Chapter 4

Characteristic exploratory behavior persists as individual *Drosophila* become hungry

4.1 Summary

Here we show early attempts to demonstrate the individuality of adult *Drosophila* reared and observed in homogeneous conditions. Using a simple machine vision strategy to track the movements of single flies within model environments, we describe a characteristic structure in the movements of individuals making up their exploration and dispersal. The characteristic structure persists over the period of hours and is robust to systematic shifts in the movement of these flies over this time that are presumably due to entrained crepuscular activity and changes in their hunger state.

4.2 Introduction

We have previously suggested that hunger overrides the visual and olfactory cues from food, driving the fruit fly, *Drosophila melanogaster*, to disperse from inaccessible food

patches. To describe this movement, which seems largely driven by the animal's changing internal physiological state, in terms of behavioral algorithms, we mounted single cameras above the same environmental chambers previously used and developed a simple machine vision strategy to reconstruct the 3D trajectory of single, isolated flies moving within these chambers. We started recording the movements of the flies just after they had been removed from food, and therefore we captured onto digital video the change in the behavior as the flies became hungry. Upon analyzing the search movements of flies near a water resource as they shift to exploring and then to exiting from the chamber to an adjoining chamber, we noticed a surprising non-uniformity in their movement (see Figure 4.1).

As a starting point to determine whether flies exhibit *individualistic* exploratory movement over the period of hours, we learn a function that inputs a quantitative description of the behavior of a fly during one time period and predicts this description during another period. We then show for a number of behavioral statistics describing the movement of exploring flies, that the error in this prediction is significantly lower than for a control experiment, in which we try to predict the behavior of a fly given the behavior of a different fly.

Within this terse introduction of the project, we analyze 1 of 4 collected data sets: (1) 34 males and 34 females over 6 hours that may freely move between the chambers. We are also working with, but largely do not mention, results from the three other data sets: (2) 10 females over 6 hours blocked from moving between chambers by a visually transparent window allowing the flies to see out of the chamber, (3) 16 males and 14 females over 12 hours that are shut within the first chamber with a plug made to appear as near a possible as more of the chamber wall, effectively acting as a *single* chamber, and (4) 12 hours of behavior for 9 females that have been food deprived for 3 hours

and 9 females that have been food deprived for 12 hours that may pass back and forth between the connected chambers.

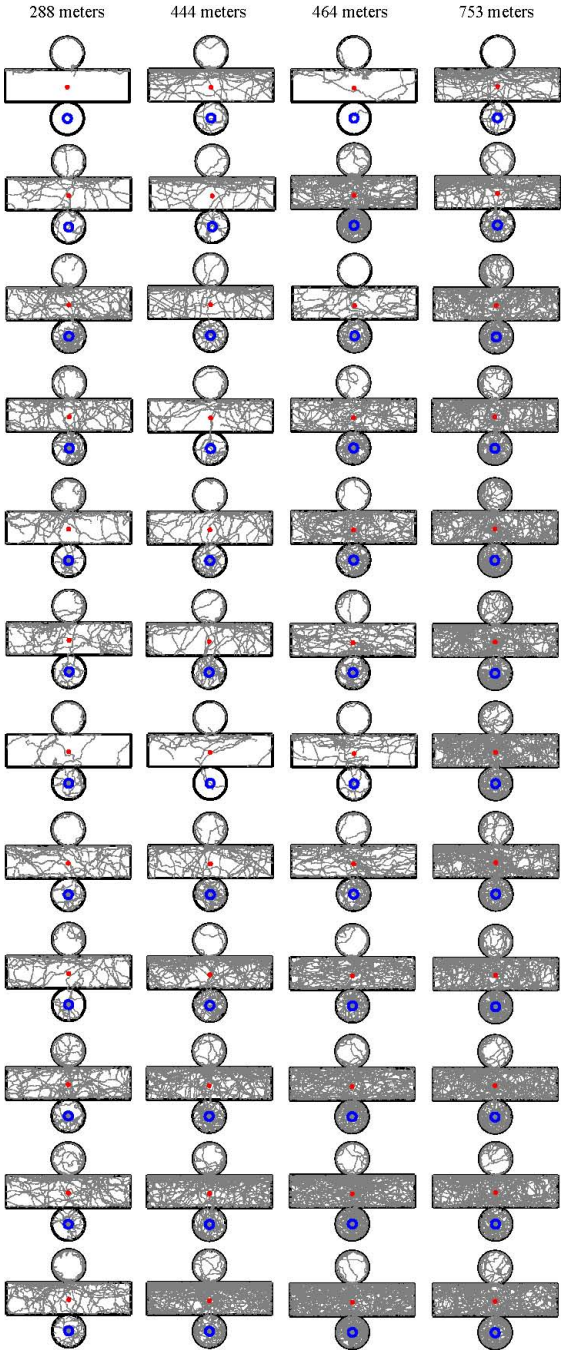

Figure 4.1: Individual trajectories of isolated, single flies moving within a single cylindrical chamber for 12 hours. Shown are the hourly movements of four individual flies throughout the 12 hour trials (top to bottom). The trajectories from two flies come from trials where flies walked for a medium total distance (444 and 464 meters), whereas the other two come from flies that had walked two standard deviations shorter (288 meters) or further (753 meters) than the medium distance. To help illustrate the trajectory of a fly, we unwrapped its 3D positions within the experimental chambers and report its movement in a flattened representation. A patch of agar (blue circle) embedded within the center of floor prevented flies from dehydrating. The exit (red dot) leading to a connected second chamber was blocked during these particular trials.

4.3 Materials and Methods

4.3.1 Animals and their handling

We performed experiments on 4-day-old adults from a laboratory colony of the fruit fly, *Drosophila melanogaster* (Meigen), descended from a wild-caught population of 200 females. We reared, entrained, and tested all flies on a 16 h: 8 h light: dark photoperiod. Transitions between light and dark were immediate. The light-on phase started at 7AM PST. We maintained fly stocks at $\approx 25\,^{\circ}\mathrm{C}$ and at a relative humidity of $\approx 30\%$ on Lewis food medium in standard 250 mL bottles (Lewis, 1960). We introduced individual flies from stock vials directly into the experimental chambers with a mouth pipette. The stock vials were kept within a controlled density of flies containing an equal mixture of males and females of comparable age, which were provided *ad libitum* access food.

4.3.2 Long-duration recordings of movement within environmental chambers

To test the exploratory behavior of individual flies, we introduced single flies fed *ad labitum* into the first of a pair of connected environmental chambers where flies could move freely between the chambers through narrow tubes, as described previously. Each day we ran two trials, simultaneously observing a single male and a single female. We switched back and forth each day which pair of chambers contained the male or the female, and after every experiment, we washed down the chambers with water and dilute ethanol. All chambers provided access to a 2 mL plug of 0.5% agar that was embedded into the center of the floor to prevent dehydration. We introduced flies into chambers at 9AM and started observing their movements immediately, until 4PM or 10PM. Flies

were free to move back and forth between the adjacent chambers, unless in the particular trial the exit to the second chamber was blocked. To record the change in behavior as a fly became hungry, we mounted digital cameras above the first chambers and recorded the fly's movement at a rate of 15 s^{-1} (Fig. 4.2A). To avoid large video files, instead of saving a full record of their movement, we extracted and saved only the x,y coordinates of the fly as observed within the 2D image plane (Fig. 4.2B), a corresponding cropped image containing just the region surrounding the fly for each of these coordinates (Fig. 4.2C), and a single median background image calculated from XXX frames from the video containing a view of the entire chamber (Straw and Dickinson, 2009). With this data we have developed software capable of reconstructing a high spatial and temporal 3D representation of the fly's movement that is cross-indexed to each original video. Data in this form allow us to confirm the quality of tracking and also provide an efficient means to extract movie clips of interesting behaviors, or over specified time windows, for further analysis.

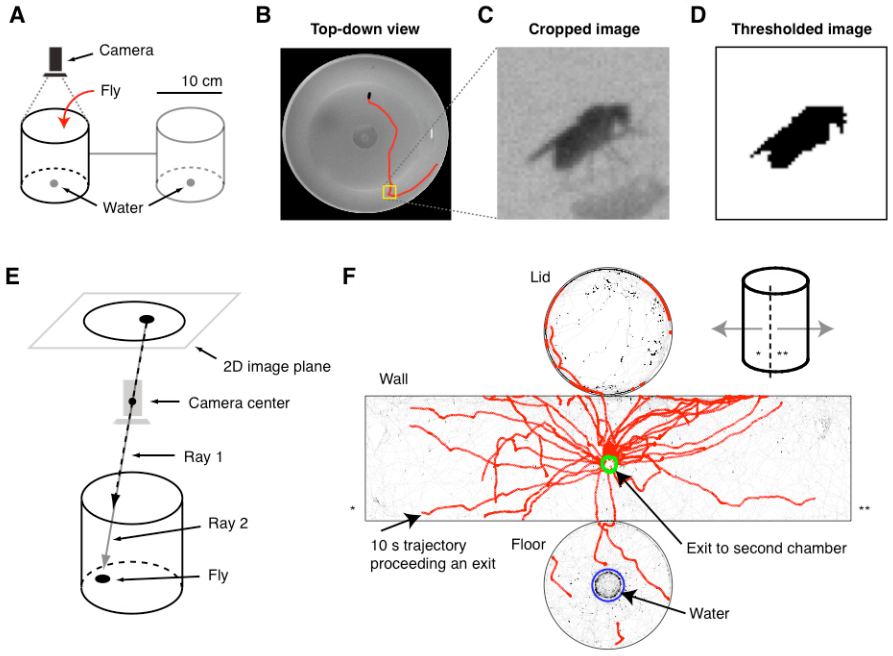

Figure 4.2: Technology devised to study the movement of *Drosophila* within a controlled sensory environment. (A) An individual fly is introduced to the first of two connected chambers, both containing only water. A single camera is mounted above the first chamber. (B) An example of the 2D position of a fly obtained from digital video using custom software from the laboratory. (C) Cropped image of the fly from this video. (D) Binerized threshold image of the fly in C. (E) Cartoon illustrating the two possible locations for a fly from the perspective of the camera. *Ray 1* represents the possible location of a fly on the underside of the chamber lid (dashed; black arrow); *Ray 2* indicates the true location of the fly in this illustration (solid; gray arrow), which sits on the chamber floor. (F) Reconstructed trajectory of a fly filmed for 6 hours as it became hungry (black). We highlighted 10 second segments for all trajectories proceeding exits (red) into the second chamber (green circle). To help visualize and compare the variability in movement between individual flies, we rotated and unwrapped their trajectories so that the exit hole is aligned directly opposite from the readers. We made a vertical slice down the front section of the chamber (dotted line) and then folded the lid up, the floor down, and walls apart as shown. Note * and ** denote corresponding sections of wall; see inset and reconstructed trajectory.

4.3.3 Single-camera strategy for three-dimensional video-tracking

To reconstruct the 3D positions of a fly as it moved throughout a chamber, we calculated the total pixel area representing a fly from a thresholded image (Fig. 4.2D) for each cropped image. We used this information together with 2D coordinates to deduce the location of the fly throughout the length of a video. If the pixel area representing a fly was greater than a specific computed amount, we would assume that the fly at this point in time was closer to the camera and therefore on the underside of the chamber lid; conversely, if the pixel area was less than this amount, we would assume that the fly was farther from the camera, either on the wall or floor of the chamber (Fig. 4.2E). Our strategy assumes in accordance with our observations that single flies introduced to our experimental chambers spent the majority of their time on the surface of the chamber, rather than flying within its volume. From over 1200 hours of video, isolated flies remained on the surface of the chamber for more than 99% of the time. Using this strategy, we could build up a fly's trajectory frame-by-frame over 6 or 12 hours. We calibrated the projection between the 2D coordinates of a fly and its 3D positions using known anchor points. The points were assigned within an image of the experimental chamber corresponding with known positions within the chamber, using a direct linear transformation (standard DLT). Finally, we estimated the most likely sequence of positions (lid vs. wall or floor) for the fly between each video image and used this estimate to reconstruct the trajectory for a fly (Viterbi optimization). For example, the probability of a fly transitioning between a location on the lid to a location on the wall is quite low if the fly is in the center of the chamber; this transition is more probable if the fly is on the lid near the wall. We wrote custom code in Matlab (Mathworks Natick, MA, USA) for transforming, optimizing, and analyzing all data.

4.3.4 Quantitative descriptors of exploratory behavior

Each behavioral statistic describes the behavior of a fly during a one-hour interval. There are 4 types of statistics: those describing the behavior of flies while walking, flying, searching near water, and dispersing from the chamber through the exit into the second chamber. In Figures 4.5–4.21, we plot the interval number versus one of various statistics for each of the 20 selected flies, and that were sorted based on the total distance traveled during the 6 hour trial. For each statistic and selected fly, we plot the interval number versus the statistic value. We plot the raw statistic in part (A) of each figure. In many of the statistics, we can see temporal dependencies throughout the day. As the flies grow hungrier with time since last feeding, various measurements of their locomotor movements increase. Their behaviors then decrease during the middle of the day, and for the 12 hour trials, the flies' behavior then increases again around dusk near the end of the experiment (for example see 4.3). In part (B) of each figure, we plot the z-scored statistic. That is, we compute the mean and standard deviation for each interval and statistic over all flies, then plot the number of standard deviations from the mean the statistic is for a given fly. This manipulation largely removes the temporal dependencies of the statistic. For some of the statistics analyzed, the statistic is only computed from a few observed values, and thus will be noisy. For example, Figure 4.4 shows values on which the noise in a particular statistic depends in (A), and the standard error of the median-based estimates in (B).

4.3.4.1 Walking statistics

We segment a sequence of a trajectory in which the fly is both (1) in the chamber and (2) not flying into subsequences in which the fly is walking or stopped using a variant

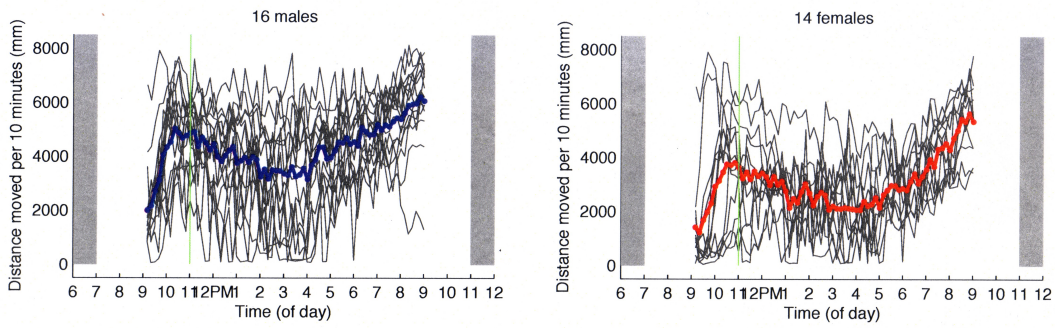

Figure 4.3: Movement duration for individuals as they become hungry. Total distance traversed for isolated, single flies for each successive 10 minute interval over the period of 12 hours. Collective means for male (blue) and female (red) flies are noted. The transition from dark (gray boxes) to light are indicated. The green vertical line denotes 2 hours into experiment.

of the Viterbi algorithm (Cormen et al., 2001). We model the probability of a sequence of walking/stopped states as a first-order, binary hidden Markov model. The intuition behind the chosen model is as follows. First, there is a higher probability that the fly will remain in the same state, i.e., either remain walking or stopped, than switch to the other state. Second, if the fly's speed in the current frame is small, there is a higher probability that the fly is stopped than walking. Conversely, if the fly's speed is high, there is a higher probability that the fly is walking. We use dynamic programming to find the sequence of states with globally maximal probability.

More formally, let $s_t = 1$ represent the classification of frame t as walking and $s_t = 0$ the classification of frame t as stopped. Using a first-order Markov assumption, we can write the probability of a sequence of hidden states $\mathbf{s_{1:t}}$ for frames 1 through t given the observed speeds $\mathbf{v_{1:t}}$ recursively as

$$P(\mathbf{s_{1:t}}|\mathbf{v_{1:t}}) \propto \mathbf{P}(\mathbf{s_t}|\mathbf{s_{t-1}})\mathbf{P}(\mathbf{v_t}|\mathbf{s_t})\mathbf{P}(\mathbf{s_{1:t-1}}|\mathbf{v_{1:t-1}}).$$

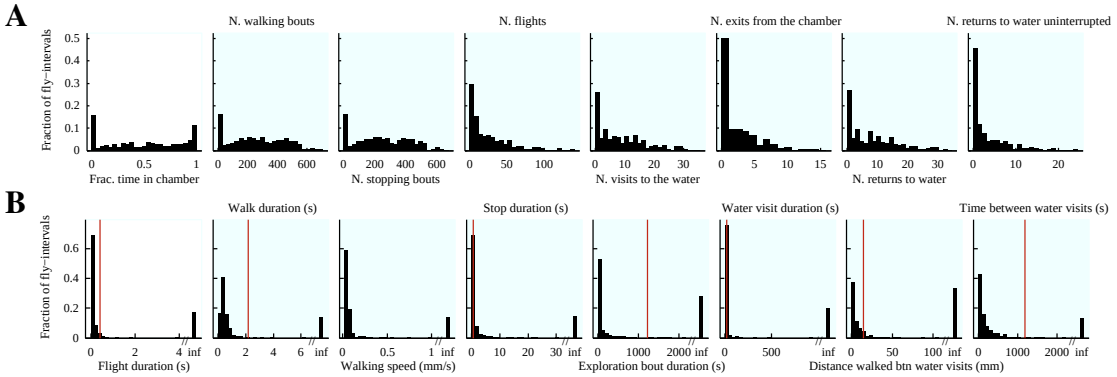

Figure 4.4: Uncertainty in estimates of per-interval statistics. (A) We histogram the value on which the noise of various statistics depends. (B) We show the standard error of the median estimate, assuming the data is normally distributed. In both, infinity corresponds to intervals with no data available. (B) for comparison we plot the mean value of the statistic in red (for walking speed, the mean value is 8.5, which is off the plot). We see that for the rarer events – water visits, return trips to the water – the noise is high. "inf" stands for infinity and corresponds to n = 0, complete uncertainty.

The transition probability $P(s_t|s_{t-1})$ is set to 0.98 if the state remains the same, $s_t = s_{t-1}$, and 0.02 if the state changes $s_t \neq s_{t-1}$. The likelihood of observing speed v_t (in mm/s), given that the fly is walking, is assumed to be proportional to

$$P(v_t|s_t = 1) \propto \exp(-(v_t - 1)^2/(2 \cdot .15)),$$

that is, proportional to a Gaussian distribution with center 1 mm/s and variance .15 (cm/s)^2. The likelihood of observing speed v_t (in mm/s), given that the fly stopped, is assumed to be proportional to

$$P(v_t|s_t = 0) \propto \exp(-(v_t - 0)^2/(2 \cdot .005)),$$

that is, proportional to a Gaussian distribution with center 0 mm/s and variance .05 (cm/s)^2.

The fraction of time walking (*fractimewalking*) statistic reports the fraction of time the fly is in the observed chamber that it is classified as walking (Figure 4.5). It is a unitless quantity. The less time the fly spends in the chamber during the interval, the noisier this statistic is. If the fly is not in the observed chamber at all during the interval, then this statistic is completely unknown.

The distance traveled (*disttraveled*) statistic reports the total distance in centimeters the fly travels while it is classified as walking in the observed chamber (Figure 4.6). This value is normalized by the number of seconds the fly is in the observed chamber, thus the units reported are $\text{cm} \cdot \text{s}^{-1}$. The noise in this statistic also depends on the amount of time the fly spends in the chamber during the interval.

The frequency of walk onsets (*freqwalk*) is the number of times the fly begins a bout of walking during the interval, normalized by the number of seconds the fly spends in the observed chamber (Figure 4.7). The units reported are therefore onsets per second. If a walking bout crosses the division between two intervals, we choose the interval in which the middle frame of the bout falls. We follow this policy with all other bout-related properties. The noise in this statistic also depends on the amount of time the fly spends in the chamber during the interval.

The mean walking speed (*meanwalkspeed*) is the mean speed of the fly over all frames in which the fly is classified as walking in the first chamber (Figure 4.8). The units reported are centimeters per second. The noise in this statistic depends on the amount of time the fly spends walking in the observed chamber during the interval.

The median duration of walking bouts (*walkdur*) is the median duration of sequences during which the fly is classified as walking while in the chamber (Figure 4.9). This statistic is reported in seconds. The noise in this statistic depends on the number of bouts of walking the fly performs in the observed chamber during the interval.

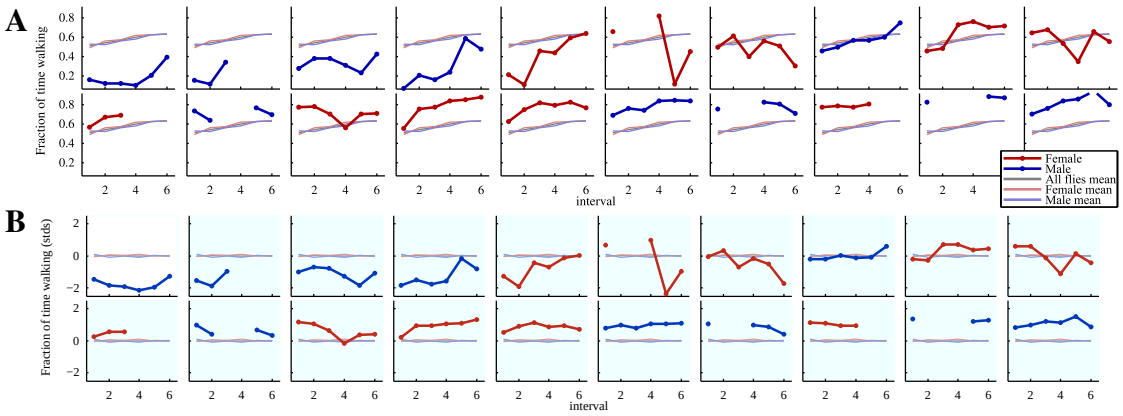

Figure 4.5: Interval vs. fraction of time walking for selected individuals. We select 20 flies whose ranks based on total distance travelled are evenly spaced. The top left fly walks the least of all flies and the bottom right fly walks the most of all flies. We plot the fraction of time walking for each interval, where the interval length is 1 hours. Missing points indicate intervals for which the statistic cannot be computed. (A) We plot the interval number vs. the raw statistic, while in (B) we plot the interval number vs. the number of standard deviations from the mean statistic value for the given interval (i.e. the data in each interval has been z-scored. In gray, we plot the mean value over all flies. In light red, we plot the mean value over all female flies. In light blue, we plot the mean value over all male flies. We plot the value for the selected fly in dark red if it is female and dark blue if it is male.

The median duration of stop bouts (*stopdur*) is the median duration of sequences during which the fly is classified as stopped while in the chamber (Figure 4.10). The statistic is reported in seconds. The noise in this statistic depends on the number of bouts of stopping the fly performs in the observed chamber during the interval.

4.3.4.2 Flying statistics

The fraction of time flying (*fractimeflying*) is the fraction of time the fly is in the chamber that it is classified as flying, i.e., the fly's velocity was $>$0.6 cm/s (Figure 4.11). It is a unitless quantity. The less time the fly spends in the chamber during the interval, the noisier this statistic is.

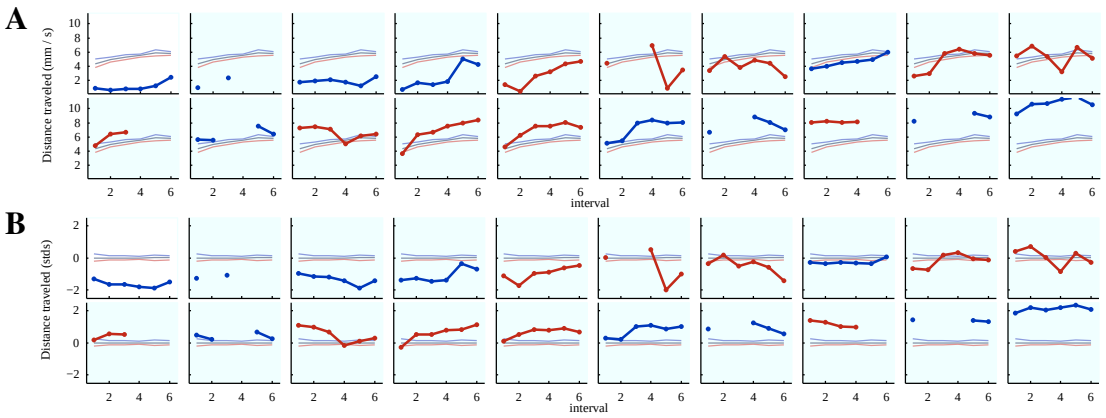

Figure 4.6: Interval vs. distance traveled for selected individuals. See Figure 4.5 for a more complete description.

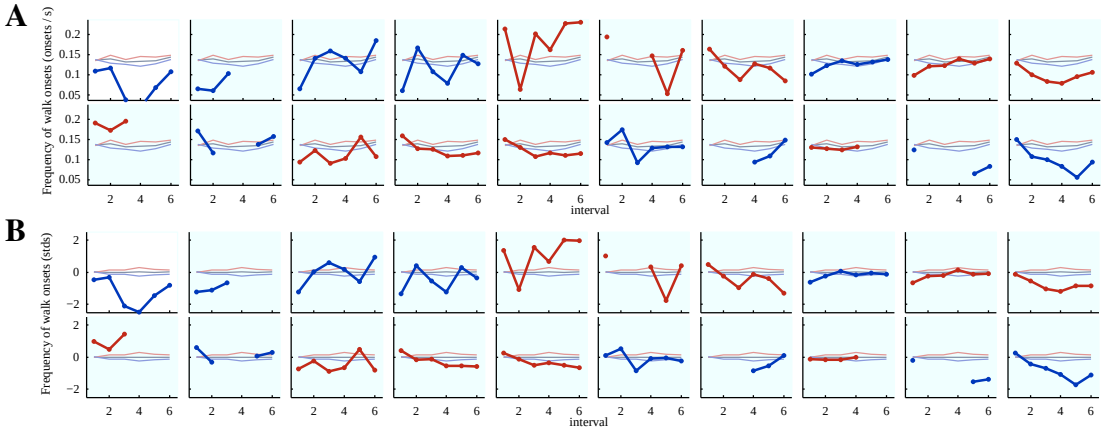

Figure 4.7: Interval vs. frequency of walk onsets for selected individuals. See Figure 4.5 for a more complete description.

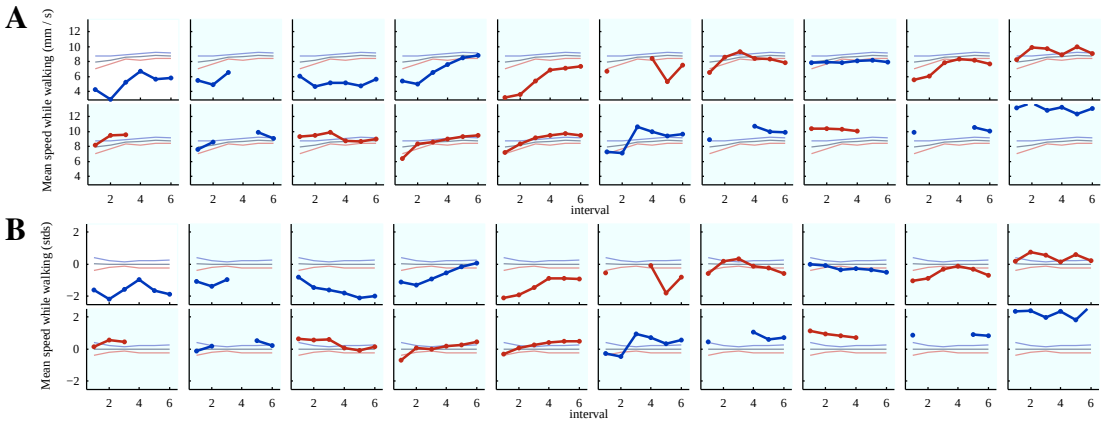

Figure 4.8: Interval vs. mean speed while walking for selected individuals. See Figure 4.5 for a more complete description.

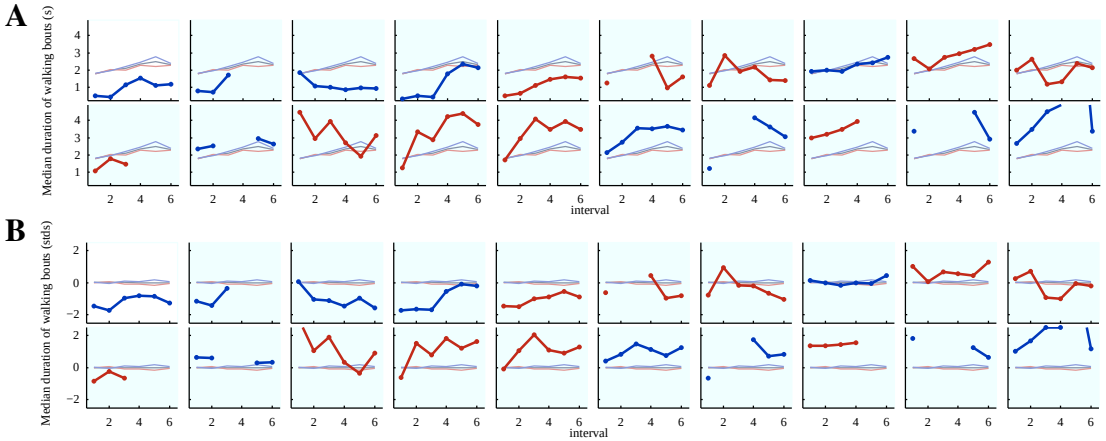

Figure 4.9: Interval vs. median duration of walking bouts for selected individuals. See Figure 4.5 for a more complete description.

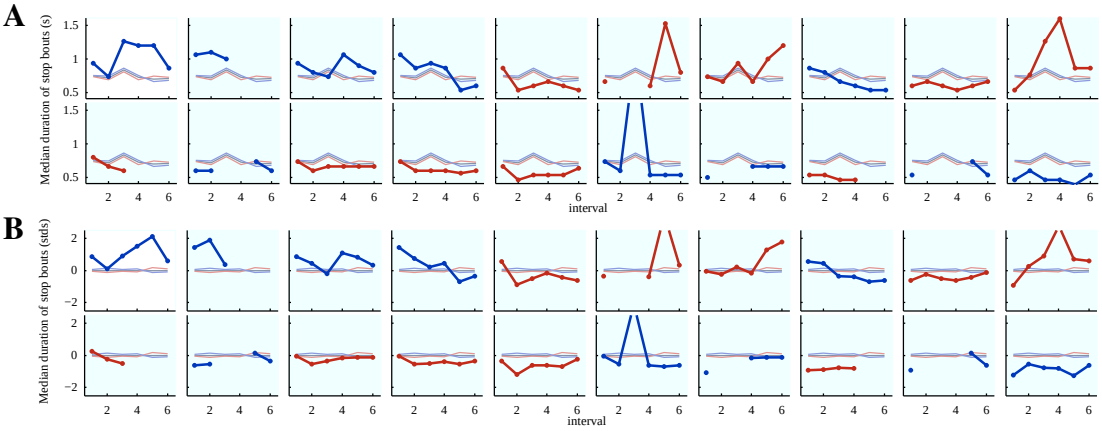

Figure 4.10: Interval vs. median duration of stop bouts for selected individuals. See Figure 4.5 for a more complete description.

The frequency of flights (*freqflight*) is the number of times the fly begins a bout of flying during the interval, normalized by the number of seconds the fly spends in the observed chamber (Figure 4.12). The units reported are therefore onsets per second. The noise in this statistic depends on the amount of time the fly spends in the chamber during the interval.

The median duration of flights (*flightdur*) is the median duration of sequences during which the fly is classified as flying while in the chamber (Figure 4.13). This statistic is reported in seconds. The noise in this statistic depends on the number of flights the fly performs in the observed chamber during the interval.

4.3.4.3 Local search near water statistics

We segment the trajectory of a fly into sequences in which it is either visiting or not visiting the patch of agar, a source of water. A fly is considered visiting the water if it is ≤ 0.1 cm from the water's edge, or if it is ≤ 0.3 cm from the edge and within 2 frames from a frame in which the fly is ≤ 0.1 cm from the edge. We based this classification

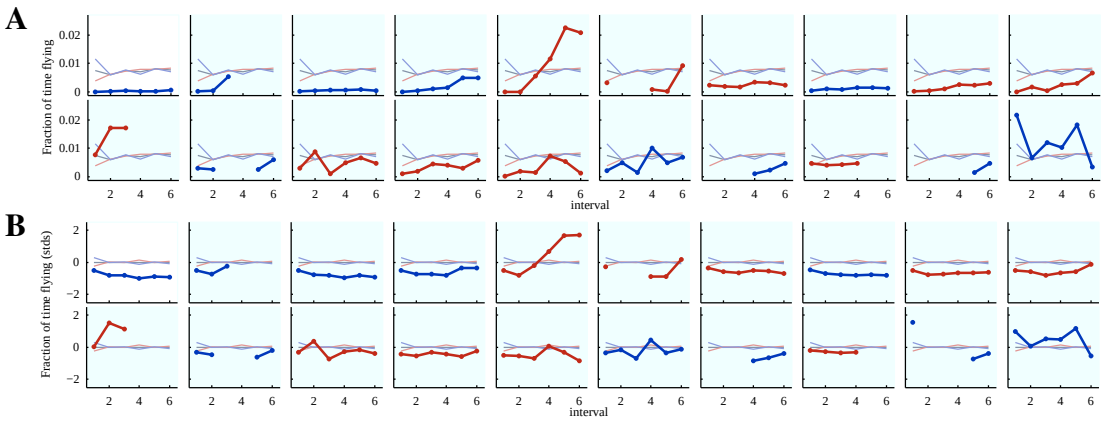

Figure 4.11: Interval vs. fraction of time flying for selected individuals. See Figure 4.5 for a more complete description.

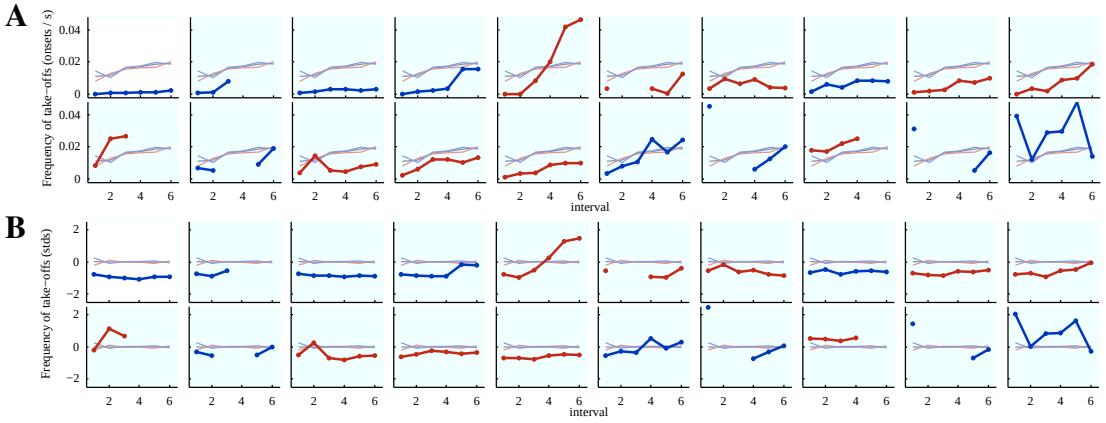

Figure 4.12: Interval vs. frequency of take-offs for selected individuals. See Figure 4.5 for a more complete description.

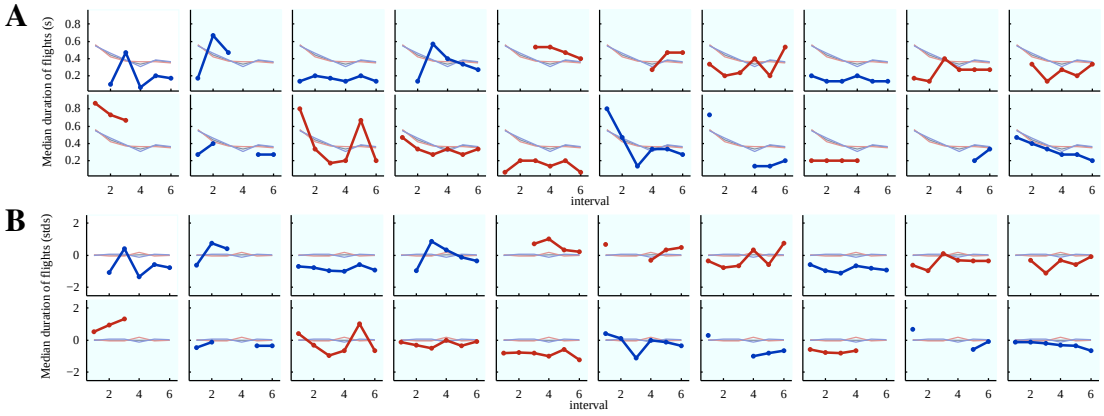

Figure 4.13: Interval vs. median duration of flights for selected individuals. See Figure 4.5 for a more complete description.

from average transit probabilities from 125 flies as they search near the agar (see Figure 4.30). This definition results in an average of 9 visits to the water resource per interval, with 20% of fly-intervals having no visits to the water.

The fraction of time near the water (*fractimenearwater*) is the fraction of frames the fly is in the chamber that it is classified as near the water resource (Figure 4.5). It is a unitless quantity. The noise in this statistic depends on the amount of time the fly spends in the chamber during the interval.

The frequency of visits to water (*freqvisitwater*) is the number of continuous sequences of frames in which the fly is classified as visiting the water, normalized by the number of seconds the fly spends in the observed chamber (Figure 4.6). The units reported are therefore sequences per second. The noise in this statistic depends on the amount of time the fly spends in the chamber during the interval.

The median duration of visits to water (*watervisitdur*) is the median duration of sequences during which the fly is classified as near the water resource (Figure 4.7). This statistic is reported in seconds. The noise in this statistic depends on the number of times the fly visits the water resource during the interval.

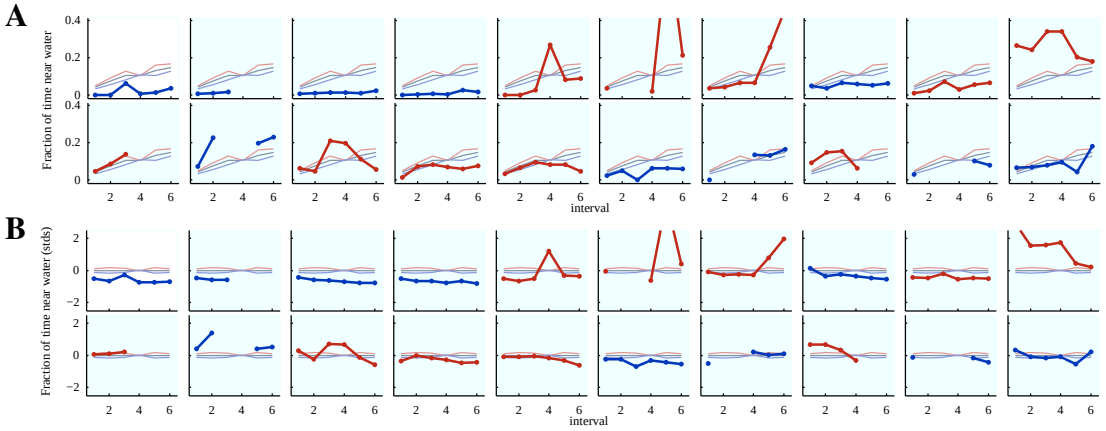

Figure 4.14: Interval vs. fraction of time near water for selected individuals. See Figure 4.5 for a more complete description.

The median time between visits to water (*timebtnwater*) is the median duration of sequences which begin when the fly leaves the water resource and end when the fly returns to the water resource (Figure 4.8). As with other types of sequences, if a return trip crosses the division between two intervals, we choose the interval in which the middle frame of the trip falls. This statistic is reported in seconds. The noise in this statistic depends on the number of return trips to the water resource during the interval, which of course is related to the number of times the fly visits the water resource.

The median length of the walking path between visits to water (*pathlengthbtnwater*) is the median distance traveled in *uninterrupted* sequences that begin when the fly leaves the water resource and end when the fly returns to the water resource (Figure 4.9). By uninterrupted, we mean that the fly does not leave the chamber or fly during the return trip, as we cannot measure the distance traveled during these periods. This statistic is reported in centimeters. The noise depends on the number of uninterrupted return trips.

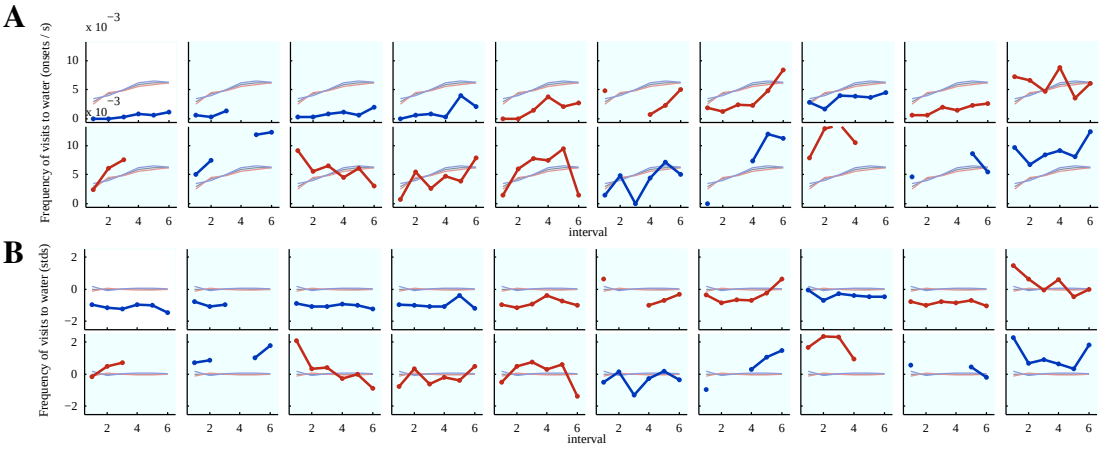

Figure 4.15: Interval vs. frequency of visits to water for selected individuals. See Figure 4.5 for a more complete description.

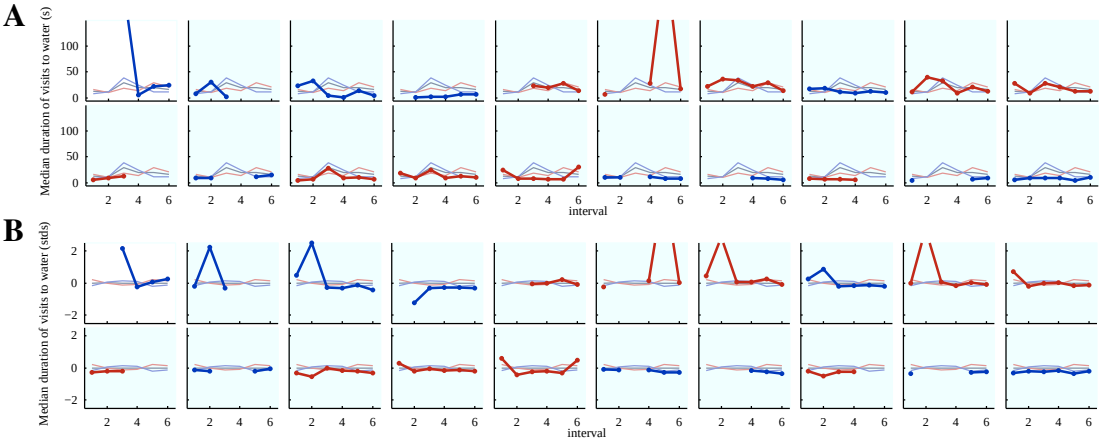

Figure 4.16: Interval vs. median duration of visits to water for selected individuals. See Figure 4.5 for a more complete description.

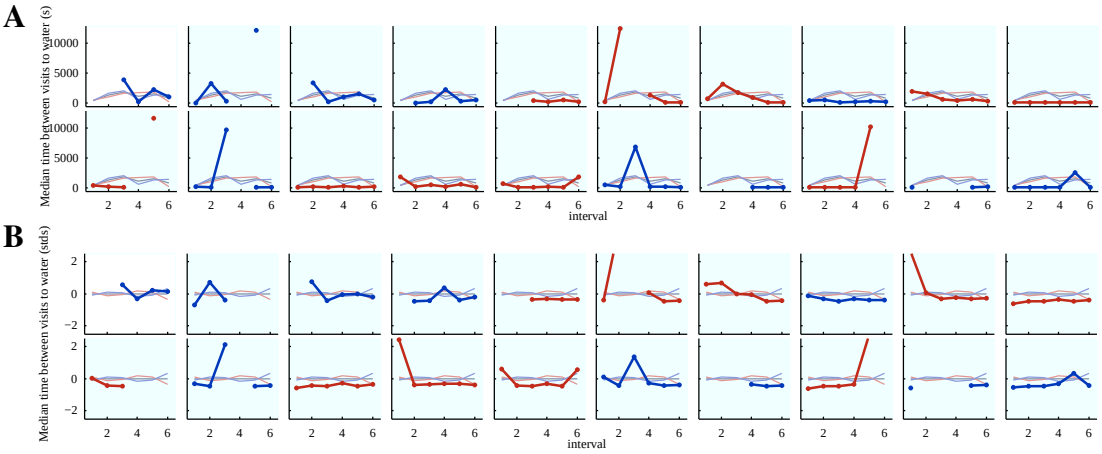

Figure 4.17: Interval vs. median time between visits to water for selected individuals. See Figure 4.5 for a more complete description.

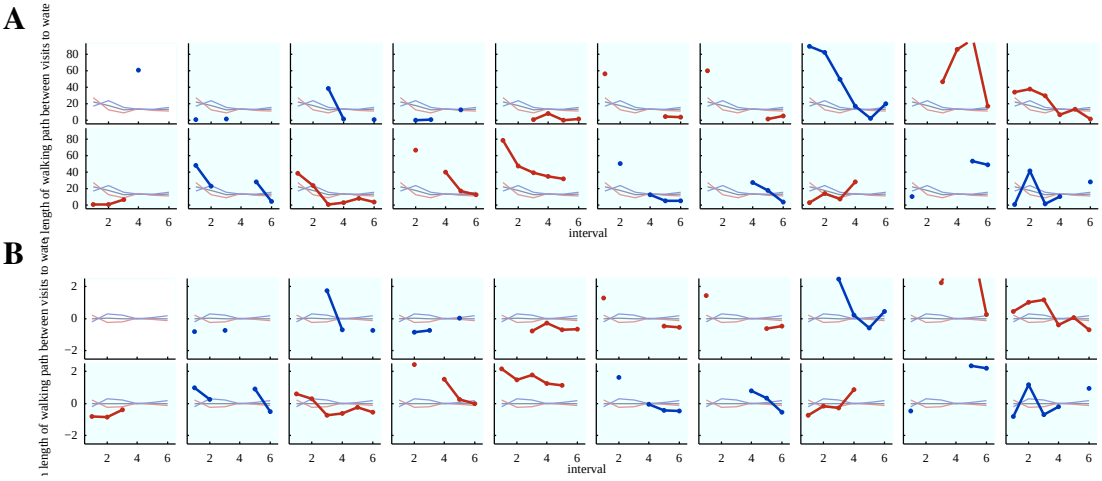

Figure 4.18: Interval vs. median length of walking path between visits to water for selected individuals. See Figure 4.5 for a more complete description.

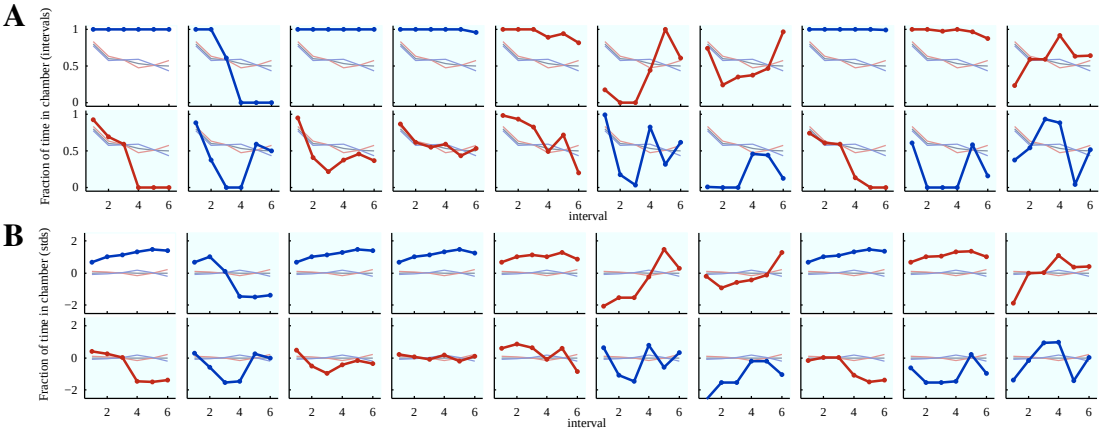

Figure 4.19: Interval vs. fraction of time in chamber for selected individuals. See Figure 4.5 for a more complete description.

4.3.4.4 Dispersal from chamber statistics

The fraction of time in the chamber (*fractimeinchamber*) is the fraction of the interval the fly spends in the main, observed chamber (Figure 4.19). This is a unitless quantity. The noise in this statistic depends only on the interval length. The classification of when flies left the chamber was primarily based on when the pixel area of a thresholded video image dropped to zero (see Figure 4.32), and (see Figure 4.33 for details).

The frequency of exits from the observed chamber (*freqexits*) is the number of times the fly exits the interval, normalized by the number of seconds the fly spends in the interval (Figure 4.20). The units reported are therefore onsets per second. The noise in this statistic depends on the amount of time the fly spends in the chamber during the interval.

The median duration of exploration bouts (*awaydur*) is the median duration of visits to the second, unobserved chamber. This statistic is reported in seconds. The noise in this statistic depends on the number of times the fly visits the second chamber.

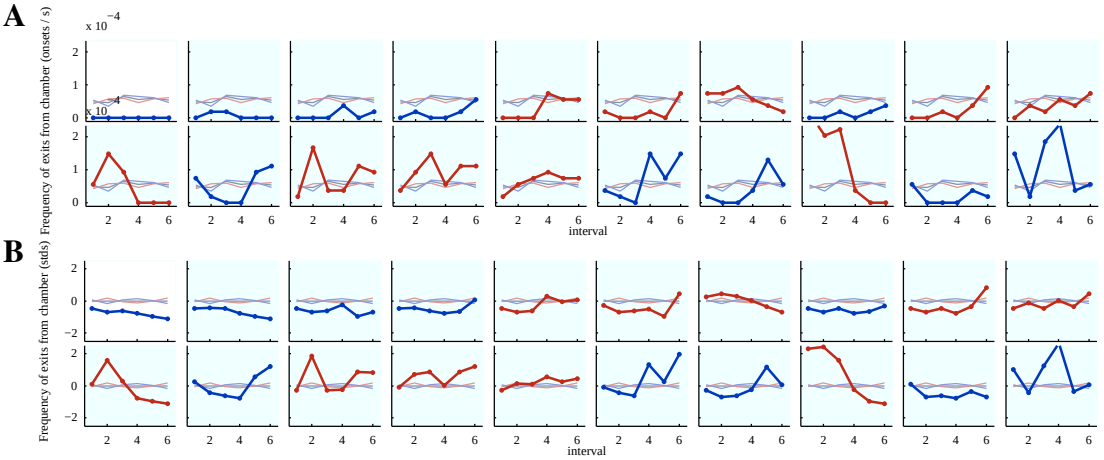

Figure 4.20: Interval vs. frequency of exits from chamber for selected individuals. See Figure 4.5 for a more complete description.

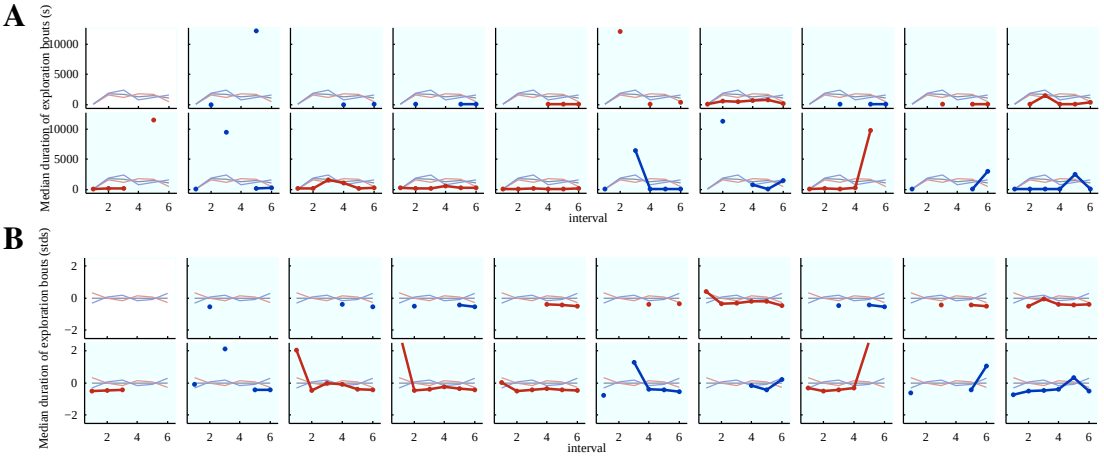

Figure 4.21: Interval vs. median duration of exploration bouts for selected individuals. See Figure 4.5 for a more complete description.

4.3.5 Training algorithms and strategy for predicting behaviors

For a given statistic and a given interval, we learn a function that inputs the values of this statistic for a subset of the remaining intervals and predicts the value of this statistic in the given interval. For example, we learn a function that inputs the 5-D vector consisting of the distance traveled in intervals 1–3 and 5–6, and outputs an estimate of the distance traveled in interval 4. We do this for each statistic and each of the 6 intervals in the 6-hour experiment. We also consider different lengths of buffers between the input and predicted intervals. A buffer length of 0 intervals corresponds to predicting the given interval using all the remaining intervals. A buffer length of 1 interval corresponds to predicting the given interval using all remaining intervals except for those adjacent to the given interval (e.g., predicting interval 4 from intervals 1–2 and 6). More generally, a buffer length of n intervals corresponds to predicting the given interval using all remaining intervals except those within $n+1$ intervals of the given interval. The larger the buffer, the less data we have to predict the given interval, thus we expect our estimates to be less accurate. However, we consider these larger buffers to demonstrate that the individuality effects can be seen over larger time frames.

We use a form of regularized linear regression to learn the predictor function. Let $\mathbf{x_{ij}}$ represent the input statistic vector for statistic i and fly j, and the predicted interval, and y_{ij} the true value of statistic i and fly j for the current predicted interval. Any linear predictor of y_{ij} given $\mathbf{x_{ij}}$ can then be represented as

$$f(\mathbf{x_{ij}}) = \mathbf{c_i}^\top \mathbf{x_{ij}} + \mathbf{c_{i0}},$$

where $\mathbf{c_i}$ is a constant vector of coefficients and c_{i0} is a constant offset for statistic i and all flies. For ease of notation, let us append the input data vector $\mathbf{x_{ij}}$ with an element

that is always 1, and fold the constant offset c_{i0} into the vector of coefficients $\mathbf{c_i}$.

In ordinary linear regression, we choose the coefficients $\mathbf{c_i}$ that minimize the mean-squared error:

$$J_o(\mathbf{c_i}) = \frac{\mathbf{1}}{\mathbf{n}} \sum_{\mathbf{j=1}}^{\mathbf{n}} \left(\mathbf{c_i^\top x_{ij}} - \mathbf{y_{ij}} \right)^{\mathbf{2}},$$

where the sum is over the data for the n flies in the training set.

Because we had limited amounts of data compared to the amount of noise in the statistics, particularly for statistics such as the median water visit duration, we used a regularized form of linear regression. First, we z-score the inputs $\mathbf{x_{ij}}$ and outputs y_{ij} using the mean and standard deviation computed from the training set only. That is, we subtract the training set sample mean and divide by the training set sample standard deviation for each statistic and interval. This manipulation takes into account much of the temporal dependencies of the statistics. Note that z-scoring the data is itself a linear transformation, thus it would not affect the results of ordinary linear regression. Let $\mathbf{x_{ij}}$ and y_{ij} now represent the z-scored data. The linear regression can then be thought of as a weighted mean of the statistics for the given intervals. We will most likely want to give higher weights to intervals closer to the predicted interval, or perhaps to give lower weights to intervals that are less reliable. Based on the assumption that the coefficients for different statistics but the same interval will be somewhat similar, we consider the following regularized criterion:

$$J_r(\mathbf{c_1}, ..., \mathbf{c_m}) = \frac{\mathbf{1}}{\mathbf{m}} \sum_{\mathbf{i=1}}^{\mathbf{m}} \frac{\mathbf{1}}{\mathbf{n}} \sum_{\mathbf{j=1}}^{\mathbf{n}} \left(\mathbf{c_i^\top x_{ij}} - \mathbf{y_{ij}} \right)^{\mathbf{2}} + \lambda \frac{\mathbf{1}}{\mathbf{m}} \sum_{\mathbf{i=1}}^{\mathbf{m}} \left(\mathbf{c_i} - \frac{\mathbf{1}}{\mathbf{m}} \sum_{\mathbf{k}} \mathbf{c_{k=1}^m} \right)^{\mathbf{2}}.$$

There are two differences between this regularized criterion, J_r, and the ordinary least-squares criterion, J_o. First, this criterion is a function of the coefficients for all the

statistics, rather than just one statistic. Thus, we are solving for the coefficients for all statistics simultaneously. The first term in J_r is the same as the first term in J_o, except that we are summing over all the statistics. The second difference between J_r and J_o is the inclusion of the second term, the regularization term. This term penalizes differences between a coefficient vector for one statistic and the mean over all statistics. The constant λ weights the data term and the regularization term. We only experimented with setting $\lambda = 1$. To improve robustness to outliers, we threshold all inputs and training outputs at 3 standard deviations. Figure 4.34 shows the coefficients learned using ordinary linear regression in (A) and regularized linear regression in (B) for the z-scored data. Figure 4.35 shows a comparison of the mean-squared error for regularized linear regression to other regression algorithms. We compare to linear and quadratic regression with the ordinary least-squares criterion, quadratic regression for the proposed regularized criterion, and linear and quadratic regression using iteratively reweighted least-squares with the bisquare weighting function, implemented with the `robustfit` function in Matlab. The regularized linear regression was usually the best performing method, and the most reliable when training data was scarce, particularly in data sets with smaller numbers of flies, not reported here.

While J_o can be minimized by a simple matrix inversion for each statistic, J_r is slightly more difficult to minimize, as the coefficients for all statistics must be simultaneously selected. However, J_r is convex, so we can choose an arbitrary initialization, perform a gradient descent, and be guaranteed to find the global optimum. We initialize with the ordinary least-squares regression coefficients. At each iteration of our algorithm, we hold the coefficients for all statistics except one constant, then find the optimal values for the coefficient vector for this single statistic. For efficiency, we order the statistics whose coefficients we will optimize based on their sum-squared error (J_o).

The optimal $\mathbf{c_i}$ for fixed $\{\mathbf{c_k}, \mathbf{k} \neq \mathbf{i}\}$ can be found in closed form as

$$\mathbf{c_i} = \left(\frac{\mathbf{1}}{\mathbf{N}} \mathbf{X_i^\top X_i} + \lambda \left(\mathbf{1} - \frac{\mathbf{1}}{\mathbf{M}} \right) \mathbf{I_{P \times P}} \right)^{-\mathbf{1}} \left(\frac{\mathbf{1}}{\mathbf{N}} \mathbf{X_i^\top y_i} + \lambda \frac{\mathbf{1}}{\mathbf{M}} \sum_{\mathbf{k} \neq \mathbf{i}} \mathbf{c_k} \right).$$

Here P is the dimensionality of the input vector $\mathbf{x_{ij}}$, X_i is the $N \times P$ matrix in which row j is $\mathbf{x_{ij}^\top}$, $\mathbf{y_i}$ is the $N \times 1$ vector in which element j is y_{ij}, and $I_{P \times P}$ is the $P \times P$ identity matrix. Note that most of the quantities involved do not change from one iteration to another, thus the iterative optimization is efficient.

Many of the statistics are often undefined, e.g., if the fly does not spend any time in the observed chamber, or the fly does not perform a certain behavior. If we remove all flies for which the statistic for some interval is undefined, then we will lose a lot of data. Instead, we only remove flies for which the predicted interval and statistic is undefined, and set the undefined input statistics to the sample mean over the training data.

In all our experiments, we use hold-one-out cross validation. That is, we learn the regression coefficients from all flies except one, then compute the error on this held-out fly. We do this for each fly and therefore learn a different regressor for each fly. In this way, we keep the training and test data independent for all parts of the learning.

4.3.6 Control data

To see the effects of individuality on the regression error, we create semi-synthetic data sets which should not have any effects of individuality. Within each interval, we randomly permute the identities of the fly. For example, we may end up with statistics for fly 10 in interval 1, fly 29 in interval 2, fly 7 in interval 3, etc. To control for effects due to gender, we only permute identities within gender. Thus, the fly identities chosen for a given vector will all have the same sex as the identity in the first interval.

4.4 Results

4.4.1 Behavioral statistics of individual flies persist over time

Figures 4.22–4.25 graphically compare the true statistics with the predictions from the learned regressors. We plot the true versus predicted statistics for each fly for selected statistics, intervals, and buffer sizes. Each point on each plot corresponds to one fly. The x-axis corresponds to the true statistic, while the y-axis corresponds to the prediction from the learned regressor. Each figure corresponds to a different statistic. These statistics were chosen to span the range of normalized, mean-squared generalization error for buffer length = 0 intervals. Part (A) corresponds to the real data. Part (B) corresponds to the semi-synthetic control data; however, note that we generated 20 control sets for these plots, thus there are 20 times more points in the control plots. The left column (i) corresponds to buffer length = 0 intervals; the right column (2) corresponds to buffer length = 3 intervals. We plot the true versus predicted statistics for intervals 1, 2, 5, and 6, for these are the only intervals that can be predicted for buffer length = 3 intervals. So that we could use the same axes for each statistic and interval, we plot the number of standard deviations from the mean on each axis. If predictions were perfect, the data points would lie on the line of slope 1 through (0,0). We see that the real data does indeed look correlated for the well-predicted statistics, but less-so for the poorly predicted statistics. The buffer length = 0 data also looks more correlated than the buffer length = 3 data. In all the control plots, we see no correlation.

We can quantitatively compare the accuracy of the predictions for the real data to the accuracy for the control data. We measure error as the square-root of the mean-squared error (the square-root of J_o). So that errors on the different statistics can be directly compared, we normalize the error by the standard deviation of the statistic and

interval over all flies. In these normalized units, the standard deviation of the data is 1, hence the error of always predicting the mean statistic for an interval will be 1. The expected hold-one-out, cross-validation performance of the mean statistic for flies in the training set will be slightly more than 1. In Figure 4.26 (A), we plot the normalized square-root of the mean-squared error for the real and control data for each statistic. The x-axis corresponds to the statistic and the y-axis to the error. We plot the per-interval error (thin lines) as well as the mean error over all intervals (thick lines). Each plot corresponds to a different buffer length, with (i) corresponding to buffer length = 0 and (v) corresponding to buffer length 5. We generated 100 sets of control data. We observe that the error for the control data is indeed near 1 for all statistics and intervals – thus it is not performing better than just the sample mean – the statistics from other intervals are not useful for predicting the statistic for a given interval. For many statistics, the error for the real data is less than 1. The statistics on the x-axis are ordered by the mean error over all intervals for buffer length = 0 (i).

We can determine whether the error for the real data is significantly less than the error for the control data by computing the fraction of randomly generated control sets that have an error as good as the real data. This is an empirical measure of the probability of achieving the error computed for the real data regressors if there were no effect of individuality. This p-value is plotted in Figure 4.26 (B). The x-axis again corresponds to the statistic and the y-axis to the log p-value. The gray horizontal line corresponds to $p = .05$. We see that for many of the statistics, the effects of individuality are highly significant.

If the real data produces regressors with significantly less error than the control data, then we have observed the effects of individuality. That is, a positive result indicates an effect. Conversely, a negative result does not necessarily imply that there is no effect of

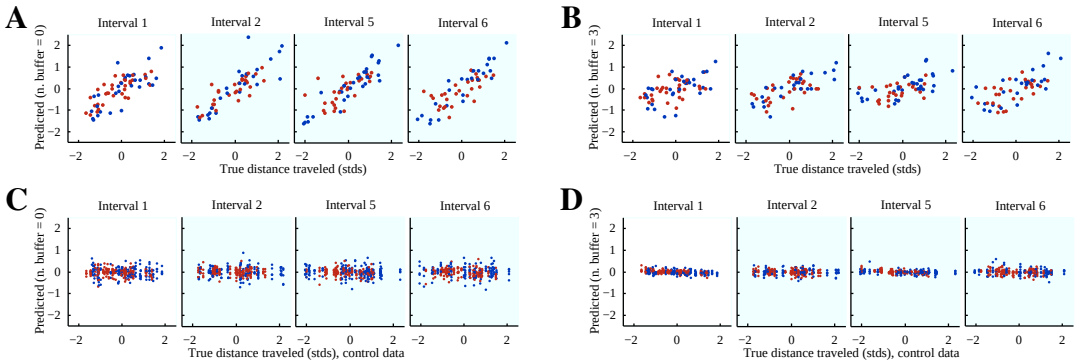

Figure 4.22: True vs. predicted distance traveled. Each axis corresponds to a different interval, where the interval length is 1 h. Each point corresponds to a fly. Male flies are plotted in blue, females in red. For each fly, we plot the true distance traveled versus the distance traveled predicted by the regression (measured in standard deviations from the mean). In the perfect regression, all points would be along the diagonal. In **A**–**B**, we create this plot for the real data. Each column corresponds to a different number of buffers maintained between the predicted and predicting intervals. In **C**–**D**, we create these plots for control data created by randomly permuting the identities independently in each interval (preserving sex). For each of the 5 control sets, we learn a regressor as with the real data, and plot the true control vs. the predicted control.

individuality for the statistic. No significant difference could result from three possible cases. First, there could be a true lack of an effect of individuality. Second, it could be that regularized linear regression does not fit the data well. Finally, it could be that there is too much noise and not enough data to accurately learn the regressor (or, a combination of the latter two). In particular, the errors greater than 1 correspond to overfitting the data, and are evidence that there is not enough data to overcome noise and learn a proper fit. To determine which of these cases apply, one would need to repeat the experiment with larger numbers of flies.

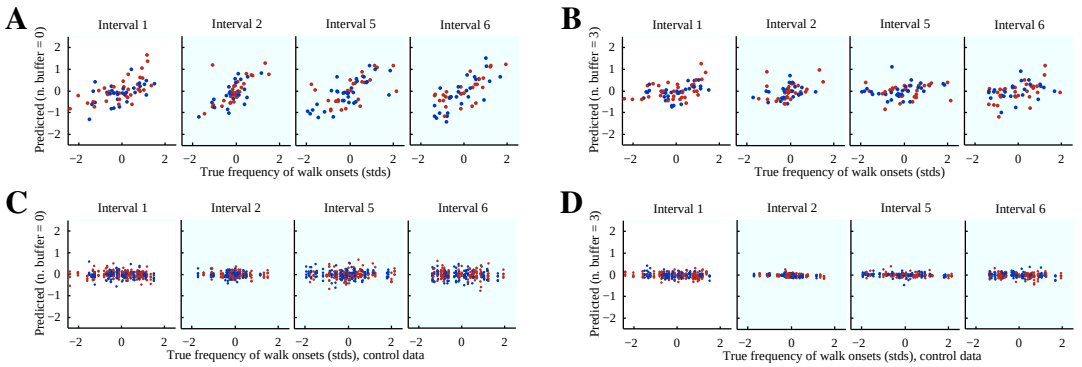

Figure 4.23: True vs. predicted frequency of walk onsets. See Figure 4.22 for a more complete description.

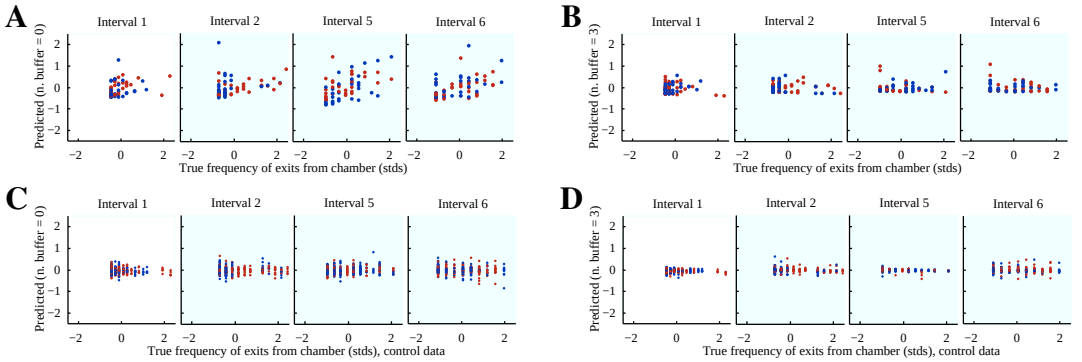

Figure 4.24: True vs. predicted frequency of exits from chamber. See Figure 4.22 for a more complete description.

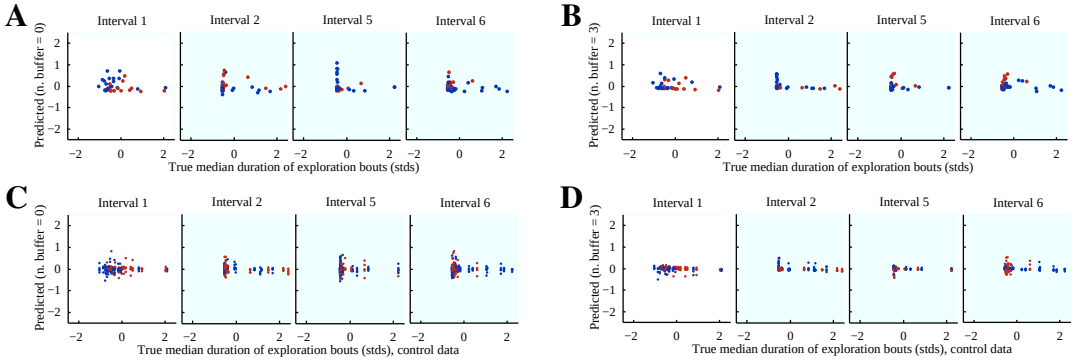

Figure 4.25: True vs. predicted median duration of exploration bouts. See Figure 4.22 for a more complete description.

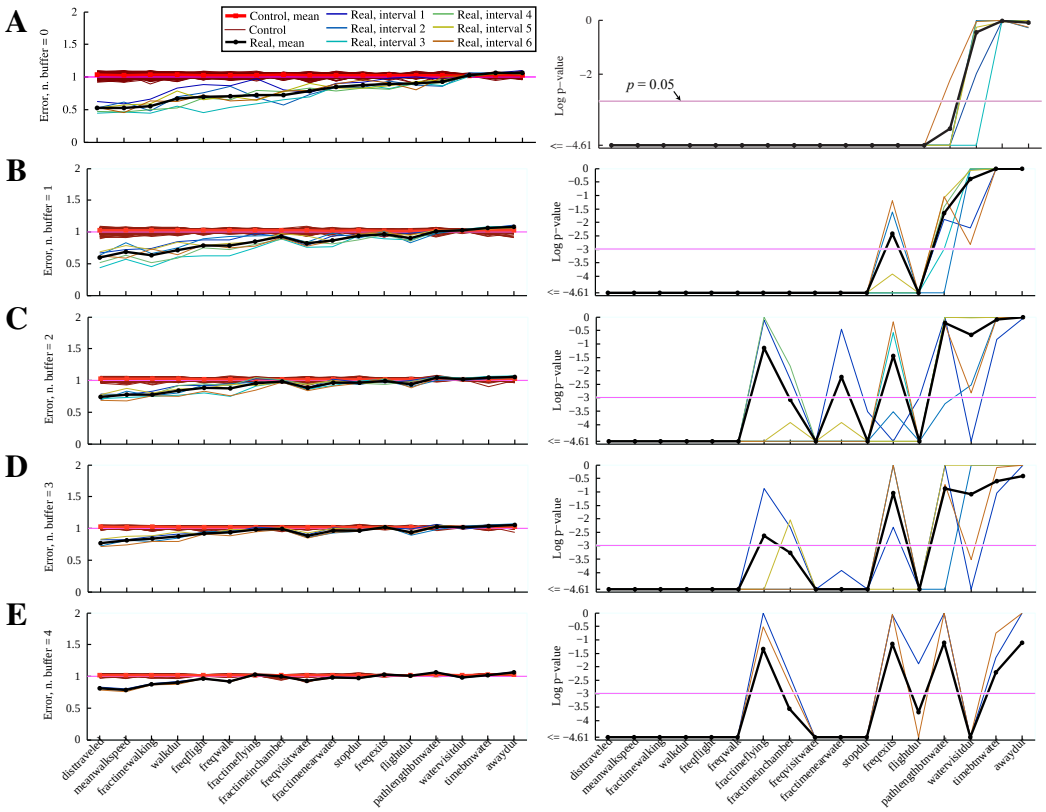

Figure 4.26: Mean squared-error for real and control data for interval length = 1 h. Each row of plots corresponds to a different numbers of buffers maintained between the predicted and predicting intervals. In the left column, we plot the square root of the mean squared error between the true and predicted behavioral statistic. For each statistic, we normalize the error by the standard deviation of the statistic so that errors are comparable between different statistics. Because of this normalization, a regressor that always predicts the mean statistic will have a mean normalized error of 1 (horizontal magenta line). The red lines correspond to the randomly permuted data, all other lines correspond to the real data. The thin lines correspond to the errors for single intervals, while the thick lines correspond to the mean over all intervals. Note that there are many thin red lines because we plot 100 control sets. In the right column, we plot the log of the fraction of the 100 control set errors that are less than the real errors. This is an empirical estimate of the probability that we would observe an error as low or lower than the real residual if the behavioral statistic was independent of identity. As in the left column, the thin lines correspond to per-interval p-values, while the thick line corresponds to the mean over all intervals. The magenta line corresponds to $p = 0.05$.

4.4.2 Dimensionality reduction analysis

We used principal component analysis (PCA) to examine what structures in the data may be used to predict a given behavioral statistic. For each of the 12-most predictable statistics (for buffer length = 0), we examine the 6-dimensional vector composed of the per-interval statistics (e.g. element 1 is the distance traveled in the first hour, element 2 is the distance traveled in the second hour, ...). We perform PCA on the z-scored data set consisting of these 6-D vectors for all flies (thresholding outliers at 3 standard deviations) to find the 6-D directions of greatest variance.

Figure 4.27 shows the results. For each statistic, there are three plots. In the top row, we plot the projection of the per-interval statistics on the first two principal components, the highest variance 2-D linear subspace. For no statistic do we see clearly clusterable data.

In the middle row, we plot the error of the projections onto increasing numbers of principal components. As emphasized in Figure 27, we see that the usefulness of the first principal component corresponds with predictability of behavior statistics based on individuality (Figure 4.26).

In the bottom row, we show the directions of the first and second principal components. For all the statistics plotted, the first principal component is flat across all intervals, implying that the first principal component represents the average value of the statistic across all intervals. Thus, for instance for the distance traveled statistic, the highest-variance direction corresponds with the average distance traveled in all intervals some flies walk far and some flies do not. In addition, for all the statistics, the second principal component increases nearly monotonically with time. To emphasize the similarities between statistics, we flip the sign of the component to be increasing it is the

monotonicity of the change that is relevant. At interval 3 and below, the coefficient is usually negative, while above it is positive. Thus, the second principal component corresponds to how the statistic increases or decreases with time, implying that flies differ in how their behaviors change over time. In addition, the coefficients seem to level off in the last 23 intervals, perhaps relating to the change in behavior due to hunger.

Based on the observation that the first and second principal components for different statistics are similar, we found the average first and second principal components (where signs are set as above). We project the data onto these 2 principal components for each statistic, resulting in 24-dimensional vectors. We repeated the analysis in Figure 4.27 on this new 24-dimensional data set. That is, we performed PCA on these statistics of all 12 plotted behavior statistics. The results are shown in Figure 4.29. As in Figure 4.27, the first plot shows the projection of the 24-dimensional data on the first two combined principal components, that is, the highest variance 2-D linear subspace of the combined data. Again, we see no clear clusters. Male flies appear to be more extreme in the first dimension, and female flies more extreme on the second dimension. In the second plot, we show the error of the projections onto increasing numbers of principal components. In the third and fourth plots, we show the directions of the first 3 principal components. The sign is chosen so that the mean element is positive. The third plot shows the elements of the combined principal component corresponding to the first per-statistic principal components, while the fourth plot shows the elements of the combined principal component corresponding to the second per-statistic principal components. We can attempt to interpret the first combined principal component; the second and third have no obvious interpretations. We see that the first combined principal component is close to 0 for all the second per-statistic components, and the absolute weight of the first per-statistic component decreases with the error of the regression. For most of the behavior

statistics, the weight is positive. The statistics with positive weight correspond to those we associate with an active fly (with the exception of the fraction of time near the water, which has a small but positive weight). The two negative weight statistics fraction of time in the observed chamber and median stop duration, would be associated with a more sedentary fly. Thus, we can interpret the first principal component as a measure of how active the fly is.

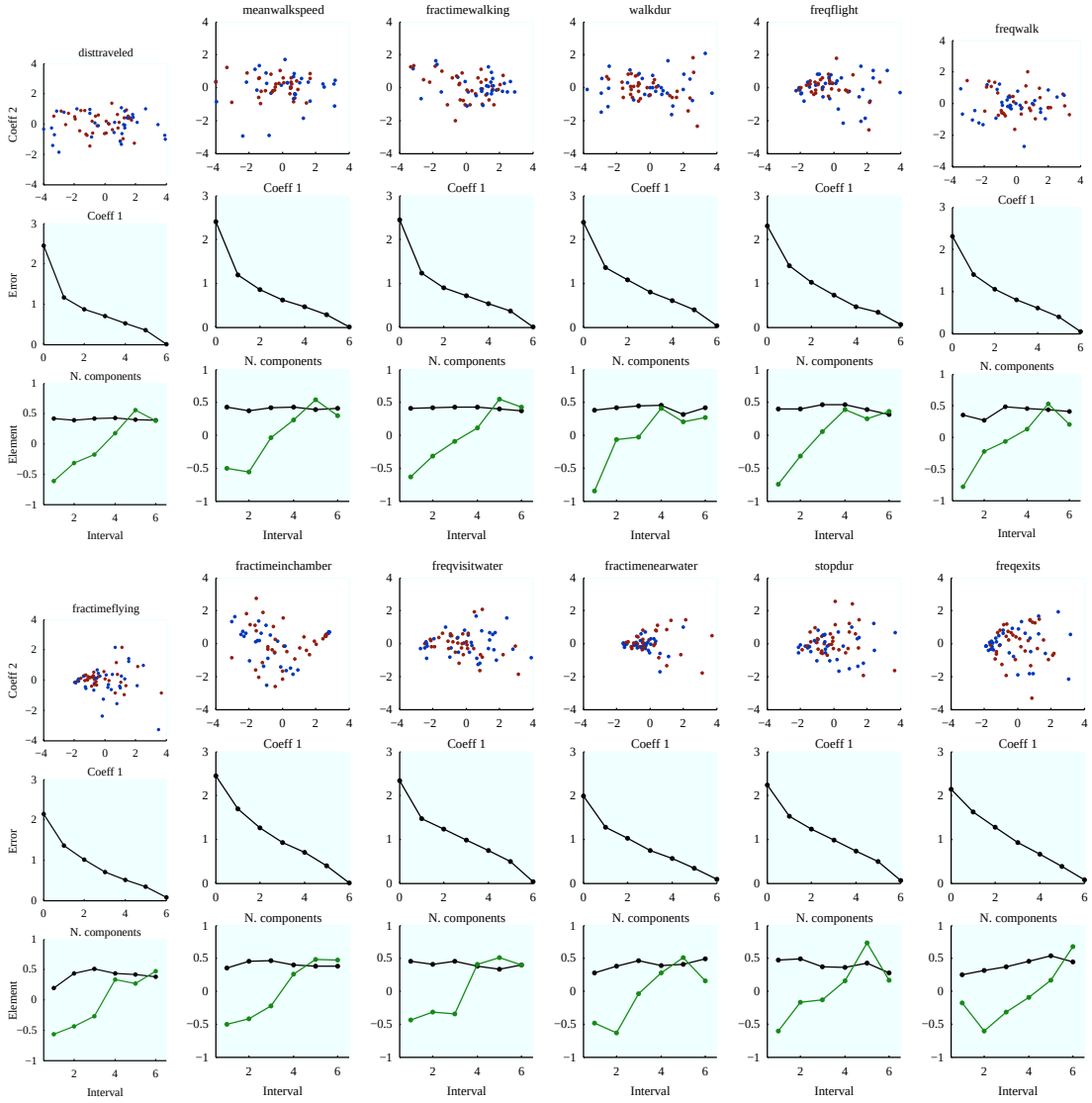

Figure 4.27: Principal component analysis per statistic. For each statistic, in the top row we plot the projection of the 6-D vector of per-interval statistics onto the first 2 principal components. There is a point for each fly; males are plotted in blue, females in red. In the middle row, we plot the error (square root of the average sum-squared error) of the reconstruction of the z-scored per-interval statistic vector with varying numbers of principal components. N. components = 0 corresponds to just using the mean, while N. components = 6 corresponds to using all principal components, and thus will always have error = 0. In the bottom, we plot the first (black) and second (green) principal components. For many of the statistics, the first component is an average over all intervals, while the second measures change in the statistic over the trial. We flip the sign of the first principal component so that its average element is positive, and the second principal component so that the last element is bigger than the first element to emphasize these trends.

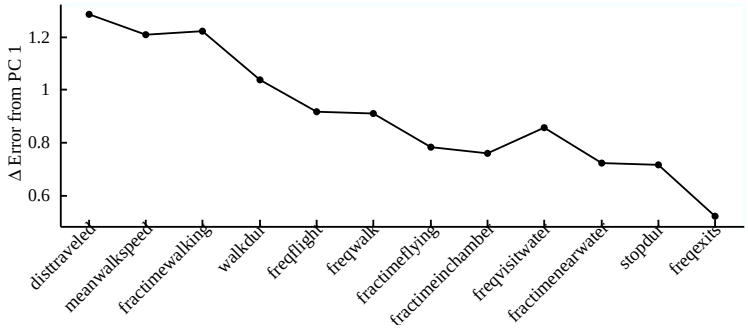

Figure 4.28: Decrease in reconstruction error by including the first principal component. For each statistic, we compute the decrease in error between the mean-based reconstruction and the one-dimensional principal component reconstruction. The statistics are sorted by error in the regression for buffer length = 1 (see Figure 4.26). Note that the decrease in error decreases monotonically as regression error increases.

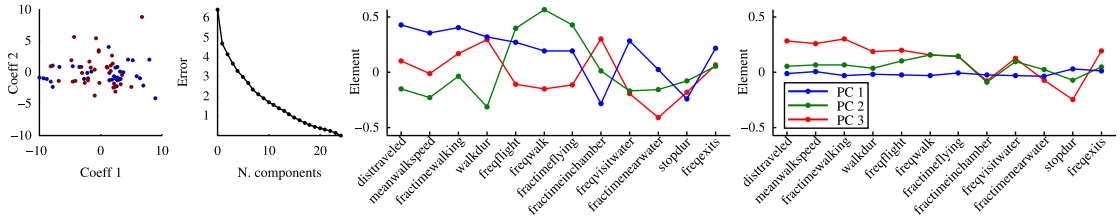

Figure 4.29: Principle component analysis of first 12 statistics combined. Following the observation that the first and second principal components for the first 12 statistics were similar, we found the average first and second principal component over all plotted statistics. We projected the per-interval vectors for each of these 12 statistics onto these first two mean principal components, resulting in 68 flies $\times$ 24-dimensional vectors. We then performed the same analysis as in Figure 4.27 on these new vectors. In the first plot, we show the projection onto the first two principal components. There is a point for each fly; male flies are blue, female are red. In the second plot, we show the error (square root of the average sum-squared error) of the reconstruction with varying numbers of principal components. In the third and fourth plot, we illustrate the first 3 principal components. The third plot shows the principal component elements for the elements derived from the first per-statistic principal component. The fourth plot shows the principal component elements for the elements derived from the second per-statistic principal component.

4.5 Discussion

Behavioral flexibility is often proposed as an adaptation that allows individuals to maximize their fitness within the multifaceted environments an animal may encounter over its lifetime (Dingemanse and Réale, 2005). However, it has been reported that the majority of the time individuals exhibit very limited behavioral plasticity (Sih et al., 2004a,b) and also that animals exhibit consistent differences in their reaction towards the same sensory stimuli in their local environment (See references within (Dingemanse and Réale, 2005)). In order to observe the degree of behavioral plasticity and the consistency of differences of an individual's behavior to the same environmental stimuli, we introduced single flies that were reared and handled in a similar manner into homogenous model environments, and observed their movement over the period of hours. We were particularly interested in the exploratory movement patterns of individuals in relation to a source of water and the exit into an adjacent chamber. We quantified various basic measures of walking and flying that contribute to exploration, and also several higher-order measures of the the flies' exploratory movements. We tested the following predictions: (1) would individuals show markedly different degrees of exploration?, (2) would the difference an individual's exploration persist over time?, and (3) would the various measures of exploration be independent and thus not merely a consequence of a more general phenomena such as activity level?

We report significant differences among individuals in their exploratory movements. The differences among individuals were apparent in our raw observations of their movement, e.g., the position of each fly throughout the time course of the experimental trial (see Figure 4.1), or the total distance traversed during each successive 10 minute internal (see Figure 4.3). The differences among individuals were also salient in simple

per-frame statistics describing their walking and flying movements (see Figures 4.5–4.13), and in higher-order quantitative descriptions of their exploratory and dispersal behaviors, e.g., search near water and movements that resulted in leaving the chamber (see Figures 4.14–4.21).

At the time of writing up this dissertation, we have only examined the 6-hour data set. Within this data set, the characteristic structure in exploratory movements of individuals that were observed was stable for greater than over 5 hours in all behavioral descriptors observed except in the following measurements: (1) the fraction of time flies spent flying (*fractimeflying*), two related measurements of local search near water: (2) the time flies spend between water visits (*timebtnwater*) and (3) the travel length of movement between water visits (*pathlengthbtnwater*), and also two measurements related to dispersal from the chamber: (4) the number of exits from the chamber (*freqexits*) and (5) the total time spent in outside the chamber (awaydur). It is likely, however, that these measurements would exhibit persist characteristic structure if we had a larger sample size, for these five descriptors of exploratory behavior are made up of the other descriptors that *did* exhibit persistent structure, and also there was a significant amount of error observed in the measurements of these descriptors due to some movements quantified by descriptors never or very infrequently occurring. Finally, from our the dimensionality reduction, we suggest that many of differences seemed related and were due to activity and a component related to time of day (see Figures 4.27,4.28,4.29).

4.6 Supplementary materials

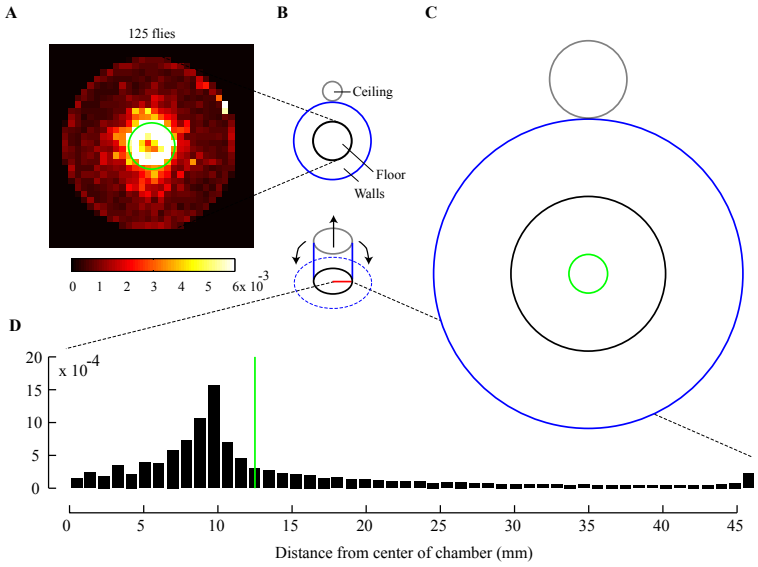

Figure 4.30: Collective transit probabilities and individual local searching movements near a source of water. (A) Collective transit probabilities for 125 flies on the floor of chambers near a source of water (green circle). (B) Illustrations to help visualize the flattened projection for displaying the individual movements of flies near water shown in C and also to show the cross-section transect (red) for the probability histogram in D. (C) To help illustrate the trajectory of a fly, we unfold its 3D positions within the experimental chambers and report its movement in a flattened representation. A patch of agar (green circle) embedded within the center of floor prevented flies from dehydrating.

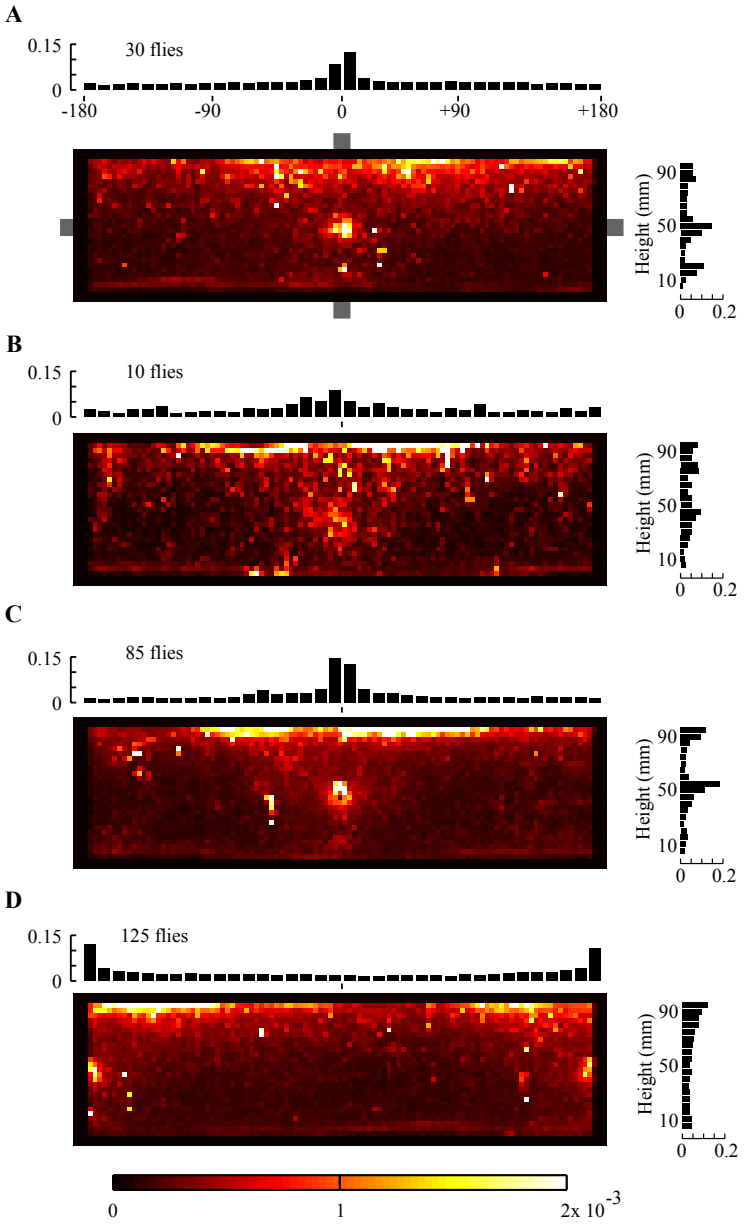

Figure 4.31: Tactile and visual cues are salient features of the exit leading between chambers. Collective transit probabilities for flies on the wall of chambers when the exit is (A) blocked, (B) covered with transparent material allowing light to pass through, (C) open to a second connected chamber, and (D) rotated 180 degrees, for all trials from A-C. Histograms of the transit probabilities calculated from one centimeter horizontal and vertical strips (gray bars, as denoted in A) are shown above and on the side of each panel.

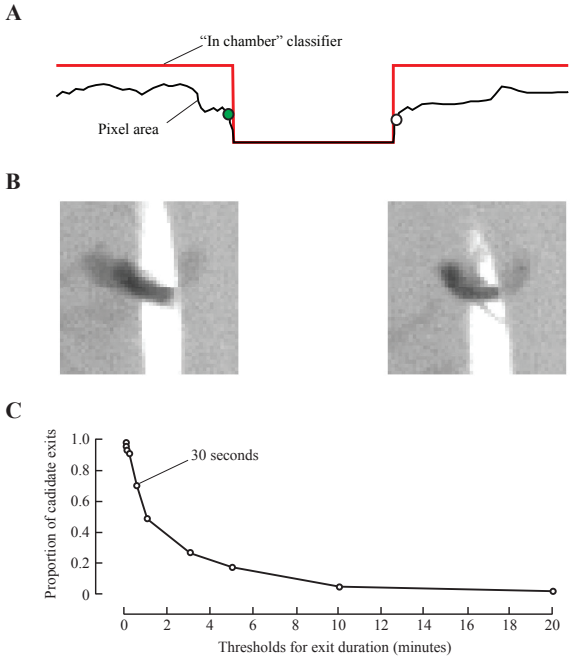

Figure 4.32: Graphic illustrating the classification of exits from a chamber. (A) Flies were *nominated* as exiting the chamber at a particular image frame when at that frame the pixel area of the binerized difference between the image and its corresponding background image dropped to zero. (B) Example image frames for a fly just proceeding to an exit (filled green circle in A) and just after returning from the second chamber (open circle in A). (C) Proportion of classified exits from total number of candidate exits as a function of the shortest time away that constitutes an exit. A fly was required to have left the chamber for at least 30 second to be considered an exit; this criteria, in one particular data set, restricted the number of leaving events that were classified as exits to 860/1222.

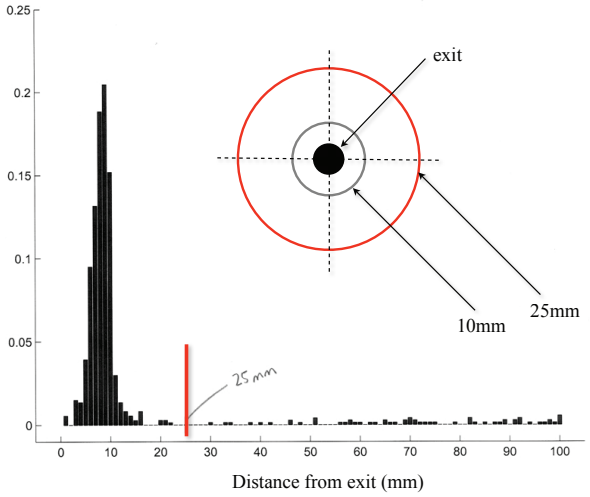

Figure 4.33: Frequency histogram of distances from known 3D exit locations for candidate exits events. Candidate exits initially classified by pixel area were excluded if their distance was greater than 25 mm from the center of the known exit location (red line; red circle within inset). False exits were rare, e.g., 66/1222 for a particular set of trials, and generally due to an adaptive thresholding error when a fly had exited and was currently outside the chamber.

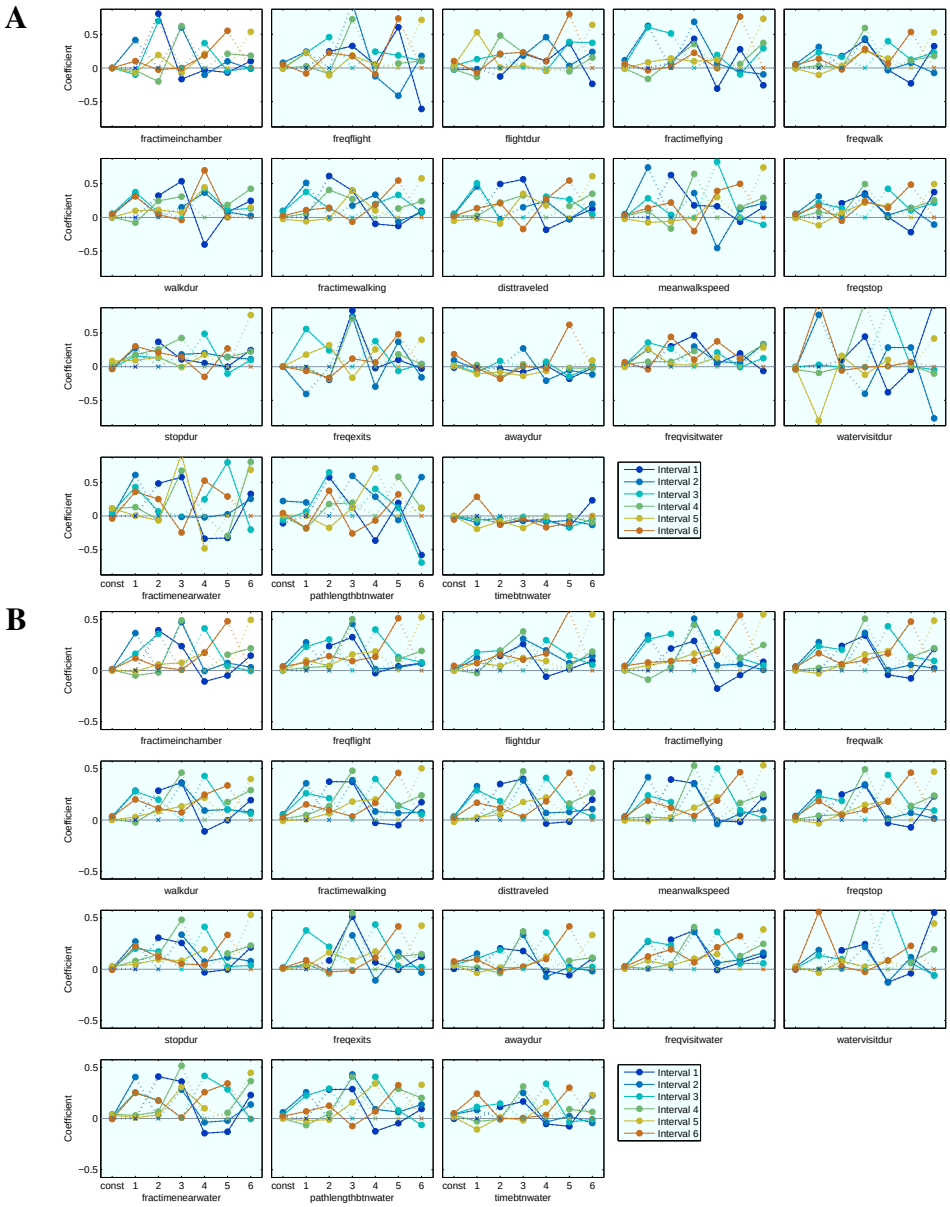

Figure 4.34: Coefficients of regressors learned. Each plot corresponds to a different statistic. The x-axis describes to which input vector element the coefficient corresponds, either the constant offset 1 or the statistic for one of the input intervals. The y-axis corresponds to the value of the coefficient for that input element. There is a line for each of the intervals predicted. The dashed lines and X's indicate intervals that are not input. The buffer length for these regressions is 0. (A) shows the coefficients learned with ordinary linear regression, while (B) shows the coefficients learned with regularized linear regression.

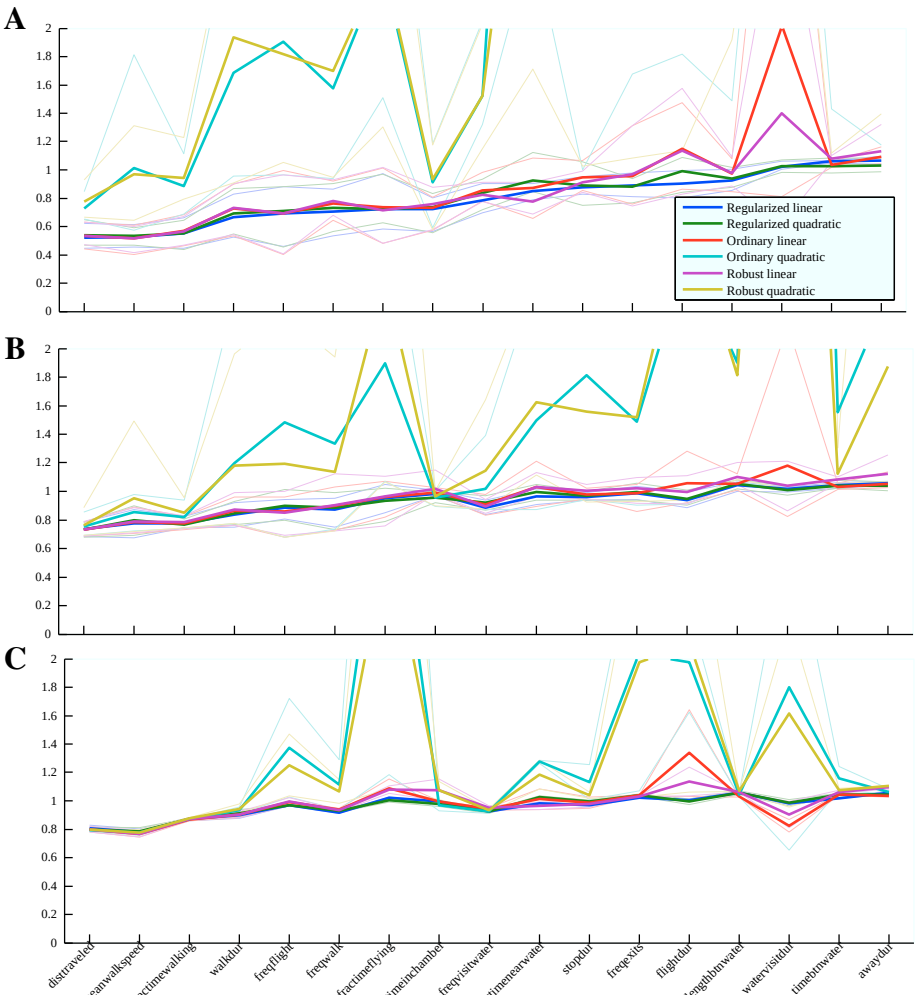

Figure 4.35: Comparison of different learning algorithms. Each plot shows the square root of the mean squared error between the true and predicted behavioral statistic, as in Figure 4.26. Each color corresponds to a different type of learning algorithm/regression criterion optimized. The thick lines show the mean error over all intervals for a particular regression type, and the thin lines show the minimum and maximum over all intervals for a particular regression type. Each plot corresponds to a different buffer length: (A) buffer length = 1 intervals, (B) buffer length = 3 intervals, (C) buffer length = 5 intervals.