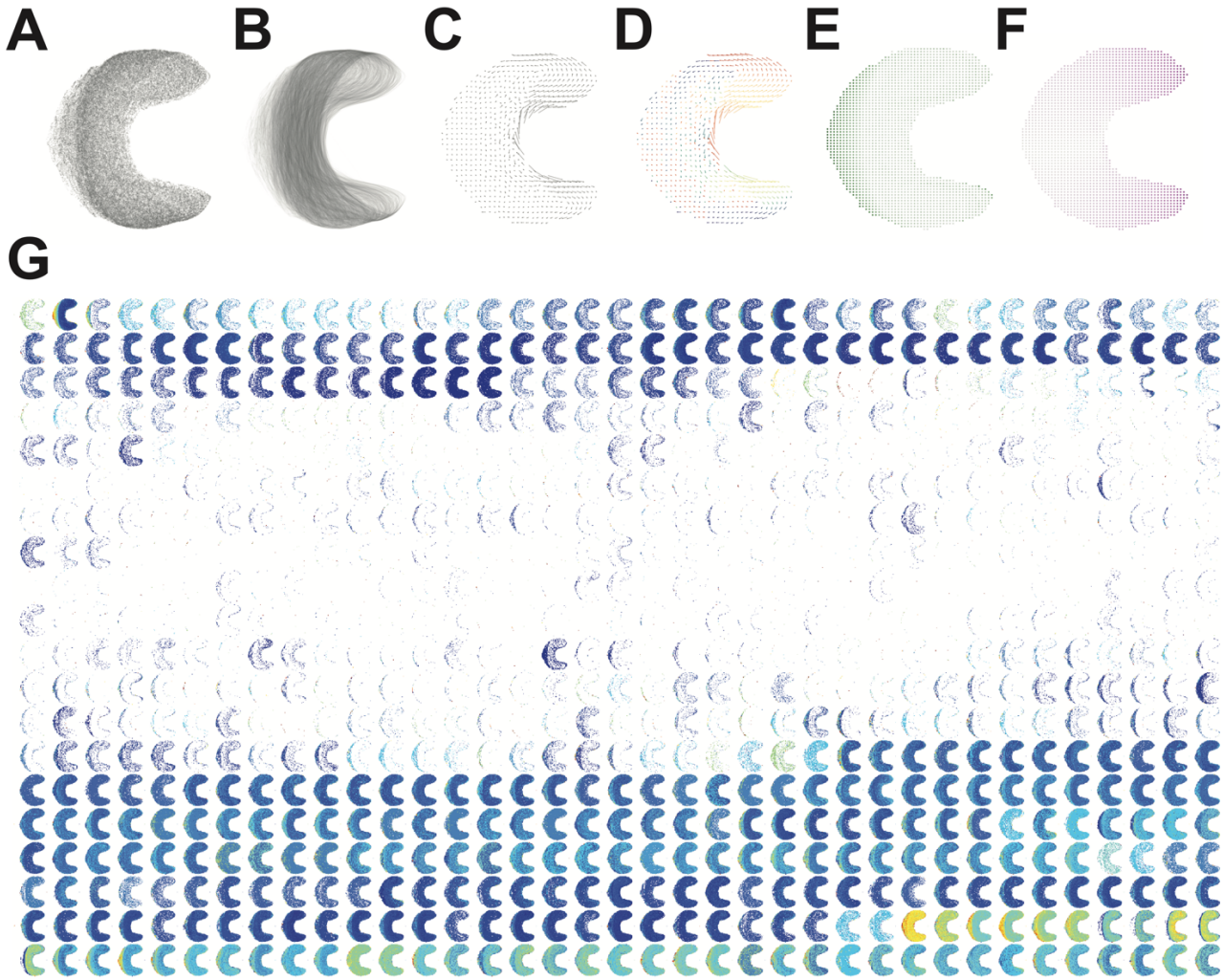


Figure S5: The association of mouse behavior space and MEC neuron activity

(A) Mouse locomotor behavior space, each point corresponds to a temporal window.

(B) Pathways through mouse locomotor behavior space, produced by connecting temporally adjacent windows with partially transparent lines.

(C) Mouse locomotor behavior space represented as a vector field. Arrow direction and magnitude correspond to the angle and mean direction taken after visiting each bin.

(D) Arrow direction and magnitude correspond to the angle and mean direction taken after visiting each bin. Arrows are colored by the degree of the direction vector (corresponding to circle in upper left-hand corner).

(E) Distribution of translational velocity across mouse behavior space (darker green corresponds to higher values).

(F) Distribution of angular velocity across mouse behavior space (darker purple corresponds to higher values).

(G) 2-d tuning curves of neural activity (1 second bins) for all MEC neurons. The ordering of tuning curve position was determined by hierarchical clustering (i.e. more similar tuning curves are placed adjacent to each other).

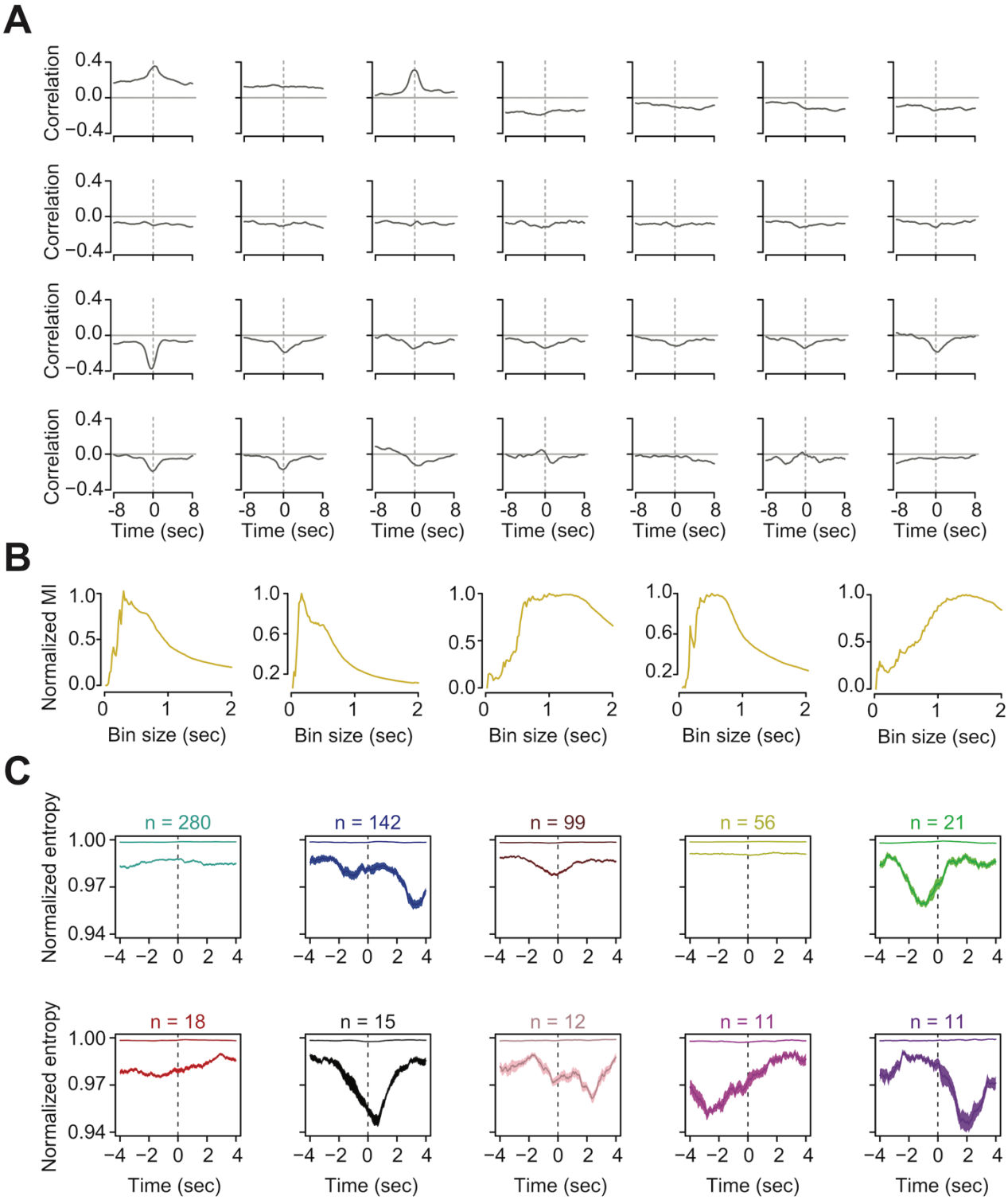

A
Correlation
0.4
0.0
−0.4
-8 0 8
Time (sec)
B
Normalized MI
1.0
0.5
0.0
0.6
0.2
0 1 2
Bin size (sec)
C
Normalized entropy
1.00
0.97
0.94
−4 −2 0 2 4
Time (sec)
n = 280
n = 142
n = 99
n = 56
n = 21
n = 18
n = 15
n = 12
n = 11
n = 11

Figure S6: Variation in temporal and rate-based associations between MEC neurons and behavior

(A) Example cross-correlations between behavior space position and MEC neuron activity over a 16 second time window. The dotted grey line corresponds to time point zero. The solid grey line reflects a correlation coefficient of zero.

(B) Example mutual information (MI) tuning curves as a function of binning size of neural activity. For comparison purposes, MI is presented as a normalized value in which each distribution has been divided by its maximum value.

(C) Mean entropy distributions of ten clusters identified using dynamic tree trimming. Plotted are the distributions for the 75% (top) and 99% (bottom) cutoffs. Mean and standard error are plotted as well as sample sizes, as in Figure 7E.