Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Katherine Overman Date

Long Time Scales and Hierarchical Structure in Drosophila Behavior

By

Katherine Overman
Doctor of Philosophy

Physics

Gordon Berman, Ph.D.
Advisor

Justin Burton, Ph.D.
Committee Member

Ilya Nemenman, Ph.D.
Committee Member

Sam Sober, Ph.D.
Committee Member

Daniel Weissman, Ph.D.
Committee Member

Accepted:

Kimberly J. Arriola, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Long Time Scales and Hierarchical Structure in Drosophila Behavior

By

Katherine Overman
B.S., Augusta University, GA, 2014

Advisor: Gordon Berman, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Physics
2022

Abstract

Long Time Scales and Hierarchical Structure in Drosophila Behavior
By Katherine Overman

Animal behavior is fundamentally a biological process with complex dynamics. However, to fully understand this process, an interdisciplinary approach is essential. In this thesis, we used methods from physics and machine learning to study behavior. We aim to quantify *Drosophila melanogaster* fruit flies' stereotyped behaviors, to build a model that is capable of reproducing complex, hidden behavioral features, specifically the long time scales and hierarchical structure that we see in our data, and to offer explanations for what internal processes are modulating these features, both in general and in the context of aging. The data we are using are hour long recordings of fruit flies imaged from above in a featureless dish that restrict their ability to fly but otherwise allows them to move freely. Using methods from Berman (2014) [8], we are able to take these recordings, identify the set of stereotyped behaviors and embed them into a 2D plan known as the behavioral map. This indirectly gives us a time series of behavioral states, and we have two different frameworks of this data that we use.

The first consists of 300 fruit flies, half male and half female, with ages ranging from 0 to 70 days, which is the typical lifespan of these flies. With this data set, we measure behavioral changes as a function of age and find that the changes differ between the males and females. As youths, the female flies are very active, performing mostly locomotive behaviors or behaviors that require lots of movement. As they age, they become more lethargic and perform more idle or "lazy" behaviors. On the other hand, male flies are lazy as youths, perform more active behaviors in mid-life when it is optimal mating time, and finally return to laziness in their late life. We find that many of the hidden dynamics that we measure from data, show no change. This includes the entropy of the behavioral map, the stereotypy (or repeatability) of the behaviors, the long time scale dynamics, and the hierarchical structure in the data. However, we also find that a major contributor to the change in behavior is a change in the amount of energy available to the flies, or rather, that the flies have a changing energy budget that modulates their behavior.

The second data set has 59 young, male fruit flies. We use the time series from this data to train a recurrent neural network so that it is capable of reproducing the data and the data's hidden dynamics. It has been shown by previous work that the internal states of an RNN act like a dynamical system with fixed points, and that the interactions between the fixed points lend some clarity on how neural networks make their computations [117]. By doing so with our network, we find that the fixed points of our network behave as a multi-well plane with basins containing one or more behavioral states. The plane shifts over time which changes the barriers between our basins, and we show that this paradigm is capable of modulating the dynamics in our data.

In conclusion, this thesis serves to offer explanations for the types of internal mechanisms that are necessary to produce the dynamics that are inherent to behavioral

data across long time scales, even as long as a fruit fly's lifetime. Common behavioral models are currently incapable of doing this, thus we also propose neural networks are a better class of model to be considered for future behavioral studies. These models could be crucial to unlocking what is modulating behavior within an organism.

Long Time Scales and Hierarchical Structure in Drosophila Behavior

By

Katherine Overman
B.S., Augusta University, GA, 2014

Advisor: Gordon Berman, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Physics
2022

Acknowledgments

Seven years ago I joined Gordon's lab in his first year of being a professor at Emory. At the time, I underestimated the risk I was taking in doing so, however I'm very grateful I took that risk as working for Gordon has been an excellent experience. Throughout my time as his graduate student, I've never doubted that I could come to him with any problem, whether work-related or personal. I cannot stress enough the importance of that quality in an advisor, especially as a woman in a male dominated field. Thank you for everything, Gordon. It's been a pleasure to be a part of the lab and see it develop and grow. I'd also like to thank the folks who have joined the lab after me that have contributed to a supportive and enjoyable environment.

To my support system outside of the lab, both friends and family, I could not have survived the past seven years without y'all. One of the best outcomes of me moving to Atlanta for grad school is the friendship it has helped cultivate with my sister, Pamela. Growing up with a 6-year age gap made friendship challenging at other stages in life, but now she is my closest friend and fiercest supporter. Massive thanks to both my parents who have always been there for me and offered nothing but support and love. To everyone else who has been with me on this journey: I may have not always been my best self but thank you for being with me and supporting me along the way.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Animal behavior is a biological process that occurs in response to a stimulus, whether internal, external, or both. For example, a cat hearing the shake of a treat bag will invariably run over to the source of the sound in the hopes of being fed. This is a response to an external stimulus (the sound of the bag) and internal stimulus (hunger, or more likely, greed for food).

While studying animal behavior is inherently a biological problem, many of the questions asked about behavior require a multidisciplinary approach using methods and ideas from chemistry, mathematics, physics, and computer science. Behaviors are complex and dynamic by nature, driven by multiple factors like genetic and developmental processes within an organism, external stimuli, and the surrounding environment. How those processes interact and react to the external factors can also have behavioral impacts. Understanding which biological processes modulate animal behavior is an important question that many disciplines have tried to answer dating back to the ancient philosophers [33]. There have been improvements in behavioral studies thanks to advancements in genetics [109], developmental biology [47], and neuroscience [19]. It has been shown that many behaviors occur as a reaction to a physiological change via the nervous system - eating when hungry, sleeping when

tired, etc [18]. There are still many remaining behaviors that have no such explanation though. For example, what explanation can be offered for a dog getting the zoomies?

With all these complexities, it is unsurprising that studying behavior is both difficult and fraught with limitations. In an ideal world, we could collect data on any organism in its natural environment. That data would contain information on both the observable behaviors, as well as readings from internal mechanisms that are known to modulate behavior. We would be able to build informative and intuitive models that are capable of making important predictions and shed light on how internal mechanisms communicate to produce behavior. From a technological standpoint, there are multiple complications. Working in a lab environment alters individual behaviors and behavioral repertoires as a whole. For example, in the lab setup for the data used in this thesis, the flies' movements are restricted in a way so that they are unable to fly. The use of markers is common for tracking purposes, but animals have been known to pick up perturbations in their motions even after acclimating to the marker. Some animals never acclimate to having a marker attached to them. Other issues arise from the nature of behavior itself. For example, to fully capture behavior and to understand it, data would need to include information about every behavior across the animal's whole lifetime as behavior spans multiple time scales. These time scales range from seconds (e.g. grooming, locomotion, blinking), to hours (e.g. circadian rhythms, hunger) to years (e.g. aging, puberty, mating seasons). These are just some of the challenges that one faces when studying animal behavior.

This thesis offers two major advancements to the field. First, we investigate a data set from recordings of male and female fruit flies across their lifespan and quantify how their behavior changes with age. Further, we find that a major contributor to this change in behavior is a change in the overall energy budget available to the fly at a given age. Second, we propose a model to explain how internal mechanisms result

in behavior. We do this by training a recurrent neural network (RNN) on behavioral data from fruit flies and studying the internal dynamical system of that RNN. We find these dynamics are comparable to a particle in a multi-well energy landscape with barriers changing under Arrhenius behavior, and that these dynamics produce the long time scale and hierarchical structure seen in the data.

There exists a wide range of animal behavior studies and how they are conducted. However, there are three major themes that, although varying, can be seen throughout the field: defining behavior, measuring behavior, and modeling behavior.

1.1 Defining Behavior

How we define behavior is an important question. There already exists several varying and (often) opposing definitions of behavior ranging from "the total of movements made by the intact animal" [119] to considering anything an organism does as its behavior [26]. Some definitions establish a connection to the organisms environment [52], while others posit internal processes dictate external movements and consider the combination of internal and external factors to be behavior [63].

Notably, Tinbergen (1963) [120] proposed four questions to ask of animal behavior that established how to think about studying animal behavior.

- **Adaption**: This question addresses why the animal does a behavior and specifically how it contributes to the animal's survival (or fitness). For example, species that migrate during a season do so as a mean's of survival.

- **Phylogeny**: We know natural selection modifies behavior over a long period time. This question asks how behaviors evolve. The answer usually comes from looking at comparable behaviors in species that are closely related.

- **Mechanism**: Which external stimuli or physiological mechanism(s) caused the observed behavior to be performed? Examples include eating when hungry,

seeking out a mate when certain pheromones or hormones are detected, or a cat running under the couch in response to a knock at the door.

- **Ontogeny**: The final question investigates how behavior changes over the animal's lifetime - often a result of learning, like baby birds and their song.

These four questions can be classified into two different categories. Adaption and phylogeny, also referred to as function and evolution respectively, require extended perspectives in time to explain the behavior we currently observe and are classified as ultimate explanations. As for mechanism and ontogeny (or causation and development), they classify as proximate explanations looking only at immediate causes for a behavior. In order to fully understand all the necessary components behind a behavior, both classifications must be explained and all four questions answered [120].

Another important contribution from Tinbergen was the idea of hierarchical structure in behavior, where the same behavior can have different meanings depending on the context. An example of this is demonstrated in the context of reproductive instinct (diagrammed in Figure 1.1), you may have bouts of fighting, building, mating, and care for offspring. These bouts will consist of different actions or behaviors, and there may be overlap in the actions seen. This concept leads into a way of building up long time scale structures in behavior which will be discussed more in later sections of this thesis [119].

Others studied behavior from a more experimental approach rather than theoretical musings. A historically famous behaviorist B.F. Skinner believed you could only study observed behavior, and that all behavior was either random or a product of the environment. To support this claim, he designed an experiment where he placed a rat in a box with a lever that is connected to a food dispenser. By pressing the lever, the rat receives some food and eventually learns to associate pressing the lever with getting food. This set the precedent for teaching animals new behaviors through operant conditioning [107]. The experimental setup he designed, known as the Skinner

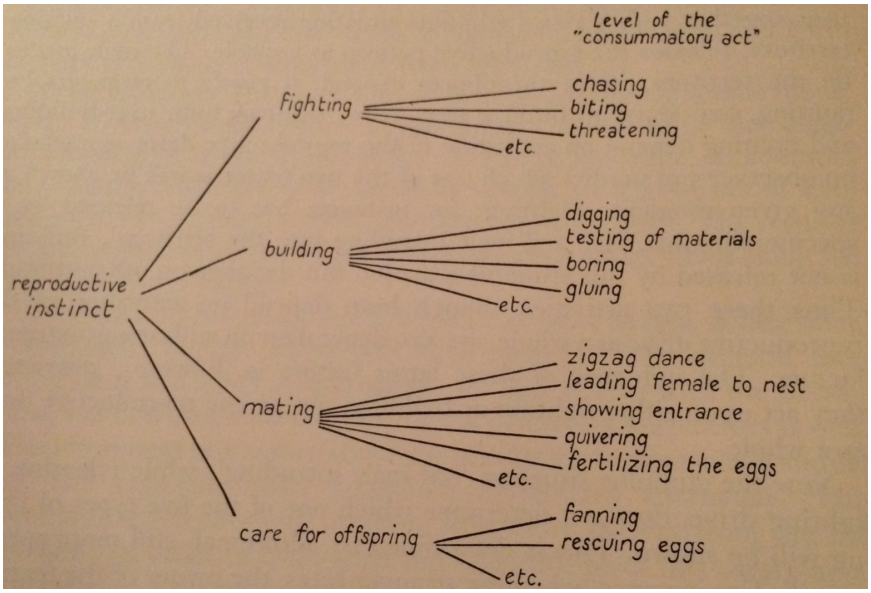


Figure 1.1: Hierarchical and long time scale structure in reproductive instinct behaviors. Figure taken from Tinbergen (1951)[119]

box, is still used in labs today to manipulate and teach behaviors.

An additional aspect to defining behavior is quantifying behaviors. How do you assign numbers to behavior in a meaningful way and how do you interpret those numbers? This consideration has gained importance in the field relatively recently thanks to advancements in technology and other fields such as mathematics and engineering. Further, what behaviors do you consider relevant to your study? For many it is sufficient to study one particular action in individuals such as the gait of a mouse [71], or in large groups such as the overall movement of a group of baboons [114] or the geometry of how bees will cluster [95]. Other studies will look at simultaneous recordings of behavior (arm movement [111], facial expressions [31], birdsong vocalizations [91]) and their corresponding neural spikes. On the other end of the scale, several studies have been done to measure the full repertoire of behaviors [8, 112].

1.2 Measuring Animal Behavior

Once a definition of behavior has been selected, the task is to measure that behavior, if possible. Despite some technological advancements, measuring animal behavior remains a challenge. There currently exist multiple approaches to measuring animal behavior, but perhaps the most prevalent approach is image based analysis, which can range from tracking multiple animals in a 2D plane [29] to using multiple cameras to track animals in a 3D arena[115]. Using computer vision to automatically detect and track lab subjects has had a lot of development through algorithms such as DeepLabCut [78] and LEAP (LEAP Estimates Animal Pose) [97], which train networks on a few hand-labeled frames - assigning markers to distinct parts of the animal's body. These algorithms are looking at the overall shape of the animal, however, some other approaches track movements of the animals by detecting their center of mass. This approach can be useful particularly if you're studying social interactions - ignoring any bouts where the center of masses are too far apart to be interacting [25].

When tracking more than one animal at a time, maintaining the identity of each animal throughout the data set can be an issue. Animals interact with each other, and these interactions will often confuse an algorithm, resulting in an identity swap. Further, depending on the recording setup, you may lose sight of the animal for a period of time which can also result in identity confusion. However, several studies have been done to advance tracking abilities. One such study (idTracker) using videos of a school of zebrafish found that each fish has a unique "fingerprint". This fingerprint is generated by taking each pixel on the fish and calculating it's distance to every other pixel on the fish. The resulting 2D array successfully creates a unique identity for each fish. Using the fingerprints as a means to maintain identity eliminates nearly all identity switches and is robust to occlusion [98]. More recently, SLEAP (Social LEAP Estimates Animal Pose) is a LEAP successor that is capable of tracking multiple animals. Like LEAP, it uses human input to label body parts in various frames in

order to track the animals in the remaining frames while maintaining each animal's identity [96].

Measuring limb movements is another popular approach to measuring behavioral changes, particularly in animals that have rigid or constantly visible legs, such as fruit flies [80]. Although, studies like this have been done in other species as well, like in mice. Results show that gait size in mice can determine overall body size, and gait patterns can predict various neurological disorders [71].

Machine learning has become a useful tool for measuring animal behavior, especially as the field becomes more advanced. Machine learning algorithms can be classified into two different types - supervised and unsupervised. The distinguishing feature between these learning styles is whether or not the data is labelled. As the names might imply, supervised algorithms get labelled data or some sort of human input and can solve problems dealing with classification and regression. Unsupervised learning analyzes unlabelled data with no human intervention to cluster, make associations, or reduce the dimensionality of the data set. Both have their advantages and are used in the field regularly.

A popular supervised learning algorithm is the Janelia Automatic Animal Behavior Annotator (JAABA) [55]. This algorithm automates behavioral classification. It presents the user with short snippets of videos from data of fruit flies moving around and asks the user to classify the behaviors seen frame by frame. Using the human classifications, it then makes predictions on the rest of the frames in the data set. The user can then step through those predictions and agree or disagree with JAABA's classifications. Another supervised learning example is MiceProfiler [28], which automates tracking of two mice. Then, based off of the orientation of the two mice, classifies different types of interactions as a function of proximity. In the past, data sets like these would have to be hand labelled so being able to automate this process saves a tremendous amount of time. However, there remain limitations - certain

frames are hard to classify for both the algorithm and a human. Without knowing all the possible categories, humans struggle with classification problems. Also, by asking for human input, human error and bias are introduced into the analysis.

On the other hand, unsupervised learning has become a useful tool in automating behavior identification without human input. Stephens et al. (2008) [112] measured the curvature of *Caenorhabditis elegans* moving around freely in a featureless dish and, by looking at the space of natural postures in the worm, found the dynamics to be low dimensional. In fact, the space can almost entirely be explained by four "eigenworms", qualitatively corresponding to forward crawling, reversals, Omega-turns, and pauses. Any posture the worm produces can be broken down into some linear combination of these eigenworms. Within the postural space and their determined equations of motion, they found a set of attractors. This paper established a language to describe worm behavior that is generalizable to other species.

Another landmark unsupervised example is MoSeq[130]. This study builds off a previous work[129] that measured 3D motion in mice and found that every behavioral bout was composed of subsecond modules. These modules could be treated like syllables and fed into language processing models. Ultimately, they found that every type of behavior the mouse expressed could be described by varying these modules, and that this method is robust to genetic and neural mutations. MoSeq tests the robustness of this method on the behavioral effects of drugs. It can distinguish between the differences in behavior even when the mice have been given closely related drugs and identify which drug caused behavioral changes regardless of the actual behaviors displayed, which reveals information about drug and gene perturbations that was previously unknown and even still not totally understood.

For this thesis, we will be using an unsupervised algorithm to measure the full repertoire of a fruit fly and quantify that repertoire as a 2D behavioral space developed by Berman (2014) [8], and it will be an important tool used throughout this thesis.

1.2.1 Behavioral Spaces

This method has many benefits over other methods. First, being able to measure the full repertoire of behavior (in a lab context) as opposed to singular actions provides rich information about behavior and behavioral dynamics - how do behaviors interact over time? The unsupervised approach saves time by not requiring a human to make and assign labels for the data set. This also eliminates any human bias or error. Finally, the method allows for comparisons across individuals or classification of individuals (males vs. females, difference species, different age groups, etc.). By comparing male and female fruit fly behavioral spaces, it was discovered that the two groups perform left-wing grooming in slightly different ways, and the algorithm was able to distinguish them as unique behaviors rather than lumping them together into one state. Thus, the method is also robust to minor variances in how behaviors are performed. These spaces are generated from videos of freely moving fruit flies using an unsupervised approach that employs computer vision, dimensionality reduction, and various other computational techniques.

1. **Image segmentation and registration**: To start, edge detection is used to isolate the fly in each frame from the raw data set of videos. While the fly is always in frame, its location and orientation in the frame varies. In order to get a consistent image across all the frames, the flies are rotationally algined using polar cross-correlation with a template image and translationally aligned using a sub-pixel cross-correlation. Additionally, to address variation in the size of the flies' bodies, the images are rescaled so that they are all the same size or cover the same number of pixels on average. In this data set, there were $4\text{x}10^4$ frames after image processing.

2. **Postural decomposition**: Projecting the pixel values of these images onto a 50-dimensional Euclidean space explained over 90% of the variance in the flies'

postures. After a Radon transform, which gives a sparse parameterization of the images, principal component analysis (PCA) is applied. PCA is a common dimensionality reduction technique. Here, the outcome of PCA is called the postural modes or eigenflies. Given a fly's posture, it can be described by a combination of these modes, but the instantaneous values of these modes does not complete the picture for the dynamics of behavior.

3. **Spectrogram generation**: Using the amplitudes of Morlet transforms by measuring the power frequency associated with the postural modes within a window of time forms a more complete picture of the modal dynamics across multiple time scales. This gives a spectrogram representation with 25 frequency channels.

4. **Nonlinear embedding**: With 25 frequency channels for each of the 50 postural modes, each point in time is described by a 1250-dimensional feature vector. Given strong correlations between the frequency channel, the actual dimensionality of the postural dynamics is much lower. Prioritizing local structure rather than more distant structure, t-distributed stochastic neighbor embedding (t-SNE) [125] embeds the data points into a space while preserving the Markovian transitions. By applying a Gaussian smoothed probability density function (PDF) over the embedded points, structured peaks and valleys emerge. The red peaks correspond to stereotyped behaviors, and the blue valleys are non-stereotyped behaviors.

5. **Spatial segmentation**: In order to isolate the peaks in the space, the pdf undergoes a watershed transform. This extracts the behavioral states. It is possible to inspect the data within each watershed region and doing so confirms that the states contain unique, stereotyped behaviors.

A visual representation of these steps are laid out in the data pipeline shown

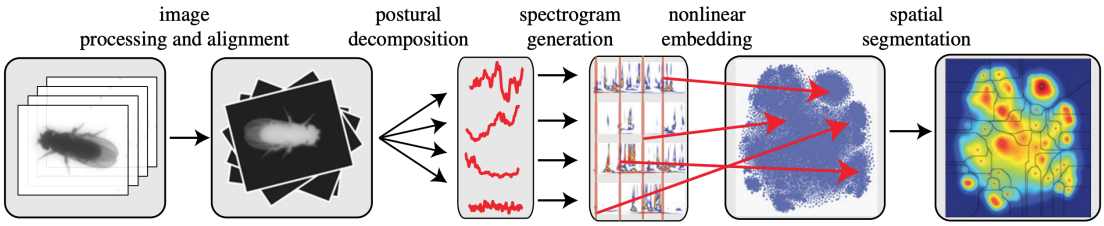


Figure 1.2: Behavioral space data pipeline. Raw images of the fruit flies (*D. melanogaster* are isolated through background subtraction, rescaled and aligned to a consistent size and direction. The resulting stack of images then undergo PCA, which gives a relatively low-dimensional set of time series. A spectogram for the postural modes is generated by applying a Morlet wavelet transform to these time series. Using t-SNE, each point in time is embedded into two dimensions. After applying a Gaussian-smoothed density over the points, the individual peaks are isoloted via a watershed transform.

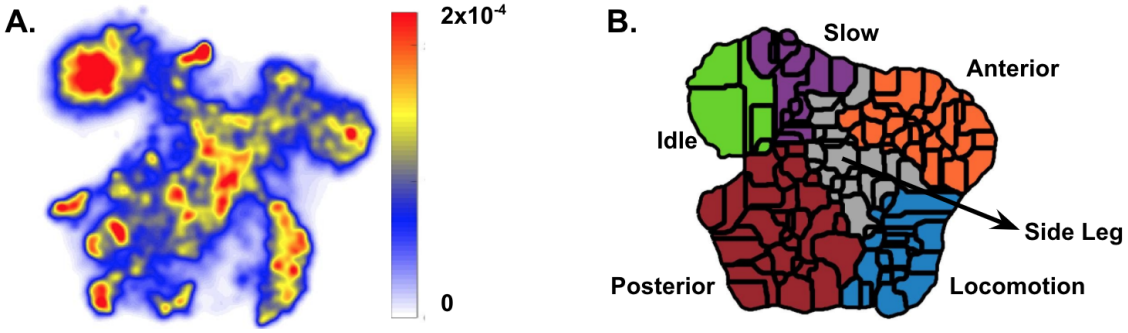


Figure 1.3: Final steps of data pipeline. A. Probability density function generated from t-SNE embedded points. B. Watershed regions with broad descriptions of the stereotyped behaviors in each region. These descriptions were defined by visually inspecting movies from the data.

in Figure 1.2. In Chapter 3, we use data from Berman (2014) [8] to reproduce the behavioral space following this pipeline. That replicated space is provided here in Figure 1.3A.

An important feature of these spaces is that posturally similar behaviors are embedded near each other in the space, which was not inherently built into the algorithm. Thus, by visually inspecting the data from each of the states, it is possible to create clusters on the space with broader behavioral classifications. This is plotted in Figure 1.3B. The black lines, here, are the watershed regions outlining the area containing a behavioral state.

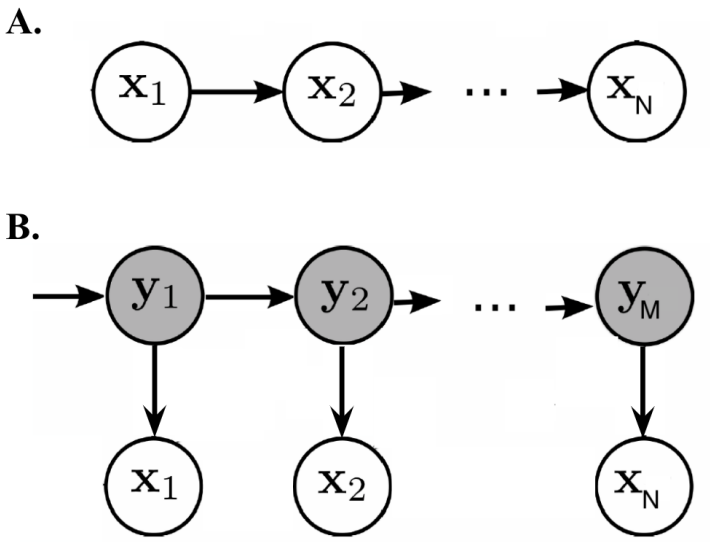


Figure 1.4: Graphical structures: A. Markov Model and B. Hidden Markov model.

The behavioral spaces also contain rich information about behavioral transitions, time scales, and hierarchy which are all important themes to this thesis. We will go into more detail about those features, how we used them, and what they tell us in the following sections and chapters.

1.3 Modeling Animal Behavior

Now, with behavior defined and measured through postural movements, it is possible to ask how to model behavior and what features a model need in order to produce behavioral dynamics.

Historically, Markov models have been the preferred model of choice and even more recently, hidden Markov models [70]. These models are stochastic and make transitions depending on the current state that the system is in. Graphical representations of a Markov model and a Hidden Markov model can be seen in Figure 1.4. In this figure x_i's are observable states, and y_j's represent hidden states. The transitions between these states are determined by probability tensors.

Mathematically, these probability tensors evolve through time according to:

$P(x_{t+1} = i|x_t = j) = T_{ij}$ for Markov models, where t indicates a time step, i and j represent all possible state values, T is the transition matrix containing transition probabilities, P, which dictate how likely state j is going to transition to state i. The primary difference between the Markov model and hidden Markov model is the introduction of the set of hidden states. There is a layer of hidden states transitioning between each other and emit an observable state at each time step. Thus, we have a probability matrix, S. This matrix contains the emission probabilities, or how likely a hidden state will produce an observable state as an output. In addition, there is a transition matrix that determines how the hidden states interact through time. These tensors evolve with $P(y_{t+1} = i|y_t = j) = T_{ij}$ modulating the hidden states and $P(x_t = k|y_t = j) = S_{jk}$ controlling the observable states.

Although these models are desirable for their structural and mathematical simplicity, studies have revealed that any model that restricts dynamics to a single time scale (like the Markov models) fail to reproduce some of the more complex dynamics we see in behavioral data and have significantly less predictive capabilities [9]. For example, if we measure the transition matrix of our data at different time steps into the future, multiple long time scales emerge. Taking similar measure from a Markov model result in significantly shorter time scales, by orders of magnitude even [9]. While it is possible to increase the complexity by adding extra dependencies to the transition, it has been shown that even with a nearly infinite history, the Markov model still performs poorly[2]. We will go into more detail on measuring time scales later in the text, but it is clear that in order to extract complex dynamics, a more complex model is needed.

1.3.1 Recurrent Neural Networks

While some researchers have employed methods like latent state models with feedback [135] and agent based models [105], we have found recurrent neural networks (RNNs)

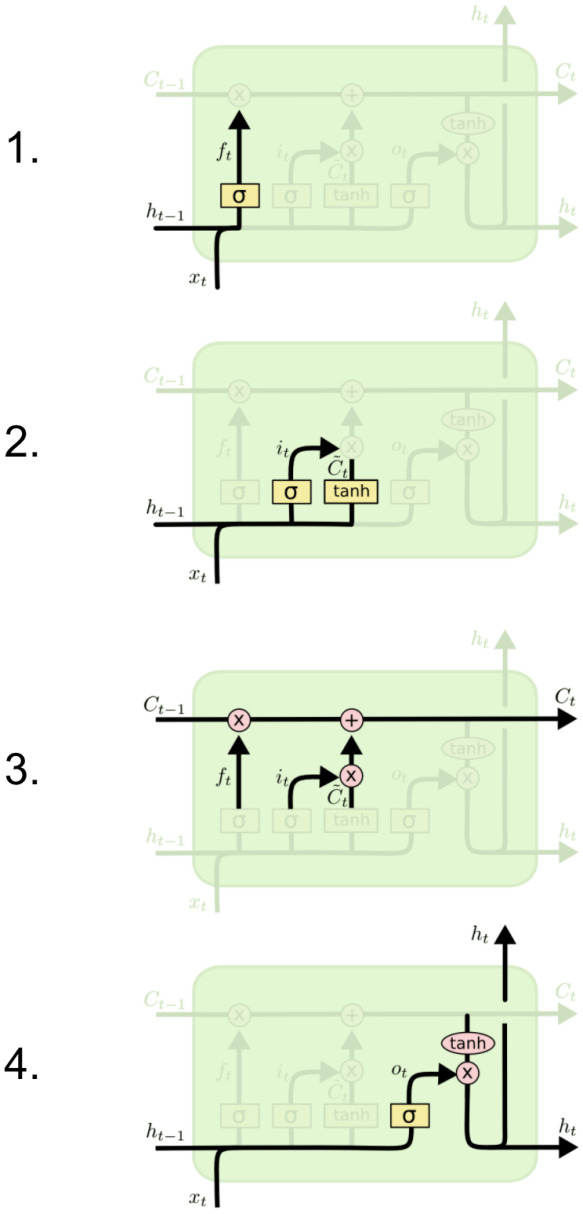
to be useful in studying animal behavior. Neural networks aim to recognize patterns, underlying dynamics, and relationships in data sets and make computations the same way that the brain makes decisions. The basic concept of neural networks has been around since the 1940s starting in the form of an electric circuit that was meant to replicate how neurons in the brain work [79], which eventually led to the invention of the perceptron in the late 1950s, which is the oldest neural network still in use [103]. We now know that perceptrons are only able to act as a binary (or linear) classifier, but when it was discovered that they are unable to predict many classes of patterns, neural networks fell out of interest for several years [75]. Modern day neural networks have many different architectures, but they all have a handful of essentials: input, layer (or layers) of neurons (also called nodes) each with their own weight and bias, and an output. The shapes of the input, layers, and outputs take many forms, but these are the essential elements to the networks. Applications of neural networks include weather prediction, natural language processing [40], facial recognition [59], and recommendation systems [67] just to name a few.

These networks can be trained to learn nonlinear relationships in a data set [134, 22, 77]. To avoid overfitting, where the network learns exactly the data set provided and is unable to make predictions on new data, the data must be split into a training and test set. As the name implies, the training set is used during training, and the test set is used to validate that the network is learning the overall dynamics and not just the data it can see. During the training process, the network is learning to map inputs to outputs by finding a set of weights within the network that successfully produce the desired output or something close enough to the output. This process is iterative making small updates to the model weights in order to improve performance. Further, this process is an optimization problem - typically done through stochastic gradient descent (SGD) and backpropagation [104]. Each iteration the weights are updated with backpropagation of errors and the loss is minimized through SGD.

The basic concept behind this learning algorithm is exploring the plane of errors given different combinations of weights and finding the weights that minimize that error by computing the gradient of the loss function with respect to each weight [94]. While this may sound straightforward, two problems can arise when computing the gradients. In some cases, the gradient may go to zero or infinity exponentially fast, resulting in a vanishing or exploding gradient respectively. When this occurs, the network cannot learn [7].

Long short-term memory (LSTM) [48] networks are a flavor of RNN that avoids these gradient issues. At each iteration, the network decides whether to add or remove information to the cell state C_t by performing a series of calculations. First, at time step t, the forget gate takes h_{t-1}, the hidden state, and x_t, the input, and feeds them through a sigmoid function, σ, which returns a 0 or 1 signifying what percent of the cell state to keep. Next, the new information to store is chosen. The input gate, which is also a sigmoid function, chooses the values to update, and a tanh function creates potential values that could be added to the state, $\tilde{C}_t$. By multiplying the old cell state, C_{t-1} with the output from the forget gate and adding the input gate multiplied by $\tilde{C}_t$, we update the cell state. Finally, the output gate is the result of the updated cell state after a sigmoid function, and the hidden state is updated by multiplying the output gate by the tanh of the cell state. A step-by-step diagram of this with the corresponding equations can be seen in Figure 1.5.

LSTMs are used to solve a variety of problems. We will be using LSTMs to learn and replicate the complex temporal dyanmics that our behavioral data contains. In recent years, there has been an increased interest in neural networks for measuring and modeling behavior [14, 113, 43, 32] (among numerous other applications) as they have proven to be effective at making predictions and replicating complex data sets. How they are able to do this is not clear at all, which for many purposes is fine. For our purposes, if we're able to reproduce a behavioral data set and all the dynamics



$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

$$o_t = \sigma\left(W_o \; [h_{t-1}, x_t] \; + \; b_o\right)$$

$$h_t = o_t * \tanh\left(C_t\right)$$

Figure 1.5: Computational steps for LSTM network. 1. Forget gate. 2. input gate and candidate values to be added to the state. 3. Cell state is updated. 4. Output gate and hidden state. For more details see the text. [93]

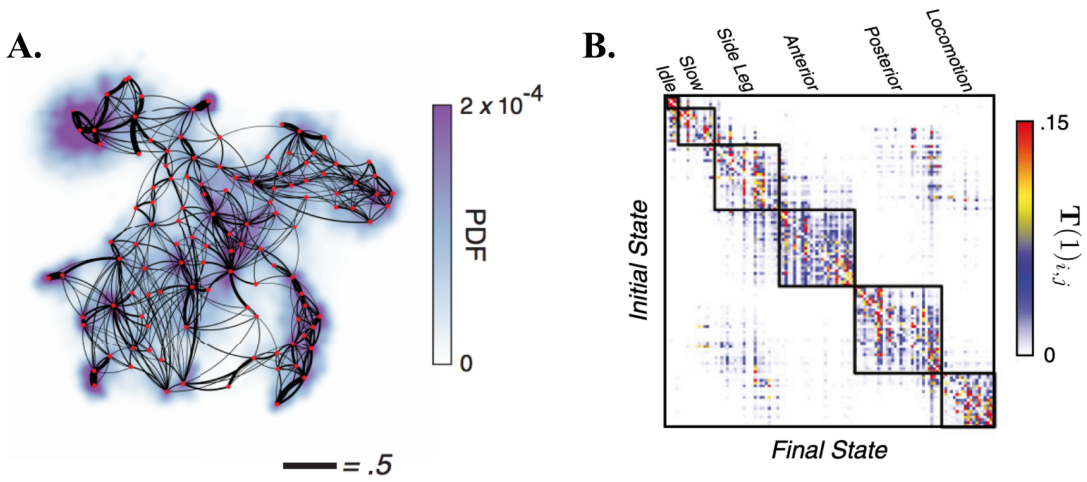


Figure 1.6: A. Transitions rates plotted on the behavioral map. Each red point represents a stereotyped behavior, and the black lines' thicknesses represent the probability of making the transition. Curvature of the lines are right-handed to indicate the direction of the transition. B. One-step Markov transition probability matrix $T(\tau = 1)_{ij}$. There are 117 states clustered into six groups with boundaries indicated with black lines given by the information bottleneck algorithm. Broad descriptions of the behaviors in each cluster are labeled above the matrix. Figure from Berman et al. (2016) [9]

of that data set, we would like to be able to say something about the inner workings of the network so that parallels could be made with the inner workings of the animal itself. Fortunately, lots of work has been done to uncover and demystify the way neural networks make computations [131, 117].

1.3.2 Time Scales and Hierarchy

Some of those dynamics that we want a model to replicate follow the analysis done in Berman et al. (2016) [9] on the behavioral space from Berman et al. (2014) [8]. By looking at the Markov transitions, which quantify the probability of transitioning to a state given the state you're currently in, and the movements from peak to peak in the behavioral spaces, some interesting structures emerge. These transition probabilities are plotted on the space in Figure 1.6A.

It is noticeable that this plot is relatively structured and not a messy hairball.

States mostly transition to other nearby states that are similar. This makes sense from the postural perspective - to go from sitting to standing there are some intermediate postures to enable the action and avoid falling. This is quantified with the transition matrix, seen in Figure 1.6B. The block diagonal structure in the matrix confirms that most transitions are staying local. As it turns out, this matrix also contains significant information about future transitions.

If the dynamics in this data set were Markovian, then iterative multiplication of the one step transition matrix would compute the transition matrix after τ steps, or mathematically, $T_{Markov}(\tau) = T(\tau = 1)^{\tau}$. Even after 100 iterations, the Markovian matrix has lost all information about the initial state (Figure 1.7A). In contrast, the measured transition matrices from the data still maintain a significant amount of structure at $\tau = 100$ (Figure 1.7B, left), even at $\tau = 1000$ (Figure 1.7B, right) some information about the initial state remains in tact, which says that the behavioral dynamics are not Markovian.

The structure of the transition matrices is then quantified by measuring the eigenvalues for different values of τ. According to the Perron–Frobenius theorem, the largest eigenvalue will always have a magnitude equal to 1, and all other eigenvalues will have a magnitude less than one and describe the loss of predictability over time. The resulting curves can be seen in Figure 1.7C. The solid lines correspond to the timescales and predictability of the data set with SEM error bars, and the dashed lines are from the Markovian matrices. This shows the Markov model failing to capture the timescale in the data set by two orders of magnitude. Further work has since been done that shows a Markov model would have to have access to a nearly infinite history in order to reproduce the structures seen in the data [2].

The data's long time scale structure implies that it is possible to cluster the behaviors in such a way that much of the information about the future is preserved [10]. More precisely, given the behavioral space $S(n)$, they want to find a representation Z

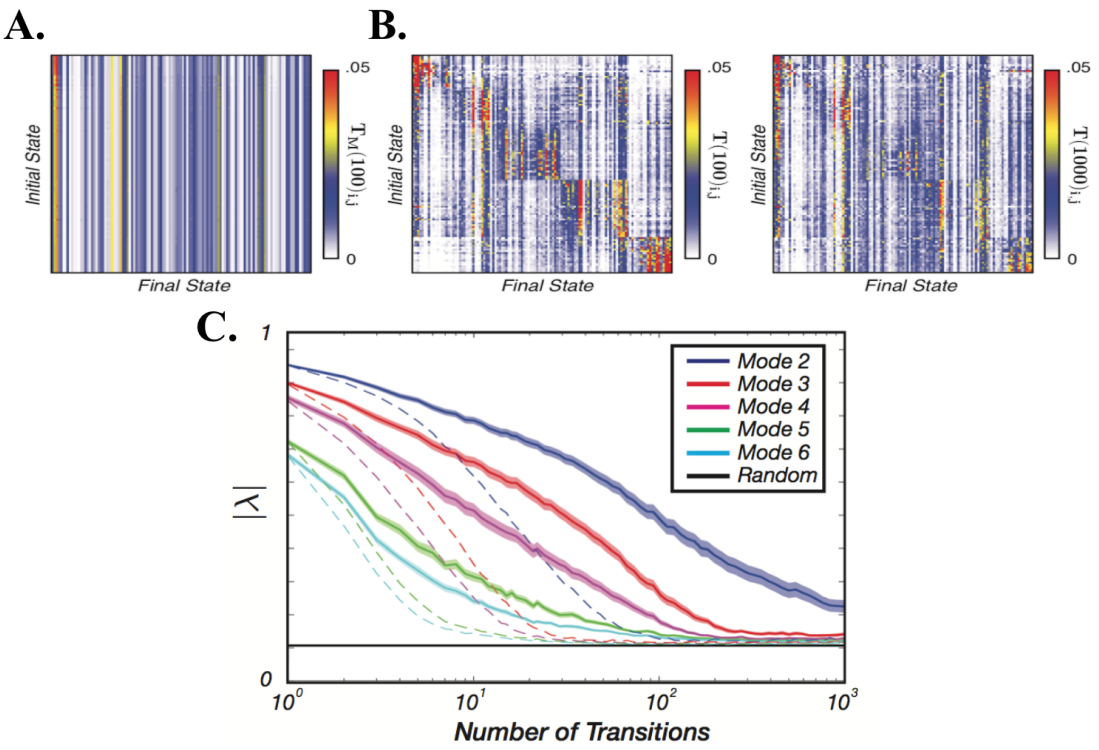


Figure 1.7: Transition matrices across long time scales. A. Markov model transition matrix for τ=100. B. Transition matrices measured from data for τ=100 (left) and τ=1000 (right). C. Absolute values of the transition matrix eigenvalues across different values of τ. The solid curves represent the average over all the flies with SEM error bars. Dashed lines are the Markovian timescales. The black line is a noise floor, the second eigenvalue of a transition matrix after randomly shuffling of the data set. Figure from Berman et al. (2016) [9]

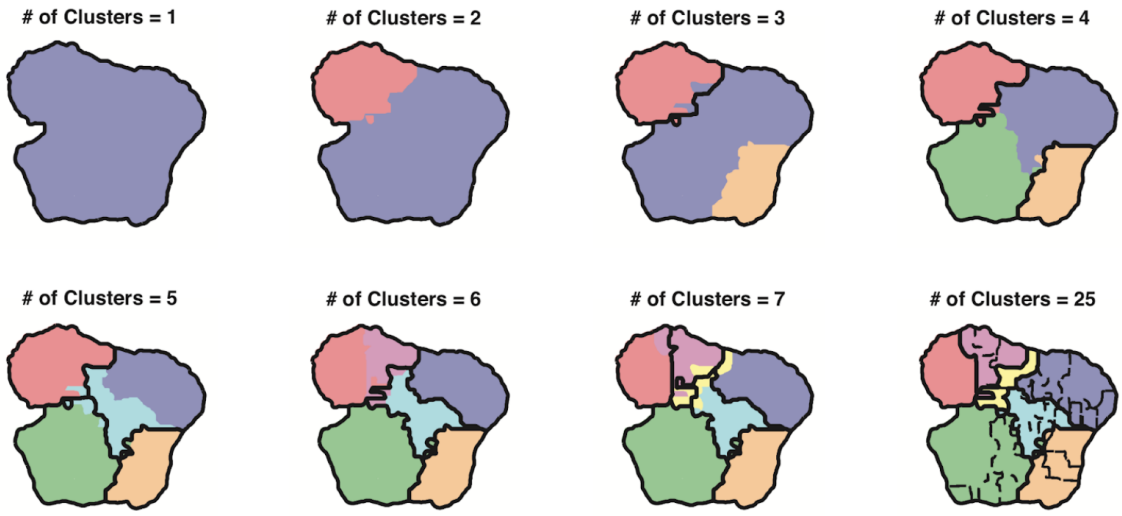


Figure 1.8: Information bottleneck clusters of behavioral space for $\tau = 67$. Previous borders are plotted in black (dashed lines for 25 clusters). Figure from Berman et al. (2016) [9]

that maximizes information τ transitions in the future, while maintaining information from the past. This turns into an optimization problem known as the information bottleneck:

$$\max \mathcal{F} = I(Z; S(n+\tau)) - \beta I(Z; S(n)),$$

where I gives the mutual information and β is a Lagrange multiplier[121].

This algorithm varies in both τ and β and results in many solutions. The optimal solutions for $\tau = 67$, which is approximately twice the Markovian timescale can be seen in Figure 1.8. This Markovian timescale refers to the maximal time scale that a Markov model is capable of predicting. Mathematically, the Markovian time scale is: $t_{Markov} = -1/\log|\lambda_2(1)|$, where $\lambda_2(1)$ referes to the second eigenvalue of the 1-step transition matrix. The clusters from the DIB algorithm are hierarchical, which is not a built in feature of the information bottleneck. The idea of hierarchy is hardly a new one, dating back to Tinbergen in the 1950s [119] as was mentioned earlier. Previous measurements of hierarchy in behavior typically limited the scope of their findings by focusing on single behavioral types like grooming [64, 65, 66, 27], or modeling the

hierarchy across a single time scale (often Markovian) [88, 53], or relying on methods that would only yield hierarchical results [45].

While Berman et al. (2016) [9] used the Information Bottleneck, in this thesis, we will be using the Deterministic Information Bottleneck (DIB) [116] instead. Ultimately, DIB returns similar clusters as IB but the clusters themselves are deterministic which is preferred for our purposes. Also, DIB is supposed to better capture the compression by using entropy of the clustered representation instead of mutual information:

$$\min \mathcal{F} = H(Z) - \beta I(Z; S(n+\tau)).$$

Here, we will be discussing behavior of fruit flies under two time scale paradigms. One is behavior measured in an hour, and the other is behavior measured in an hour but across the span of the flies' lifetimes.

1.4 Behavioral Effects of Aging

While there are many internal and external factors that impact and modulate behavior, one internal process that is known to have a significant impact across all organisms is aging. As an organism ages, its behavior will change. Sometimes the change is because a behavior no longer serves a purpose or has become morphologically impossible. For example, kangaroos at a certain age stop riding around in their mothers' pouch. Sometimes, an organism learns a new behavior, like birds learning to fly. While there are plenty of other reasons for behavioral changes throughout a lifetime, most commonly, a behavior is continually performed throughout the animal's life but changes as the animal's physical body changes. An easy example of this is the gait and posture of humans walking from infancy to midlife to old age all look very different but are considered to be the same behavior. The changes in gait observed from midlife to old age can be partially explained by senescence, the deterioration

one experiences with age, but there are likely many other factors involved as well. It is not surprising, then, that a big question in aging studies is if we can determine what system or systems are modulating the changes.

Aging effects are present across multiple scales, which can make answering that question difficult. At the molecular level, previous studies have identified the age-regulated genes and the corresponding genetic pathways' responses to aging [132]. Even certain single gene mutation can have huge impacts on longevity, suggesting that there may be only a few cellular processes that contribute to aging [54]. One such process is called programmed cell death, and aside from aging, it is also associated with aging-related diseases [123]. Development of diseases or physical deficits can also introduce perturbations affecting postures, overall longevity, and can especially affect the neural level with regards to vision, speech encoding, and dementia [110, 3, 126].

Other studies have shown how vision, sensation, strength, reaction time, and balance (all variables that change with age) are related to the optimal gait pattern for stable locomotion [81], how oxidative stress and caloric restrictions can extend lifespans [108, 76], how disrupting circadian rhythm accelerates aging [62] as well as cellular senescence, deregulated nutrient sensing, or a breakdown in mitochondrial regulations and proteostasis [68], and finally, how senescence [42] reduces an organism's ability to perform its full range of behaviors [60]. Many of these studies use fruit flies (or *Caenorhabditis elegans* worms) because they are easily contained, have relatively short life spans, can be manipulated genetically, and mutated neurally. They are a good model organism for humans as approximately 70% of genes in humans are homologous to genes in fruit flies (*Drosophila melanogaster* specifically), and most cellular mechanisms are the same between flies and humans via evolution [11, 39].

In this thesis, we will work with data from Choi (2017) [23] which recorded behavior from male and female fruit flies at different stages of life. From these recordings, we can measure the full repertoire of behaviors and quantify how those behaviors

change with age. We find that the male fruit flies experience a "mid-life crisis" where they become very active, performing a lot of locomotive behaviors, where before and after this period of time, they are mostly idle. This "mid-life crisis" corresponds to the peak mating age for fruit flies. On the other hand, the female fruit flies start life very active and gradually become lazier. After measuring many quantities that remained constant, we determine that the primary predictor of behavioral changes is the changing energy budget, or rather, the average amount of energy expended as a function of age, which is a common theory in the field.

1.5 Internal Recurrent Neural Network Dynamics

In this thesis, we will implement a basin hopping model with multiple attractors to describe the internal dynamics we observe in our recurrent neural network. The barriers between these basins will interact according to Arrhenius behavior.

1.5.1 Basin Hopping Model

Basin hopping models are a global optimization technique that transforms the potential energy plane. By randomly iterating through perturbations of the plane, it performs local optimization to accept or reject the perturbation based on a minimized energy function value [127]. These models have stochastic and deterministic dependencies. Stochasticity in a system can be a driving force behind hopping from one basin to another, and the deterministic force acts as a counter balance to resist that stochasticity. The strength of the stochastic forces present in the system directly impacts the persistence times, or the amount of time spent in a basin. Persistence times can vary dramatically across the basins, but an increase in stochasticity will overall decrease the persistence times. However, the shape of the basin can also have an impact. With a certain amount of stochastic force, a deeper, more narrow basin

will likely have a longer persistence time than a basin that is wide and shallow [30]. In fact, without any stochastic forces present, the dynamics would be Markovian. To introduce stochasticity to our system, we will use Arrhenius behavior, adapted from chemical kinetics, to describe our change in basin heights.

1.5.2 Arrhenius Behavior

Molecules each contain a certain amount of energy, collectively potential energy and kinetic energy. Properly aligned collisions between particles can cause chemical reactions when the kinetic energy of the collision is above a certain threshold. Otherwise, the particle will simply bounce off each other. This energetic threshold is called the activation energy. Obtaining the activation energy can be difficult to do with classic Arrhenius analysis. Mathematically, the activation energy is defined as:

$$E_a = -\frac{\partial \ln(k(T)}{\partial \beta},$$

and $\beta = 1/k_B T$. Here, E_a is the activation energy, $k(T)$ is the temperature dependent reaction rate, k_B is a gas constant, and T is temperature. Typically, E_a is calculated by measuring the slope from an Arrhenius plot of $\ln k(T)$ vs. 1/T. However, this approach requires calculating or measuring the reaction rate, $k(T)$ which is where the challenges arise. Temperature needs to be able to vary which may not be possible in some cases. For example, a change in temperature can lead to a phase change near a phase transition. This would change the empirical value of k as it is a phase dependent quantity. Additionally, some systems' structure are highly sensitive to temperature such as folded proteins, bilayer membranes, and self-assembled structures. Thus, the temperature range must be constrained enough to avoid these potential issues but broad enough that the changes in k are detectable either through experimental or simulation methods.

In thermodynamics, activation energy and reaction rates are fundamental characteristics of a chemical reaction. These characteristics are often referred to in the context of Arrhenius's Law, given by rearranging the above equation:

$$k(T) = A\text{exp}\left(-\frac{E_a}{k_B T}\right).$$

A is a frequency factor that is temperature independent [5, 82].

Although this discussion has so far been in the context of chemistry kinetics, the concept of activation energy is applicable to more than just chemical reaction rates. For example diffusion coefficients, reorientation times, viscosity, and dielectric relaxation times can all be described by their own, analogous Arrhenius equation. In these examples, there are smaller activation energies and therefore less dependency on temperature, but without a clear threshold, the interpretability of the activation energy also becomes less clear [99].

Arrhenius' law has been derived through principles from statistical physics as well [122, 34]. The activation energy, E_a, is generally considered to be the difference between the threshold energy for the reaction and the reactant's energy, which allows us to interpret it as a ratio of probabilities [85, 86]. By exploiting the exponential distribution, we can recover Maxwell-Boltzmann statistics for a single particle and restrict the maximal randomness in a system [84, 85]. We will be using probabilities to measure how the basins are interacting and want to quantify the probability of transitioning. To do that, we essentially need an expression for the probability of a particle's energy being at least equal to the threshold energy. We can get that expression using the probability density function (PDF). The PDF for this system has the form:

$$f(E_i) = \frac{1}{\langle E_i \rangle}\text{exp}\left[-\frac{E_i}{\langle E_i \rangle}\right],$$

where E_i is the energy for an individual particle, and $\langle E_i \rangle$ is the overall average energy. By integrating over this PDF, we can calculate the probability that a particle's energy is greater than or equal to the threshold energy, E_t.

$$P(E_i \geq E_t) = \int_{E_i=E_t}^{\infty} f(E_i), dE_i = \exp\left[-\frac{E_i}{\langle E_i \rangle}\right]$$

Then, by taking a difference between probabilities, $P(E_i \geq E_t) - P(E_i \geq E_t + 1)$, we can get an expression for $P(E_i = E_t)$:

$$P(E_i = E_t) = \frac{\exp\left[\frac{-E_t}{\langle E \rangle}\right]}{\sum_j \exp\left[\frac{-E_j}{\langle E \rangle}\right]}$$

This expression replicates Boltzmann's partition function and reaffirms that the exponential probability is fundamental to statistical mechanics. The same expression can be derived by integrating over the PDF from E_i to $E_i + 1$ [87].

In this thesis, we will be using an adapted version of this statistical mechanics approach to Arrhenius behavior. We find that the internal dynamics of our recurrent neural network are analogous to a particle in a multi-well landscape with barriers between the wells changing according to Arrhenius's law.

1.5.3 Application to Our Network

By treating the internal states of the network as a nonlinear dynamical system following the findings from Sussilo et al. (2013) [117], we can measure fixed points, both stable and unstable. These fixed points then act as our internal states, and their interactions modulate the external behavioral states. From the interactions, a basin-like structure emerges, where a set a of states act as the bottom of a basin for other states to flow to. Thus, if we treat this dynamical system of internal states as a particle in a multi-well landscape, we can see how transitions are modulated through changes in the barriers between basins. This concept is illustrated in Figure 1.9. Changes in the

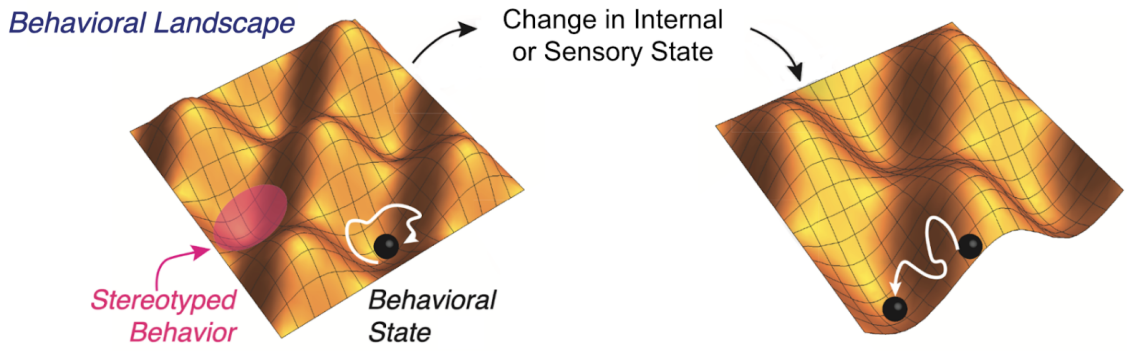


Figure 1.9: Behavioral transitions as a multi-basin landscape allowing transitions to occur between different states by shifting barriers between basins.

barrier heights could result from a number of things including changes in an internal, sensory state or external stimuli.

By assuming the barrier heights interact under Arrhenius behavior, we can derive how those barriers change with time. Although Arrhenius behavior is classically applied to reaction rates in chemistry, we have discussed how it has applications in many other contexts including thermodynamics [85] and microbial "social" decision rates [24]. Thus, we use a statistical mechanics approach to describe the Arrhenius behavior in our system. In this context, Arrhenius's Law gives us an expression for persistence time rather than reaction rate:

$$\langle \tau \rangle = A\exp[-\beta \Delta E]$$

Here, the reaction rate, $k(T)$, has been replaced with persistence time, τ. Activation energy, E_a, is analogous to our basin depth, ΔE. We can also write an expression to quantify how transitions between basins occur as a function of time as the barrier heights change:

$$T_{ij}(t) = \frac{\exp[-\beta \Delta E_{ij}(t)]}{\sum_k \exp[-\beta \Delta E_{ik}(t)]} \equiv \frac{\exp[-\beta \Delta E_{ij}(t)]}{Z_i(t)}.$$

Here, T_{ij} is the transition probability from basin i to basin j, β is a temperature, and ΔE_{ij} is the barrier height between the two basins i and j.

Understanding how the internal mechanisms of an RNN are able to produce behavioral dynamics provides a framework of how the internal mechanisms of an animal produce the same dynamics.

1.6 Thesis Outline

This thesis consists of four chapters. **Chapter 2** introduces a novel study on aging in *Drosophila melanogaster* fruit flies. The data analyzed in this study measures the full repertoire of behaviors across the flies' lifespan in both male and females, ages ranging from 0 to 10 weeks. We find a sexual dimorphism in how the repertoire changes with age and further that while repertoire of behaviors being performed changes with age, the postures required to perform the behavior remain the same without experiencing degradation with age. In fact, many quantities of our data remains constant across age, but we find that the observed changes in behavior can be predicted by the fly's overall energy budget. In **Chapter 3**, we train a recurrent neural network to predict and replicate the behavioral sequences from the Berman et al. (2014) [8] data set. This data is a subset of the data used in **Chapter 2** having only the male flies in the 0-2 week old age group. We find that the network successfully reproduces multiple long time scales as well as some other features of our data. However, we would like to have a model that is more easily interpreted. Thus, we built a model that describes the internal workings of the network. First, we treated the internal dynamics of the network as a dynamical system and observed how the fixed points of the system interact with each other. This leads us to treating the fixed point interactions similar to a set of basins with varying barriers between the basins. These changing barriers provides a framework to discuss how the network is modulating behavioral sequences

and making transitions. Finally, **Chapter 4** is a conclusion to summarize the work laid out in this thesis and to propose future works for the field.

Chapter 2

Measuring the repertoire of age-related behavioral changes in *Drosophila melanogaster*[1]

Abstract

Aging affects almost all aspects of an organism – its morphology, its physiology, its behavior. Isolating which biological mechanisms are regulating these changes, however, has proven difficult, potentially due to our inability to characterize the full repertoire of an animal's behavior across the lifespan. Using data from fruit flies (*D. melanogaster*) we measure the full repertoire of behaviors as a function of age. We observe a sexually dimorphic pattern of changes in the behavioral repertoire during aging. Although the stereotypy of the behaviors and the complexity of the repertoire overall remains relatively unchanged, we find evidence that the observed alterations in behavior can be explained by changing the fly's overall energy budget, suggesting

[1]This text is adapted from Overman, Katherine E., et al. "Measuring the repertoire of age-related behavioral changes in Drosophila melanogaster." PLoS computational biology 18.2 (2022): e1009867.

potential connections between metabolism, aging, and behavior.

2.1 Introduction

Aging is a biological process that affects nearly all organisms, resulting in profound changes to their morphology, physiology, and behavior [4, 102, 106]. While there exists variability in the precise form and timing of these alterations, stereotyped patterns of aging-related change are commonly observed at scales ranging from molecules to tissues to the entire organism [13]. However, we lack a comprehensive framework for predicting how the multifarious age-related changes at the molecular and neuronal level directly lead to behavioral changes.

While many age-related changes in behavior are due to direct reductions in an animal's capacity for movement (e.g., arthritis in humans or wing damage in flies), another commonly posited hypothesis is that aging effects in behavior can be partially understood as an alteration in an animal's energy budget [60, 73]. In other words, while the organism may still be able to physically perform most activities within its repertoire, its reduced metabolic efficiency might impose constraints on an animal's total amount of energy to expend, leading to age-related changes in its behavioral repertoire. This idea, that the available energy an animal possesses would have systemic effects on its chosen actions, is reminiscent of the "hydraulic" theory of action selection that was popularized by Lorenz and others [69] and might be related to molecular models of metabolic decline such as insulin pathway modifications [56, 1, 90].

Testing the hypothesis that age-related alterations can be understood through alterations in energy budgets, however, has proven difficult, partially due to the limitations in our ability to accurately measure full repertoires of behavior across time. Aging is a complex, dynamical process that cannot be measured at a single time-point,

but, rather, it must be characterized as a trajectory across a lifetime. Accordingly, to measure how animals' behavioral repertoires and their usage alter with age, we need to have not only a framework to measure repertoires at the timescale of single stereotyped movement (order of tens of milliseconds to seconds), but also new analysis methods to isolate the between-age-group variability from the within-age-group variability in these behaviors, finding combinations of behaviors that best describe the dynamics of aging.

In this paper, we study the age-related changes in the behavioral dynamics of the fruit fly *Drosophila melanogaster*, a common model system for the study of aging and behavior [15, 16, 36, 100]. We measure the full repertoire of behaviors that flies of varying ages perform. While previous research on aging and behavior in flies focus on how only a small number of behaviors change with age, here, by quantifying the full repertoire of behaviors that the animals exhibit in our experimental conditions, we can observe how behavioral performance, in terms of both usage frequency and context-dependent usage (e.g., transition probabilities), changes with age. To measure the animals' behavior, we use an unsupervised method that identifies the stereotyped behaviors that the fly performs without a priori behavioral definitions - behavioral mapping [8]. Our results show that (1) large changes and a sexual dimorphism in how the behavioral repertoire changes with age; (2) despite these changes, the overall complexity of the flies' behavior remains unchanged; (3) as the fruit flies age, their behavioral repertoires alter, but the behaviors are still performed with similar stereotypy; (4) we can explain most of the inter-age-group behavioral variability that we observe by using an estimation of average power consumption. Thus, we provide evidence that the energy budget that an animal has available may be a key factor in regulating its behavior with age. This result encourages further investigation into the physiological basis of aging, lending credence to hypotheses that link metabolic decline to age-related behavioral changes in animals.

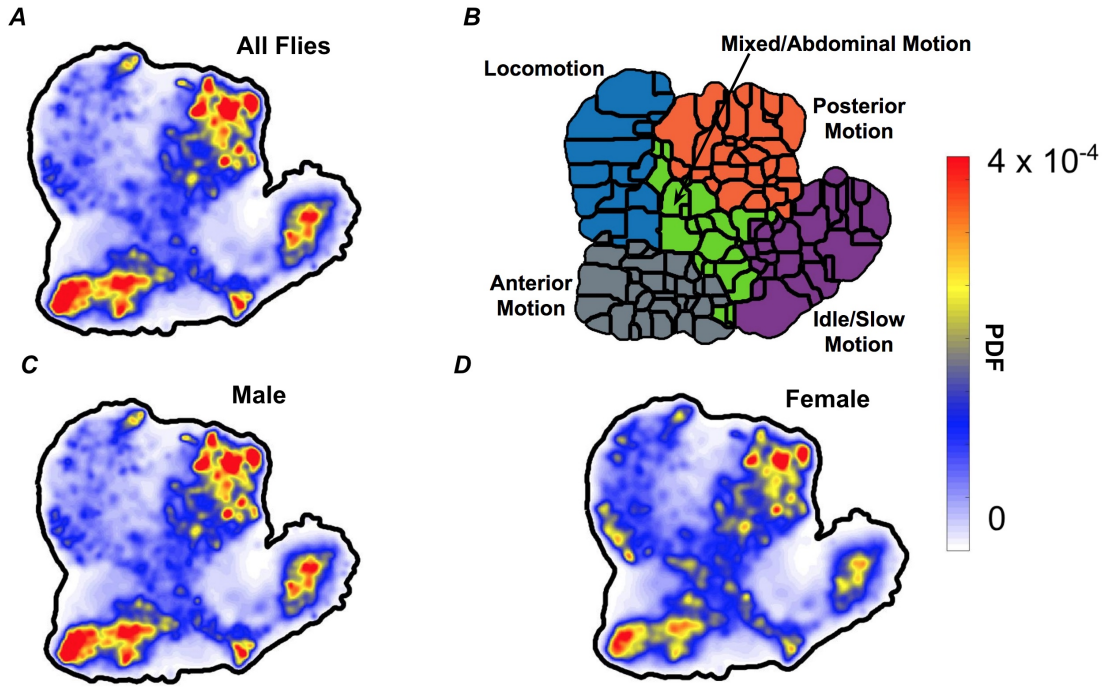


Figure 2.1: Behavioral densities for quantifying the full repertoire of behaviors of male and female fruit flies (*D. melanogaster*) with ages ranging from 0-70 days. (**A**) A behavioral density averaged across all flies in this study (both males and females). The color scale corresponds to the probability density function, where red peaks correspond to individual stereotyped behaviors. (**B**) Applying a watershed transform on the PDF from (**A**) produces boundary lines for the different behavioral states. Similar types of movement (described via manual annotation of the videos) are clustered together on the behavioral space, and broad descriptions of the type of movements in each cluster are obtained from the original videos. By taking the embedded points and sorting them by sex, we can make behavioral densities for the males and females separately. (**C**) The behavioral density averaged over all the male flies from all age groups. (**D**) Same as in (C), but averaged over all female flies.

2.2 Results

2.2.1 Experiments and behavioral densities

In order to characterize how flies' behavioral repertoires changes with age, we imaged flies (*Drosophila melanogaster*) in a largely featureless environment (see Materials and Methods for details). In total, we imaged 304 flies (155 male and 159 female), each imaged once with and age between 0 and 70 days old (the average lifespan is

60-80 days [37]). The flies were placed in the arena via aspiration and given 5 minutes to acclimate to the environment. To measure the flies' behavioral repertoires, we use the behavioral mapping approach originally described in Berman (2014) [8]. In brief, this method uses image compression techniques to measure a time series of the fly's postural dynamics, computes a continuous wavelet transform to isolate the dynamical properties of these time series (i.e., finding which parts of the body are moving at what speeds), and uses t-Distributed Stochastic Neighbor Embedding (t-SNE) to perform dimensionality reduction on the amplitudes of this transform, creating a 2-dimensional probability density function over the space of postural dynamics. We refer to the arrangement of peaks within this probability density function as our behavioral space.

Each peak within this density represents a distinct stereotyped behavior (e.g., grooming, running, idle, etc.). Thus, the relative probabilities of observing a fly within each peak in the density is a measure of the animal's behavioral repertoire, seen in Figure 2.1A. Following the procedure described in Cande (2018)[20], all flies - including all males and all females of all ages - were embedded into the same space in order to facilitate comparisons between individuals, sexes, and ages. We isolate the individual peaks by applying a watershed transform [83] to segment the density into 122 discrete states, with near-by regions corresponding to similar behaviors (Figure 2.1B). The density for all the males can be seen in Figure 2.1C, and the density for all the females in Figure 2.1D. These behavioral densities provide the foundation for our analysis, as we use them to quantify how behavioral repertoires change with age.

2.2.2 Quantifying behavioral changes with age

Dividing the males and females each into two-week-interval age groups (Figure 2.2), we observe a sexual dimorphism in how their behaviors change with age. Specifically, the younger male flies mostly perform idle behaviors. In mid-life, they perform more

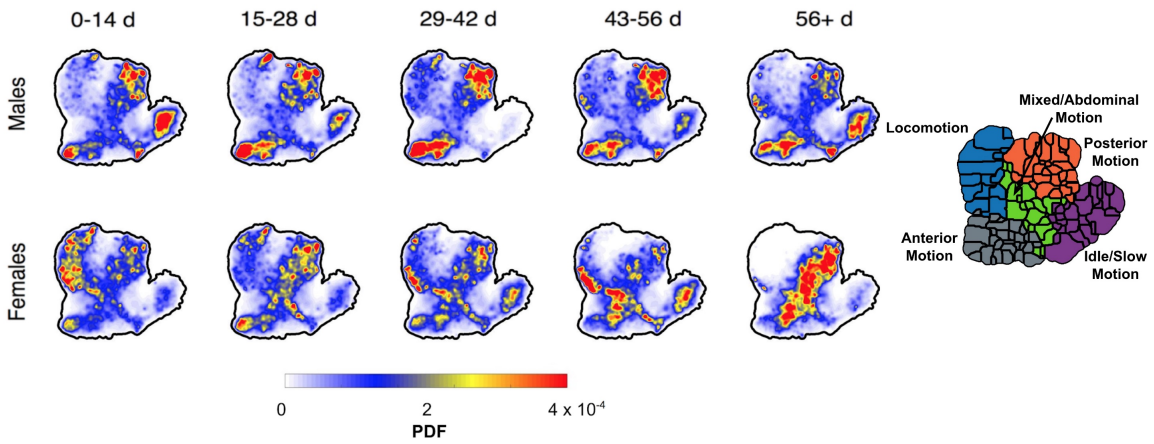


Figure 2.2: Behavioral densities as a function of age. Behavioral densities for male and female flies are shown on the left, with ages broken down into 2-week intervals. We construct these densities by separating the embedded points into subgroups of male, female, and their age. Then, we take each set of points and apply a probability density function. In this figure, we can see a broad description for behavior as a function of age emerge. Male flies mostly perform idle or slow throughout their life with the exception of mid-life, when they do more active behaviors. In contrast, females are very active when young and become more idle as they age. An annotated behavioral space from Figure 2.1 is displayed on the right.

active behaviors before again becoming lethargic in later life. Conversely, the females perform active behaviors when young, and gradually begin to perform more idle behaviors as they increase in age (excepting the last age group, which is likely under-sampled). While these results could have been found with center-of-mass tracking or other less computationally intensive methods than behavioral mapping, that our method replicates previously observed experimental results [16], provides additional confidence in the analyses to follow.

While the data plotted in Figure 2.2 displays how flies' mean behavioral profile alters with age, there also exists variance and co-variance within sex and age groups [46, 50, 6] (although no notable structure based on the precise time of imaging – see Figs. 2.9 and 2.10). Thus, we need to isolate the variance in our data that is associated with changing age, rather than from inter-age-group variability. To quantify the inter-group behavioral variance structure, we measured the behavioral

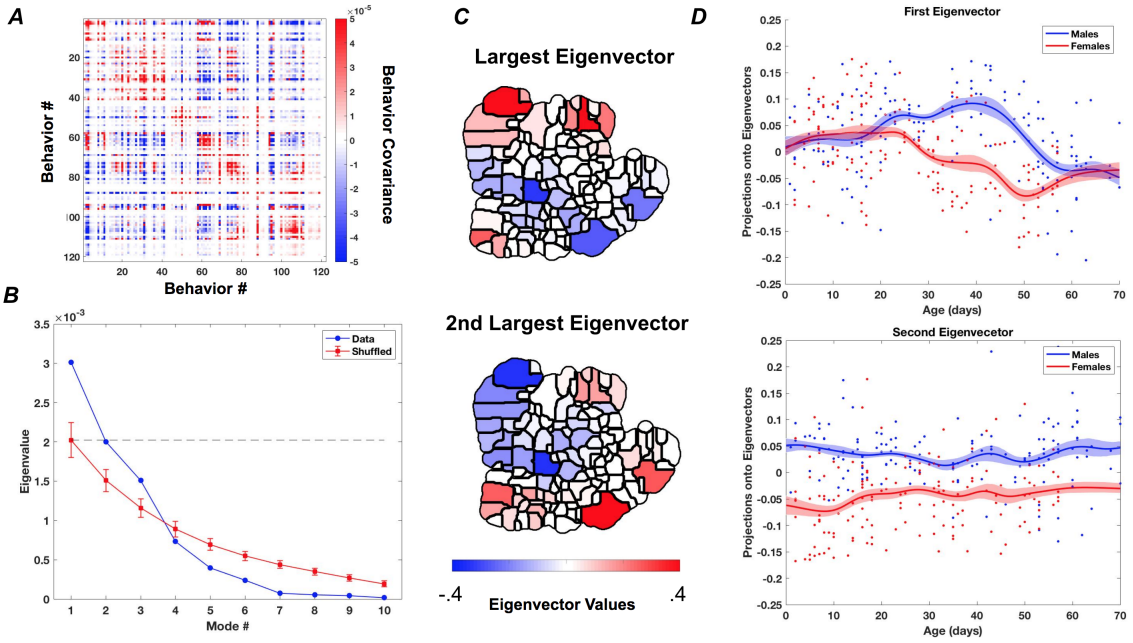


Figure 2.3: Identifying aging-specific behavioral covariances. (**A**) The covariance matrix of the mean behaviors sorted according to the clusters in in Figure 2.1B. (**B**) The eigenvalues of the covariance matrix. There are two eigenvalues (blue) that are larger or approximately equal to the eigenvalues returned from shuffling the behavioral density matrix (red, error bars are the standard deviations from many independent shuffles of the data). These two modes account for approximately 62% of the variation in the data. (**C**) The eigenvectors corresponding to the largest (top) and second-largest (bottom) eigenvalues. (**D**) Projections of the data onto the largest (top) and second-largest (bottom) eigenvectors in (C), plotted as a function of age. Here, dots are values for individual animals, and the solid lines are from smoothing the data with a Gaussian of $\sigma = 3.5$. Error bars are the standard deviations of this process as a function of age after re-calculating the curve with re-sampled data (drawn with replacement from the original data).

covariance matrix across all sex/ages, providing a quantification of the behaviors that are shifting together with age.

Our analyses here use the discretized version of the behavioral densities, using the watershed-transform-derived regions shown in Figure 2.1B. $\boldsymbol{P}^{(i)}$ is a vector of probabilities, where, $P_j^{(i)}$ is the the time-averaged probability that fly i performs behavior j during the one hour filming epoch – we call this vector our behavioral vector. Given these values, we can then calculate the average behavioral density for all individuals within each sex/age group. We define this group-specific mean behavioral vector to be $\boldsymbol{\mu}_k^{(z)}$, where $z \in \{\text{male}, \text{female}\}$ and k is the age group. From these means, we can then compute the covariance matrix of the set of mean behavioral vectors, $M \equiv \left[\boldsymbol{\mu}_1^{(\text{male})} \cdots \boldsymbol{\mu}_5^{(\text{male})} \boldsymbol{\mu}_1^{(\text{female})} \cdots \boldsymbol{\mu}_5^{(\text{female})}\right] \in \mathcal{R}^{122\times 10}$ (5 different 2-week groups for each sex).

This covariance matrix ($C^{(M)} \equiv \text{Cov}(M)$), shown in Figure 2.3A, quantifies which behaviors are likely to increase or decrease with respect to each other across sex/age groups. To further quantify the structure within $C^{(M)}$, we calculate its eigenvectors and eigenvalues (Figure 2.3B-C), effectively performing Principal Components Analysis on the mean vectors. Because the covariance matrix is, by definition, real-valued and symmetric, all of its eigenvalues must be greater or equal to zero. We focus here on only the modes corresponding to the two largest eigenvalues, as only these two modes have eigenvalues that are significantly larger or similar in value to those from a covariance matrix derived from independently shuffling each of the columns in M. Although there is not a clear interpretation of these two eigenvectors ($\hat{v}_1$ and $\hat{v}_2$), both appear to capture the relative performance of idle and locomotory behaviors, and the first also appears to capture the relative usage of slow vs. fast locomotion. By plotting the projection of each fly's behavioral vector as a function of age and finding Gaussian-smoothed average curves (see Materials and Methods), we see how this low-dimensional space of behaviors alters as the flies age (Figure 2.3D). There

is a clear sexual dimorphism in the projections onto the first eigenmode, with the male flies exhibiting non-monotonic dynamics with age, whereas the female's average curve is largely monotonically decreasing. A similar dynamic can be observed in the second eigenmode but with a more subtle shift, as well as a sign flip. These results agree with the visual intuition from Figure 2.2 and provide a quantification of the most important changes in the flies' behavioral repertoire with age.

2.2.3 Estimated Energy Consumption Alters with Age

As stated in the introduction, a potential mechanism for the flies' observed changes in behavior could be an overall reduction in the flies' energy budget with age. While it was not possible to directly measure the power consumption from the animals in our experiments, we can instead estimate the metabolic cost of the observed behaviors with a biomechanical model.

Given that the flies are constrained to move within a two-dimensional environment, we focus our modeling efforts on estimating the cost of legged locomotion within the arena (making the assumption that non-locomotion behaviors like grooming are negligible in energetic cost compared to locomotion, see Materials and Methods for further justification). Our model of the power consumption during locomotion largely follows that of *Nishi* (2006) [92], which estimates the heat dissipation and work done during each swing and stance phase of locomotion at a given velocity using a biomechanical model of force production during legged locomotion (see Materials and Methods for details). While this model has several free parameters related to the fly morphology and how gait dynamics alter with speed, we use morphological and scaling data from the literature on legged locomotion [80, 51] to set these parameters. More precisely, we wish to calculate $R(v)$, the specific power (mechanical power per unit mass) required for the fly to move at a speed v.

From tracking the center-of-mass of each fly, we are able to measure $p_i(v)$, the

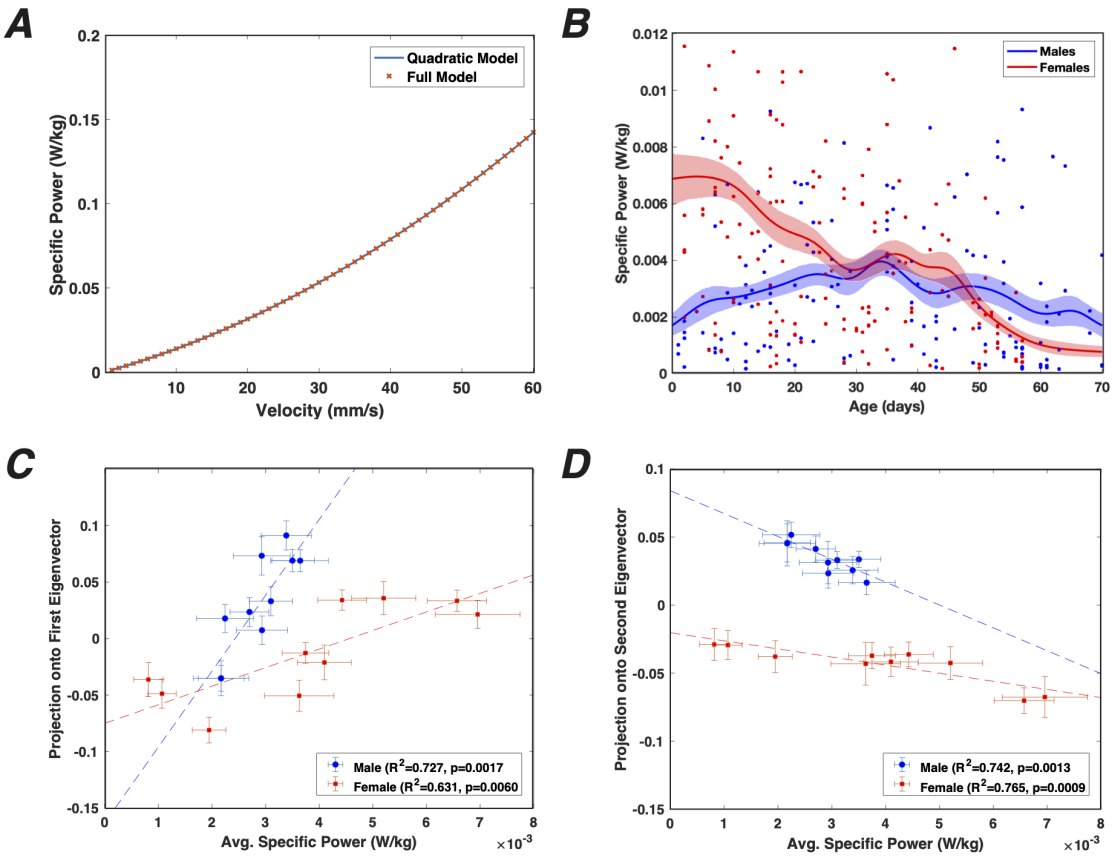


Figure 2.4: Energy usage predicts aging-specific changes in behavior. (**A**) Comparison of the quadratic model to the full model of Nishii (2006) to estimate specific power (power per unit mass) of legged locomotion in fruit flies. (**B**) Specific power as a function of age for male (blue) and female (orange) fruit flies. Each point represents an individual, and the curves are the Gaussian-smoothed means ($\sigma = 3.5$ days), with error bars generated in the same manner as Fig. 2.3D. (**C**) Average projections onto the first eigenmode (Fig. 2.3C (top)) plotted versus the average specific power consumption for both male and female flies. Each point represents the value (plus error bars) from the curves in (B) and Fig. 2.3D (top), each spaced 7 days apart. Dashed lines are the linear fits to the data. (**D**) Same as (C), but instead using projections onto the second eigenmode (Fig. 2.3C (bottom)). Note that at over 70% of the mean aging-specific variation can be explained using the first two eigenmodes.

probability density function for speed for fly i, for each animal. Given this distribution and our expression for $R(v)$, we can calculate the average specific power consumption, $\bar{R}_i$ for each animal through numerically integrating

$$\bar{R}_i = \int_0^{v_{max}} p_i(v) R(v) dv, \tag{2.1}$$

where v_{max} is the largest observed speed for the flies. To make this calculation more tractable, we find that for biologically realistic range of locomotion speeds (0-60 mm/s), $R(v)$ is well-approximated by a quadratic function ($R(v) = av^2 + bv + c$, where $a = 19.9s^{-1}$, $b = 1.17m/s^2$, and $c = .0002m^2/s^3$), as shown in Figure 2.4A.

The results of this calculation for each individual animal are shown in Figure 2.4B as a function of age. While there is significant scatter in the data (likely due to variance in the internal activity state of the flies [9, 46]), when we compute a smoothed average of the data, a clearer portrait emerges. Specifically, we observe that these curves are reminiscent of the sexual dimorphism we observed in the inter-group eigenvector projections in Figure 2.3. More quantitatively, we see that when plotting the eigenvector projections versus the group-average specific power (Figure 2.4C), we see a high degree of correlation for each of these values. As seen in the figure, we can explain at least 72% of the aging-specific behavioral variation using a linear fit to the estimated specific power consumption. Thus, these analyses imply that most of the age-related changes we observe in the animal's behavior are correlated with changes in the average energy expenditures of the flies.

2.2.4 Complexity of the Behavioral Repertoire

Although we show that most age-related changes in fly behavior are correlated with energy consumption, it still may be possible that other factors such as the complexity of the behavioral repertoire or the degradation of stereotyped behaviors might also

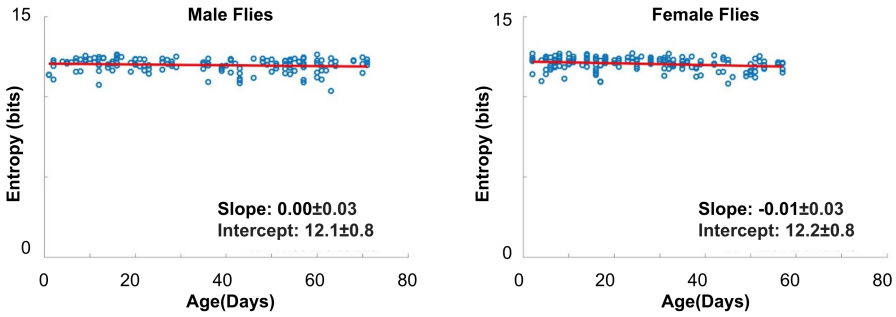


Figure 2.5: Entropy of the behavioral densities as a function of age for the males (left) and females (right) with a best fit line to estimate the value included on the plot by taking the mean value. Error bars for individual animals are smaller than the symbol size in the plot. The males have a slope of -0.00 ± 0.03 and the females have a slope of -0.01 ± 0.03.

be observed as the animals age [38, 44]. We test the former of these hypotheses by calculating the entropy of the behavioral space, using this metric as a proxy for the overall repertoire complexity.

Specifically, we measure the entropy, H_i, of each individual fly's behavioral density according to

$$H_i = -\iint \rho(x,y)\log_2\rho(x,y)dxdy, \tag{2.2}$$

where $\rho(x,y)$ is the probability distribution over the two-dimensional behavioral space. Plotting H_i as a function of the flies' ages (Figure 2.5), we see no discernible trend in entropy vs. age, with the best fit slopes showing a value of -0.00 ± 0.03 for the male flies and -0.01 ± 0.03 for the female flies. (We also took a measure of the entropy using the probability distribution over the 122 discretized behaviors and obtained comparable results – Figure 2.12.) Thus, even though the behavioral densities are dramatically changing with age, the overall complexity remains largely unaltered, and thus we cannot conclude that the complexity of the repertoire degrades with age.

2.2.5 Long Time Scales and Hierarchical Structure in Behavior with Age

While the complexity of the behavioral repertoire remains unchanged, the complexity of how the animals traverse through this space over time might still show significant deviations. Prior investigations into the complexity of fly behavioral sequences have shown that these dynamics of transitions between stereotyped behaviors exhibit long time scales and hierarchical organization [9, 46]. A hypothesis for aging-related behavioral change is that the structure of the behavioral repertoire becomes less complex with age [38, 44], and with the detailed measurements of behavior described here, we can test this idea, potentially gaining insight into changes occurring to the internal programs that may generate these patterns.

First, to assess the overall timescale structure of the flies' behavioral patterns, we measure the transition matrix at different time scales,

$$[\mathbf{T}(\tau)]_{i,j} = p(S(n+\tau) = i | S(n) = j), \tag{2.3}$$

where i and j as two stereotyped behaviors, $S(n)$ is the behavioral state of a system at transition n (note: to decouple waiting time in a state from complexity in the order of pattern of transitions between states, we measure time in units of transitions, following the methods in [9]). We can decompose each of these matrices via

$$[\mathbf{T}(\tau)]_{i,j} = \sum_{\mu} \lambda_{\mu}(\tau) u_i^{\mu}(\tau) v_j^{\mu}(\tau), \tag{2.4}$$

where u_i^{μ} and v_j^{μ} are the i^{th} right and j^{th} left eigenvectors of the matrix, respectively, and λ_{μ} is the eigenvalue with the μ^{th} largest modulus. Because the columns of each of these matrices must sum to one, $\lambda_1(\tau) = 1$ for all values of τ, and $|\lambda_{\mu>1}(\tau)| < 1$ by the Perron-Frobenius Theorem. While for a Markov Model, the eigenvalues should

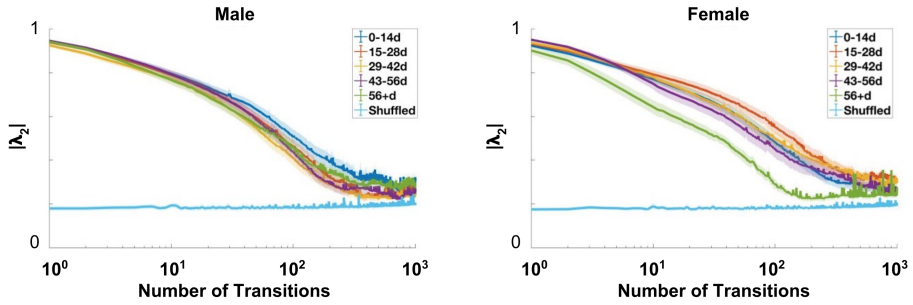


Figure 2.6: Absolute value of the second eigenvalue of the transition matrices as a function of transitions into the future, averaged over all flies in each age group with error bars corresponding to standard error of the mean for the male flies (left) and the female flies (right). The light blue line, which acts as a noise floor, is the second eigenvalue in a transition matrix calculated after shuffling our finite data set.

decay exponentially with τ, we find that flies in all sex and age groups exhibit super-Markovian time scales (Figure 2.6 shows the results for the second-largest eigenvalues in each transition matrix.The 3rd-5th eigenvalues can be seen in Figure 2.13). With the exception of the > 56 day-old females (for which we had fewer individuals in our sample), however, we found no differences larger than the standard error of the mean between the time scales across age groups.

While the complexity of the repertoire or the overall timescale might not be changing with age, the underlying structure of the behavioral transitions might still be altering. To test for this possibility, we applied a predictive clustering analysis to the space to identify groupings of behaviors that best preserve information about the long timescale structure in our data. More precisely, we would like to find a partition of our behavioral space, Z, such that this representation has a simple of a representation as possible, while still maintaining information about the future behavioral states of the animal. Here, we achieve this using the Deterministic Information Bottleneck (DIB) approach [10, 116], which minimizes the functional

$$\mathcal{F}_{DIB} = H(Z(n)) - \beta I(Z(n); S(n+\tau)), \tag{2.5}$$

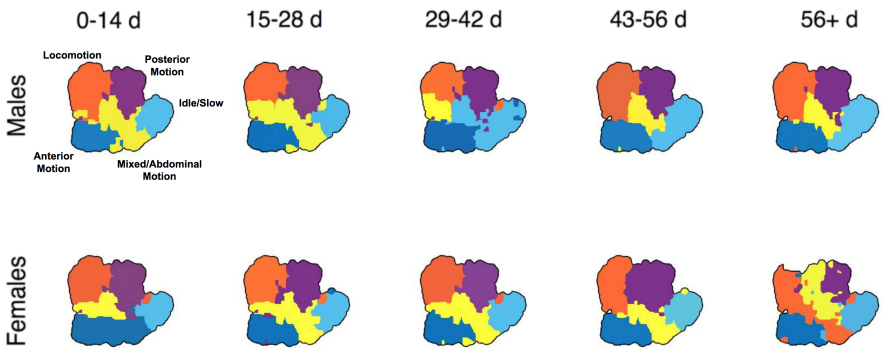


Figure 2.7: Hierarchical partitioning solutions from deterministic information bottleneck for the behavioral density with $\tau = 100$ and 5 clusters for male flies (top) and female flies (bottom) as a function of age.

where Z is our partition, $H(Z(n))$ is the entropy of the partition, β is a Lagrange multiplier that modulates the relative importance of simplicity and predictability, and $I(Z(n); S(n+\tau))$ is the mutual information between the partition and the future behavioral state at a time τ in the future. We perform this optimization for several values of τ for each age group, in all cases varying β and the number of initial clusters in Z to create a full curve of values (see Materials and Methods for details and Figure 2.14).

The resulting clusterings for τ=100 with five clusters can be seen in Figure 2.7. As with the eigenvalues in the previous plot, the clusters obtained via this approach remain nearly constant with varying age, with only small-probability behaviors flipping between regions. Thus, we lack evidence of significant alterations of the temporal complexity of the flies' behavior with age.

2.2.6 Stereotypy

Lastly, while we observe no significant changes to the flies' repertoire or temporal complexity, we still can measure if there is deterioration in how the behaviors are performed, potentially implying that the flies are undergoing a physical deterioration or some other inability to consistently perform behaviors while aging. To assess changes

in how stereotyped behaviors are performed, we measure how much the performance of individual behaviors are altered with age, quantifying a decreased stereotypy with an increase in the variance of the postural trajectories underlying the performance of these actions.

We divide the data into age groups of two week intervals, with a one week overlap (0-14 days, 8-21 days, 15-28 days, etc.), finding the postural trajectories associated with the performance of each behavior. While the details of this can be found in Materials and Methods, broadly, we use a phase-reconstruction method (based on Revzen (2008)[101]) across all of the postural modes for each time a behavior is performed. We measure the mean postural dynamics across all individuals in a given sex/age group and assess the stereotypy of each behavior (b) in each age group (κ) with our Stereotypy Index, $\chi_{b,\kappa}$, which is the fraction variance explained by the mean trajectory for that behavior. Thus $\chi_{b,\kappa} \to 1$ implies that each time the behavior is performed, its postural trajectories are exactly the same (maximally stereotyped), and $\chi_{b,\kappa} \to 0$ implies that the postural trajectories are different each time the behavior is performed (minimally stereotyped).

The values of $\chi_{b,\kappa}$ for each behavior and three different age groups are displayed in Figure 2.8A. By eye, we can see only minimal changes across the age groups (and no outside errorbar changes after accounting for multiple comparisons using Bonferroni corrections). Note that a few behaviors, while stereotyped, were not performed enough to get a good estimate of their synchronization parameters so those behaviors are listed as having a synchronization parameter of 0 (see Materials and Methods for more details).

To quantify this lack of change across the whole behavioral repertoire, we calculated the average stereotypy for each age group,

$$\bar{\chi_\kappa} = \frac{1}{N_\kappa} \sum_b \chi_{b,\kappa} \sum_{i \in G_\kappa} P_b^{(i)}, \tag{2.6}$$

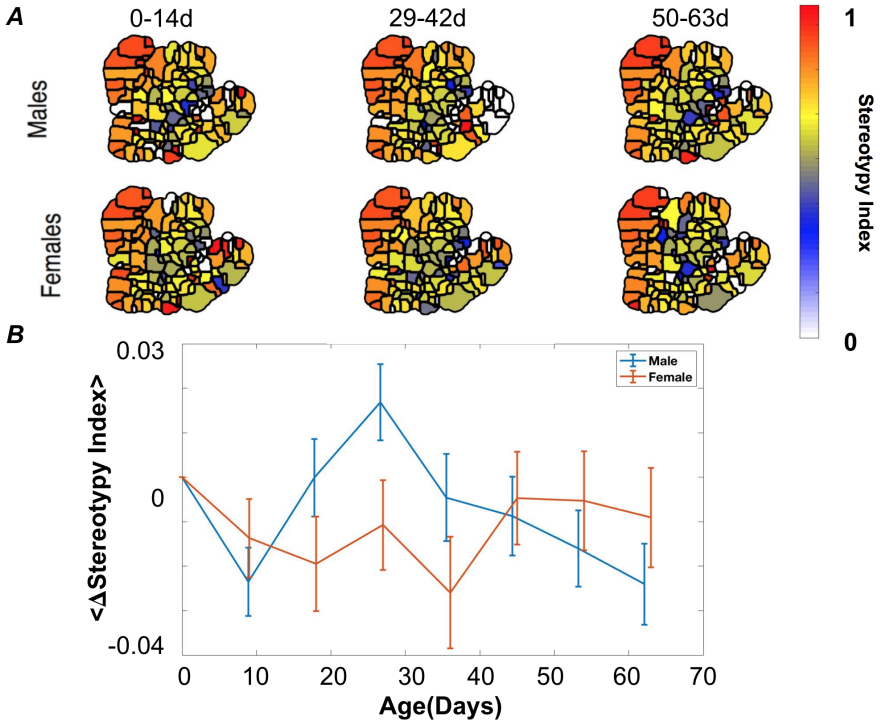


Figure 2.8: Few measurable changes in stereotypy with age. (**A**)The stereotypy of each behavior (or how stereotyped each behavior is - 1 being very stereotyped, 0 being not at all stereotyped) plotted as a function of age by calculating the maximum synchronization parameters of each behavior for the males (top) and females (bottom). (**B**) Quantification of how the synchronization parameters change between each age group and the initial age group by taking the mean difference between the synchronization parameters of each behavior. Error bars are bootstrap estimates from re-sampling individuals with replacement.

where G_κ is the set of all flies in age group κ, N_κ is the number of flies in the group, and $P_b^{(i)}$ is the fraction of time that fly i performs behavior b. We then measured the difference in the average stereotypy the youngest age group and each of the subsequent age groups for each sex ($\bar{\chi_\kappa} - \bar{\chi_0}$). Figure 2.8B shows the results of this calculation for both the males and females. Although we do observe some changes between the age group, they are within 1.5 standard deviations. Thus, although the probability of choosing a behavior changes with age, each behavior, when performed, is, on average, no less stereotyped.

2.3 Discussion

In this paper, we measured the behavior of fruit flies (*D. melanogaster*) at many points along their lifespan, aiming to isolate patterns of behavioral change with age and to make predictions about the physiological basis of these changes. Consistent with previous studies, we found a sexual dimorphism in changes in the animals' overall activity level, but we also identified subtler patterns of change with age by measuring the largest eigenvalues, and their corresponding eigenvectors, of the inter-age-group covariance matrix. Despite observing no significant changes in the repertoire complexity or stereotypy with age, we find that most of the age-specific behavioral alterations can be explained by a model of energy consumption, implying that energy budget may play an overarching role in regulating aging behavior.

This observation that energy may play a key role in aging-specific changes in behavior is in accordance with results from long-lived mutants in a variety of species, many of which have changes in gene regulation pathways that affect energy availability [57]. For example, mutations in the insulin/IGF-1 receptors or homologs, which promote food storage and cell replication, have been shown to extend lifespan in flies [118, 17], nematodes [58, 90], and rodents [49]. In addition, another long-lived

fly mutant, the E(z) histone methyltransferase heterozygous mutation, is associated with large alterations in a variety of metabolic regulation pathways [89]. In addition, these changes were found to exhibit sex-dependent effects, similar to our results as well. However, these studies do not examine how longevity affects full repertoires of behavior. There is a known inverse correlation between frequency of high energetic cost behaviors and longevity [21], so we would expect the long-lived mutants to use fast locomotion less and idle more than an unaltered fly. We further hypothesize that caloric restricted animals too would exhibit fewer of the high energetic cost behaviors.

In future efforts where behavioral repertoire and metabolic state could be simultaneously assayed (through, for example, proteomic or transcriptomic measurements), we would expect to find correlations between position along the the curves seen in Fig. 2.3D and key metabolic regulators. Through this methodology, it may be also possible to provide an effective age for each individual in a heterogeneously aging population, providing a phenotyping tool for identifying new molecules involved in increased and decreased longevity, as well as for the study of evolutionary aging dynamics. A possible follow-up study could use measurements of metabolic state and behavioral repertoire with groups of flies that are optigenetically altered to express primarily higher energy costing behaviors versus lower energy costing behaviors. This would allow some probing at the question of whether performing more lower energy behaviors leads to higher longevity or if the idleness in longer living flies is a result of a lower energy budget overall.

While the analysis framework detailed in this paper should be generalizable to other data sets, including other species [74, 61] and neuroimaging data [12], the data used in this study present several limitations that need to be studied in future work. First, despite the wide range of behaviors we observed in our assay, many natural behaviors, including courtship and flying, were not measured here. Flight in particular is known to be more common in young flies [21] and likely a large

source of oxidative stress and potential injury for the animals, likely creating more opportunities for decreased stereotypy and the degradation of behavioral performance. Additionally, due to technical constraints in our experimental set-up, we only imaged flies for one hour during their life. Future studies would benefit from having longer recording epochs – up to the animal's full lifetime – that could capture the influence of circadian rhythms and could more ably measure inter- vs. intra-individual variability across the lifespan.

Despite these limitations, this study points a way forward for using full repertoires of behavior to study aging and its physiological underpinnings. Although many of our energy budget-related analyses here could have been performed using center-of-mass tracking alone, by studying multiple actions simultaneously, it becomes not only possible to identify the age-relevant behavioral changes (here, primarily related to locomotion and slow/idle behaviors), but also to control for other possibilities such as the complexity of the animal's usage of its behavioral repertoire or behavioral degradation and to isolate covariances between and within age groups. These measurements allows us to better predict how genetic or neural manipulations may affect aging across individuals and across the lifespan, as well as to make more specific predictions as to what types of physiological factors might play a role in these changes.

2.4 Materials and methods

2.4.1 Data

The data consist of 304 flies (D. melanogaster), 150 of which are male and 154 of which are female, with ages ranging from 0 to 70 days of age (all from strain Oregon-R). Within 4 hours of eclosion, flies were isolated in a vial that was changed every other day for food (female flies were all unmated). While unmated and mated female flies

might behave differently, we decided to focus on unmated flies in order to more readily facilitate comparisons between males and females. We anticipate that behavioral differences with age between mated and unmated animals (both males and females) could potentially be different and could be the basis for future studies.

Each fly was imaged from above for an hour while contained in a featureless dish with sloped sides to prevent aerial movements, following the approach detailed in [8]. Flies were placed into the arena using aspiration and provided 5 minutes to adapt to their environment before data collection. To reduce the effect of circadian rhythms, all recordings occurred between 10:00 and 17:00 with incubator lights on from 07:00 to 19:00. (We've measured the behavioral spaces as a function of when the data was taken and calculated the corresponding projections to quantify how time of day did not have an impact on how the flies behaved. See Figures 2.9 and 2.10.) The temperature was kept constant at 25°±1°C.

	Male Flies	Female Flies
0-2 Weeks	32	46
2-4 Weeks	35	40
4-6 Weeks	20	37
6-8 Weeks	31	27
8-10 Weeks	32	4

Table 2.1: Number of Flies in Each Age Group

2.4.2 Behavioral Densities

We created our behavioral densities following the data pipeline outlined in [8]. This approach begins with image analysis (segmentation and alignment), projecting images onto postural eigenmodes, Morlet wavelet transforms [41], and a dimensionally reduced embedding via t-distributed Stochastic Neighbor Embedding [125]. We applied a watershed transform [83] to a Gaussian-smoothed ($\sigma = 1$) density containing points from all the flies in each grouping (All flies, all males, all females, 0-2 week

old males, 0-2 week old females, etc.) in order to isolate the individual peaks. We defined behavioral epochs as lengths of time lasting at least 0.05s with low speeds in the behavioral densities, again following the approach of [8].

2.4.3 Gaussian-smoothed Average Curves

For Figure 2.3D, we applied a Gaussian-smoothed average according to the following equation:

$$y(t) = \frac{\sum_{i=1}^{N} e^{\frac{(t_i - t)^2}{2\sigma^2}} \cdot X_i}{\sum_{i=1}^{N} e^{\frac{(t_i - t)^2}{2\sigma^2}}} \tag{2.7}$$

Here, t is age, X is the original value of the eigenvector projections, y is the smoothed value of X, N is the number of flies, and σ corresponds to the standard deviation of the projections. For example, Figure 2.3D is a plot of y vs. t.

Error bars for these plots are generated through a bootstrapping procedure. Specifically, the data ($\{t_i, X_i\}$) are sampled with replacement, and (2.7) is now applied to this re-sampled data set. This procedure is repeated $1,000$ times (each independently sampled), and the error bars are the standard deviations of these re-sampled curves at each point in time.

2.4.4 Synchronization Parameter

By treating the fruit flies' postural modes as a phase-locked oscillator, we use the *Phaser* algorithm [101] to estimate the behaviors' phases, providing a measure of stereotypy. For each behavior, we use the algorithm to map the individual behavioral bouts to a phase variable between 0 and 2π, providing us with a phase reconstruction of our data that we can compare to the original trajectories (the methodology is the same as in [8]). To ensure the phase-averaged orbits are aligned between individuals and bouts, we calculate the maximum cross-correlation value between orbits for ev-

Body Weight, M	2.5×10^{-6} kg [51]
Body Length, L	2.5×10^{-3} m [51]
Stance Length, S	$(.0472 \times V + .748)/1000$ m [80]
Velocity, V	$0 - 6 \times 10^{-2}$ m/s [51]
Length of Leg, l	1.3×10^{-3} m [51]
Moment of Inertia of the Leg, I	1.6×10^{-14} kgm^2 [51]
Duty Ratio, β	$\frac{t_{st}}{t_{st}+t_{sw}}$
Stance Duration, t_{st}	$11.5 + .910V$ s [80]
Swing Duration, t_{sw}	$(-0.126V + 36.56)/1000$ s [80]

Table 2.2: Parameters Used for Locomotion Energetics Calculations

ery postural mode separately, which gives our phase offset. After determining which modes contribute to each behavior (We use only modes that have mode-specific synchronization parameters of greater that 0.1 which means some behaviors will have a synchronization parameter of 0 if they don't have any modes greater than 0.1), we calculate the synchronization parameter for age group κ for each behavior b across all postural modes γ according to:

$$\mathcal{X}_{b,\kappa} = \frac{1}{N_\kappa^{(\gamma)}} \sum_\gamma \left[1 - \frac{\sigma^2(\vec{y}(\gamma)_{b,\kappa}(\phi) - \vec{\mu}(\gamma)_{b,\kappa}(\phi))}{\sigma^2(\vec{y}(\gamma)_{b,\kappa}(\phi))} \right], \tag{2.8}$$

where $\vec{y}(\gamma)_{b,\kappa}(\phi)$ contains the postural projection time series from every bout of behavior b, $\vec{\mu}(\gamma)_{b,\kappa}(\phi)$ is the phase-averaged orbits for the projection data in $\vec{y}(\gamma)_{b,\kappa}(\phi)$, $N_\kappa^{(\gamma)}$ is the number of postural modes used, and $\sigma^2(x)$ is the variance in x.

By taking the maximum value across the modes, we quantify our stereotypy for each behavior. This value ranges from 0 to 1, where 0 signifies no stereotypy and 1 signifies full stereotypy. This algorithm requires many bouts of each behavior in order to make the calculation.

2.4.5 Deterministic Information Bottleneck

The deterministic information bottleneck algorithm is an iterative algorithm that obeys a set of self-consistent equations:

$$q(t|x) = \frac{1}{\mathcal{Z}(x,\alpha,\beta)} exp\left[\frac{1}{\alpha}(logq(t) - \beta D_{KL}[p(y|x)q(y|t)])\right] \tag{2.9}$$

$$q(t) = \sum_x p(x)q(t|x) \tag{2.10}$$

$$q(y|t) = \frac{1}{q(t)} \sum_x q(t|x)p(x,y) \tag{2.11}$$

Here, $x \in S(n)$, $y \in S(n+\tau)$, $t \in Z$, $\mathcal{Z}$ is a normalizing function, and D_{KL} is the Kullback-Leibler divergence between two probability distributions. For a given $|Z| = K$ number of clusters, inverse temperature β, and random initialization of $q(t|x)$, the equations are iterated until $(\mathcal{F}_t - \mathcal{F}_{t+1})/\mathcal{F}_t < 10^{-6}$ is satisfied, where $\mathcal{F}$ is the previously defined cost function, $\mathcal{F}_{DIB} = H(Z) - \beta I(Z; S(n+\tau))$. We performed 24 replicates of the solution using a range of $\beta \in [0.01, 500]$ spaced exponentially, $K \in [2, 30]$, and $\tau \in [1, 4096]$. The optimization is done for each value of β until the convergence criterion is satisfied. The resulting solution is then used as the initial condition for the next value of β.

2.4.6 Power Estimation Model

We used the model from Nishii (2006)[92] to estimate the power consumption according to the following equations. The swing and stance phase describes the portion of motion where the leg is sweeping forward and when the leg applies pressure to the ground to propel the body forward, respectively.

Specifically, we model the power consumption using the following equations:

$$H^{st} = \gamma \int_{T^{st}} (|\tau^{st}(t)|^k + |\alpha N(t)|^k) dt = \gamma \left(\frac{M}{n}\right)^2 \frac{T}{\beta} \left(\alpha^2 + \frac{S^2}{12}\right) \tag{2.12}$$

$$H^{sw} = \gamma \int_{T^{sw}} |\tau^{sw}(t)|^k dt = \gamma \frac{2\pi^2 I^2}{l^2} \frac{\beta V^3}{S(1-\beta)^3} \tag{2.13}$$

$$W^{st} = \int_{T^{st}} f(N(t)x(t)) \frac{V}{l} dt = \frac{MS^2}{8nl\beta} \tag{2.14}$$

$$W^{sw} = \int_{T^{sw}} f(\tau^{sw}\dot{\theta}) dt = I \left(\frac{V}{l}\right)^2 \frac{1+\beta^2}{(1-\beta)^2}. \tag{2.15}$$

Here, H^{st} is the heat dissipation during the stance phase, and H^{sw} is the heat dissipation during the swing phase. Similarly, W^{st} and W^{sw} denote the mechanical work done during the stance and swing phase, respectively. In these equations, n is the number of legs, γ represents the ratio of heat dissipation to mechanical work, and α is the amplitude of the torque required to maintain a bent leg posture. The rest of the parameters are defined in Table 2.2. We use values to calculate the specific power as a function of velocity, which we called e. We calculate e by summing together the power consumed from the heat and work during the stance and swing phase according to the following equations for e_h^{st}, e_h^{sw}, e_w^{st}, e_w^{sw}:

$$e(V, \beta, S) = e_h^{st} + e_h^{sw} + e_w^{st} + e_w^{sw} \tag{2.16}$$

$$e_h^{st} = \frac{\sum_{i=1}^{n} H_i^{st}}{MVT} = \frac{\gamma M}{n\beta V} \left(\alpha^2 + \frac{S^2}{12}\right) \tag{2.17}$$

$$e_h^{sw} = \frac{\sum_{i=1}^{n} H_i^{sw}}{MVT} = \gamma \frac{2n\pi^2 I^2}{l^2 M} \frac{V^3 \beta^2}{S^2(1-\beta)^3} \tag{2.18}$$

$$e_w^{st} = \frac{\sum_{i=1}^{n} W_i^{st}}{MVT} = \frac{S}{8l} \tag{2.19}$$

$$e_w^{sw} = \frac{\sum_{i=1}^{n} W_i^{sw}}{MVT} = \frac{nI\beta}{MS} \left(\frac{V}{l}\right)^2 \frac{1+\beta^2}{(1-\beta)^2}, \tag{2.20}$$

where T is the gait cycle period.

Using this model, we can estimate the relative mechanical cost of grooming com-

pared to locomotion by the quantity $\frac{e_h^{sw}+e_w^{sw}}{e_h^{st}+e_w^{st}}$, since the animal is moving its legs but is no longer having to expend excess energy to propel itself forward during the stance phase. Across all speeds, this ratio is $\approx 10^{-7}$, justifying our treatment of all zero-velocity epochs as having the same energetic cost.

2.5 Acknowledgments

K.E.O., and K.L., and G.J.B. were supported by NIMH (R01 MH115831-01), the Human Frontier Science Program (RGY0076/2018), and a Cottrell Scholar Award, a program of the Research Corporation for Science Advancement (25999). K.E.O. was supported by the NSF Physics of Living Systems Student Research Network (PHY-1806833). Experimental work and J.W.S and D.M.C. were supported by NIH R01 GM098090 and in part by the National Science Foundation, through the Center for the Physics of Biological Function (PHY-1734030).

2.6 Supplemental Figures:

Figure S1. Behavioral maps as a function of the time of day when the data was taken for all the flies, male flies, female flies, and the 0-2 week old female flies.

Figure S2. Eigenvector projections (as calculated in Figure 2.3D for the maps in Figure 2.9) as a function of time.

Figure S3. Average number of transitions per hour as a function of age. Each data point is a different animal, and the line is the Gaussian-weighted average (error bars are standard error of the mean for the average).

Figure S4. Entropy of the behavioral probabilities as a function of age for the males (left) and females (right) with a best fit line to estimate the value included on the

plot by taking the mean value. Error bars for individual animals are smaller than the symbol size in the plot. The males have a slope of -0.00 ± 0.03 and the females have a slope of -0.01 ± 0.08.

Figure S5. The third, fourth, and fifth eigenvalue timescales for each sex and age group. Line thicknesses represent the standard errors of the mean.

Figure S6. Trade-off curves computed from the deterministic information bottleneck for each sex and age group.

Supplemental Figures:

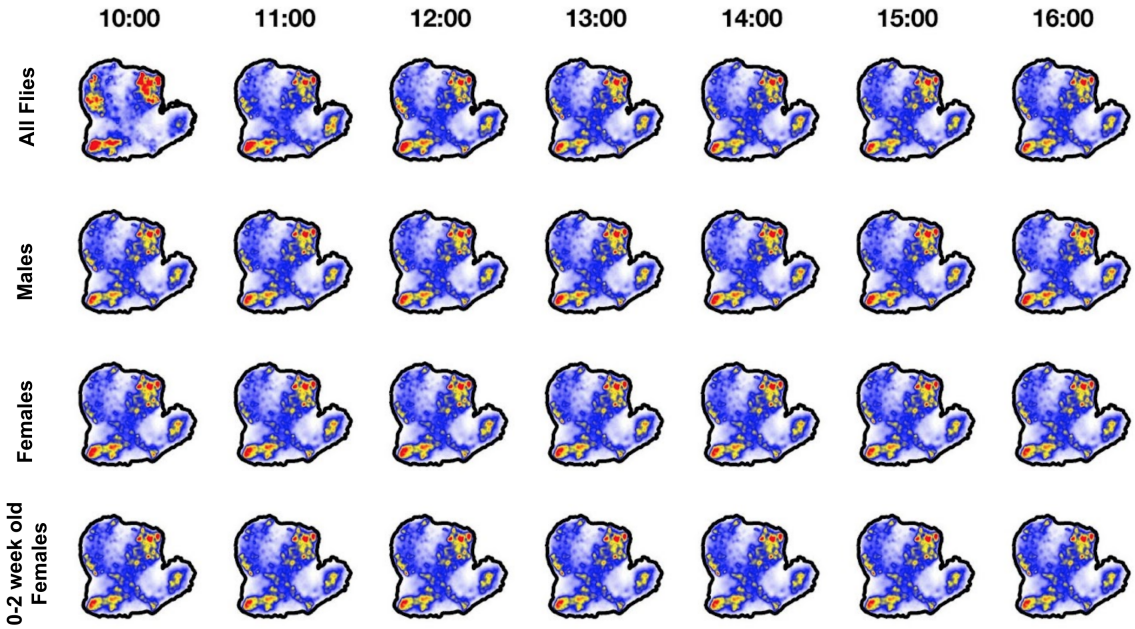


Figure 2.9: Behavioral maps as a function of the time of day when the data was taken for all the flies, male flies, female flies, and the 0-2 week old female flies.

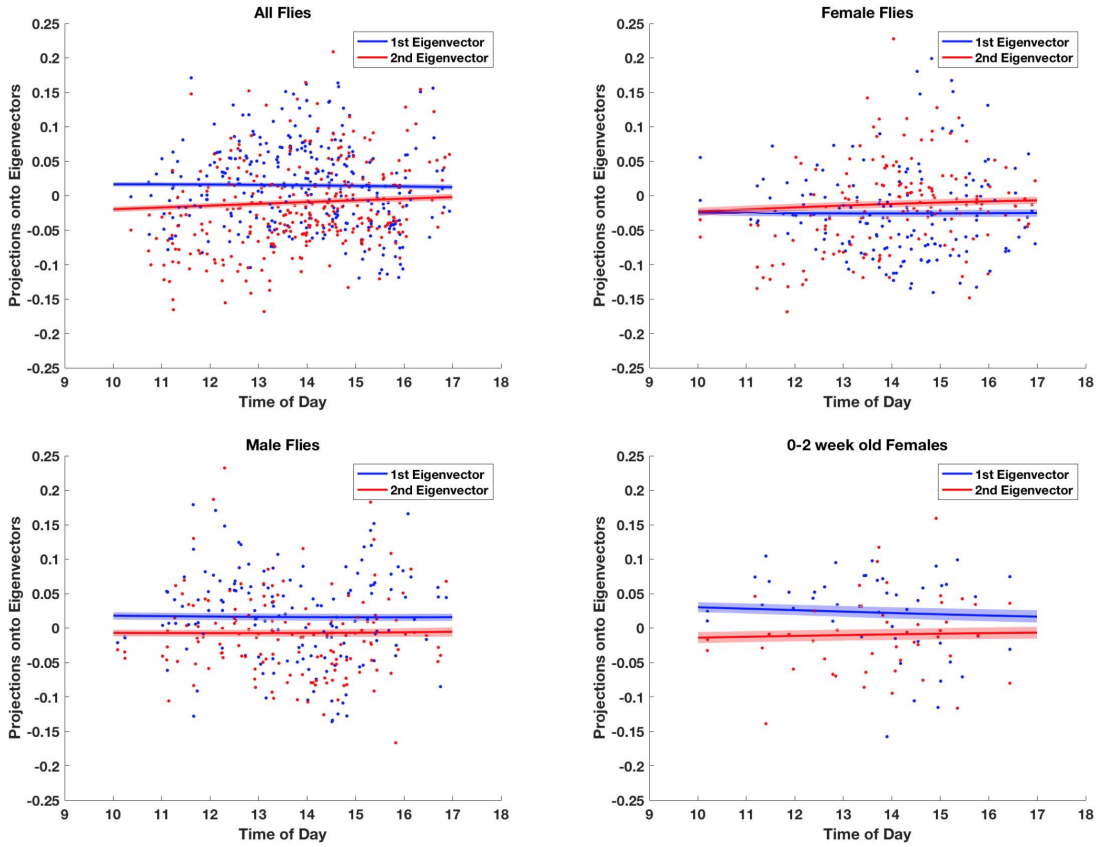


Figure 2.10: Eigenvector projections (as calculated in Figure 2.3D for the maps in Figure 2.9) as a function of time.

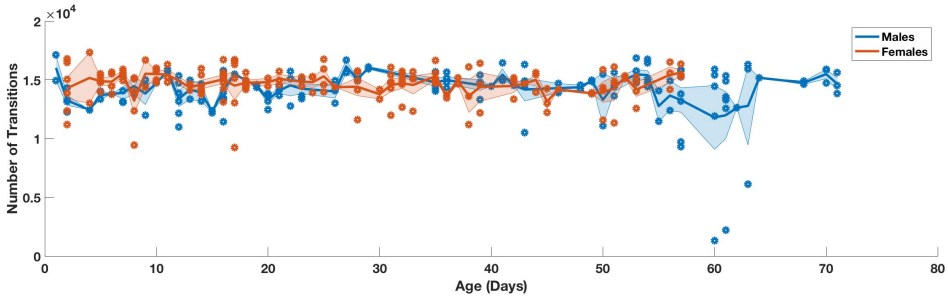


Figure 2.11: Average number of transitions per hour as a function of age. Each data point is a different animal, and the line is the Gaussian-weighted average (error bars are standard error of the mean for the average).

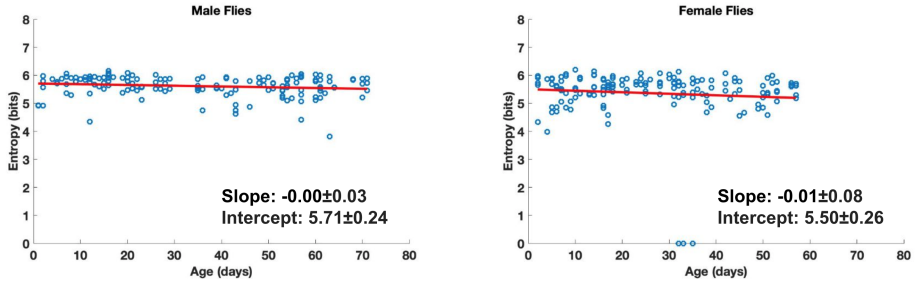


Figure 2.12: Entropy of the behavioral probabilities as a function of age for the males (left) and females (right) with a best fit line to estimate the value included on the plot by taking the mean value. Error bars for individual animals are smaller than the symbol size in the plot. The males have a slope of -0.00 ± 0.03 and the females have a slope of -0.01 ± 0.08.

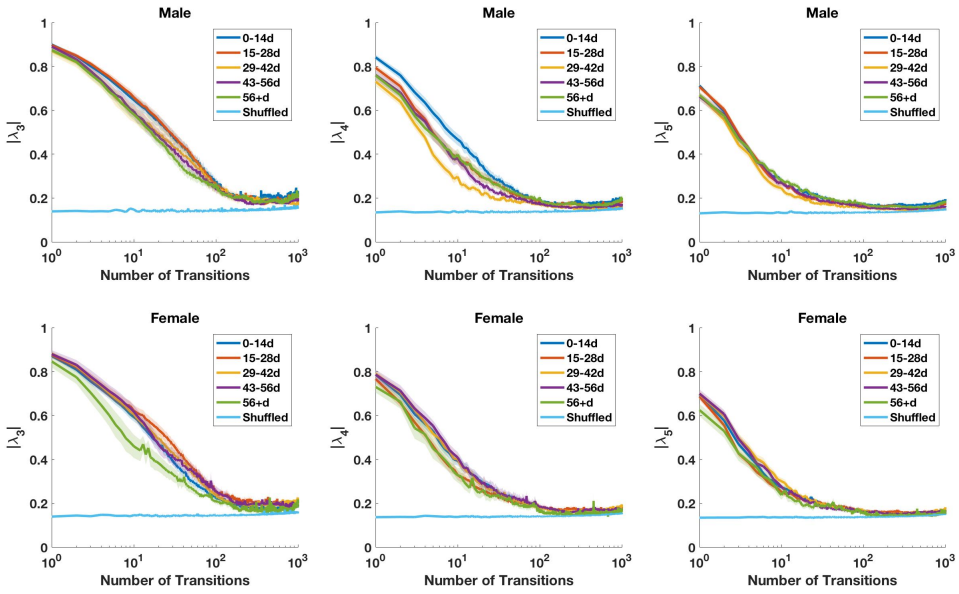


Figure 2.13: The third, fourth, and fifth eigenvalue timescales for each sex and age group. Line thicknesses represent the standard errors of the mean.

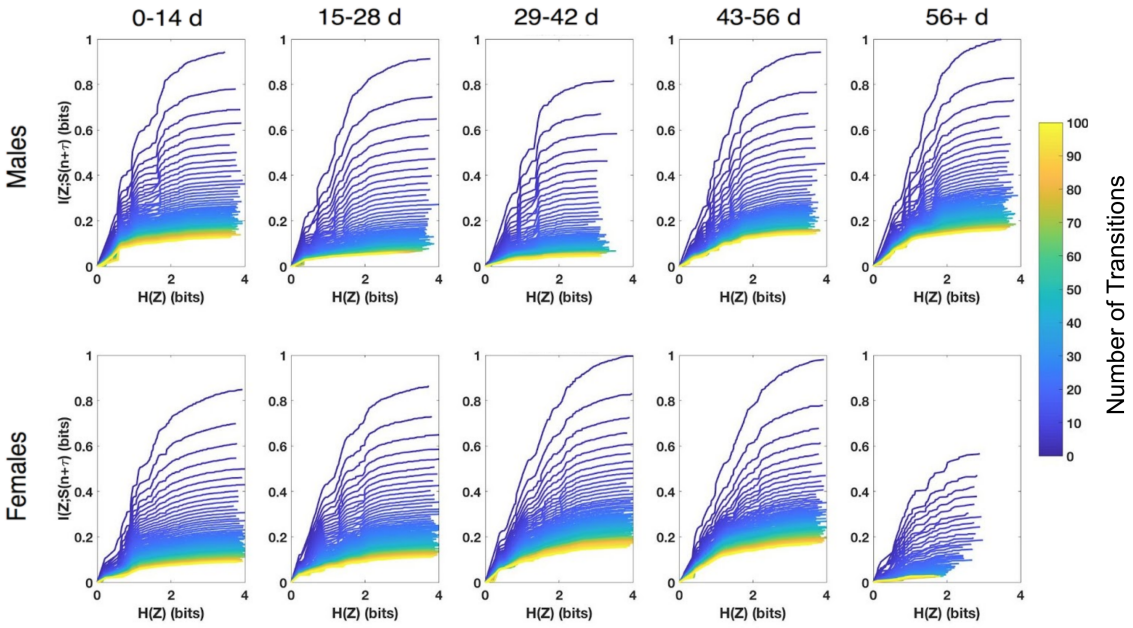


Figure 2.14: Trade-off curves computed from the deterministic information bottleneck for each sex and age group.

Chapter 3

Using Recurrent Neural Networks to Model Hidden Dynamics of Animal Behavior

Abstract

Animal behavior can be broken down into two components, the external postural actions and the internal states that drive those observable postures. For example, hunger will cause an animal to eat, circadian rhythm effects sleeping, and courtship behavior results from various hormones and puberty. Understanding how these internal states drive external behaviors has been a difficult problem to solve. In this chapter, we train a recurrent neural network on behavioral data from fruit flies, verify the network is reproducing the long time scale and hierarchical structures that we see in our data, and study how the internal states of the network interact to produce behavioral data of the network itself. We find that these internal states behave as a particle in a multi-well landscape with shifting barrier heights. This provides a framework for how behavioral sequences are modulated and could be analogous to

systems within the fly.

3.1 Introduction

Animal behavior is known to be a complex process to study. It is generally accepted that an animal's behavior results from interactions between many internal systems within the animal, such as metabolic state and neurological signals as a couple of examples. Behavioral sequences are hierarchical in nature and thus must have some memory of what occurred in the past in order to produce the repeated and non-random patterns that make up behavior. To understand behavior and how these patterns and hierarchy emerge, a deeper understanding of the interactions of the organisms' internal states is essential.

Previous work from Berman et al. (2014) [8] established a new method for categorizing behavioral data. Starting from images taken of a fruit fly from above in a featureless dish, the algorithm uses computer vision, dimensionality reduction, and a few other techniques to automatically detect the full set of stereotyped behaviors (those behaviors that were repeated consistently and frequently) and arrange those behaviors into what we will refer to as the behavioral space (see Figure 3.1A). In this space, the red peaks correspond to the stereotyped behaviors - the behaviors that are performed repeatedly in the same way. By happenstance, similar behaviors end up clustered together as can be seen by the annotated space in Figure 3.1B. Further work from Berman et al. (2016) [9] demonstrates that the behavioral time series from this data set exhibits multiple long time scales (solid lines in Figure 3.1C), and that Markov models (dashed lines in Figure 3.1C) fail to produce any comparable time scales when asked to replicate the dynamics seen in the data. Additionally, when clustering the behavioral space to optimize information about the future (using the information bottleneck algorithm, see Methods section for details), the space breaks

down hierarchically, or rather, to make a new cluster, one of the old clusters breaks into two (see Figure 3.1D). This is also a feature that is difficult to capture with a model.

Recurrent neural networks are a class of model capable of replicating complex temporal and dynamical structures with applications in language processing [128], molecular dynamics [124], image recognition [35], predicting traffic patterns[133], and anomaly detection in time series [72] as merely a few examples. A subclass of these models that are currently popular is long short-term memory (LSTM) which allows the network to have access to and remember information across longer time scales than other RNN architectures like gated recurrent units (GRU) [48]. In the context of behavioral data, these long time scales become an important feature. In this paper, we will use those innate long time scales to model behavior of freely moving fruit flies, *Drosophila. melanogaster* and replicate the long time scales and hierarchical structure mentioned earlier. We will assess whether our RNN is reproducing comparable data with four different metrics: 1) Persistence time within each behavioral state, 4) the structure of the transition matrix across forward predictions, 3) the time scales that the model generates, and 4) the deterministic information bottleneck clusters of the behavioral space (similar to those seen in Berman et al. (2016) [9]).

Although our RNN performs well under these metrics, it is a very complex model that is not easily interpretable. RNN's, although certainly proven to be useful and versatile models, have largely been a mystery in how they manage their computations, but efforts have been made to solve that mystery by treating the internal state interactions as a nonlinear dynamical system [117]. For each state, we can find a fixed point within our network, and "basins" form where the states only interact with certain subsets of states. We find that we can quantify how these basins interact with each other by using Arrhenius behavior. This forces the energy barriers between the basins to change in time, which drives the system to make transitions and produce

the complex dynamics that we see in behavioral data.

3.2 Results

3.2.1 Model Validation

In this paper, we use data from Berman et al. (2014) [8] of a fruit fly moving around freely in a dish. The data has been analyzed to identify the stereotyped behaviors as a function of time and create an embedding of those stereotyped behaviors. This embedding can be seen in Figure 3.1A.

Further work from Berman et al. (2016) [9] has shown that behavioral data exhibits complex temporal dynamics such as long time scales (Figure 3.1B) and hierarchical organization (Figure 3.1C). We want the sequences produced by our RNN to have those same dynamics.

We trained a single layer Long-Short Term Memory (LSTM) network on the behavioral data until the validation loss on our test data set had stopped decreasing after 50 epochs of training, and then chose the set of weights which minimized that loss from the prior epochs. From the training curve in 3.3A, our network appears to have completed training by the 200th epoch. We trained the network using the sequence-to-sequence paradigm with a lag in the output sequence. During the training process, we give the network sequences of data batched into shorter lengths for inputs and outputs and ask the network to learn the relationship between those sequences. The output sequences that we want to predict are the continuation of the input sequences but shifted forward in time by lag-number of steps. For example, with a sequence length of 50 and a lag of 10, the input sequence would look like: $[t_1, t_2, t_3, ..., t_{50}]$ and the sequence the network would predict would look like: $[t_{10}, t_{11}, t_{12}, ..., t_{60}]$. Our network was trained with a sequence length of 500 and a lag of 200. For more details about these parameters, see the Methods section.

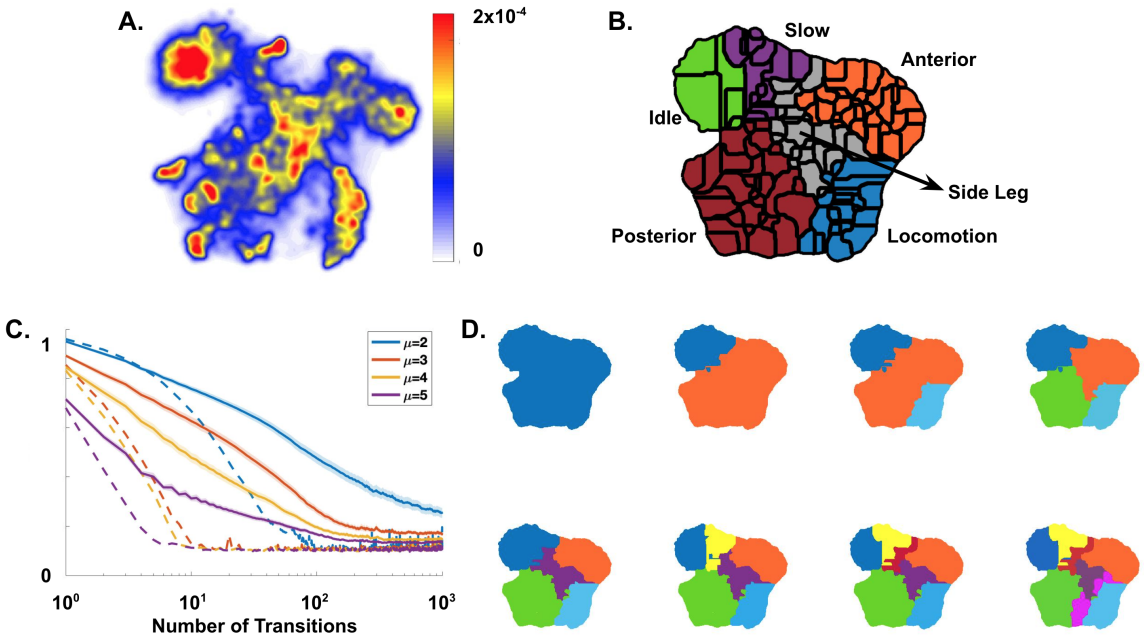


Figure 3.1: Previous Work. (**A**) behavioral space as a probability density function, where each peak represents a stereotyped movement from *D. melanogaster* fruit flies. (**B**) Broad descriptions of different regions in the behavioral space. (**C**) Long time scale dynamics produced by taking the absolute value of the leading eigenvalues of the transition matrices $T(\tau)$ for different values of τ averaged across 59 flies with SEM error bars. Dashed lines represent the same eigenvalues for a Markov model simulation. The black line represents the noise floor calculated from a transition matrix with random temporal shuffling of our finite data set. (**D**) Hierarchical breakdown of the behavioral space with the previous clusters' borders drawn in black. These clusters are calculated with the information bottleneck algorithm at $\tau = 67$, approximately twice the longest time scale in a Markov model.

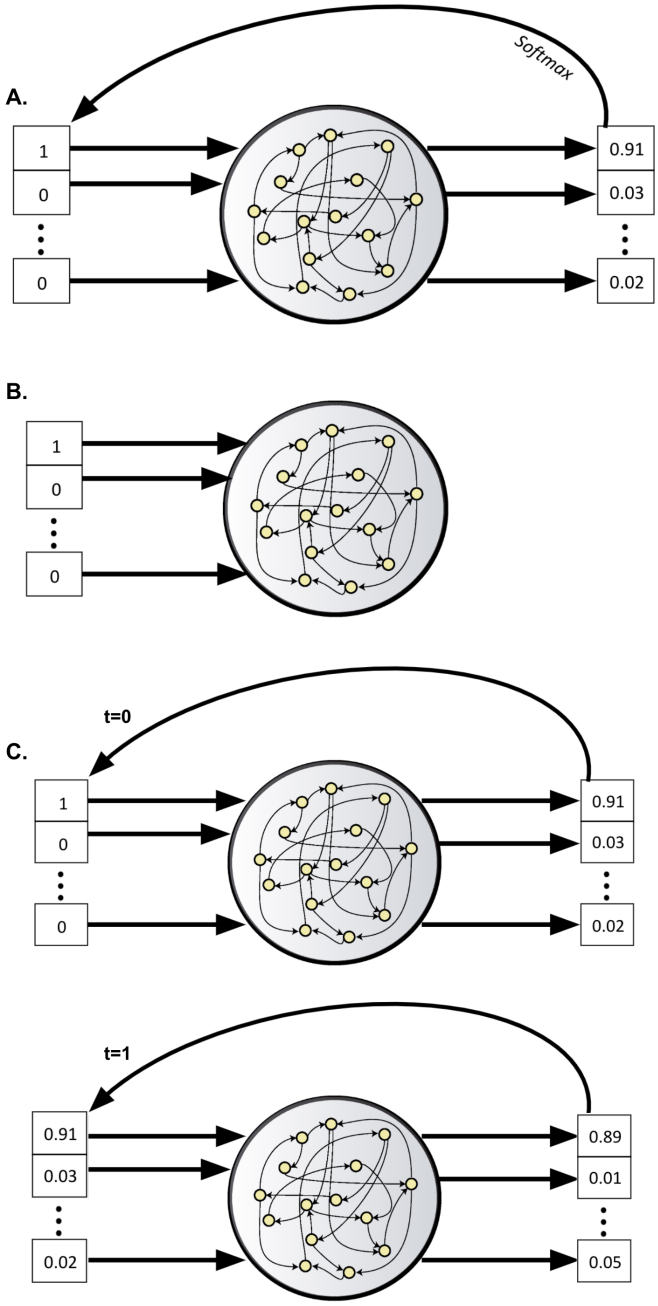


Figure 3.2: RNN modes. (**A**) Generative mode: To generate sequences, we initially give the network one-hot encoded sequences from data. Then, we feed the output through a soft max layer and use that as the input for the next time step. (**B**) Clamp mode: We use this mode to measure the fixed points of the network. Disregarding the output entirely, the network receives a constant input for multiple time steps until it has reached a steady state. (**C**) Release mode: This mode allows us to see how the fixed points interact with each other. Functionally, it is similar to generative mode, but here, we skip the softmax layer and use the output directly as the input for the next time step. This is shown for t=0 and t=1.

Once the network is trained, we are able to run it in what we call "generative" mode in order to produce a data set from the network with time series that are equivalent in length and quantity to our original data set. To initiate this mode, we pass the network data from the testing set and sample the resultant output. This sampled output is used as the initial time steps for out RNN "behavioral" sequences as well as the input to the network to acquire the next time steps. By iteratively passing the network sequences that it generated, we create the RNN "behavior" time series and are able to make comparisons to the fly behavior time series. A diagram of this process can be seen in Figure 3.2A. First, we measure the average amount of time that the RNN and flies spend in a specific state. We've plotted these quantities against each other in Figure 3.3B. By eye, we can see that they are not wholly uncorrelated but overall the correlation is a bit weak. Thus, and in order to remain consistent with the published results from Berman (2016)[9], we will largely work in transitions rather than time. The 1-step transition matrix is known to have a block diagonal structure (Figure 3.3C, right), which implies behaviors prefer to transition between posturally similar behaviors. Although we do not get the exact transition matrix from the flies, the 1-step transition matrix for our RNN sequences maintain that block diagonal structure. Measuring the eigenvalues of the transition matrices n-steps into the future quantifies the time scales of our data. We know from Berman et al. (2016) [9], that the flies produce multiple long time scales (Figure 3.3D, right), and our model successfully reproduces this dynamic. Finally, we want to see a hierarchical breakdown of the behavioral space when using the deterministic information bottleneck algorithm. In other words, to add a new cluster to the space, one of the old clusters breaks into two. The RNN (Figure 3.3E, top) produces different clusters from the flies (Figure 3.3E, bottom) initially and with larger number of clusters, but the clusters are formed hierarchically as a result of the inherent hierarchy in the time series. These hierarchical clusters and the multiple long

time scales together are what we call the summary statistics of our data set. Although the summary statistics from our network are quantitatively different, they have the same qualitative structures that we require from our model, which implies the overall dynamics are the same. Thus, we are satisfied that our network is reproducing the features from the fly data.

3.2.2 Fixed Points Dynamics

Now, that we've shown that our model is reproducing the dynamics that we see in our data, we can study the internal workings of the network, to determine the mechanisms that modulate the "behavior" of the network. Ultimately, we would like to make connections between the internal dynamics of the network to the internal dynamics of the fly itself. By treating the network as a dynamical system, we can measure the fixed points of the network's internal states (commonly referred to as c and h) and observe how those fixed points interact with each other. For more information on these internal states, refer to the Methods section.

To measure the fixed points, we "clamp" the network, giving it a constant input, until the network's internal states have reached a steady state and are no longer changing. This is diagrammed in Figure 3.2B. We do this for each behavioral state and store the final values from both c and h. Using multidimensional scaling (MDS), we get our fixed points, one for each behavioral state, and can embed these points into a 2D plane. This is shown in Figure 3.4.

Here, each of the black points represent one of the fixed points, and the clusters applied come from the 4 bottleneck cluster solution from Figure 3.3E (bottom). This tells us that the internal fixed points of the network are replicating similar structure as the original behavioral states of the flies since the clusters are mostly continuous. Since these fixed points correspond to the stereotyped behaviors, we will use the original behavioral map from Berman (2014)[8] for future visualizations of our

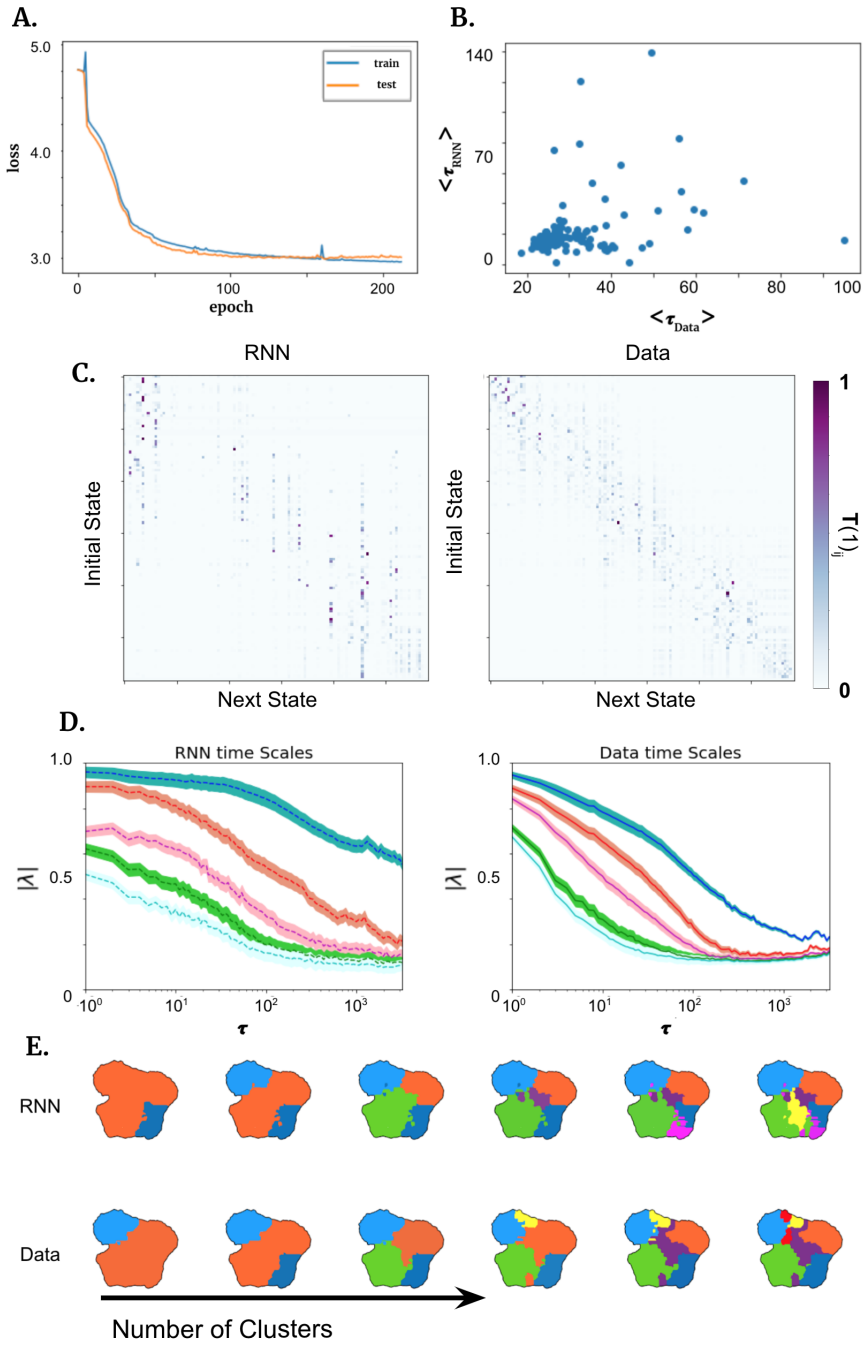


Figure 3.3: Model Summary Statistics. (**A**) Training curve for out RNN, showing both training and testing validation loss as a function of number of epochs. (**B**) Average persistence times in a behavioral state of the RNN vs. our data. (**C**) 1-step transition matrix for the data (left) and RNN (right). (**D**) Timescales from the transition matrices for the data (left) and RNN (right) sequences. (**E**) Comparisons between bottleneck clusters from sequences from the RNN (top) and the data (bottom)

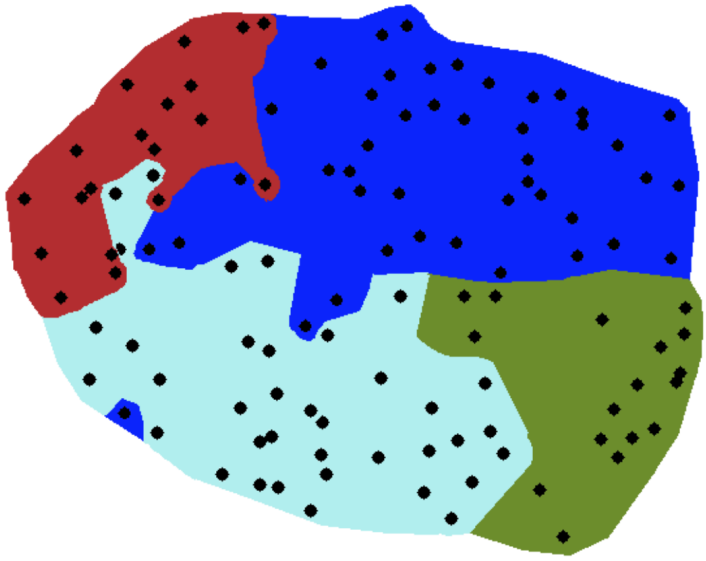


Figure 3.4: MDS embedding of LSTM's fixed points (in black) with the 4 bottleneck cluster solution from Figure 3.3E (bottom) applied.

analyses.

In order to quantify the interactions between these fixed point states, we once again "clamp" the network to initialize the network at a fixed point and then "release" the network and allow it to drive itself, meaning that at each time step we use the network's output from the previous time step as the input. This step is essentially the same as "generative" mode but without the sampling the output. To see this process, refer to Figure 3.2C. Doing this for each fixed point allows us to see where the fixed points flow to. What emerges is a basin-like structure, where a stable fixed point will act as a local minimum for a set of other states to flow to. Every state's trajectory after releasing can be seen in Figure 3.5A. In Figure 3.5B is an example of one of the basins with four fixed point states flowing to the same stable fixed point state. In both figures, each black point indicates the fixed points' location in the behavioral map. Arrows indicate where each state flows to, and the red dots are indicators of states that have been flowed to. The clusters on the map are the same bottleneck solutions from Figure 3.4 and Figure 3.3E(bottom).

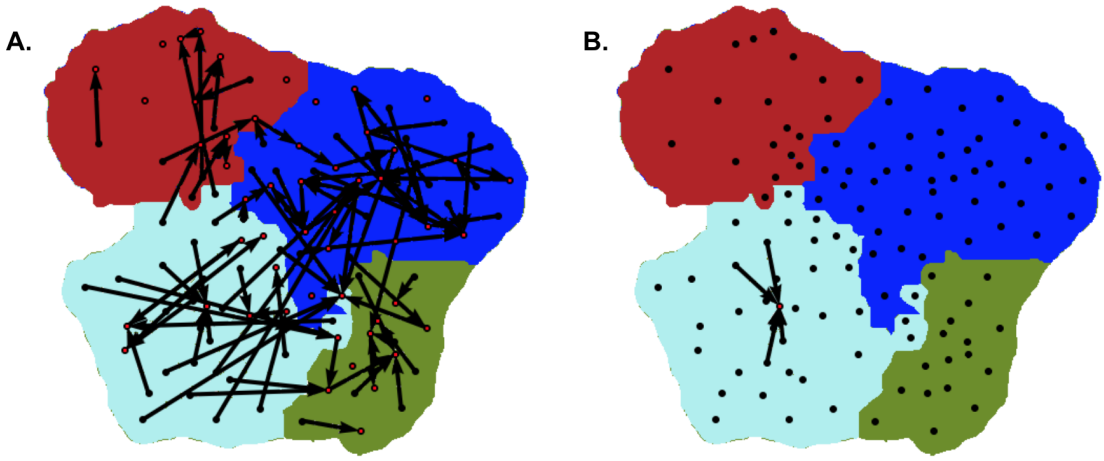


Figure 3.5: Fixed Point Basins. (**A**) Every fixed point's "release" trajectory with arrows indicating the direction of the flow. Red points indicate a state that has been flowed to. This reveals multiple, disparate basins where a set of stable fixed point states act as a local minimum for other states to flow to. (**B**) One example basin for the sake of visibility. Clusters are once again from the bottleneck solution in Figure 3.3E (bottom).

3.2.3 Basin Dynamics

With the emergence of these basins, we can coarse grain our data according to which states are in each basin. This gives us sequences that have basins as a function of time rather than behavior as a function of time. Determining the mechanisms behind how these basins interact will shed light on how behavioral sequences that produce our summary statistics are generated and could reveal similar mechanisms within the fly. After converting the behavioral sequences of both the fly and RNN to the basins paradigm, we check our summary statistics (time scales and DIB clusters) to ensure that these sequences still produce the dynamics of our data. Both RNN basin sequences and the fly data basin sequences produce multiple long time scale (see Figure 3.6A) and have hierarchical clusterings from the DIB algorithm (see Figure 3.6B). In Figure 3.6B, we are using the fixed points' embedding from Figure 3.4. We have represented each basin by using the state that acts as the bottom of the basin. These are shown on the embedding as black dots.

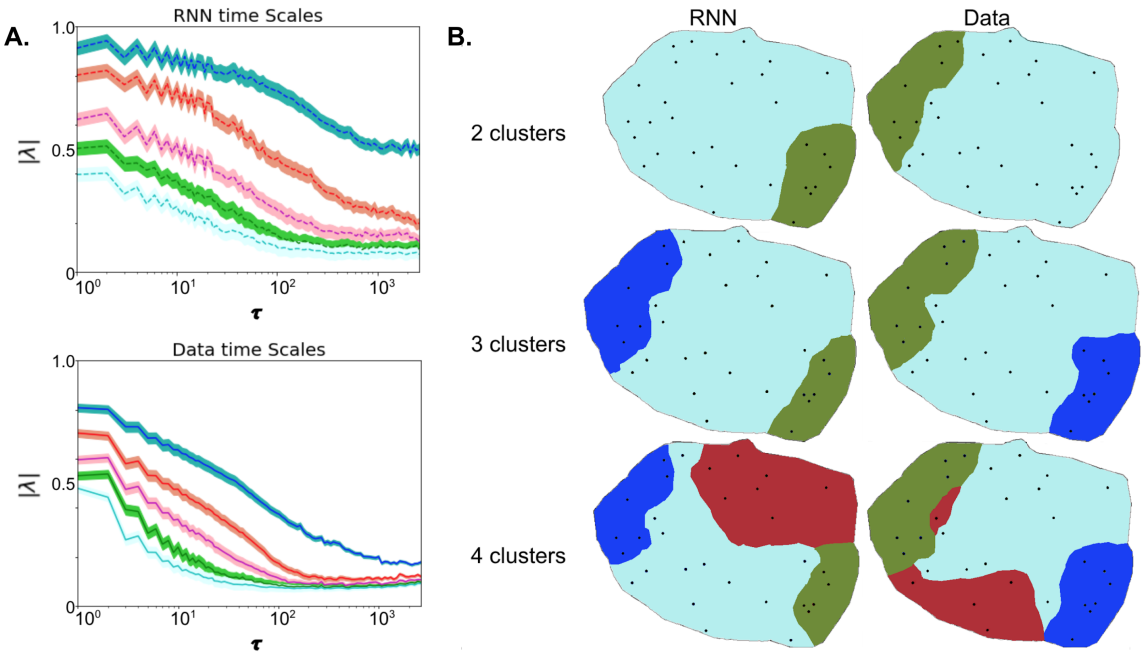


Figure 3.6: Summary Statistics in Basin Paradigm. (**A**) Multiple long time scales emerge from both RNN and fly data coarse grained to basin paradigm. (**B**) One example basin for the sake of visibility. Clusters are once again from the bottleneck solution in Figure 3.3E (bottom).

We propose that there exist energy barriers between the basins that shift with time and modulate the transitions between states. We can think of this as a particle in a multi-well energy landscape, which is illustrated in Figure 1.9. As the particle bounces around in a well, transitions between those behaviors occur. After some time, the landscape and barriers between basins can shift. This could be because of some internal state change or as a response to external stimuli. With this shift, the particle is allowed to switch basins and make transitions between the behaviors in that basin.

To simulate these changes between basin barriers, we can construct a transition matrix that changes in time depending on the local dynamics, by using a weighted average between the average transition probabilities and the local-in-time probability. That weight depends on how many of the behaviors are seen in the local-in-time probabilities.

$$T_{ij}(t) = W_i(t)\tilde{T}_{ij}(t) + (1 - W_i(t))\bar{T}_{ij}$$
$$W_i(t) = \exp\Big[\frac{-\alpha\rho_i(t)}{<\rho_i>}\Big]$$

Here, i and j denote different fixed point basins. T is the transition matrix, $\tilde{T}$ is a transition matrix that's calculated over a sliding window of 20 seconds, and $\bar{T}$ is the overall average transition probability. W is a weight factor dependent on α, which is inversely proportional to a weighting factor on the previous transition matrix in the case that $\rho =< \rho >$. ρ is the states' probability densities across that sliding window. We can further define a transition matrix that changes in time and depends on the change in the energy barriers by using concepts from thermodynamics, specifically Arrhenius behavior.

$$T_{ij}(t) = \frac{\exp[-\beta\Delta E_{ij}(t)]}{\sum_k \exp[-\beta\Delta E_{ik}(t)]} \equiv \frac{\exp[-\beta\Delta E_{ij}(t)]}{Z_i(t)}.$$

In this equation, β is a temperature and ΔE will represent the change in barrier height between basins. We can rearrange the final equation to solve for ΔE, which gives us the change in the energy barrier as a function of time. This gives us:

$$\Delta E_{ij}(t) = \frac{1}{\beta}\big[log(T_{ij}(t)) + log(Z_i(t))\big]$$

Since we know that both the flies and RNN sequences produce similar summary statistics within the fixed point basin paradigm, we can use these equations on the flies' basin sequences to determine if these changing barriers modulate long time scales and hierarchical structure. Without the assumed Arrhenius behavior, this equation returns Markovian behavior. In Figure 3.7A, we show one of the flies' ΔE's time

series for one of its basins. Each line represents the barrier to transition to another basin, and the values are restricted by W_i, or how many behaviors we see in our sliding window.

Now that we know the barriers are changing with time in our data set, we can quantify the time scales of the ΔE's by calculating their auto-correlation as a function of time into the future. Doing so on average across all individuals for the basin in Figure 3.7A, we see long time scales emerge even on the order of minutes, shown in Figure 3.7B. This confirms that the ΔE's are modulating behavior in a way that produces long time scales. However, to further show the long time scale dynamics from these ΔE's, we can simulate transitions between behavioral states and measure the eigenvalues of the resultant transition matrices. These time scales can be seen in Figure 3.8. The solid lines correspond to data, the dashed lines correspond to the simulation. Thus, the simulation from the ΔE dynamics produce long time scales.

Finally, we can check if hierarchical structure is being preserved in the ΔE framework. We do this by further coarse graining our data from basins as a function of time to basin DIB clusters as a function of time. Specifically, we coarse grain according to the DIB clusters from Figure3.6B, left. By calculating the auto-correlation time scales for each of these hierarchy levels, we see long time scales on the order of seconds. These time scales are plotted in Figure 3.9A. This is shorter than the time scales from Figure 3.7B, but that's expected since a lot of information is lost by using the clusters. However, we also see that adding more clusters lengthens the those time scales (see Figure 3.9B). This trend of increasing number of clusters (and therefore increasing the amount of information) varying directly with autocorrelation time tells us the hierarchy is being preserved in the basin paradigm.

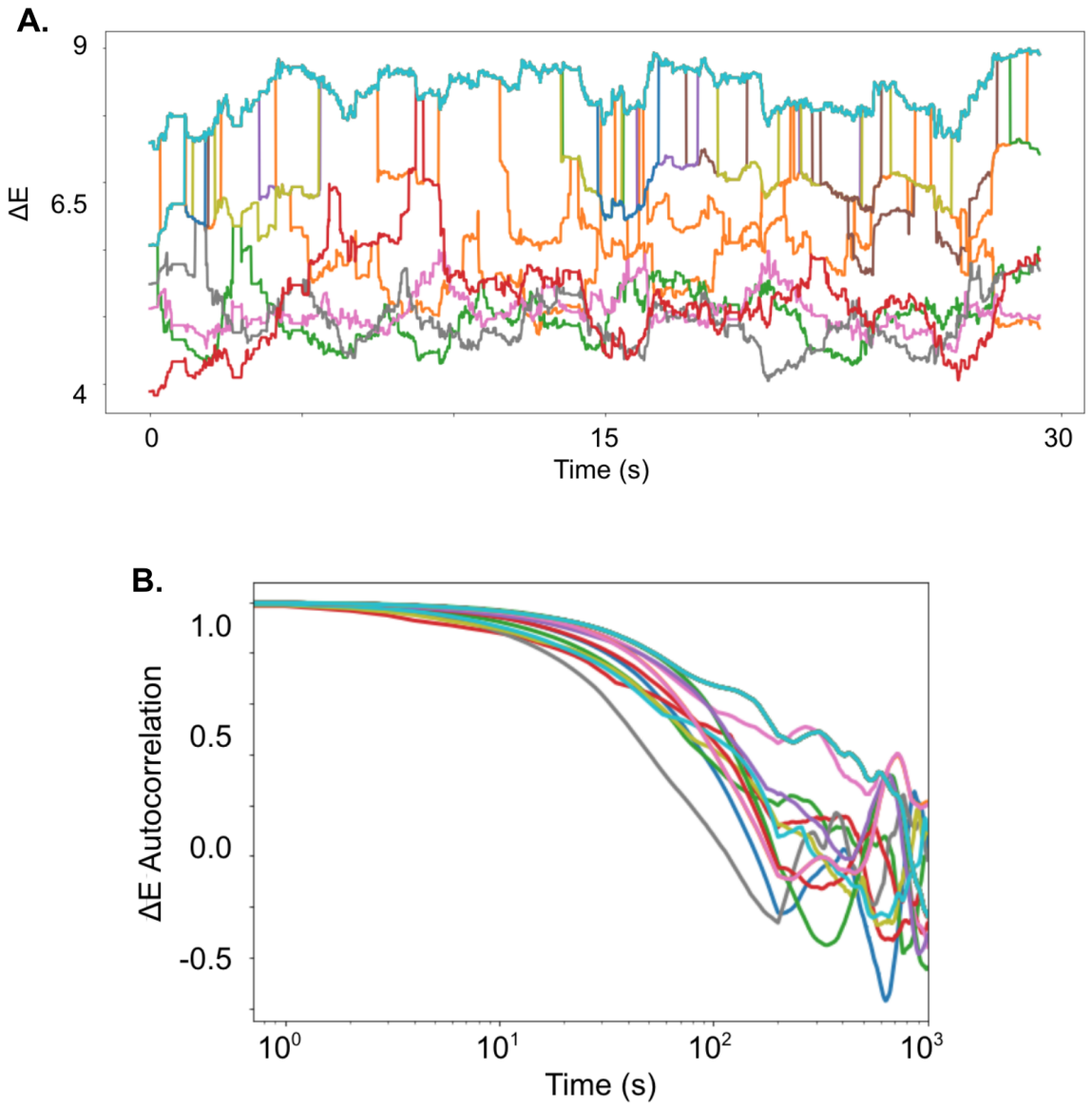


Figure 3.7: (**A**) ΔE barriers changing with time for one of the basins. Each line is the barrier to a different basin. For visibility (**B**) Example of auto-correlation time scales for all the flies combined together (dark blue line) and calculated for flies individually (dark green line).

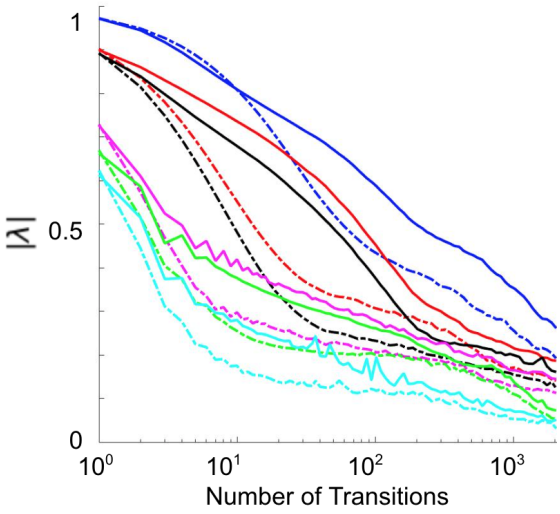


Figure 3.8: Long time scale dynamics from data (solid lines) compared to long time scale dynamics from simulated ΔE dynamics (dashed lines).

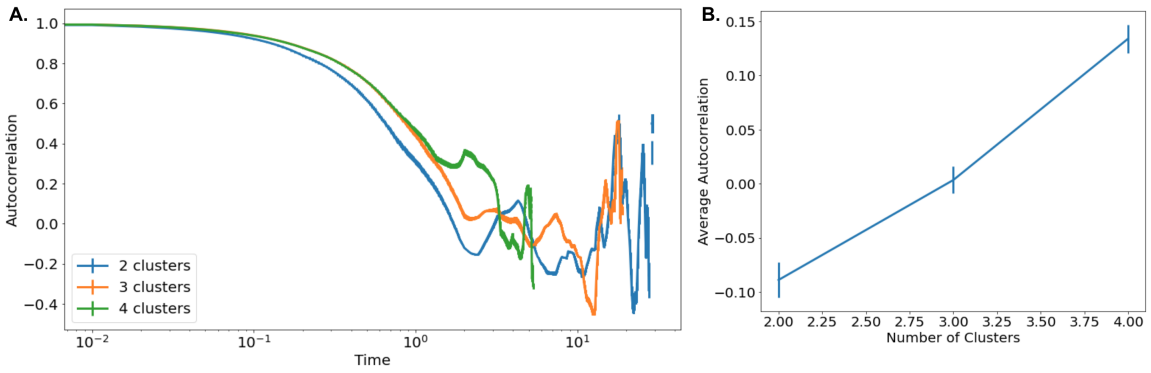


Figure 3.9: (**A**) Auto-correlation time scales calculated from one set of ΔE's that have been coarse grained into the hierarchical structure of the fixed points. (**B**) Plotted the auto-correlation of the time scales from the dashed line at t=100 as a function of number of hierarchical clusters to see an overall increasing trend.

3.3 Discussion

In this chapter, we used a recurrent neural network to replicate behavioral data from 59 fruit flies. This RNN dataset qualitatively reproduced complex structures (block diagonal transition matrices, multiple long time scales, and hierarchical clustering of the behavioral spaces) from the fly data that Markov models have been unable to reproduce. Ultimately, if we want to talk about behavior as a dynamical process, Markov and other statistical models will struggle to learn behavioral data. Thus, we used RNN's as they are inherently a dynamical class of model. These models allow us to investigate the kind of internal mechanisms that are in place to modulate behavioral sequences, by measuring the internal states and their interactions. By treating those internal state interactions as a dynamical system, we measured the fixed points of the system and observed their interactions. From those interactions, a basin-like structure emerged where each state belonging to a basin would ultimately flow to a stable fixed point that acts as a local minimum. By coarse graining our data into the basin paradigm, we confirmed that our long time scales and hierarchical structures were preserved. Thus, we proposed a mathematical model that infers how barriers between the basins change with time and how those changing barriers are capable of modulating behavioral sequences. This theoretical framework is built entirely from dynamics within the network and tested on data from the flies. For future work, this analysis should be done on data that takes simultaneous neural and behavioral recordings.

3.4 Methods

Our Data. The data consisted of 59 male fruit flies (*D. melanogaster*) all under two weeks of age. After given 5 minutes to acclimate to their environment, each fly was imaged from above for an hour while contained in a featureless dish with sloped

sides to prevent aerial movements. To reduce the effect of circadian rhythms, all recordings occurred between 09:00 and 13:00. The temperature was kept constant at 25°±1°C.

Behavioral Densities We created behavioral densities following the data pipeline outlined in [8]. This approach begins with image analysis (segmentation and alignment), projecting images onto postural eigenmodes, Morlet wavelet transforms [41], and a dimensionally reduced embedding via t-distributed Stochastic Neighbor Embedding [125]. We applied a watershed transform [83] to a Gaussian-smoothed density of the resulting points to isolate the individual peaks. We defined behavioral epochs as lengths of time lasting at least 0.05s with low speeds in the behavioral densities, again following the approach of [8].

Recurrent Neural Network. We performed sequence to sequence training using our data set on a 1-layer, stateful Long-Short Term Memory (LSTM) recurrent neural network with 512 hidden units and 500 time steps as the sequence length, and used the number of flies (59) as our batch size. LSTM's have an input gate (i_t), forget gate (f_t), and output gate (o_t) that obey the following set of equations, respectively:

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \tag{3.1}$$

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \tag{3.2}$$

$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o) \tag{3.3}$$

In these equations, w_x contains the weights for the respective gate neurons, h_{t-1} represents the output of the lstm block at the previous time step, x_t is the input of the current time step, and b_x corresponds to each gates' biases.

The internal cell states (c and h) are updated according to:

$$c_t = f_t * c_{t-1} + i_t * \tanh(w_c[h_{t-1}, x_t] + b_c) \tag{3.4}$$

$$h_t = o_t * \tanh(c_t) \tag{3.5}$$

By running the network in generative mode, we created an equivalent dataset of "behavioral" sequences that were then used to make comparisons between our network and the fly.

Deterministic Information Bottleneck The deterministic information bottleneck algorithm is an iterative algorithm that obeys a set of self-consistent equations:

$$q(t|x) = \frac{1}{\mathcal{Z}(x, \alpha, \beta)} \exp\left[\frac{1}{\alpha}(\log q(t) - \beta D_{KL}[p(y|x)q(y|t)])\right] \tag{3.6}$$

$$q(t) = \sum_x p(x)q(t|x) \tag{3.7}$$

$$q(y|t) = \frac{1}{q(t)} \sum_x q(t|x)p(x, y) \tag{3.8}$$

Here, $x \in S(n)$, $y \in S(n + \tau)$, $t \in Z$, $\mathcal{Z}$ is a normalizing function, and D_{KL} is the Kullback-Leibler divergence between two probability distributions. For a given $|Z| = K$ number of clusters, inverse temperature β, and random initialization of $q(t|x)$, the equations are iterated until $(\mathcal{F}_t - \mathcal{F}_{t+1})/\mathcal{F}_t < 10^{-6}$ is satisfied. We performed 24 replicates of the solution using a range of $\beta \in [0.01, 500]$ spaced exponentially, $K \in [2, 30]$, and $\tau \in [1, 4096]$. The optimization is done for each value of β until the convergence criterion is satisfied. The resulting solution is then used as the initial condition for the next value of β.

Hidden State Fixed Points. Previous work from Sussillo et al. (2013) [117] has shown that the internal dynamics of an LSTM behaves like a dynamical system. By

feeding the network a constant input for a long enough time (usually on the order of 100 time steps), the c and h cells reached a stable fixed point. We call this clamping the network. We did this for each possible input (the different behavioral states), and got a set of fixed points one for each behavior state. To see how these fixed points interacted with each other, we made a 2D embedding using t-Distributed Stochastic Neighbor Embedding (t-SNE)/Multidimensional scaling (MDS). Then, we let the network drive itself by making the previous time step's output the current time step's input, and tracked the trajectory of the internal states within our embedded space. We call this releasing the network. While most of trajectories remain stationary at their corresponding fixed points, a handful of "basins" develop where different states will flow to each other and to other states. Each network of non-stationary trajectories forms a basin, and each stationary trajectory acts as its own basin.

Chapter 4

Conclusion

Animal behavior is a complex biological process, but studying it can involve concepts and methods from many other disciplines. In this thesis, we studied behavior of *Drosophila melanogaster* fruit flies. We measured the full repertoire of stereotyped behaviors using unsupervised learning methods from Berman et al. (2014) [8]. In addition, using the resulting behavior classification, we constructed sequences of behavior as a function of time. These sequences have known complex features that were referenced throughout the thesis. First, there are multiple long time scales. We measured these time scales by computing the eigenvalues of the transition matrices across different time steps into the future. Each subsequent time scale decreases in magnitude, however common behavioral models, such as Markov models, produce time scales that are shorter by an order of magnitude or more. Second, the structures that are capable of predicting these long time scales are hierarchical. We clustered the behavioral map using the deterministic information bottleneck algorithm, which finds the clustered structure that minimizes entropy while maximizing information the representation has about the future. These are important features to our data set as time scales and hierarchy are inherent features to behavior. For example, behavioral time scales can range from seconds in actions like grooming to hours like in circadian

rhythm. As for hierarchy, many behaviors like grooming can be broken down into simpler actions as grooming is largely composed of various limb movements. Thus, these two quantities are crucial to the work in this thesis.

In Chapter 2, we quantified how behavior changes with age in male and female fruit flies with ages spanning the typical *Drosophila* lifetime. Female fruit flies begin life very active and gradually become lazier as they age. On the other hand, male flies are inactive in their youth, experience a mid-life boost in activity before again becoming inactive in their old age. We showed that, as flies age, they have a changing energy budget which regulates their preferred behaviors. However, the long time scales and hierarchical structures remain constant across the flies' lifespans.

In Chapter 3, we trained a recurrent neural network on behavioral sequences from 59 males fruit flies. The network was capable of generating sequences that had long time scales and hierarchical structures similar to our data set. However, we wanted to explain what kinds of internal dynamics are modulating behavior. Ultimately, by treating the internal states of the network as a dynamical system and measuring how the fixed points of the system interact, we were able to determine that a particle in a multi-well system with changing barrier heights between the wells is an effective paradigm at producing multiple long time scales and hierarchy in behavior.

Both of these studies could be enhanced by recording data with more information. This could include an experimental set up that allows flight, that allows for social interactions, that is capable of recording neural activity, that measures metabolic state, or some combination of all of these options. Data sets with such rich information about the flies' internal states and with a fuller scope of their behaviors would greatly benefit the field of animal behavior and could help bridge the gap between animal behavior and neuroscience.

Bibliography

[1] Abimbola Akintola and Diana van Heemst. Insulin, Aging, and the Brain: Mechanisms and Implications. *Frontiers in Endocrinology*, 6:13, 2015. doi: 10.3389/fendo.2015.00013.

[2] Vasyl Alba, Gordon Berman, William Bialek, and Joshua Shaevitz. Exploring a strongly non-markovian animal behavior. *arXiv preprint arXiv:2012.15681*, 2020.

[3] Samira Anderson, Alexandra Parbery-Clark, Travis White-Schwoch, and Nina Kraus. Aging affects neural precision of speech encoding. *Journal of Neuroscience*, 32(41):14156–14164, 2012.

[4] Robert Arking. *Biology of aging: observations and principles*. Oxford University Press, 2006.

[5] Svante Arrhenius. Über die reaktionsgeschwindigkeit bei der inversion von rohrzucker durch säuren. *Zeitschrift für physikalische Chemie*, 4(1):226–248, 1889.

[6] Julien Ayroles, Sean Buchanan, Chelsea O'Leary, Kyobi Skutt-Kakaria, Jennifer Grenier, Andrew Clark, Daniel Hartl, and Benjamin de Bivort. Behavioral idiosyncrasy reveals genetic control of phenotypic variability. *Proceedings of the National Academy of Sciences*, 112(21):6706–6711, 05 2015. ISSN 0027-8424. doi: 10.1073/pnas.1503830112.

[7] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *The Institute of Electrical and Electronics Engineers: Transactions on Neural Networks*, 5(2):157–166, 1994.

[8] Gordon Berman, Daniel Choi, William Bialek, and Joshua Shaevitz. Mapping the stereotyped behaviour of freely moving fruit flies. *Journal of The Royal Society Interface*, 11(99):20140672, 2014.

[9] Gordon Berman, William Bialek, and Joshua Shaevitz. Predictability and hierarchy in drosophila behavior. *Proceedings of the National Academy of Sciences*, 113(42):11943–11948, 2016.

[10] William Bialek, Ilya Nemenman, and Naftali Tishby. Predictability, Complexity, and Learning. *Neural Computation*, 13(11):2409–2463, 2001-11. ISSN 0899-7667. doi: 10.1162/089976601753195969.

[11] Ethan Bier. Drosophila, the golden bug, emerges as a tool for human genetics. *Nature Reviews Genetics*, 6(1):9–23, 2005.

[12] Jacob Billings, Alessio Medda, Sadia Shakil, Xiaohong Shen, Amrit Kashyap, Shiyang Chen, Anzar Abbas, Xiaodi Zhang, Maysam Nezafati, Wen-Ju Pan, Gordon Berman, and Shella Keilholz. Instantaneous brain dynamics mapped to a continuous state space. *NeuroImage*, 162:344–352, 08 2017. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2017.08.042.

[13] Nicholas Bishop, Tao Lu, and Bruce Yankner. Neural mechanisms of ageing and cognitive decline. *Nature*, 464(7288):529–535, 2010. ISSN 0028-0836. doi: 10.1038/nature08983.

[14] Johan Bolhuis. The development of animal behavior: from lorenz to neural nets. *Naturwissenschaften*, 86(3):101–111, 1999.

[15] Eric Le Bourg. Patterns of movement and ageing in Drosophila melanogaster. *Archives of Gerontology and Geriatrics*, 2(4):299–306, 1983. ISSN 0167-4943. doi: 10.1016/0167-4943(83)90003-1.

[16] Eric Le Bourg. The rate of living theory. Spontaneous locomotor activity, aging and longevity in Drosophila melanogaster. *Experimental Gerontology*, 22(5): 359–369, 1987. ISSN 0531-5565. doi: 10.1016/0531-5565(87)90034-9.

[17] Susan Broughton and Linda Partridge. Insulin/IGF-like signalling, the central nervous system and aging. *Biochemical Journal*, 418(1):1–12, 2009. ISSN 0264-6021. doi: 10.1042/bj20082102.

[18] Lindsay Barton Browne. Physiologically induced changes in resource-oriented behavior. *Annual Review of Entomology*, 38(1):1–23, 1993.

[19] Edward Callaway. A molecular and genetic arsenal for systems neuroscience. *Trends in neurosciences*, 28(4):196–201, 2005.

[20] Jessica Cande, Shigehiro Namiki, Jirui Qiu, Wyatt Korff, Gwyneth M Card, Joshua Shaevitz, David Stern, and Gordon Berman. Optogenetic dissection of descending behavioral control in Drosophila. *eLife*, 7:e34275, 06 2018. doi: 10.7554/elife.34275.

[21] James Carey, Nikos Papadopoulos, Nikos Kouloussis, Byron Katsoyannos, Hans-Georg Müller, Jane-Ling Wang, and Yi-Kuan Tseng. Age-specific and lifetime behavior patterns in drosophila melanogaster and the mediterranean fruit fly, ceratitis capitata. *Experimental gerontology*, 41(1):93–97, 2006.

[22] Sheng Chen and Stephen Billings. Neural networks for nonlinear dynamic system modelling and identification. *International journal of control*, 56(2):319–346, 1992.

[23] Daniel Choi. *The Study of Complex Behavioral Changes with Age in* Drosophila melanogaster *Using a Quantitative Framework for Behavioral Description*. PhD thesis, Princeton University, Princteon, NJ, 2017.

[24] Kevin Clark. Arrhenius-kinetics evidence for quantum tunneling in microbial "social" decision rates. *Communicative & integrative biology*, 3(6):540–544, 2010.

[25] Heiko Dankert, Liming Wang, Eric Hoopfer, David Anderson, and Pietro Perona. Automated monitoring and analysis of social behavior in drosophila. *Nature methods*, 6(4):297–303, 2009.

[26] David Davis. *Integral animal behavior*. Macmillan, 1966.

[27] Marian Dawkins and Richard Dawkins. Hierachical organization and postural facilitation: Rules for grooming in flies. *Animal Behaviour*, 24(4):739–755, 1976.

[28] Fabrice De Chaumont, Renata Dos-Santos Coura, Pierre Serreau, Arnaud Cressant, Jonathan Chabout, Sylvie Granon, and Jean-Christophe Olivo-Marin. Computerized video analysis of social interactions in mice. *Nature methods*, 9 (4):410–417, 2012.

[29] Anthony Dell, John Bender, Kristin Branson, Iain Couzin, Gonzalo de Polavieja, Lucas Noldus, Alfonso Pérez-Escudero, Pietro Perona, Andrew Straw, Martin Wikelski, and Ulrich Brose. Automated image-based tracking and its application in ecology. *Trends in ecology & evolution*, 29(7):417–428, 2014.

[30] Brian Dennis, Laila Assas, Saber Elaydi, Eddy Kwessi, and George Livadiotis. Allee effects and resilience in stochastic populations. *Theoretical Ecology*, 9(3): 323–335, 2016.

[31] Nejc Dolensek, Daniel Gehrlach, Alexandra Klein, and Nadine Gogolla. Facial expressions of emotion states and their neuronal correlates in mice. *Science*, 368(6486):89–94, 2020.

[32] Magnus Enquist and Stefano Ghirlanda. *Neural networks and animal behavior*. Princeton University Press, 2013.

[33] Aaron Ericsson, Marcus Crim, and Craig Franklin. A brief history of animal modeling. *Missouri medicine*, 110(3):201, 2013.

[34] Henry Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115, 1935.

[35] Yin Fan, Xiangju Lu, Dian Li, and Yuanliu Liu. Video-based emotion recognition using cnn-rnn and c3d hybrid networks. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, pages 445–450, 2016.

[36] Jose Fernández, Michael Grant, N. Tulli, Laura Karkowski, and Gerald McClearn. Differences in locomotor activity across the lifespan of Drosophila melanogaster. *Experimental Gerontology*, 34(5):621–631, 1999. ISSN 0531-5565. doi: 10.1016/s0531-5565(99)00040-6.

[37] Miguel Angel Fernández-Moreno, Carol Farr, Laurie Kaguni, and Rafael Garesse. Drosophila melanogaster as a model system to study mitochondrial biology. In *Mitochondria*, pages 33–49. Springer, 2007.

[38] Nicolas Gauvrit, Hector Zenil, Fernando Soler-Toscano, Jean-Paul Delahaye, and Peter Brugger. Human behavioral complexity peaks at age 25. *The Public Library of Science: Computational Biology*, 13(4):e1005408, 04 2017. ISSN 1553-734X. doi: 10.1371/journal.pcbi.1005408.

[39] Marc Gistelinck, Jean-Charles Lambert, Patrick Callaerts, Bart Dermaut, and Pierre Dourlen. Drosophila models of tauopathies: what have we learned? *International Journal of Alzheimer's Disease*, 2012, 2012.

[40] Yoav Goldberg. A primer on neural network models for natural language processing. *Journal of Artificial Intelligence Research*, 57:345–420, 2016.

[41] Pierre Goupillaud, Alexander Grossmann, and Jean Morlet. Cycle-octave and related transforms in seismic signal analysis. *Geoexploration*, 23(1):85–102, 1984. ISSN 0016-7142. doi: 10.1016/0016-7142(84)90025-5.

[42] Michael Grotewiel, Ian Martin, Poonam Bhandari, and Eric Cook-Wiens. Functional senescence in drosophila melanogaster. *Ageing research reviews*, 4(3): 372–397, 2005.

[43] Daniel Gutierrez-Galan, Juan Dominguez-Morales, Elena Cerezuela-Escudero, Antonio Rios-Navarro, Ricardo Tapiador-Morales, Manuel Rivas-Perez, Manuel Dominguez-Morales, Angel Jimenez-Fernandez, and Alejandro Linares-Barranco. Embedded neural network for real-time animal behavior classification. *Neurocomputing*, 272:17–26, 2018.

[44] Justin Halberda, Ryan Ly, Jeremy Wilmer, Daniel Naiman, and Laura Germine. Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, 109(28):11116–11120, 2012. ISSN 0027-8424. doi: 10.1073/pnas.1200196109.

[45] Walter Heiligenberg. Random processes describing the occurrence of behavioural patterns in a cichlid fish. *Animal Behaviour*, 21(1):169–182, 1973. ISSN 0003-3472. doi: 10.1016/s0003-3472(73)80057-0.

[46] Damián Hernández, Catalina Rivera, Jessica Cande, Baohua Zhou, David Stern,

and Gordon Berman. A framework for studying behavioral evolution by reconstructing ancestral repertoires. *eLife*, 10:e61806, 2021. doi: 10.7554/elife.61806.

[47] Eric Tatt Wei Ho. *Automated Handling of Drosophila for Biological Investigation*. PhD thesis, Stanford University, 2013.

[48] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[49] Martin Holzenberger, Joëlle Dupont, Bertrand Ducos, Patricia Leneuve, Alain Géloën, Patrick Even, Pascale Cervera, and Yves Le Bouc. IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature*, 421(6919): 182–187, 2003. ISSN 0028-0836. doi: 10.1038/nature01298.

[50] Kyle Honegger and Benjamin de Bivort. Stochasticity, individuality and behavior. *Current Biology*, 28(1):R8–R12, 01 2018. ISSN 0960-9822. doi: 10.1016/j.cub.2017.11.058.

[51] Alexander Isakov, Sean Buchanan, Brian Sullivan, Akshitha Ramachandran, Joshua Chapman, Edward Lu, Lakshminarayanan Mahadevan, and Benjamin de Bivort. Recovery of locomotion after injury in Drosophila melanogaster depends on proprioception. *Journal of Experimental Biology*, 219(11):1760–1771, 06 2016. ISSN 0022-0949. doi: {10.1242/jeb.133652}.

[52] Richard Jessor. The problem of reductionism in psychology. *Psychological Review*, 65(3):170, 1958.

[53] Dezhe Jin and Alexay Kozhevnikov. A compact statistical model of the song syntax in bengalese finch. *The Public Library of Science: Computational Biology*, 7(3):e1001108, 2011.

[54] Brad Johnson, David Sinclair, and Leonard Guarente. Molecular biology of aging. *Cell*, 96(2):291–302, 1999.

[55] Mayank Kabra, Alice Robie, Marta Rivera-Alba, Steven Branson, and Kristin Branson. Jaaba: interactive machine learning for automatic annotation of animal behavior. *Nature methods*, 10(1):64–67, 2013.

[56] Rachel Kaletsky and Coleen Murphy. The role of insulin/IGF-like signaling in C. elegans longevity and aging. *Disease Models & Mechanisms*, 3(7-8):415–419, 2010. ISSN 1754-8403. doi: 10.1242/dmm.001040.

[57] Cynthia Kenyon. The Plasticity of Aging: Insights from Long-Lived Mutants. *Cell*, 120(4):449–460, 2005. ISSN 0092-8674. doi: 10.1016/j.cell.2005.02.002.

[58] Cynthia Kenyon, Jean Chang, Erin Gensch, Adam Rudner, and Ramon Tabtiang. A C. elegans mutant that lives twice as long as wild type. *Nature*, 366 (6454):461–464, 1993. ISSN 0028-0836. doi: 10.1038/366461a0.

[59] Suleman Khan, Hammad Javed, Ehtasham Ahmed, Syed Shah, and Syed Umaid Ali. Facial recognition using convolutional neural networks and implementation on smart glasses. In *2019 International Conference on Information Science and Communication Technology (ICISCT)*, pages 1–6. IEEE, 2019.

[60] Thomas Kirkwood and Steven Austad. Why do we age? *Nature*, 408(6809): 233, 2000.

[61] Ugne Klibaite, Mikhail Kislin, Jessica Verpeut, Xiaoting Sun, Joshua Shaevitz, and Samuel Wang. Deep Behavioral Phenotyping of Mouse Autism Models using Open-Field Behavior. *bioRxiv*, page 2021.02.16.431500, 2021. doi: 10.1101/2021.02.16.431500.

[62] Natraj Krishnan, Doris Kretzschmar, Kuntol Rakshit, Eileen Chow, and Jadwiga Giebultowicz. The circadian clock gene period extends healthspan in aging drosophila melanogaster. *Impact Journals: Aging*, 1(11):937, 2009.

[63] Daniel Laurier. Explaining behavior: Reasons in a world of causesfred dretske cambridge, ma, mit press, 1988, xi, 165 p. *Dialogue: Canadian Philosophical Review/Revue canadienne de philosophie*, 32(3):629–634, 1993.

[64] Louis Lefebvre. Grooming in crickets: timing and hierarchical organization. *Animal Behaviour*, 29(4):973–984, 1981.

[65] Louis Lefebvre. The organization of grooming in budgerigars. *Behavioural processes*, 7(2):93–106, 1982.

[66] Louis Lefebvre and Robert Joly. Organization rules and timing in kestrel grooming. *Animal Behaviour*, 30(4):1020–1028, 1982.

[67] David Zhan Liu and Gurbir Singh. A recurrent neural network based recommendation system. In *International Conference on Recent Trends in Engineering, Science & Technology*, 2016.

[68] Carlos López-Otín, Maria Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. The hallmarks of aging. *Cell*, 153(6):1194–1217, 2013.

[69] Konrad Lorenz. The comparative method in studying innate behavior patterns. *Society for Experimental Biology*, pages 221–268, 1950.

[70] Iain Macdonald and David Raubenheimer. Hidden markov models and animal behaviour. *Biometrical Journal*, 37(6):701–712, 1995.

[71] Ana Machado, Dana Darmohray, Joao Fayad, Hugo Marques, and Megan Carey. A quantitative framework for whole-body coordination reveals specific deficits in freely walking ataxic mice. *Elife*, 4:e07892, 2015.

[72] Pankaj Malhotra, Lovekesh Vig, Gautam Shroff, and Puneet Agarwal. Long short term memory networks for anomaly detection in time series. In *Proceedings*, volume 89, pages 89–94. Presses universitaires de Louvain, 2015.

[73] Todd Manini. Energy expenditure and aging. *Ageing Research Reviews*, 9(1): 1–11, 2010. ISSN 1568-1637. doi: 10.1016/j.arr.2009.08.002.

[74] Jesse Marshall, Diego Aldarondo, Timothy Dunn, William Wang, Gordon Berman, and Bence Ölveczky. Continuous Whole-Body 3D Kinematic Recordings across the Rodent Behavioral Repertoire. *Neuron*, 109(3):420–437.e8, 2021. ISSN 0896-6273. doi: 10.1016/j.neuron.2020.11.016.

[75] Minsky Marvin and Papert Seymour. Perceptrons, 1969.

[76] Edward Masoro. Overview of caloric restriction and ageing. *Mechanisms of ageing and development*, 126(9):913–922, 2005.

[77] Sami Masri, Andrew Smyth, Anastasios Chassiakos, Thomas Caughey, and Norman Hunter. Application of neural networks for detection of changes in nonlinear systems. *Journal of Engineering Mechanics*, 126(7):666–676, 2000.

[78] Alexander Mathis, Pranav Mamidanna, Kevin Cury, Taiga Abe, Venkatesh Murthy, Mackenzie Weygandt Mathis, and Matthias Bethge. Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, 21(9):1281–1289, 2018.

[79] Warren McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[80] César Mendes, Imre Bartos, Turgay Akay, Szabolcs Márka, and Richard Mann. Quantification of gait parameters in freely walking wild type and

sensory deprived Drosophila melanogaster. *eLife*, 2:e00231, 01 2013. doi: {10.7554/elife.00231}.

[81] Hylton Menz, Stephen Lord, and Richard Fitzpatrick. Age-related differences in walking stability. *Age and ageing*, 32(2):137–142, 2003.

[82] Michael Menzinger and Richard Wolfgang. The meaning and use of the arrhenius activation energy. *Angewandte Chemie International Edition in English*, 8 (6):438–444, 1969.

[83] Fernand Meyer. Topographic distance and watershed lines. *Signal Processing*, 38(1):113–125, 1994-07. ISSN 0165-1684. doi: 10.1016/0165-1684(94)90060-4.

[84] Denis Michel. Simply conceiving the arrhenius law and absolute kinetic constants using the geometric distribution. *Physica A: Statistical Mechanics and its Applications*, 392(19):4258–4264, 2013.

[85] Denis Michel. A probabilistic rate theory connecting kinetics to thermodynamics. *Physica A: Statistical Mechanics and its Applications*, 503:26–44, 2018.

[86] Denis Michel. Test of the formal basis of arrhenius law with heat capacities. *Physica A: Statistical Mechanics and Its Applications*, 510:188–199, 2018.

[87] Denis Michel. Reductionist approach to chemical rate constants using conditional energy probabilities. *European Journal of Physics*, 40(5):055103, 2019.

[88] George Miller. The cognitive revolution: a historical perspective. *Trends in cognitive sciences*, 7(3):141–144, 2003.

[89] Alexey Moskalev, Mikhail Shaposhnikov, Nadezhda Zemskaya, Liubov Koval, Eugenia Schegoleva, Zulfiya Guvatova, George Krasnov, Ilya Solovev, Maksim Sheptyakov, Alex Zhavoronkov, and Anna Kudryavtseva. Transcriptome Analysis of Long-lived Drosophila melanogaster E(z) Mutants Sheds Light on the

Molecular Mechanisms of Longevity. *Scientific Reports*, 9(1):9151, 2019. doi: 10.1038/s41598-019-45714-x.

[90] Coleen Murphy and Patrick Hu. Insulin/insulin-like growth factor signaling in C. elegans. *WormBook*, pages 1–43, 2013. doi: 10.1895/wormbook.1.164.1.

[91] David Nicholson, Todd Roberts, and Samuel Sober. Thalamostriatal and cerebellothalamic pathways in a songbird, the bengalese finch. *Journal of Comparative Neurology*, 526(9):1550–1570, 2018.

[92] Jun Nishii. An analytical estimation of the energy cost for legged locomotion. *Journal of Theoretical Biology*, 238(3):636–645, 01 2006. ISSN 0022-5193. doi: 10.1016/j.jtbi.2005.06.027.

[93] Christopher Olah. Understanding lstm networks, Aug 2015. URL `http://colah.github.io/posts/2015-08-Understanding-LSTMs/`.

[94] Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. On the difficulty of training recurrent neural networks. In *International conference on machine learning*, pages 1310–1318. PMLR, 2013.

[95] Orit Peleg, Jacob Peters, Mary Salcedo, and Lakshminarayanan Mahadevan. Collective mechanical adaptation of honeybee swarms. *Nature Physics*, 14(12): 1193–1198, 2018.

[96] Talmo Pereira, Nathaniel Tabris, Arie Matsliah, David Turner, Junyu Li, Shruthi Ravindranath, Eleni Papadoyannis, Edna Normand, David Deutsch, Yan Wang, Grace McKenzie-Smith, Catalin Mitelut, Marielisa Diez Castro, John D'Uva, Mikhail Kislin, Dan Sanes, Sarah Kocher, Samuel Wang, Annegret Falkner, Joshua Shaevitz, and Mala Murthy. Sleap: A deep learning system for multi-animal pose tracking. *Nature methods*, 19(4):486–495, 2022.

[97] Talmo D Pereira, Diego E Aldarondo, Lindsay Willmore, Mikhail Kislin, Samuel S-H Wang, Mala Murthy, and Joshua W Shaevitz. Fast animal pose estimation using deep neural networks. *Nature methods*, 16(1):117–125, 2019.

[98] Alfonso Pérez-Escudero, Julián Vicente-Page, Robert Hinz, Sara Arganda, and Gonzalo De Polavieja. idtracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature methods*, 11(7):743, 2014.

[99] Zeke Piskulich. *Beyond Arrhenius: Fluctuation Theory for Dynamics*. PhD thesis, University of Kansas, 2021.

[100] Valeriya Privalova, Ewa Szlachcic, Lukasz Sobczyk, Natalia Szabla, and Marcin Czarnoleski. Oxygen Dependence of Flight Performance in Ageing Drosophila melanogaster. *Biology*, 10(4):327, 2021. doi: 10.3390/biology10040327.

[101] Shai Revzen and John Guckenheimer. Estimating the phase of synchronized oscillators. *Physical Review E*, 78(5):051907, 2008. ISSN 1539-3755. doi: 10.1103/physreve.78.051907.

[102] Angela Ridgel and Roy Ritzmann. Insights into age-related locomotor declines from studies of insects. *Ageing Research Reviews*, 4(1):23–39, 01 2005. ISSN 1568-1637. doi: 10.1016/j.arr.2004.08.002.

[103] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

[104] David Rumelhart, Geoffrey Hinton, and Ronald Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

[105] Jeffrey Schank. The development of locomotor kinematics in neonatal rats: an agent-based modeling analysis in group and individual contexts. *Journal of Theoretical Biology*, 254(4):826–842, 2008.

[106] Rachael Seidler, Jessica Bernard, Taritonye Burutolu, Brett Fling, Mark Gordon, Joseph Gwin, Youngbin Kwak, and David Lipps. Motor control and aging: Links to age-related brain structural, functional, and biochemical effects. *Neuroscience & Biobehavioral Reviews*, 34(5):721–733, 2010. ISSN 0149-7634. doi: 10.1016/j.neubiorev.2009.10.005.

[107] Burrhus Skinner. The experimental analysis of operant behavior. *Annals of the New York Academy of Sciences*, 291(1):374–385, 1977.

[108] Rajindar Sohal and Richard Weindruch. Oxidative stress, caloric restriction, and aging. *Science*, 273(5271):59–63, 1996.

[109] Marla Sokolowski. Drosophila: genetics meets behaviour. *Nature Reviews Genetics*, 2(11):879–890, 2001.

[110] Peter Spear. Neural bases of visual deficits during aging. *Vision research*, 33 (18):2589–2609, 1993.

[111] Sergey Stavisky, Jonathan Kao, Paul Nuyujukian, Chethan Pandarinath, Christine Blabe, Stephen Ryu, Leigh Hochberg, Jaimie Henderson, and Krishna Shenoy. Brain-machine interface cursor position only weakly affects monkey and human motor cortical activity in the absence of arm movements. *Scientific reports*, 8(1):1–19, 2018.

[112] Greg Stephens, Bethany Johnson-Kerner, William Bialek, and William Ryu. Dimensionality and dynamics in the behavior of c. elegans. *The Public Library of Science: Computational Biology*, 4(4):e1000028, 2008.

[113] Ulrich Stern, Ruo He, and Chung-Hui Yang. Analyzing animal behavior via classifying each video frame using convolutional neural networks. *Scientific reports*, 5(1):1–13, 2015.

[114] Ariana Strandburg-Peshkin, Damien Farine, Iain Couzin, and Margaret Crofoot. Shared decision-making drives collective movement in wild baboons. *Science*, 348(6241):1358–1361, 2015.

[115] Andrew Straw, Kristin Branson, Titus Neumann, and Michael Dickinson. Multi-camera real-time three-dimensional tracking of multiple flying animals. *Journal of The Royal Society Interface*, 8(56):395–409, 2011.

[116] DJ Strouse and David Schwab. The deterministic information bottleneck. *Neural computation*, 29(6):1611–1630, 2017. doi: {10.1162/NECO_a_00961}.

[117] David Sussillo and Omri Barak. Opening the black box: low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural computation*, 25 (3):626–649, 2013.

[118] Marc Tatar, A. Kopelman, Diana Epstein, Meng-Ping. Tu, Chih-Ming Yin, and Robert Garofalo. A Mutant Drosophila Insulin Receptor Homolog That Extends Life-Span and Impairs Neuroendocrine Function. *Science*, 292(5514): 107–110, 2001. ISSN 0036-8075. doi: 10.1126/science.1057987.

[119] Niko Tinbergen. The study of instinct, 1951.

[120] Niko Tinbergen. On aims and methods of ethology. *Zeitschrift für tierpsychologie*, 20(4):410–433, 1963.

[121] Naftali Tishby, Fernando Pereira, and William Bialek. The information bottleneck method. *arXiv preprint physics/0004057*, 2000.

[122] Richard Tolman. Statistical mechanics applied to chemical kinetics. *Journal of the American Chemical Society*, 42(12):2506–2528, 1920.

[123] John Tower. Programmed cell death in aging. *Ageing research reviews*, 23: 90–100, 2015.

[124] Sun-Ting Tsai, En-Jui Kuo, and Pratyush Tiwary. Learning molecular dynamics with simple language model built upon long short-term memory neural network. *Nature communications*, 11(1):1–11, 2020.

[125] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[126] Devin Wahl, Samantha Solon-Biet, Victoria Cogger, Luigi Fontana, Stephen Simpson, David Le Couteur, and Rosilene Ribeiro. Aging, lifestyle and dementia. *Neurobiology of disease*, 130:104481, 2019.

[127] David Wales and Jonathan Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, 1997.

[128] Shuohang Wang and Jing Jiang. Learning natural language inference with lstm. *arXiv preprint arXiv:1512.08849*, 2015.

[129] Alexander Wiltschko, Matthew Johnson, Giuliano Iurilli, Ralph Peterson, Jesse Katon, Stan Pashkovski, Victoria Abraira, Ryan Adams, and Sandeep Robert Datta. Mapping sub-second structure in mouse behavior. *Neuron*, 88(6):1121–1135, 2015.

[130] Alexander Wiltschko, Tatsuya Tsukahara, Ayman Zeine, Rockwell Anyoha, Winthrop Gillis, Jeffrey Markowitz, Ralph Peterson, Jesse Katon, Matthew Johnson, and Sandeep Robert Datta. Revealing the structure of pharmacobehavioral space through motion sequencing. *Nature neuroscience*, 23(11):1433–1443, 2020.

[131] Jason Yosinski, Jeff Clune, Anh Nguyen, Thomas Fuchs, and Hod Lipson. Understanding neural networks through deep visualization. *arXiv preprint arXiv:1506.06579*, 2015.

[132] Jacob Zahn, Rebecca Sonu, Hannes Vogel, Emily Crane, Krystyna Mazan-Mamczarz, Ralph Rabkin, Ronald Davis, Kevin Becker, Art Owen, and Stuart Kim. Transcriptional profiling of aging in human muscle reveals a common aging signature. *The Public Library of Science: Genetics*, 2(7):e115, 2006.

[133] Zheng Zhao, Weihai Chen, Xingming Wu, Peter Chen, and Jingmeng Liu. Lstm network: a deep learning approach for short-term traffic forecast. *The Institution of Engineering and Technology: Intelligent Transport Systems*, 11(2):68–75, 2017.

[134] Man Zhihong, Hong Ren Wu, and Marimuthu Palaniswami. An adaptive tracking controller using neural networks for a class of nonlinear systems. *The Institute of Electrical and Electronics Engineers: Transactions on Neural Networks*, 9(5):947–955, 1998.

[135] Walter Zucchini, David Raubenheimer, and Iain MacDonald. Modeling time series of animal behavior by means of a latent-state model with feedback. *Biometrics*, 64(3):807–815, 2008.