

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE FÍSICA



**Ciências**  
**ULisboa**

**Quantifying Joint Behavioral States in Zebrafish (*Danio rerio*)  
Dyadic Contests through Interpretable Variables**

Rafael Franco Freire Rocha Almada

**Mestrado em Engenharia Biomédica e Biofísica**

Dissertação orientada por:  
Prof. Alexandre Andrade  
Prof. Greg Stephens

# Acknowledgements

I would like to acknowledge the aid given by the Erasmus programme, giving me the opportunity to do this research at a different institution and to pursue a path in academia.

I would also like to thank the Vrije Universiteit Amsterdam and their staff for having provided a welcoming stay at their institutions for the length of the project.

I would like to thank Greg Stephens and Liam O'Shaughnessy for their orientation, guidance and inspiration given throughout the length of the project, and for giving me motivation to continue my pursuit in academia. Our nice evening chats helped me a lot, not only on the project, but also gave me a better perspective into how to be a scientist.

Liam developed the processing pipeline used to convert the experimental data into data used on the analysis conducted. His effort is why this project was possible. Working on the same project as I was but from a different perspective, Liam was the first one I would go to for help, and he was always available. Greg has a fascination with the unknown in science, and over the past year, I've developed my own fascination for it, thanks to his influence, and for that, I thank him profoundly.

I would like to thank Tatsuo Izawa for our nice chats, for the ideas we shared, for the help he gave to the project. Our results would not be possible without him, and his understanding of the experiments, his insights were helpful towards reaching the conclusions in this project.

I would also like to thank Antonio Costa, for the chats, the ideas he gave throughout the length of the project, his perspective on the project being different from the rest of the group made for a nice contrast, and helped in the consideration of alternative paths when reaching roadblocks.

I would like to extend my thanks to everyone else on the group, for our fruitful group discussions, eliciting several new ideas and learning a lot from all our different perspectives. I want to thank the group for teaching me in general that science is a collaborative process, where the discussion amongst a group helps in the understanding of a concept, but allows for its

expansion into deeper, more concrete questions.

We may have only worked for a little over an year, but more than colleagues, I gained friends that I can rely on in the future, and for that I thank you all very deeply.

Gostaria também de agradecer ao prof. Alexandre Andrade por seu apoio e esclarecimento em dúvidas, e sua acessibilidade no decorrer do projecto.

Gostaria de agradecer aos meus professores que me ensinaram e me deram uma base de conhecimento que me permitiu entrar num projeto de forma coerente, sem estar às cegas, e de me motivarem ao longo do meu percurso académico.

Gostaria de mandar um abraço enorme aos meus amigos e colegas, que mesmo longe, me forneceram companhia e diversão, uma distração bem-vinda nestas circunstâncias.

E finalmente, gostaria de deixar um agradecimento enorme aos meus pais e avós que sempre estiveram a cuidar de mim, mesmo à distância, mesmo durante uma pandemia que ocorre uma vez a cada século, sempre a acreditarem que eu iria suceder nos meus projetos, e dar-me o seu apoio emocional.

*“Don’t compete! — competition is always injurious to the species, and you have plenty of resources to avoid it!”*

Pyotr Kropotkin, *Mutual Aid: A Factor of Evolution*

# Summary (English)

Animal behavior is a fascinating area from a physical perspective, and yet there are several challenges in making physical models about it. A specific challenge is in coming up with models derived from data, removing human-centric bias when defining behaviors. A good example of the complexity associated with it can be seen in animal contests, such as in zebrafish (*Danio rerio*) fights. Being well understood, makes it a good first example to explore these complex interactions. Our main objective consists of attempting to derive a set of behavioral states directly from the data.

We define these states for the pair, rather than for each fish individually, based on our assumption that social interactions are joint behaviors and cannot be fully described for an individual element of the interaction. We do so through the definition of simple variables with analogues in classical physical models. These variables are respectively distance, alignment of the heading and acceleration vectors, the average and deviation of the tail beating rate, and speed, and these form a 6 dimensional system. We define compound behavioral states in function of the variables defined previously. We embed the 6-dimensional system, after a normalization process in a 2-dimensional representation through a nonlinear embedding technique. We assign labels to the different density peaks through a user-defined classification tree. We build a symbolic sequence from the density peaks, and do spectral analysis on the computed transition matrix associated with the sequence, to have a first picture of the dynamics of the interaction.

We are able to recover a connection between different clusters with different behaviors, forming “super-clusters” and giving an overall structure to the fight. Spectral analysis also suggests the dynamics between these larger clusters is non-markovian and occurs in longer timescales, relative to behaviors within clusters.

These results suggest that it is possible to recover the behavioral structure of a contest from simple variables, and there are dynamical features we can recover from data. This opens up opportunities to explore dynamics by using less interpretable and more accurate variables.

**Keywords**— Behavior, *Danio rerio*, Clustering, Methods

# Resumo (Português)

O comportamento animal é uma área fascinante do ponto de vista físico, no entanto ainda existem vários desafios associados à construção de modelos ou ao desenvolvimento de teorias do comportamento em física. Um dos desafios é desenvolver modelos diretamente dos dados, eliminando o viés antropocêntrico que existe na definição de estados comportamentais. Um bom exemplo da complexidade associada ao comportamento pode ser encontrado em interações sociais, nomeadamente interações agonistas entre peixes-zebra (*Danio rerio*). Estas interações são bem compreendidas e estereotípicas, e existem catálogos a descrever os estados comportamentais associados a cada fase da interação. Isto e a versatilidade genética a que o peixe zebra se encontra associado, tornam esta interação ideal para o nosso estudo.

O nosso objetivo principal consiste na tentativa de derivar um conjunto de estados comportamentais diretamente a partir dos dados experimentais obtidos, sendo estes estados definidos para o conjunto, e não individualmente. Fazemos isso sob a assunção de em interações sociais, estados comportamentais dependem dos elementos envolvidos nessa interação (neste caso, são peixes-zebra) e que esta não é completamente descrita, exceto se levar ambos em conta simultaneamente.

Os dados são esqueletos tridimensionais dos 2 peixes-zebra num volume. O processo de aquisição desses dados consiste na aquisição de imagens em 3 planos bidimensionais com câmaras de alta definição, e um pipeline de processamento, que combina várias redes neuronais para a identificação de pontos corporais, a atribuição de identidade temporal aos peixes envolvidos, e a interpolação das imagens nos diferentes planos. Este processo permite a conversão de vários vídeos em sinais temporais que podem ser manipulados e processados de forma adequada.

Usamos variáveis interpretáveis, no caso, a distância, os alinhamentos de direcção e aceleração, e os ritmos de batimento de cauda. Essas variáveis embora sejam simples, podem dizer bastante informação sobre a natureza do comportamento, sendo úteis numa exploração inicial.

Definimos estados comportamentais compostos (colecção de vários comportamentos efec-

tuados pelos peixes ao longo de um determinado período de tempo) e exploramos a dinâmica de uma luta nesta descrição simplificada.

O sistema que resulta das variáveis definidas possui 6 dimensões, projectamos esse sistema para um plano bidimensional para melhor análise.

Efectua-se um histograma das novas variáveis, e ter uma estimativa da densidade de probabilidade através da convolução do mesmo com uma gaussiana bidimensional. Detecta-se os picos de densidade, que neste sistema podem ser interpretados como estados comportamentais.

Com essa descrição é possível gerar uma sequência simbólica que representa a dinâmica da interacção como sendo a transição entre vários estados comportamentais discretos. Constrói-se uma matriz que representa a transição entre os vários estados, e por decomposição espectral pode-se observar o comportamento dos valores próprios em função do número de transições e é possível decompor os estados em vários conjuntos através dos vectores próprios, cuja dinâmica entre eles é representada pelo valor próprio associado. Através da sequência simbólica é possível uma descrição da interacção entre os elementos, tendo inclusive informação sobre a escala temporal associada à dinâmica entre esses estados.

Ao associar os clusters aos diferentes estados comportamentais compostos definidos previamente, é possível ver que certos clusters se encontram associados, e apresenta uma certa estrutura, que pode ser representativa da dinâmica real.

Também é possível determinar a escala temporal de interações entre diferentes conjuntos de clusters. Foi possível determinar que os comportamentos ocorrem em escalas temporais maiores do que a escala típica para processo de Markov, e a escala temporal mais elevada se encontra associada a transição entre estados associados à agressão entre o par, e estados associados aos períodos entre lutas.

Mostramos que é possível obter uma estrutura comportamental da luta entre dois peixes-zebra utilizando as variáveis simples que definimos. Isto é um framework que permite explorar a dinâmica da sua interacção em maior detalhe, a utilizar variáveis ou representações mais precisas, que podem não ser interpretáveis.

**Keywords**— Comportamento, *Danio rerio*, Clustering, Métodos



# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Summary (English)</b>	<b>iv</b>
<b>Resumo (Português)</b>	<b>vi</b>
<b>List of Figures</b>	<b>x</b>
<b>List of Tables</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Quantifying Natural Behavior . . . . .	1
1.2 Animal contests . . . . .	10
1.2.1 Zebrafish dyadic contests . . . . .	12
1.3 Timeline . . . . .	13
1.4 Objectives . . . . .	14
<b>2 Methods</b>	<b>15</b>
2.1 Experimental setting . . . . .	15
2.1.1 Understanding the development pipeline . . . . .	15
2.1.2 Data Analysis and Processing . . . . .	17
2.2 Joint Variables . . . . .	19
2.2.1 Notation and Definition of Variables . . . . .	20
2.3 Compound Behavioral States . . . . .	24
2.4 2D Projection and Clustering . . . . .	27
2.4.1 UMAP Embedding . . . . .	27
2.4.2 Kernel density estimate . . . . .	28
2.4.3 Peak Detection and Clustering . . . . .	29
2.4.4 Ethogram Computation and Symbolic representation . . . . .	31

2.5 Behavioral Labelling . . . . .	32
2.6 Transition Matrix and Spectral Analysis . . . . .	32
2.7 Supplementary Material . . . . .	34
<b>3 Results and Discussion</b>	<b>35</b>
3.1 Results . . . . .	35
3.1.1 Kinematic description of a contest . . . . .	35
3.1.2 Behavioral Map . . . . .	36
3.1.3 Symbolic and Spectral Analysis . . . . .	38
3.2 Discussion . . . . .	41
3.2.1 For future projects . . . . .	44
<b>4 Conclusion</b>	<b>46</b>
4.1 Appendix - Recovering Almost Invariant Sets in Dynamical Systems . . . . .	48
4.1.1 Lorenz system . . . . .	49
4.1.2 Chua's circuit system . . . . .	52
4.1.3 Discussion . . . . .	54
<b>Bibliography</b>	<b>55</b>

# List of Figures

1.1	Development of a Behavioral Model	3
1.2	Densities in behavioral space during stereotyped (a) and non-stereotyped (b) behavior	6
1.3	<i>C. Elegans</i> ' motility, described by three quantities: speed $s$ , body orientation $\psi$ and alignment $\Delta\psi$	7
1.4	Sample tracking results compared to the original images for a turning motion	8
1.5	Zebrafish contest ethogram with transitions	12
1.6	Using computer vision and neural networks to detect zebrafish attacks from variables.	13
2.1	Experimental setup used: A $30 \times 30 \times 30$ cm <sup>3</sup> tank with three cameras oriented in perpendicular fashion, with planar LED backlights in the opposite side of the tank.	16
2.2	Processing pipeline to recover 3D postural information from a pair	18
2.3	Classifier used for Behavioral Labelling	32
2.4	Analysis pipeline from the derived variables to the results.	34
3.1	Average distance over a fighting experiment, with the highlighted areas corresponding respectively to a time when the fish are fighting and after the fight has ended.	35
3.2	Distribution of joint variables in two intervals of a fight, during a fight and after the fight.	36
3.3	Left - An estimate of the 2D probability density function for the two embedded variables. The width (standard deviation) of the gaussian kernel used was 10. The variables were obtained through UMAP embedding. Right - Clusters around density peaks, recovered through the Hessian equations, and connected components algorithm. The density peaks may be different behavioral states.	37
3.4	A fight "ethogram" based on the clusters recovered from the KDE estimate.	37
3.5	Averages of the joint variables for each cluster.	38
3.6	Cluster Labelling and association to compound behavioral states through a thresholding process	38
3.7	Transition matrix built from the symbolic representation of a fight	39

3.8	The 3 largest eigenvalues below 1 as a function of the transitions.	39
3.9	Values of the eigenmodes associated with the second, third and fourth largest eigenvalues.	40
3.10	Cluster projection of eigenvectors for the eigenvalues respectively associated.	41
4.1	Representation of three almost-invariant sets	49
4.2	KDE of the XZ projection for the Lorenz System, and the recovered 2 dimensional projection	50
4.3	Time series of the X coordinate in the Lorenz system; "Ethogram" of the UMAP- embedded Lorenz system	50
4.4	Eigenmode projection of the Lorenz System for the second largest eigenvalue	51
4.5	Eigenmode projection of the Lorenz System for the fourth largest eigenvalue	51
4.6	KDE of the XZ projection for the Chua's System, and the recovered 2 dimensional projection	52
4.7	Time series of the Z coordinate in the Chua's system	53
4.8	"Ethogram" of the UMAP-embedded Chua's system	53
4.9	Eigenmode projection of the Chua's System for the fourth largest eigenvalue	53

# List of Tables

2.1 Datasets used in analysis . . . . .	18
3.1 Test of distribution differences, during and after a fight section . . . . .	36

# Chapter 1

## Introduction

Ethology is a field in biology that studies animal behaviour. For a long period of time, it has been a qualitative science, based in direct observation by humans, classifying behaviour in classes that are determined by the observer. This may introduce several possible errors: the results may be biased by the observer; possible intrinsic properties of behaviour may be ignored; and since the observer classifies the behavior statically, the whole dynamic of the system is incomplete. In order for ethology to become a quantitative science, there needs to be a change in the method of observation. [1] Recent developments in machine learning and image tracking have led to the development of what is known as computational ethology, which allows for quantitative analysis and physical interpretation of behaviour. [1] There are examples of this, for physical modelling of collective motion using fluid dynamics [2] or to model behavioral patterns in single animals using random walk and stochastic processes [3]. It's possible to combine the quantitative behavioral model with neural or genetic models to address more interesting questions about the nature of behavior. [1] And yet, there are several questions that have not been addressed in ethology from a physical perspective, and a quantitative description may be helpful into ultimately deriving a physical theory of behavior [4].

### 1.1 Quantifying Natural Behavior

At first, several cases are presented as examples where computational methods and physical theory are used to describe certain behavioral processes. These methods are reliant on the experimental procedures used. A experimental setting may have a controlled environment where a constrained set of behaviors are observed. [5] A different approach is to focus on measuring behavior on the organism scale rather than focusing on the environment with minimal artificial constraints. There are several experimental methods possible as:

- Posture tracking - Behavioral patterns can be represented by postural changes over time. An image directly displays animal behavior which can be later converted into a more abstract representation. It used to be done by *manual frame by frame analysis* of video footage, which is ineffective for long high framerate videos. In recent years, several techniques have been developed for better and more accurate tracking.
  - **Stopgap techniques** are inspired by previously developed human motion capture, where small dye spots are introduced on the animal's limbs. It may introduce possible confounds, and the system may be biased towards human-like patterns. [6]
  - **Markerless techniques** addresses previous difficulties by using internal reflections to characterize contact points between an animal and a substrate. However it can't track portions of the animal not interacting with or separated from the substrate. [7]
  - **Model-based image analysis** uses known biomechanical degrees of freedom to find the posture that would produce images with the highest likelihood. [8]
  - Besides these techniques, another possibility is to do automatic annotations through deep learning and neural networks, which would still require human-annotated training sets. [9]
- Field imaging - Recent development of cheap equipments and open source animal tracking algorithms have made this option increasingly popular. [10]
- Real-time feedback and brain imaging - it varies in computational intensity, with computationally lighter approaches being implemented in real time. Observations can be done in two fronts, stimulating the response of behavior by neuronal input, or predicting upcoming behaviors through neuronal signals. There are challenges associated with this method, mainly in recording animals for long periods of time, and in registering volumetric neural imaging data. [1]

The resulting data can be described in a theory with basis in consistency, fidelity, interpretability and scalability. Fidelity and interpretability are in opposition, as models that are more easily interpreted tend to be less accurate. This suggests a continuum of solutions, typically dependent on the researcher's representation of behavior, and selecting that representation is dependent on the interests of the research: [4]

- **Quantification associated with the experimental apparatus** - having the advantage of consistency but being too low-dimensional and having an unnatural dynamic. [11]
- A coarse variable like **mean velocity** can also be measured, being more naturalistic, however that only allows for dynamics at a single scale, not being able to capture more precise patterns of behavior that may occur at different scales. [12]

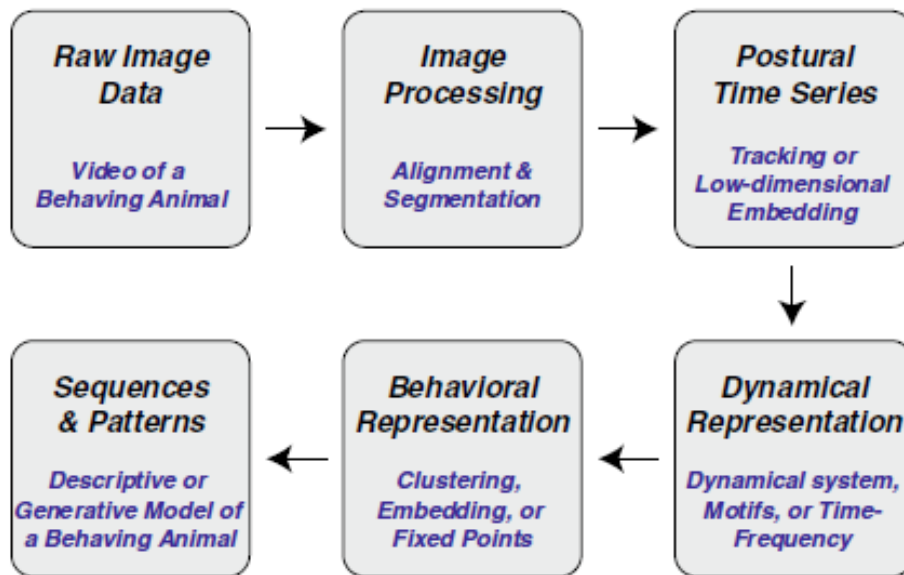


Figure 1.1: Development of a Behavioral Model (from [4])

- **Human-defined classification system by a trained observer** - This approach as a more detailed description, it is labor intensive and user-specific (the behavioral classes are defined intuitively), limiting reproducibility. Finally, it assumes behavior states are part of a discrete system of states and the data doesn't seem to show this as a good representation. [4]
- Advances in automation have increased the throughput of the first two options, and **supervised machine learning techniques** have improved the repeatability and decreased the manual labor required for human-defined classification.

But the fundamental difficulties presented by each of these cases still remain. There is a need for a general principle and approach towards behavioral classification. [4]

Bio-locomotion is an area where the focus to study behavior has been directed. Usually behavior is studied through dynamic motion trajectories, like center-mass, body-bending and/or limb trajectories. These behaviors are stereotyped, since physical constrains allow for fewer degrees of freedom in movement execution, and during locomotion, their dynamics are in a lower dimension than the number of degrees they have available. [13] Based on these observations, recent developments have been pushed for **data-driven** and **unsupervised analysis** of animal behavior, showing that a large fraction of animal movements are low dimensional compared to their degrees of freedom. [14] There are different approaches that have several similarities between them used to identify stereotyped behavior from videos (not considering other modalities). The general structure is described in fig [1.1].

The first step after data collection is to define a postural time series. Several principles may be considered when dealing with postural time series.

- **Stereotypy and discretization** - behavior is treated as a sequence of discrete states. The similarity



between instances of a state in that sequence is what is defined as stereotypy, and it can be quantified by determining how long the trajectory in posture phase space diverge. [15] Non-stereotyped behavior is considered to be part of a continuous transition between the states [16].

- **Hierarchical organization** - While it has been shown as being very effective to produce complex adaptive behaviour in robotics [17] with little computational power, there is a lack of clear demonstrations of such hierarchical organization in non-human animals. But it is still widely used as a principle when considering a representation of behavior. [18] These can be constructed as nested representations with multiple levels of nesting, using dictionary based compression algorithms [19] or cluster analysis [20].
- **Low postural dynamics** - animal behavior generally has many degrees of freedom, however when considering correlations and independence between these degrees allows for a lower dimensional representation that maintains the information about the posture. This holds not only for simple animal structures but for more complex ones.

Low dimensionality, stereotypy and hierarchy can all emerge in postural phase space. [21]

The main objective is to turn the high-dimensional data (millions of pixels) into a lower-dimensional representation that can describe the animal's posture in its entirety. The number of variables required are different for different animals, however, it is small compared to the dimension of the acquisition data. That can be done directly through posture tracking and or indirectly by image segmentation and principal component analysis for dimensionality reduction.

To measure stereotyped behaviors, we need to create a dynamical representation that describes how the postural time series are changing. This can be achieved by fitting a **differential equation** to the data or through features that incorporate dynamics to segments of data. An appropriate dynamical representation can emerge from this, describing the postural time series. [22], [21]

Another approach is to fit a **statistical model** to the data, for example by an Auto Regressive-Hidden Markov Model. This would allow for a dynamical and behavioral representation, but it requires a parameter that determines the timescale for behavioral states, and it is known that behaviors may occur in different time scales [23].

To complement this, some representations are multi-scalar in nature, through **motifs** (posture patterns that occur commonly) of various lengths [24] or using a **time-frequency analysis** approach [16], showing behavior at multiple scales.

Capturing motifs may not be robust to slight changes in postural dynamics, like changes in frequency. While time-frequency analysis gives information in both amplitude and phase, typically only uses amplitude, which helps with the robustness, and the behavioral representation can be obtained through

**clustering** [25], or **low-dimensional embedding** of the resulting vector of amplitudes [16], [26]. The embedding generates a space with anisotropic density, and the behavioral representation is either the density or the peaks.

The several representations discussed above can be grouped into being either discrete or continuous. In general, discrete representations can be derived from continuous ones.

**Continuous representations** may be interesting because they allow to display non-stereotyped behavioral dynamics.

If the data has clusters, however, it is ideal for clustering in **discrete representations** [25]. While statistical methods allow for the development of a dynamical model, it works under an assumption on the timescale for the data.

We can look at several cases where the different methods discussed were explored and see what conclusions they have reached.

Lee et al. did a series of tests in zebrafish to determine their orientation mechanisms, based in direct observation of video footage and analysis (statistical and Bayesian). The results showed that they navigate by determining directions based on distance relationships between visible surfaces. They use the perceive distance between two 3-d surfaces to determine their relative position and direction. [27]

In a follow-up study, they reach the conclusion that zebrafish rely mostly on boundaries to orientation and spatial mapping, and other features as local cues towards a certain location. This suggests a similar neural structure and function across several species. Environmental boundaries may play a major role in mapping location in an environment for other animals, and these results suggest that the mechanism behind this are related. [28]

Berman et al. used time-frequency analysis (through Morlet wavelet analysis [29]) and spatial embedding (through t-distributed stochastic neighbour embedding [30]) to develop representations for stereotyped behaviors in flies (*Drosophila Melanogaster*). This allowed for a low-dimensional representation where stereotyped behaviors could be described by peaks in density. The results they have suggest that about half the behaviors displayed by the flies are non-stereotyped. This method also allows to find subtle behavioral sex-specific differences. [16]

*Caenorhabditis Elegans*, or *C. elegans* in short, is a nematode that is very used in the development of behavioral models, not only for the simplicity of its structure but also because it is the first multicellular organism with both a complete genome [31] and connectome [32], which allows for studies not only at the genetic level, but at the neuronal level in conjunction with behavioral models.

Baek et al. developed an automated tracking and imaging system and used acquired image data to quantify the motion patterns for wild and mutant worms. They use image segmentation to have a binary mask which they use to develop a skeleton of the image (through thinning and pruning of the

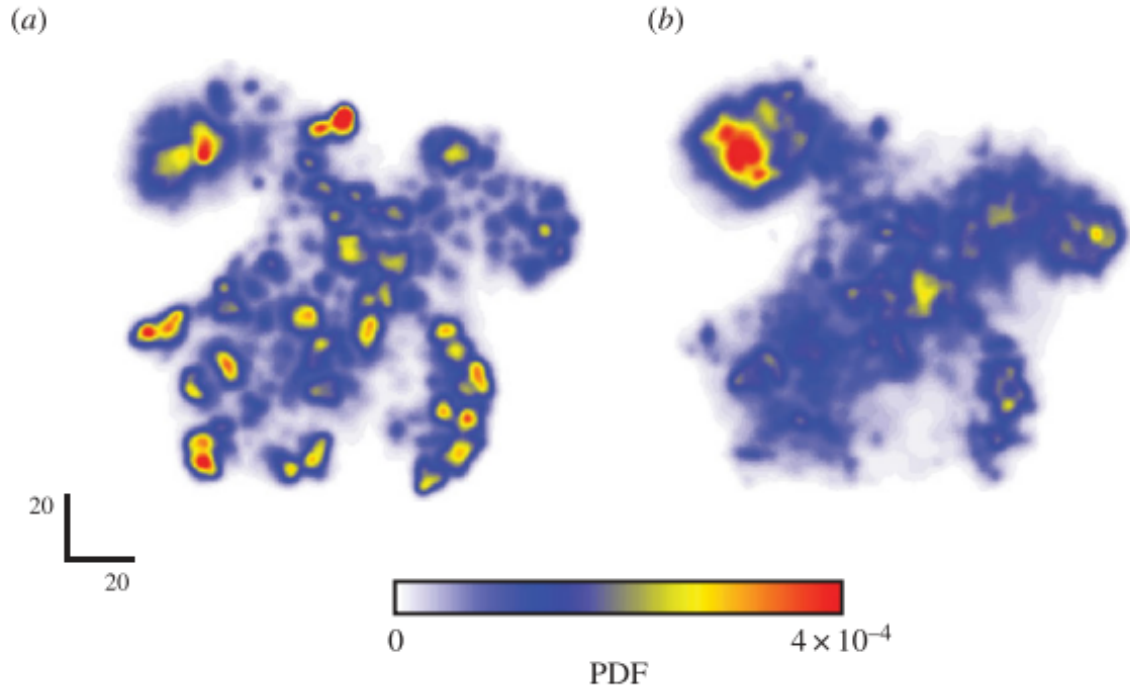


Figure 1.2: Densities in behavioral space during stereotyped (a) and non-stereotyped (b) behavior (from [16])

image [33]), being an easier structure for feature extraction. They then use the extracted features to train a classification and regression tree analysis [34]. The results show that mutant animals can be identifiable based on a small number of image features collected. Also, these methods have the advantage of giving the possibility to determine phenotypes in a more quantitative, objective manner, more reliable than previous methods. Skeleton decomposition is now a standard for *C. Elegans* behavioral analysis. [35]

Sznitman et al. used image processing and tracking to develop a skeleton representation, and combine the kinematic data with a linearized model to study the two dimensional motion of *C. Elegans* at low Reynolds regimes. The model used was a linearized model (a slender filament at low  $Re$ ) for the torque of the system with the momenta based in the Voigt model [36], and by some approximations, simplify the model, giving origin to a biharmonic equation of the form:

$$\frac{\partial y^4}{\partial x^4} + \xi \frac{\partial y}{\partial t} \quad (1.1)$$

where  $y(x, t)$  is the body displacement of the nematode,  $t$  is the time,  $x$  is the position in the  $x$  axis and  $\xi$  is a constant that depends on the nematode's material properties and the fluid's drag coefficient. Equation [1.1] can be solved analytically, and fitting the data to this model makes it possible to determine the material coefficients describing the physical motion. The results show there is a periodic swimming behavior that propagates from the nematode's head to its tail, being able to estimate the nematode's tissue material properties. Due to its non-invasive nature, it is an ideal method for genetic and small molecule

screening applications. [37]

Keaveny et al. had a similar study, wanting to better understand the mechanical relationship between animal and its surrounding environment, and to develop a model to predict the trajectories of *C. elegans*. They used a linear model for the forces applied on the worm, similar to the one used by Sznitman et al., and a nonlinear variant [38] of the model that can be reduced to the linear model with the appropriate parameters. Linear resistive force theory proved itself effective in predicting the nematode's path based on a sequence of body postures. Differences were found between nematodes captured in the wild, and those grown in the lab, when considering the drag anisotropies. The nonlinear model, while providing good predictions, does not solve the discrepancy between the generated parameters and experimentally measured ones. The linear resistive model can be used for applications as whole-animal simulations and advanced tracking, but the nature of interaction between the nematode and its lab environment remains unsolved. [39]

Helms et. al investigate the motile behavior of nematodes and possible variations due to several factors. Focusing on trajectory dynamics on timescales spanning from ballistic to diffusive movement, they find that features of the motility statistics are captured by a random walk model with independent dynamics in the speed (defined as an Ornstein-Uhlenbeck process [40]), turning (determined by body orientation, as a drift and diffusion process [40]) and reversal events (determined by alignment, a random telegraph process [41]).

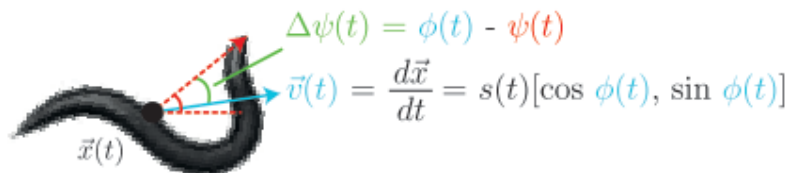


Figure 1.3: *C. Elegans*' motility, described by three quantities: speed  $s$ , body orientation  $\psi$  and alignment  $\Delta\psi$ (from [3])

The parameters change across species in a correlated low-dimensional manner suggesting a common mode of behavioral control and a trade-off between exploration and exploitation. The distribution of phenotypes shows there are behavioral differences across strains, suggesting they employ contrasting bet-hedging strategies for foraging. [3]

Stephens et al. extracted curvature (using Frenet equations [42]) that describes a posture at a low dimensional space than the motion. Similar results have been obtained for motor control in humans. Their results showed the posture as a combination of small set of eigenworms (template motor behavior), whose coefficients vary continuously. These are not correlated linearly but are nonlinear and coupled. It also shows that some responses are stochastic due to not considering other behavioral variables, and that said inclusion may uncover deterministic responses. With the eigenworms a phase space is derived, and

there equations of motion can be built (using the Langevin equation [43]) where the set of attractors can be described as rigorous behavioural states. This allows for a quantitative description of worm behavior and shows for evidence of more subtle behavioral states. [21]

They exploit the reduced space of *C. elegans* postures in a following article to develop a novel tracking algorithm which captures both simple shapes and self-occluding coil, which have yet to be studied in 2D worm behavior. They apply previously developed algorithms to show that the complex coiled shapes are a superposition of two simpler patterns.



Figure 1.4: Sample tracking results compared to the original images for a turning motion(from [44])

They discover a dichotomy in spontaneous, large-amplitude coils by studying the escape response of the nematodes, showing that these reorientations not only occur at the classical Omega postures but also larger delta-postures, that occur independently from one another, suggesting a distinct triggering mechanism, being the analog of a left-right step. These turns occur with approximately equal rate and adapt to environmental conditions in similar timescales. [44]

In Costa et al., they use adaptive locally linear analysis [45] based on a likelihood-based hierarchical clustering using Ward's minimum variance criterion [46], which simplifies the complex global dynamics into simpler local dynamics with linearization, and find the local eigenvalues, applying it to *C. elegans* postural dynamics, and whole brain imaging. The full spectrum of eigenvalues, not only gives information about oscillatory behavior but also stability and criticality [47]. The results here suggest the behavioral repertoire is far more complicated than the canonical representation of forward, reversal and turn. Tests in brain dynamics show that the global brain dynamics are damped away by a decrease in oxygen in instability boundary. This shows interesting dynamics to be studied at a critical region [48]

From these different results we can see several approaches being explored with varying results, giving insight onto the nature of behavior. One key point to take however is that none of these approaches is necessarily the best. They give access to different forms of information that may lead to different questions. It is likely this project will be the same, not giving the best approach, but a different approach to tackling behavior. It doesn't seem we are close to having a best approach, but this contributes to the

wealth of information about behavior.

## 1.2 Animal contests

Within the vast array of social interactions in animals, contests are of particular interest, due to their asymmetric nature. The outcome of a fight will have a winner and a loser, two different behavioral states for the animals involved. Animal contests have passive and active approaches, and in some cases a history of these contests involves a combination of these, being therefore of critical importance to our understanding of contests in animals to understand said approaches. Game theory provides several frameworks from which to look at these interactions, considering the interplay between the outcome of a contest and a contestant's ability or cost. Three particular models have been extensively used to describe contests.

- Sequential assessment - in the SA model, the contest is seen as a consistent sequence of behaviors, where each behavior is executed with the purpose of obtaining the maximum amount of information about an opponent fighting ability (referred to as resource holding potential or RHP), with the least amount of cost, and the interactions between the contestants allow a better estimation of the opponent's RHP over time (this implies a minimization of the cost/information ratio). There are some assumptions made by this model, like the assumption that the contestant doesn't make decisions based on past interactions. And this model predicts that the outcome of a contest is determined by the asymmetry in the contestants RHP, and that there is no escalation within behavioral stages. [49]
- War of attrition - WOA models are self assessment models, since it assumes the contestants don't take into account the opponents abilities into their decisions. It takes into account the contestant's endurance, and the length of interaction. The behaviour is determined based on the individual cost they have, and the outcome of the interaction is determined by the maximization of the pay-off (obtained the most resource value while spending the least effort.) This cost is dependent on both energy spent by the contestant and the length of the interaction, and has potential for escalation. [50]
- Cumulative assessment - in CA models, the outcome of a fight is based on a cumulative sum of a contestant's adversaries actions, and the physical cost those incur, with the fight ending after some threshold is crossed. The contestant is unable to detect the effect of its attacks on its opponents, but is able to distinguish fighting or fleeing scenarios. [51]

Overall it is possible to distinguish each model based on some characteristics, with CA and SA relying on an initial asymmetry of the contestants RHP to determine the outcome of a fight, while for a WoA model, the contestants start at a similar fighting ability. CA has an escalation within its behavior stages, something not predicted by SA models [52]. There are several experimental cases where these

models fit the description of animal fights, but there is no universal model that accurately describes an animal contest [53]. Part of the challenge may be due to the lack of proper data in said contests, and hopefully advances in tracking algorithms and data analysis will aid the development of new models or the exploration of these models from different perspectives.



## 1.2.1 Zebrafish dyadic contests

Zebrafish (*Danio rerio*) are another species that are models in ethology due to their extensive use in biomedical research and their genetic versatility, as well as neurological species (mostly on larval zebrafish that are transparent). It also has a consistent behavioral terminology, making observations more consistent. [54]

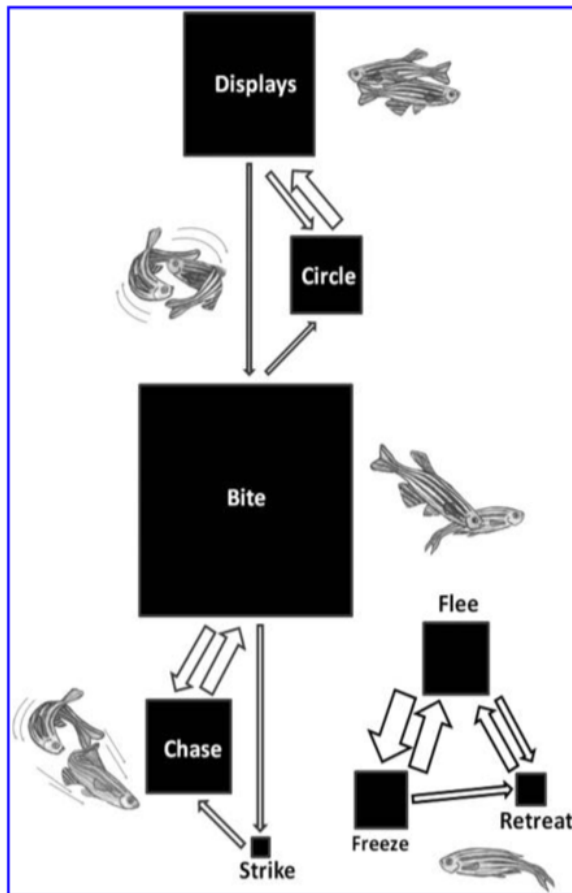


Figure 1.5: Zebrafish contest ethogram with transitions (from [55])

Zebrafish fights are a sequence of well characterized behaviors with the purpose of establishing dominance between the fish. The behavioral categories are classified into groups as displays, physical attacks, chases, and freezing (see fig 1.5). And the progression of the fight usually follow a transitions across these different categories, as shown by [55].

Due to its versatility in lab environment Zebrafish is an ideal animal model to be explored. In particular, due to its stereotyped fighting behavior being well established, it is a good set of behavioral classes to be explored quantitatively. And being an animal model in genetics means it is possible to see how this behavior is affected by mutations in particular genes [56]. This introduces questions over the robustness of this behavior to genetic modifications and to environmental conditions and it is an opening into the exploration of collective behavior in small groups of animals.

There are examples of modern day tools being used to describe zebrafish fights: Applying computer vision, it is possible to detect zebrafish trajectories automatically, preserving identity [57]. From the resulting trajectories, a neural network is trained to detect aggressive interactions between the pair. From the results, there is evidence suggesting motor coordination between the pair, giving rise to two distinct phases of aggressive interaction, a symmetric and asymmetric phases, correspondent to a contest and resolution phases. Observations also suggest a CA model is more adequate to describe the fighting interaction, relative to the other examples. [58].

A more in-depth description of the fight (in kinematic terms) will be explored further in this document.

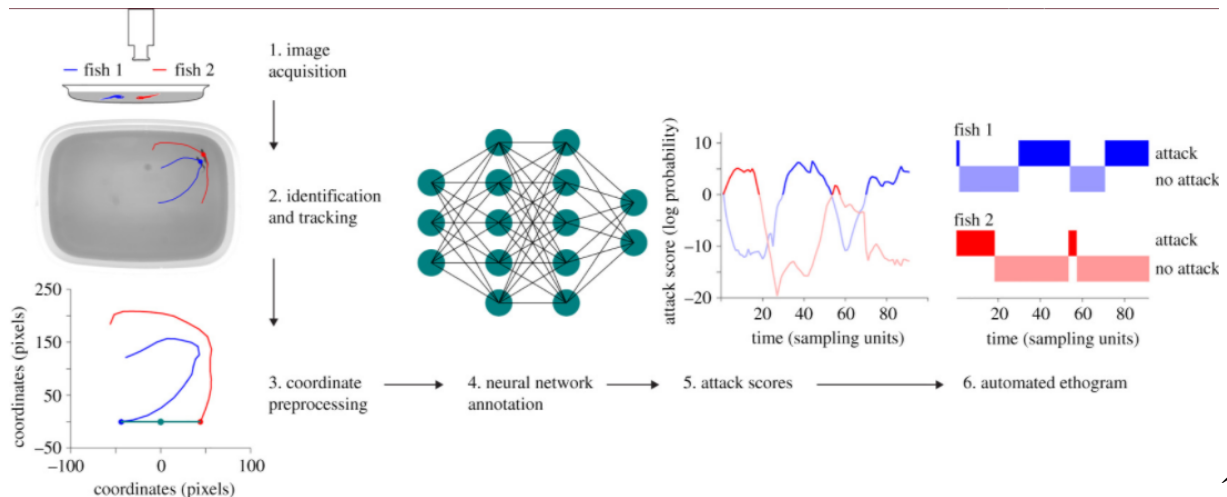


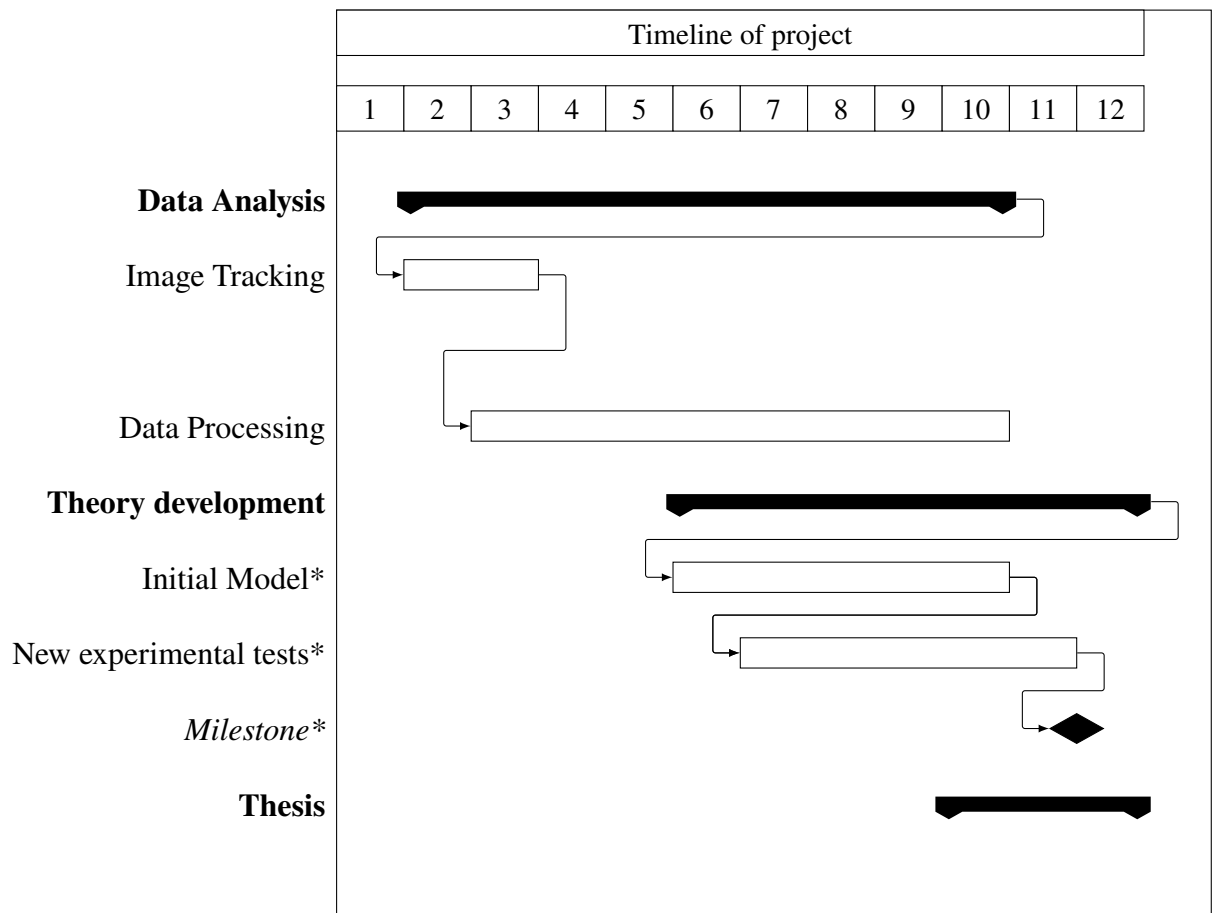
Figure 1.6: Using computer vision and neural networks to detect zebrafish attacks from variables. Source: [58]

### 1.3 Timeline

The project began in February 2020, and lasted 12 months, having all definitive results by February 2021. In the first month, the focus was directed at data analysis, following which we developed some models to describe the observed results.

Different experiments were conducted, to determine the robustness of the developed model.

In the beginning, this was the main pathway expected during the development of this project, however there were some changes, due to the CoVID-19 pandemic outbreak.



\*subject to changes

## 1.4 Objectives

As shown in this plan, there is a need to quantify behavior in order to develop mathematical models that can be better conjoined with already existent genetic and neuronal results for a more holistic description of animal behavior. There are several methodologies in experimental procedures and model developments that can be useful in different fronts and subjects, all tackling different questions, with several properties emerging depending on the method used. Choosing the most adequate approach is a challenge in both ethology and theoretical biophysics and we attempt an approach that is simplified and yet quite informative on the nature of a fight.

Our main objective is to find a description of the fight that is ethologically useful directly from the variables. In our specific case, we want to find a description from variables defined for the pair, rather than from individual fish behavior. And we want to do so from variables that are interpretable, as these allow for a first analysis and is of easy translation to an ethological description.

# Chapter 2

## Methods

### 2.1 Experimental setting

Although strictly experimental work won't be the focus of this dissertation plan, it is important to point out the details that may be relevant for data analysis and model development.

The experiments are carried out by part of the group at OIST (Okinawa Institute of Science and Technology) Graduate University, in Okinawa, Japan, which is working jointly with the group at VU Amsterdam and several others.

We did the experiment using zebrafish pairs of the wild-type variety. The observations are done in a cube of  $30\text{ cm}^3$  with three cameras (Chameleon3 USB3 Model CM3-U3-13Y3M-S-BD) perpendicular to the 3 cartesian planes (x,y and z) and at the opposite side of each camera there is a LED plane that covers the face of the cube to ensure that the luminosity is uniform (it helps in image segmentation). The output are videos from the different planes, with resolution of  $1024 \times 1280$  pixels.

#### 2.1.1 Understanding the development pipeline

As this is experimental data where no model is developed, the best procedure is to take simple analytical measures that are robust to some of the tracking errors.

Therefore it is important to understand the 3D tracked data itself, and the pipeline developed to obtain it. This pipeline has been developed over the past few years, and it has been improved.

The trajectories obtained come from images taped of a cubic fish tank. 3 cameras simultaneously capture footage from three perpendicular directions.

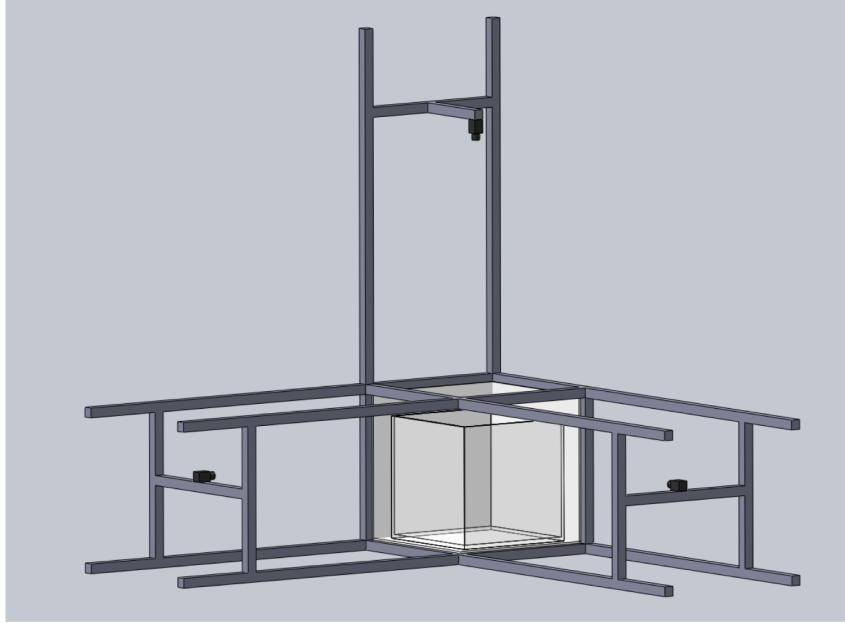


Figure 2.1: Experimental setup used: A  $30 \times 30 \times 30 \text{ cm}^3$  tank with three cameras oriented in perpendicular fashion, with planar LED backlights in the opposite side of the tank.

#### 2.1.1.1 Automatic Label Assignment

These footages are cropped and pass through a supervised convolutional neural network (CNN). Each CNN is trained for each particular direction, and for every frame, there can be one or two images per frame (depending on the distance between the fish). The outputs of the networks are skeletonized versions of the fish posture (3 Bodypoints and the edges connecting them). The process through which these are defined is through part affinity fields, detailed in [59]. A network application to determine postures in single frames is discussed in [60].

#### 2.1.1.2 Cross-camera identity assignment

To assign fish identity across different planes, a process was developed by the group where skeletons from two planes are used to compute the third one, and difference between the prediction and the network results are compared. The network result with the smallest error is chosen as the fish correspondent to the skeletons of the other frames, this error being a sum of euclidean distances between the three bodypoints. This process is repeated for all three planes, meaning that the process is a permutation across the different planes. The permutation with the lowest mean error is chosen.

### 2.1.1.3 Temporal identity assignment

To maintain temporal identity across the movie, idTrackerAI, an algorithm developed for multi animal tracking is used in the XY plane. [57]

The end result is a collection of 3D trajectories for 3 bodypoints for 2 fishes. However, it is not a perfect process. Errors in temporal assignment can happen and the calibration process that leads to the combination of 3 2D images to get one 3D coordinate does have some errors. The network itself has some associated predictions errors (since it does not give a bodypoint necessarily, but a 2D gaussian distribution around the location of said bodypoint). When those errors are combined across different steps of the pipeline, they end up multiplying. Thus far, a possible solution for this is to just ignore any frame where certain criteria are not met (for instance, there only being one bodypoint across all three camera views, or the errors in third step being higher than a certain threshold).

Two problems were a major concern for our analytical efforts.

- Fish swapping errors - While idTracker is great to assign temporal identity, it is not without flaws, specially when fish are close to each other. While these don't occur too frequently, the temporal integrity of the datasets are compromised and analytical approaches that rely on it, are somewhat affected by it. While we don't know how much it is affected, we would rather use techniques not affected by said problem.
- Missing frames - Due to a lack of training data from some regions in the tank, the network may fail to properly label the fishes, leading to missing frames, which also causes challenges for methods that require temporal identity to be consistent. While Kalman filtering has been used, it was found to be ineffective when dealing with a large number of consecutive missing frames, or large gaps. The best solution for this problem was selecting more training data from the correspondent regions in the tank. There were still some missing section remaining but their number was reduced enough that interpolation was a feasible tool to use.

For the rest of the analysis we used data from 6 experiments, involving two fish pairs:

## 2.1.2 Data Analysis and Processing

When it comes to data analysis, we used a hybrid approach for quantification based in some coarse variables and posture tracking with unsupervised machine learning techniques for image

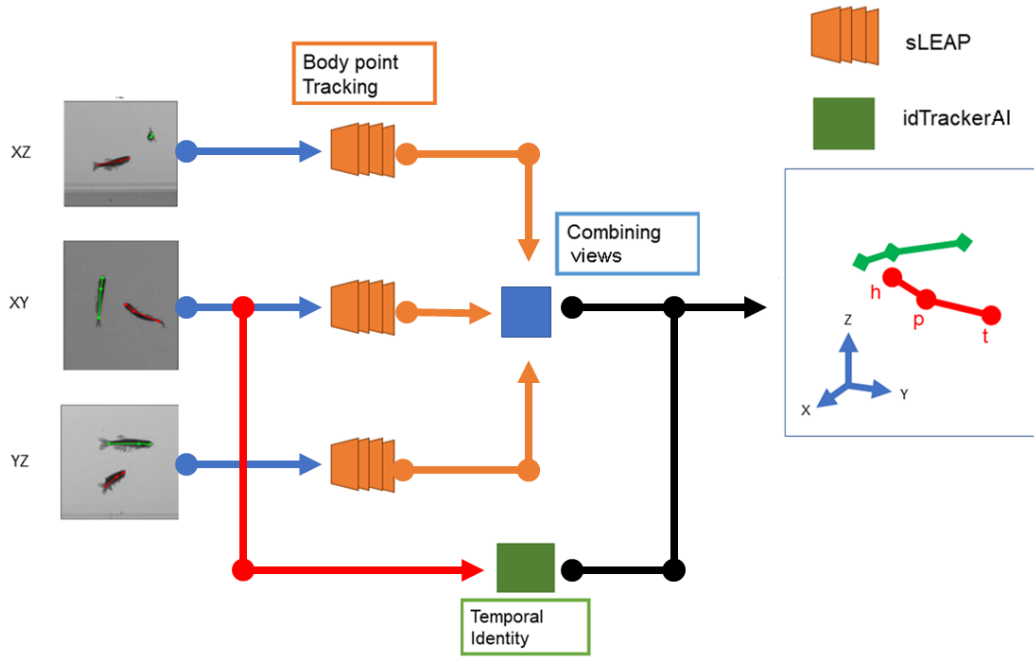


Figure 2.2: Processing pipeline to recover 3D postural information from a pair

Date of experiment	Pair	Number of frames	Missing frames (%)	Gap size $n > 20$
<b>January 29th</b>	1	769848	1.5	21
<b>January 30th</b>	2	501943	1.3	6
<b>February 14th</b>	2	495610	0.9	7
<b>February 18th</b>	1	537880	1.7	4
<b>March 27th</b>	2	561010	0.9	5
<b>March 31th</b>	1	746434	1.3	6
			Sampling rate (Hz)	100

Table 2.1: Datasets used in analysis

tracking. We want to build a physical model, therefore we will use biomechanical degrees of freedom and constraints coupled with the ML techniques, such that the obtained variables are easily interpretable.

## 2.2 Joint Variables

One of the biggest challenges when deriving a physical description of a behavioral state (i.e. posture) is the appropriate representation of said state. Often, there is a trade-off between fidelity and interpretability with more accurate representations being less interpretable and vice-versa. [4].

It is possible to obtain representations by unsupervised methods, through short time dynamic analysis and other cases as the examples showcased below:

- Time delay embedding - relevant time delays would be recovered under a minimization criteria (minimization of the mutual information between the variables). Through a neighbour embedding method, it is possible to recover a high dimensional representation of behavior.
- Machine Learning methods
  - **Autoencoder networks** on the variables would allow for nonlinear dynamics to be recovered, through information compression with the appropriate cost function. There is the issue of being a black-box approach that will give non-interpretable representations.
  - SVD would recover a higher dimensional representation of the variables and a non-linear decomposition of variables.
- **Local linear models** will give a short time linear dynamic description of the system, with the models being defined through a dissimilarity metric. This approach is useful for dynamic analysis, including the recovery of unstable eigenvalues and emergence of chaos.
- PCA decomposition which is limited by recovering linear modes and dynamics in behaviors, but that is fairly straight forward and is a good choice for a first approach for dimensionality reduction. Some improved versions can allow the recovering of nonlinear and time delayed elements of the dynamics.

Any of these techniques or a combination of them are good approaches to recover a non-human defined representation of a behavioral state. Whatever there is the issue of most of these modes not being interpretable from an ethological or physical perspective, and are best applied to a system that is better understood.



## 2.2.1 Notation and Definition of Variables

To simplify expressions and for compactness, the following notation will be used:

- $b_i^k(n)$ ,  $k = \{0, 1, 2\}$ ,  $i = \{0, 1\}$  - this is the bodypoint  $k$  for fish  $i$  at frame  $n$ , a 3-D point. So each  $b_i^k(n)$  has three values (corresponding to the x,y and z planes).
- When only one of the coordinates axis is being considered (e.g. the z-axis), it will be displayed explicitly ( $z_i^k(n)$ ).
- **Differences** - When a vector is being displayed (between two bodypoints of the same fish or between different fish) it will be represented by  $\Delta^{jk}b_i(n) = b_i^j(n) - b_i^k(n)$  for two different bodypoints on the same fish and  $\Delta_{i \rightarrow j}b^k(n) = b_i^k(n) - b_j^k(n)$  for two different fish. A difference in time is described by  $\Delta_n b_i^k(n) = b_i^k(n+1) - b_i^k(n)$ .
- $f_s$  is the sampling frequency in the experiment (100 Hz).

We are looking at this system in an exploratory fashion, so it is more helpful for us to easily interpret the results obtained, at the cost of a somewhat less accurate model. With that in mind, we use human-defined single fish variables (dependent only on a single fish), based on known body constraints and previous ethological studies, and through these variables we are able to recover the fish's trajectory and posture.

- **Heading vector**, a 3D vector embedded in a 2-sphere, the different between the head and pectoral points normalized. This variable is defined under the assumption that the fish has rigid body between the head and pectoral points as:

$$h_i = \frac{b_i^{01}}{\|b_i^{01}\|}$$

This is a reasonable assumption taking into account the anatomical structure of adult zebrafish. [61]

- **Speed**, (which is time dependent) defined as

$$s_i(n) = \|\Delta_n b_i^k(n)\| f_s$$

- **Tail beating rate**, a tail frequency defined on a 2-sphere, described in further detail ahead.

- **Vertical deflection of the tail** which is defined as

$$\phi_i(n) = \arcsin\left(\frac{\Delta^{21} z_i(n)}{\|\Delta^{21} b_i(n)\|}\right)$$

considered more relevant than the horizontal angle, as there is evidence that this postural feature is relevant in submissive behavior.

From these variables it is possible to recover a trajectory and a posture conformation for the fish, under the assumption that the horizontal coordinates of the tail are not particularly relevant (whether the fish is bending its tail to the right or to the left).

While the variables defined above will describe single fish postures, since we want to describe social interactions, there will be other constraints to be considered, giving rise to a new set of variables, composite variables. Composite variables, also human-defined, depend on the pair of fishes and can't be defined in isolation:

- Joint variables
  - The **average distance** between a and b
  - **Alignment**, defined as the dot product of the heading vectors of the pair
  - **Acceleration alignment**, the dot product of the normal acceleration vectors of the pair.
  - **Average tail beating rate**
  - **Average speed**
- Deviation variables
  - **Coefficient of variation of the tail beating rate**

The variables above are considered joint or deviation variables, based on whether or not they capture the properties of the pair as a whole entity or if they capture the differences between them.

The composite variables are defined based on ethological descriptions suggesting some symmetries, but those symmetries are not rigorously defined.

The composite variables defined above are the ones used when developing behavioral states for the pair, as they don't require picking any particular fish as a reference, albeit losing information that is relevant to specific questions about a contest, such as the winner of a fight.

The next few subsections will clarify some of the variables above

### 2.2.1.1 Distance

Distance is a common variable when looking at social behaviour and interaction in general, being easily interpretable and obtainable. There are two particular cases in zebrafish fight where distance plays a role: in display and circling where both fish are close to each other, and bites, strikes and other physical attacks, where the head of one fish is close to the body of the other fish. To be able to distinguish contacts, a variable denominated heading distance from fish  $i$  to fish  $j$  is defined as

$$d_{i \rightarrow j} = \min_k (|b_i^0 - b_j^k|) \quad (2.1)$$

In this case, the mean distance is the average of the heading distances of both fish, normalized by the average fish bodylength  $d = \frac{1}{2L_b} (d_{1 \rightarrow 2} + d_{2 \rightarrow 1})$ .

### 2.2.1.2 Relative alignment

Relative orientation is another variable that is studied when discussing collective behavior. It is defined as the difference between the orientation of different animals and it's a proxy variable for alignment (equivalent to polarization in magnetic systems).

It is possible to define a relative orientation between both fish using the previously defined heading vectors, given by

$$p = h_1 \cdot h_2 \quad (2.2)$$

To define the acceleration alignments, the formulation is equivalent but instead of  $h_i$  we use the acceleration defined as the change in alignment

$$a_i^n = \Delta_n h_i f_s$$

$$p^a = a_1^n \cdot a_2^n \quad (2.3)$$

### 2.2.1.3 Tail beating

Quick changes in speed and direction in fish's motion are associated with the change in fish's posture displayed by a tail beating. Therefore it is interesting to see how often these fish change

direction and more important, if there is a connection between these and the onset and outcome of a fight. To define an appropriate variable we first consider our fish's representation, a 3 body-point-skeleton, or two vectors from the pectoral point to the head and tail points respectively, ( $b_i^{01}$  and  $b_i^{21}$ ). As we are not concerned with the fish's body length but with the bending, we normalize the tail-pectoral vectors, and from the new  $\hat{b}_i^{21}$  vector we compute tail beating rate.

From a continuous analogue:

$$\omega_i = \frac{1}{2\pi} \left\| \frac{d}{dt} \hat{b}_i^{21} \right\| = \frac{1}{2\pi} \sqrt{(\dot{\theta} \cos \phi)^2 + \dot{\phi}^2} \quad (2.4)$$

An interesting observation we can make is that when one of the angles is constant (for a purely vertical or horizontal oscillation), the tail beating rate is equal to the frequency of oscillation for that direction. Of course from our case, we will be working with finite differences rather than derivatives, but the formulation is equivalent.

## 2.3 Compound Behavioral States

From ethological descriptions, there is an escalation that happens in a fight. To capture that escalation we must look within the fighting interaction. As some behaviors are similar in terms of their objective within a contest, we can look at them as being part of a larger category or compound behavioral state. They are called compound because, rather than defining these behavioral states for single fish, they are defined for the pair, and may include several individual fish behaviors within:

- **Displays** In this section of a fight the pair are measuring each other's fighting ability through indirect means. In general the pair begins their interactions at this stage, usually escalating into more aggressive interactions, but sometimes the contest is resolved at this stage if the fighting ability gap is large enough. Since we controlled the experiments for pairs with roughly the same fighting ability, that is not the case for most of the experiments ran. This compound behavior has two distinct phases.
  - Passive display - In this phase, the pair exhibits display behaviors (The pair is aligned, moving slowly or stopped).
  - Active display - This phase includes a characteristic circling behavior (The pair is aligned in antiparallel fashion, and they move in a circular pattern).
- **Aggression** - There is an escalation of the interaction and the pair is showing aggressive behavior. This usually occurs right after displaying but there may be shorter bouts that occur outside of the general fight interactions, usually attacks from the winner over the loser, usually interchanged with freezing behavior by the loser. This stage includes chasing, striking and biting, and the distinction is done based on the relative, rate of these events.
  - Symmetric aggression - Both fish attack at a similar rate.
  - Asymmetric aggression - One fish attacks at a significantly higher rate, likely the winner of the contest.
- **Freezing** - An asymmetric stage of the fight, where one fish displays freezing behavior (is at the bottom of the tank, not moving with a droopy tail). At this stage, the resolution of the contest has been settled, with the loser fish exhibiting this behavior. Usually occurs at the end of an aggressive stage, and becomes more common the longer the fish are interacting with each other.

Analysing the fight through these compound behavioral will allows a more detailed view of a fight that is still interpretable.

To describe the dynamics of the fight, we must understand these compound behavioral states through the variables we have defined. These compounds behavioral states are defined by the different behaviors contained within them. Exploring these behaviors over a fight will give us a better insight into the compound behavioral states themselves.

### 2.3.0.1 Display

We can think of the display state as having two distinct phases, a passive and active phase. Within those different phases, specific behaviors are exhibited by the fish which have an effect on their general property. We detect these different behaviors by looking at average values over a time window of 100 frames (1 second) and comparing them to thresholds. This done under the assumption these behaviors occur at these timescale, or that their effects on the variables are visible at this timescale.

For the passive display phase, we expect the fish pair to exhibit mostly **display behavior**, characterized by these variables:

- $\langle d|d \rangle < 2 \text{ bl}$
- $\langle |p||p| \rangle \approx 1$
- $\langle s|s \rangle < 3.5 \text{ bl/s}$

For active display, **circling behavior** is the most relevant and distinct exhibited by the pair.

- $\langle p|p \rangle \approx -1$
- $\langle p^a|p^a \rangle \approx -1$

within a short time window, something also shown on the experiment. This is the key distinction between this stage and a general display.

These events are in general short lived, but they are all contained within the display state of a fight, as these do not require physical contact between the fish.

### 2.3.0.2 Aggression

The key distinction of the aggressive state relative to the display state is that the fish in this state exhibit **aggressive behaviors**.

There are several different aggressive behaviors contained within it, but they share characteristics as physical contact and high intensity, which can be thought of in simplified fashion as collisions event between the pairs.

If we think of an aggressive interaction as a collision between the fish, we can detect it based on its speed and its distance, in this case, the distance from the head of the aggressor to the body of the other fish. As the interaction becomes more aggressive it becomes more intense which is manifested in higher speed and closer distance, which may be captured over a time window.

- $\langle d_{i \rightarrow j} | d_{i \rightarrow j} \rangle < 0.25 \text{ bl}$
- $\langle s_i | s_i \rangle > 3.5 \text{ bl/s}$

To distinguish between a symmetric and asymmetric phase of the aggressive section, we can compare the rate of aggressive behaviors over a longer window of time. In this case, collisions are a proxy for aggressive behavior, so we look at the rate of collisions over time.

### 2.3.0.3 Freezing

Within the freezing state, one of the fish exhibits **freezing behavior**. The most relevant variable when discussing freezing behavior are their speed. The fish tend to be at the bottom of the tank and a speed close to zero (but not necessarily zero, as the fish may drift).

- $\langle s_i | s_i \rangle < 1 \text{ bl/s}$

## 2.4 2D Projection and Clustering

The joint variables defined above form a 6-dimensional system. We are limited by our 3-dimensional view of the world, so there is a need for lower dimensional projection. We will be analyzing multiple experiments at once, in order to build a more robust representation. So we use 6 experiments from 2 different fighting pairs, attaching the experiments together in a single trajectory. The lower dimensional projection is done through an embedding process.

Consider the 6-d system as being part of a topological space  $X$ , and the lower dimensional projection being in a topological space  $Y$ . The embedding process  $\Gamma$  is described by

$$\Gamma : X \rightarrow Y \tag{2.5}$$

where  $\Gamma$  is an injective and continuous map (an homeomorphism) between the two spaces. There are several challenges with implementing this process numerically, but there are several embedding techniques that have these properties.

The following approach is inspired by a similar pipeline used in [16], with changes in the kind of processing and variables used, the embedding techniques and the clustering process.

### 2.4.1 UMAP Embedding

#### 2.4.1.1 Theoretical foundation

This embedding is done under a couple of assumptions for our data:

- The data can be uniformly distributed in a manifold
- This manifold, if it exists, is locally connected
- The primary purpose of the embedding process is to preserve topological structure.

A manifold  $M$  is locally connected if every point it contains is locally connected, that is, the neighbourhood around any point in  $M$  contains a connected open set. [62]

These assumptions are the motivation behind the methods present in UMAP. Under the first assumption, any volume  $V$  in the manifold should contain the same number of points in every location in the manifold. This can also imply that for  $k$ -th nearest neighbours, there is a fixed volume for every point in the manifold. Thus distances between points in the manifold can be normalised by the distance from a point to its  $k$ -th nearest neighbour. The working principle of



UMAP consists of approximating local manifolds around every point and using their fuzzy simplicial set representations to build topological representation of the high dimensional data. From an initial low dimensional representation, the latter part of the method consists in finding an equivalent topological representation minimizing the cross-entropy between the high and low-dimensional representations, through stochastic gradient descent. [63]

UMAP is computationally implemented by:

- Computing a weighted k-neighbourhood graph over the initial high dimensional system, with the connections corresponding to the neighbours of a particular point under a distance metric.
- Defining a function preserving characteristics of the graph (global properties) and finding a low dimensional representation that optimizes it.

The choice of hyperparameters has an influence on the representation of variables, and for our application we used the default options for the metric, the euclidean metric, with 15 neighbours for each point. We used a learning rate of 7, and the number of output components we choose is 2 (the dimension we want to visualize our system in). [63]

#### 2.4.1.2 Minimum square difference sampling

Consider a  $M \times N$  matrix  $A$  representative of the time series, with  $M$  variables and  $N$  points, and we want to select a set of  $n$  points from said matrix,  $v$ . From window of length  $w = \lfloor \frac{N}{n} \rfloor$ , we select the point  $j$  in the  $k$ th window such that:

$$v_j = \min_{j \in w_k} \left( \sum_{1 \leq i \leq M} (A_{ij} - \hat{A}_k)^2 \right) \quad (2.6)$$

with  $\hat{A}_k$  being a  $M \times 1$  vector corresponding to the average values in the  $k$ th window. This ensures the sample we get is more representative of the average properties of the whole time series.

#### 2.4.2 Kernel density estimate

To visualize the structure of the resulting projection, we do a kernel density estimate (KDE) over the two variables by computing an histogram and doing a convolution of this two dimensional histogram with a gaussian kernel.

### For one dimension:

Consider an histogram  $H$  with  $n$  bins of some random variable  $v$  as a sum of delta functions:

$$H(v) = \sum_{i=1}^{n-1} c_i \delta(v - \hat{v}_i)$$

where  $\hat{v}_i$  is midpoint of bin  $i$  (the point in the middle of the interval  $[v_i, v_{i+1}]$ ), and  $c_i$  is the value of the histogram in that interval. The kernel density estimate  $g_h(v)$  is defined as:

$$g_h(v) = (H * k_h)(v) = \int_{-\infty}^{\infty} H(u) k_h(v - u) du = \sum_{i=1}^{n-1} c_i k_h(v - \hat{v}_i) \quad (2.7)$$

with  $k_h(v)$  being the kernel function with bandwidth  $h$ . This result can also be generalized for two dimensions.

This approximates the probability density function to a better extent than the regular histogram with the risk of oversmoothing and losing relevant detail. Choosing the appropriate parameters can help avoid oversmoothing, such as the number of bins in the histogram and the bandwidth of the gaussian kernel used (determining a resolution for the representation). The number of bins used is proportional to the square root of the number of samples used in the embedding. The bandwidth of the kernel is determined by the user (In this case it is equal to 10).

The approximate probability density function gives a "behavioral map" where the peaks in density (probability modes) can be interpreted as behavioral states. (see figure [3.3-a](#))

### 2.4.3 Peak Detection and Clustering

Detecting peaks in the probability density is equivalent to finding the maxima in a scalar function of 2 variables. The gradient is a good method when trying to find extrema in a multivariate function, but it has some caveats. It doesn't give a distinction between minima and maxima, so while it would be necessary it is not a sufficient method.

For a scalar function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , the  $n \times n$  Hessian Matrix  $\mathbf{H}$  is defined as the Jacobian of the gradient of  $f$   $\mathbf{J}(\vec{\nabla} f)$ , of the form:

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_i} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_i \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_i^2} & \cdots & \frac{\partial^2 f}{\partial x_i \partial x_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_i} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

This definition is for real numbers, but it can be extended to the complex numbers as well [64].

From this definition,  $\exists H(f) \iff \exists \frac{\partial^2 f}{\partial x_i \partial x_j}$  continuous  $\forall x_i, x_j \in \mathbb{R}^n$ . Under Schwarz's theorem,  $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ , which implies  $\mathbf{H}(f)$  is a symmetric matrix.

For simplicity, we will show the following results for an Hessian of a function defined in  $\mathbb{R}^2$ , but they are generalized for higher dimensions. The Hessian is used often for second derivative tests [65], for the detection of extremal points (and determining whether they are maxima or minima).

This is done through Hessian equations:

$$L(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (2.8)$$

$$D(x, y) = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 \quad (2.9)$$

The first equation is the Laplacian of  $f$ , and the second equation is called the Hessian determinant equation, and these correspond respectively to the trace and the determinant of the Hessian, turning the analytical problem into an algebraic problem, that can be solved computationally.

For an extremal point of  $f$   $p = (x_0, y_0)$ :

- if  $D(x_0, y_0) > 0$  and  $L(x_0, y_0) > 0$  then  $p$  is a local minimum
- if  $D(x_0, y_0) > 0$  and  $L(x_0, y_0) < 0$  then  $p$  is a local maximum
- if  $D(x_0, y_0) < 0$  then  $p$  is a local saddle point
- if  $D(x_0, y_0) = 0$  the second derivative test is not conclusive, and there is a need for a higher order derivative test.

That itself gives us information about the location of the maxima in a particular function or scalar field. Another observation to make is that the neighbour around said point also satisfies this conditions, since the Hessian is equivalent to the shape factor (curvature) of a surface defined in a Euclidean space (as the matrix representation of the metric is the identity). Therefore a detection of clusters can be seen as the detection of regions satisfying the same conditions as the maxima they enclose, (regions of positive curvature, reflected in the negative Laplacian).

Under the assumption these results remain valid for finite differences, we can use the Hessian to find the probability modes. We not only want to find the extrema, but also want to find the neighbourhood of said extrema. The Laplacian can work as an edge detection method, and combined with the determinant of the H, it's possible to find regions in the map with positive curvature (the neighbourhood surrounding a local maximum).

Having found those probability modes, we binarize the maxima and a neighbourhood around them, and apply morphological transformations on this binary representation, labelling the resulting clusters with a connected component algorithm [66], ensuring we capture the modes and their neighbourhood, and the inclusion of more data points in our analysis. (see figure 3.3-b)

For each cluster we find locations in time satisfying the condition (being contained within the cluster). From these temporal locations we can build a symbolic representation of the fight through the clusters, again, under the assumption than the clusters obtained are equivalent to behavioral states. (see figure 3.4)

#### 2.4.4 Ethogram Computation and Symbolic representation

From the clusters recovered from KDE we can compute the ethogram through the following process:

Consider a fight sequence  $\vec{s} = (s_1, s_2)$  and a set of clusters  $C = \{C_i\}, i = 1, \dots, n$ .

We define the ethogram  $e$  as an n-dimensional binary sequence:

$$e_i(t) = \begin{cases} 1 & \text{if } \vec{s}(t) \in C_i \\ 0 & \text{otherwise} \end{cases} \quad (2.10)$$

The n-dimensional binary ethogram can also be symbolic represented by a 1-dimensional symbolic sequence:

$$S_t = \sigma_i, \text{ for } i \text{ such that } e_i(t) = 1 \quad (2.11)$$

This symbolic sequence can be interpreted as a stochastic process, that can be later analysed.

## 2.5 Behavioral Labelling

From the temporal ethogram, we can learn some properties of the recovered clusters, in this case, the average values of the joint variables within the clusters. By knowing the average properties of the clusters, it is possible to label them as being associated with different compound behavioral states, based on phenomenological conditions. With this in mind, we use a labelling process and classification based on phenomenological conditions.

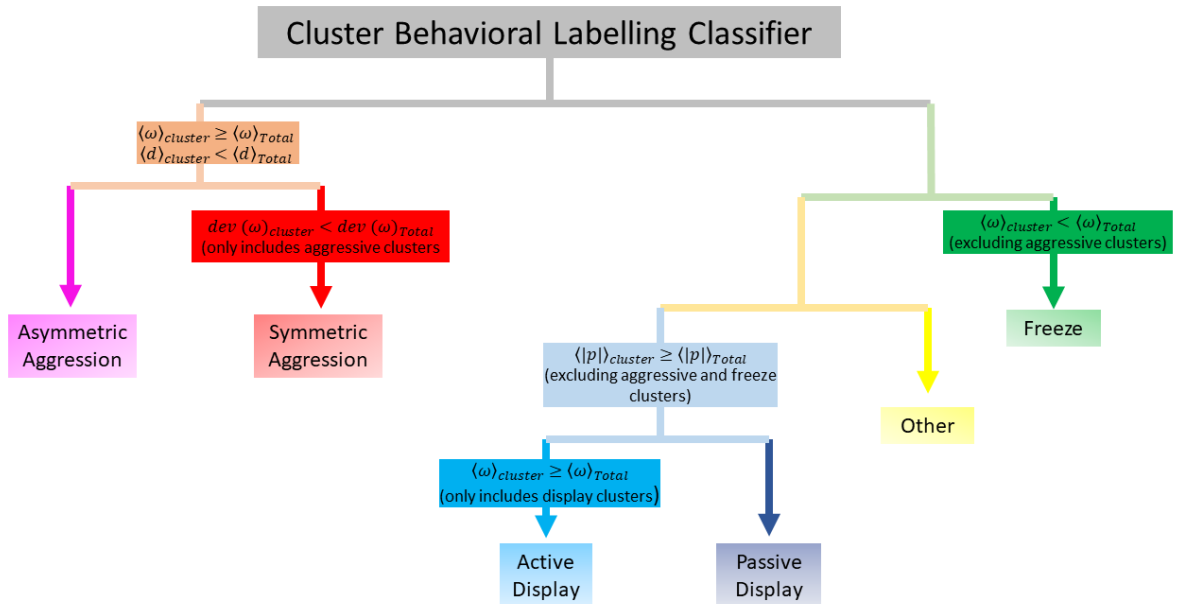


Figure 2.3: Classifier used for Behavioral Labelling

A cluster labelled with this process does not necessarily correspond to the compound behavior, rather, it has characteristics associated with that compound behavior.

## 2.6 Transition Matrix and Spectral Analysis

Having a symbolic sequence allows us to make a transition matrix possible describing the dynamics of the fight through transition across these different clusters.(see figure 3.7)

Consider a sequence  $S = S_1 S_2 S_3 \dots$  where  $S_\tau$  is the state occupied by the pair at transition time  $\tau$  (to distinguish from time  $t$ ), and consider the set of states  $\sigma = \sigma_i, i = 1, \dots, n$ , and a

transition matrix describing a change of states from  $\tau$  to  $\tau + 1$ . If at a transition  $\tau$ ,  $S_\tau = \sigma_i$ , the entry in in the transition matrix is equal to the probability of the next state being  $\sigma_j$ , given by:

$$A_{ij} = P(S_{\tau+1} = \sigma_j | S_\tau = \sigma_i) \quad (2.12)$$

The computation of the elements of the matrix is done through a recursion process over the finite sequence.

We compute eigenvalues of transition matrices for different number of transitions, for a better understanding of their properties.

The transition matrix is used to describe transitions between states in a Markov Chain. If our process is Markovian, the transition matrix describes it completely. A Markov Chain is a stochastic property that exhibits the Markov property, presented here only as a statement:

*Given a present state  $S_\tau$  and a past state  $S_{\tau-1}$ , the future state  $S_{\tau+1}$  is only dependent on  $S_\tau$ .*

A more detailed description of this property is given by [67], but it is used in general to describe memoryless stochastic processes. While not being accurate in the description of behavior, it is a useful framework to understand the dynamics between the different clusters.

What the Markov property shows is that if the transition from  $S_\tau$  to  $S_{\tau+1}$  is given by a transition matrix  $A_1$ , then in general:

$$S_{t+\tau} = A_\tau S_t = (A_1)^\tau S_t \iff A_\tau = (A_1)^\tau \quad (2.13)$$

For  $\tau$  transitions, the transition matrix is a function of the transition matrix for one transition. If  $A_1$  is diagonalizable, this expression becomes:

$$V^{-1}D_\tau V = V^{-1}(D_1)^\tau V$$

With  $D$  being a diagonal matrix where the nonzero elements correspond to the eigenvalues of  $A_1$ .

From the Perron-Frobenius theorem, we know that if we have a nonnegative square matrix, the largest eigenvalue  $\lambda_M$  is positive, with all remaining eigenvalues  $|\lambda_i| < \lambda_M$ . So  $\lambda_M$  determines the spectral radius of the matrix.

With  $\lambda_M = 1$  and all other eigenvalues contained within the radius  $|\lambda_i| < 1$ , all other eigenvalues approach zero as the number of transitions ( $\tau$ ) approach infinity.

So the eigenvector associated with this eigenvalue describes the steady-state distribution of the states as  $\tau$  approach infinity, therefore it has nonnegative values.

The other largest eigenvalues  $\lambda_i < 1$  are associated with dynamic properties of the process, which are described by the associated eigenvectors [68].

The clusters of behavior are a discretization of the behavioral phase space recovered, and if extended to a continuous representation, the transition matrix becomes the infinite dimensional Perron-Frobenius operator. The spectrum of this operator are modes equivalent to the eigenvectors of the transition matrix, eigenfunctions. Since all eigenvalues are bound to a spectral radius  $\lambda_M = 1$ , these provide a temporal hierarchy for relaxation towards the eigenfunction describing the equilibrium distribution (with the eigenvalue  $\lambda = 1$ ). [69]

The eigenvectors can be interpreted in a similar way as describing different hierarchies of dynamics between clusters tending towards the equilibrium distribution.

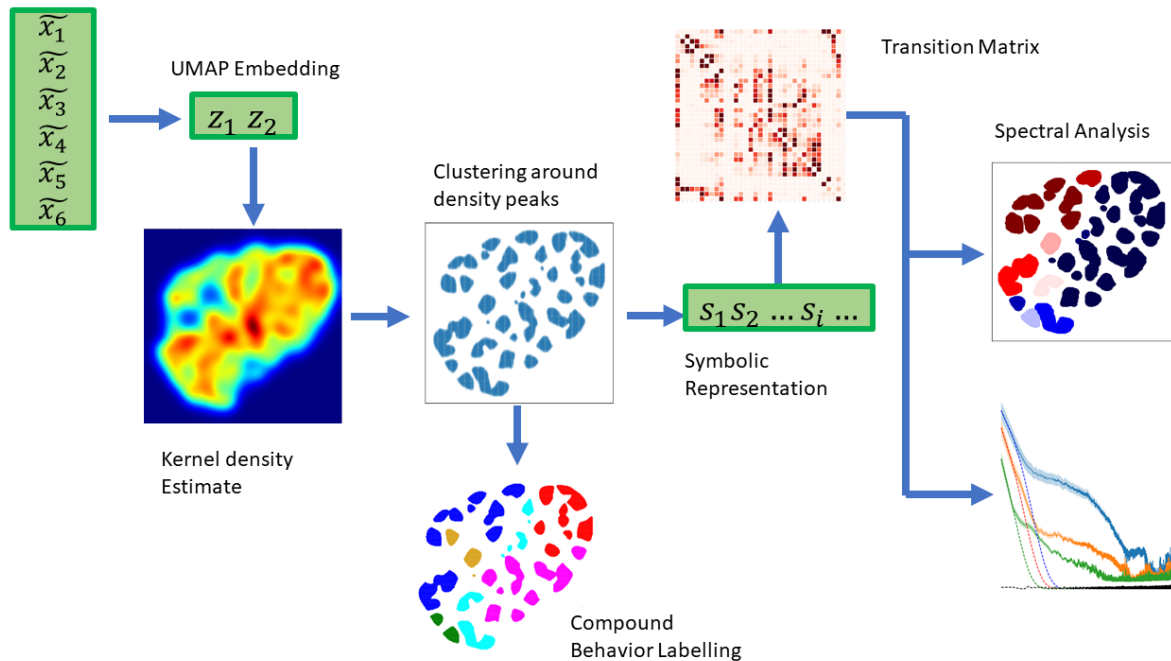


Figure 2.4: Analysis pipeline from the derived variables to the results.

The processing pipeline will be discussed in further detail in the next sections, and it does not include side analysis done on individual variables, as this was to tackle the main objective of the project.

## 2.7 Supplementary Material

The code developed in the project can be found at the GitHub repository <https://github.com/rafaelalmada240/zebrafishfightanalysis>.

# Chapter 3

## Results and Discussion

### 3.1 Results

#### 3.1.1 Kinematic description of a contest

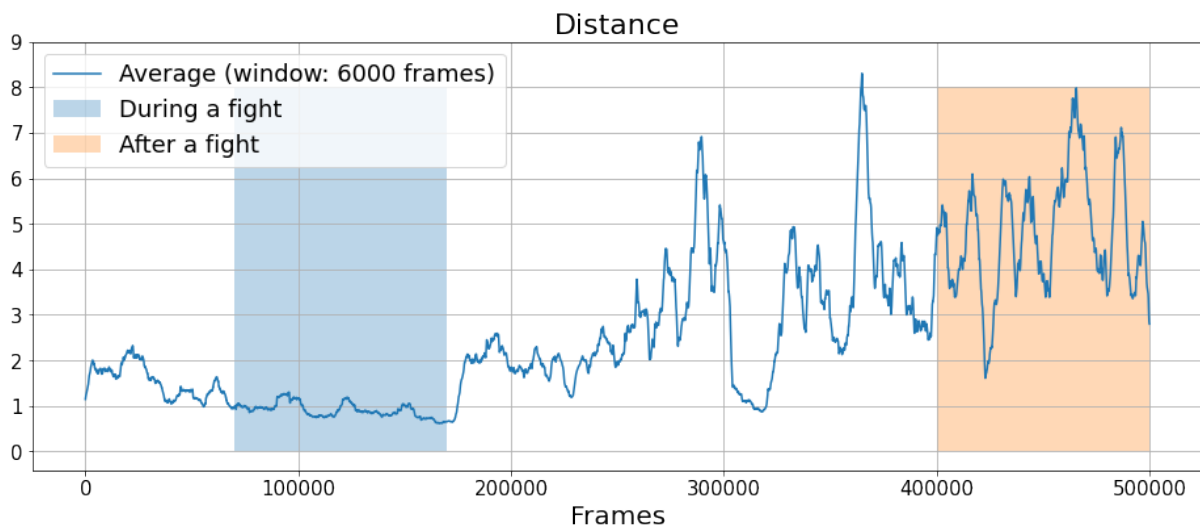


Figure 3.1: Average distance over a fighting experiment, with the highlighted areas corresponding respectively to a time when the fish are fighting and after the fight has ended.

From dataset 1, we looked at the distribution of joint variables during and after a fight, using the average distance as a measure to determine the proper time windows to consider (see figure 3.2).

A Kolmogorov-Smirnov test was done on the histograms, to have an empirical measure of the difference between distributions. Below we have a table displaying the difference between the distributions.



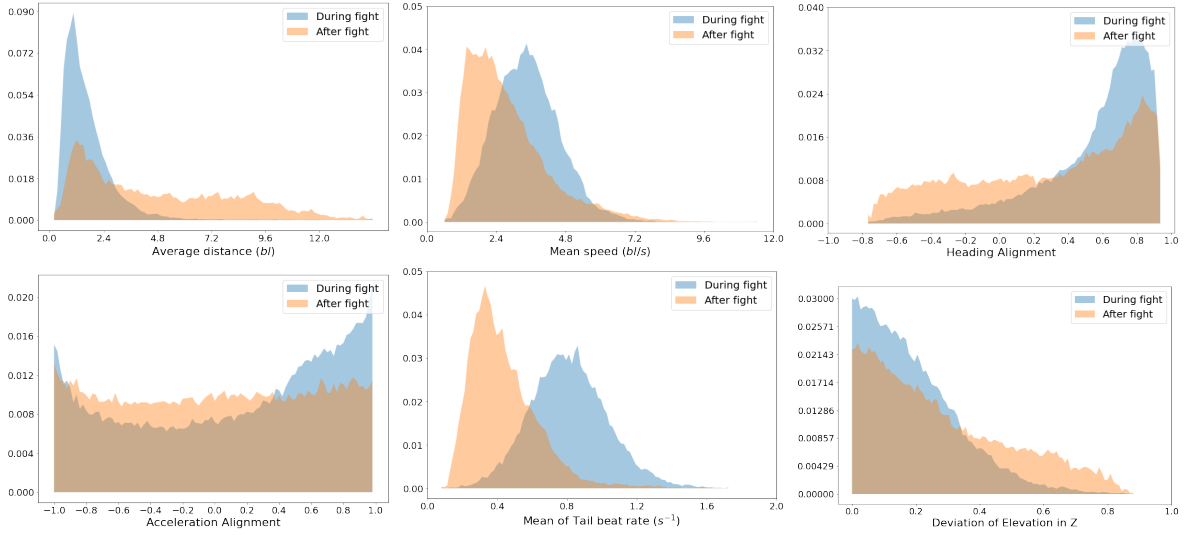


Figure 3.2: Distribution of joint variables in two intervals of a fight, during a fight and after the fight.

Variable	Test statistic (100 bins)	p-value
Distance	0.49625	$< 10^{-16}$
Alignment	0.25717	$< 10^{-16}$
Acc. Alignment	0.12454	$< 10^{-16}$
Speed	0.31731	$< 10^{-16}$
Mean Tail Beating	0.65807	$< 10^{-16}$
Deviation Tail Beating	0.18885	$< 10^{-16}$
Deviation of Elevation	0.36728	$< 10^{-16}$
	$D_\alpha = 0.00728$	$\alpha = 0.01$

Table 3.1: Test of distribution differences, during and after a fight section

## 3.1.2 Behavioral Map

### 3.1.2.1 State variables

We first combined all of the six datasets into a single time series, so we could analyse them all at once. We used a 2D projection of the original 6D system made from the joint variables, acquired using the UMAP embedding technique previously described in the methods section.

From the resulting 2D projection of the system, we compute its KDE and detect the density peaks, using the Hessian matrix, the process of which is also discussed in the aforementioned section. (see figure 3.3)

We assigned temporal labels to these clusters, through the process described in the previous section. This results in our "ethogram" of the fight. (see figure 3.4)

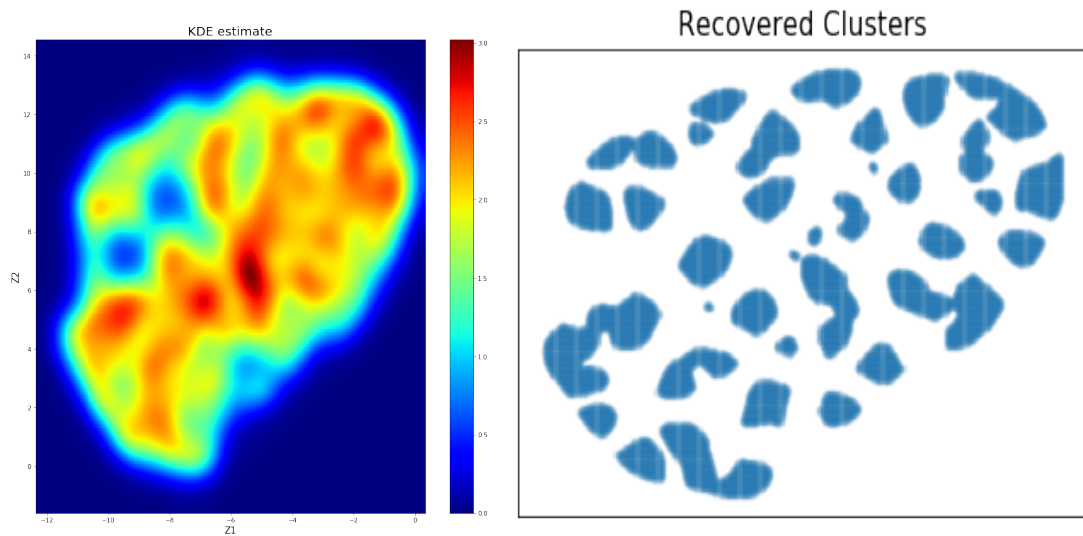


Figure 3.3: Left - An estimate of the 2D probability density function for the two embedded variables. The width (standard deviation) of the gaussian kernel used was 10. The variables were obtained through UMAP embedding. Right - Clusters around density peaks, recovered through the Hessian equations, and connected components algorithm. The density peaks may be different behavioral states.

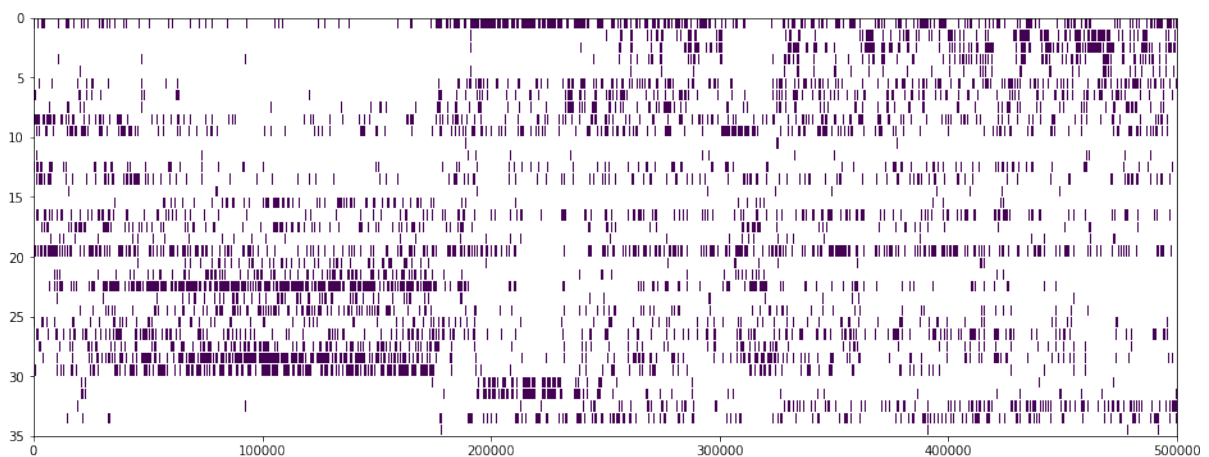


Figure 3.4: A fight "ethogram" based on the clusters recovered from the KDE estimate.

### 3.1.2.2 Cluster properties and overall structure

We looked at average values for our 6 variables in the clusters using the computed ethogram, and compared these with the average over the whole dataset. (see figure [3.5](#))

We labelled the clusters obtained based on the values of their average variables, as either display, aggressive behaviors or freezing behaviors, using thresholds for these variables. We characterized them as either passive or active display, and as symmetric or asymmetric aggressive behaviors. (see figure [3.6](#))

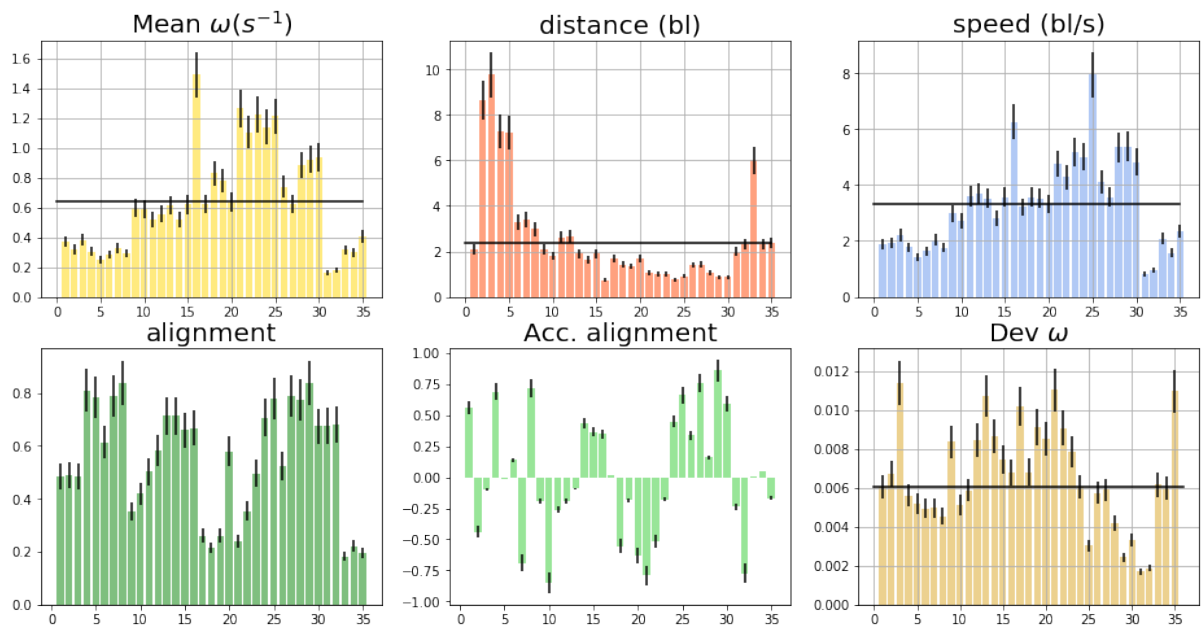


Figure 3.5: Averages of the joint variables for each cluster.

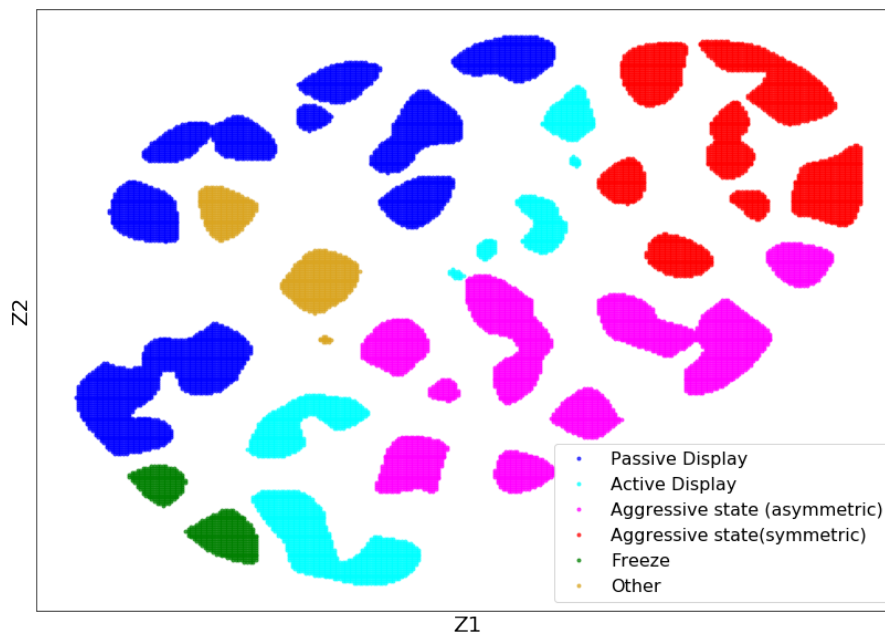


Figure 3.6: Cluster Labelling and association to compound behavioral states through a thresholding process

### 3.1.3 Symbolic and Spectral Analysis

From the computed ethogram, we obtained a symbolic representation, from which we computed the transition matrix for a dyadic contest. (see figure [3.7](#))

Having a transition matrix, gave us the opportunity to do spectral analysis. First we did spectral decomposition of the matrix, and later we computed the eigenvalues for other number

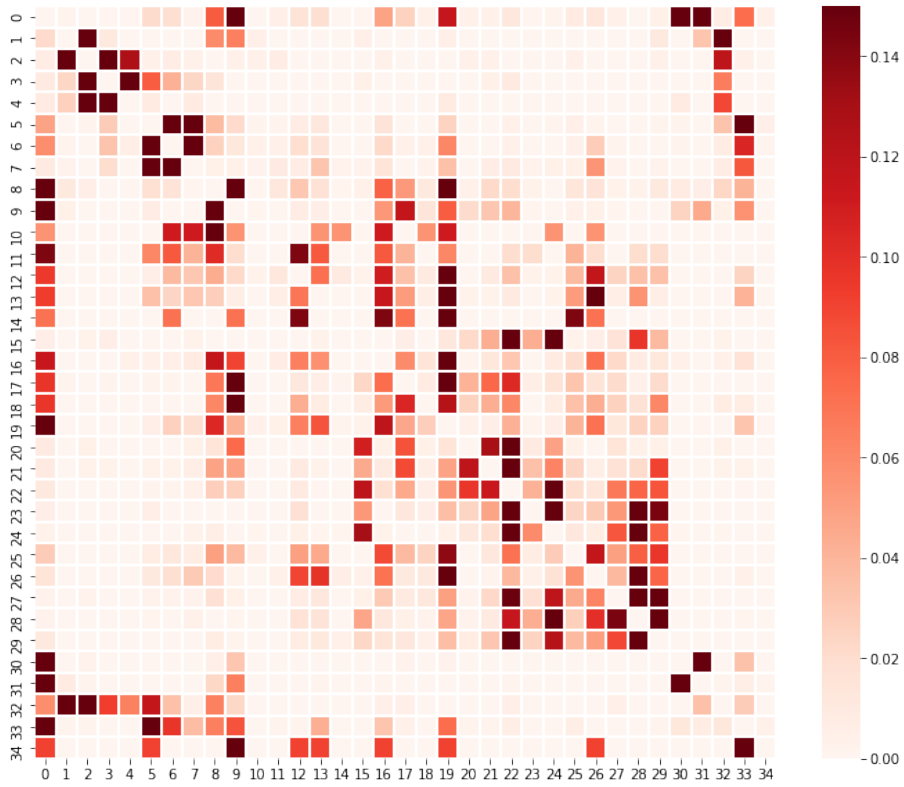


Figure 3.7: Transition matrix built from the symbolic representation of a fight

of transitions, to have a better insight over the dynamics of the fight. (see figure [3.8](#))

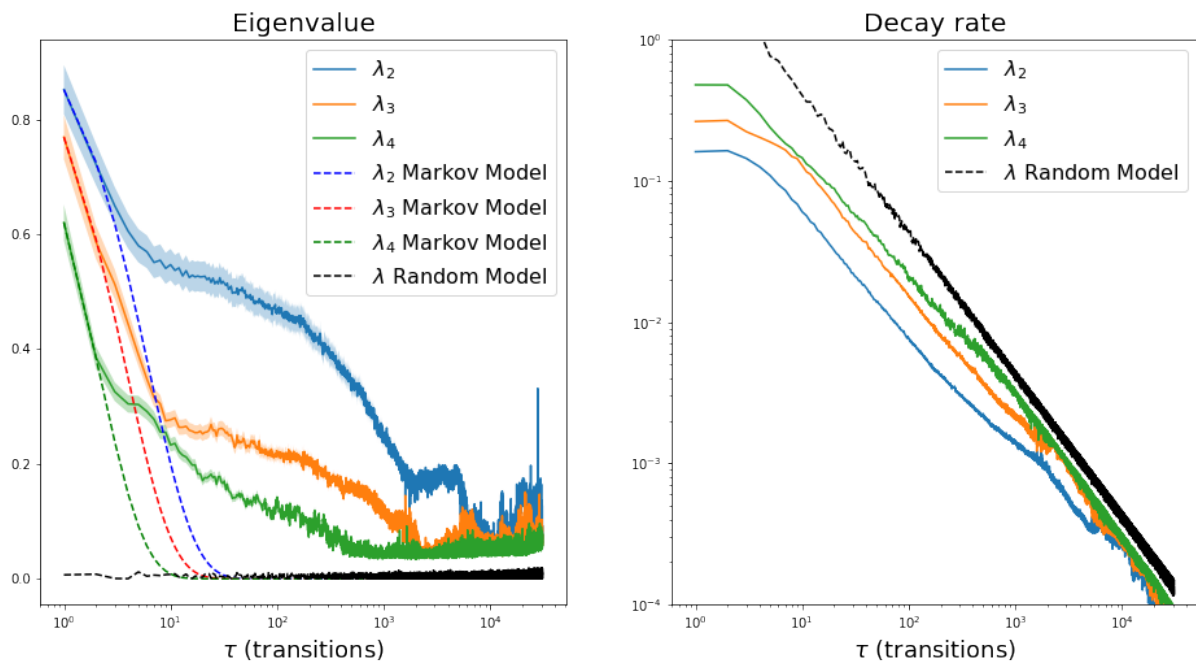


Figure 3.8: The 3 largest eigenvalues below 1 as a function of the transitions.

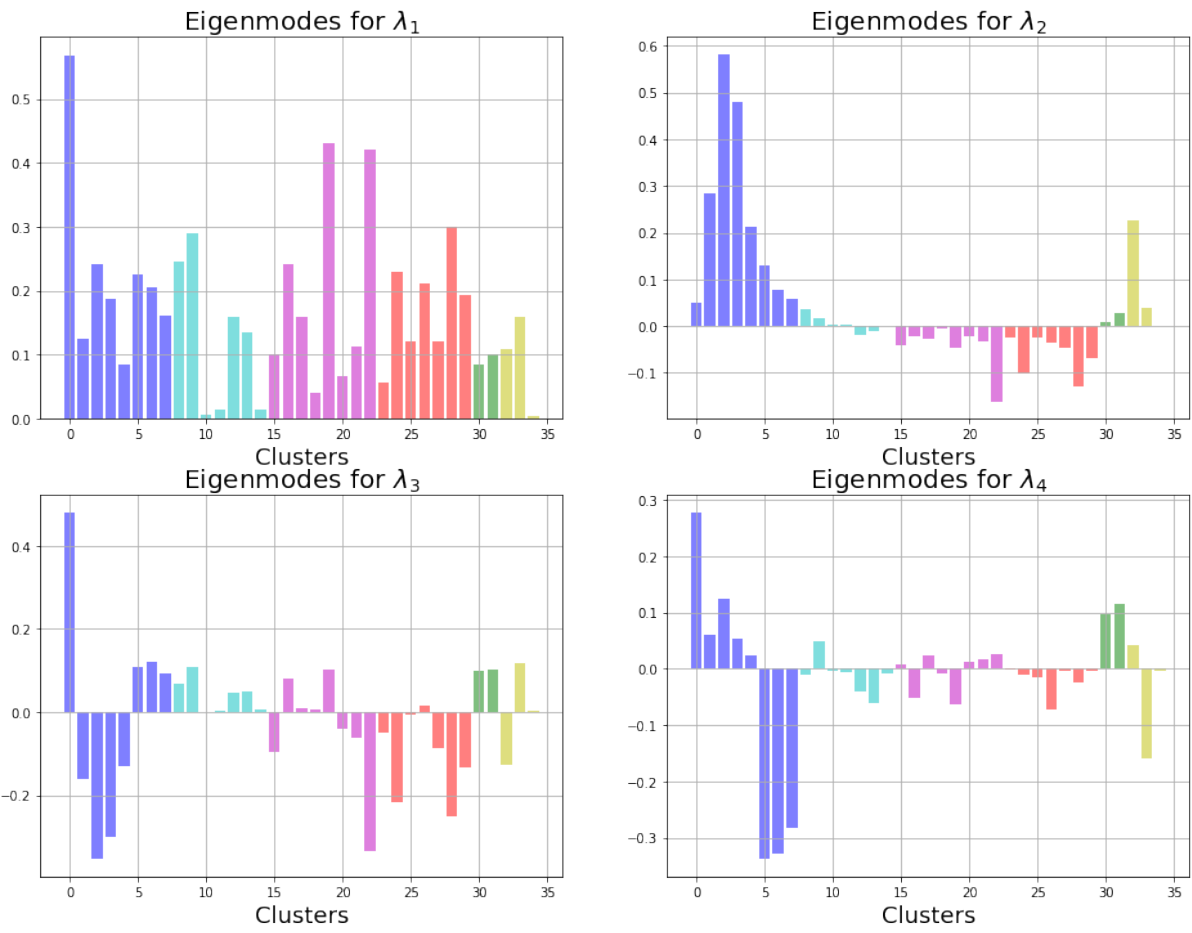


Figure 3.9: Values of the eigenmodes associated with the second, third and fourth largest eigenvalues.

We did a projection of the eigenvectors associated with the three largest eigenvalues over the clusters with which the values of their coefficients are associated, to better visualize them in context of the fight. (see figure [3.10](#))

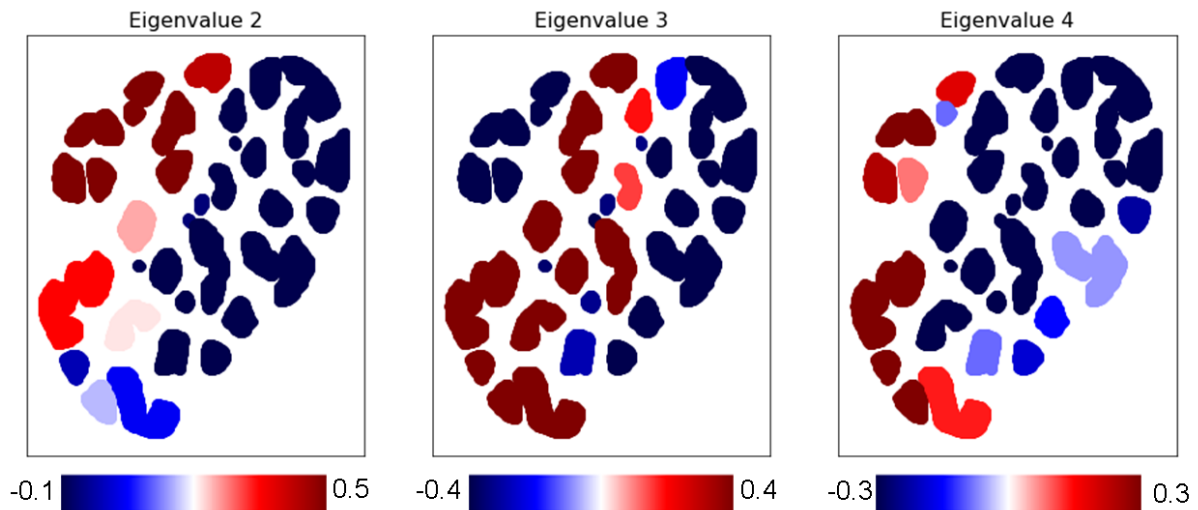


Figure 3.10: Cluster projection of eigenvectors for the eigenvalues respectively associated.

## 3.2 Discussion

Initially we looked at the distance time-series over an experiment. We observed a trend for the pair to be close during a fight bout, increasing rather quickly after the interaction is over. (see figure [3.1](#)).

After this initial analysis we compared the distribution of the six joint variables in two different instances of the experiment, during a fight bout and after the fight is finished. (see figure [3.2](#)).

As expected from the time series, the distribution of distances (fig [3.2-a](#)) during the fight concentrates in general in lower distances, while it is broad after the fight, suggesting the pair is not interacting with each other as strongly and the fish are exploring more of the tank rather than engaging with one another.

The difference in distribution (fig [3.2-b](#)) for the speeds shows an higher average speed of the pair during the fight, suggesting an higher activity, a result reflected on the distribution of mean tail beat rates (fig [3.2-e](#)), reinforcing the idea that the tail beating rate may be a good proxy measure for the activity of the pair.

The distribution of the heading and acceleration alignments (figs [3.2-c](#)-[3.2-d](#)) shows the fish are more aligned with each during the fight, which may be due to the fight including chases and displays, behaviors in which the pair is more aligned. The lower average alignment outside of a fight suggest a greater degree of independence in their interaction at least on a first order.

The distribution of deviation of tail beating rate (fig [3.2-f](#)) shows an increase after the fight. A possible way to interpret deviation is to consider the deviation as a measure of

synchronization(the closer the two tail beating rates are, the more synchronized the pair is) this can be seen as a break in said order or symmetry. It may be a way to distinguish between symmetric and asymmetric phases.

From the Kolmogorov Smirnov tests, it is possible to see significant differences between the distributions of these variables in different instances.

These results tell us that there is a process of escalation in a fight, higher activity (from the higher tail beat rate and speed), higher level of interaction (from the closer distance and higher alignments), and that there is a break in synchronization (from the increase in tail beat deviation after the fight).

This description of the fight is a reductive view, but it gives us a general interpretation of the dyadic contest that can be further explored by trying to capture the dynamics directly from our variables.

By combining 6 datasets from two different fish pairs in a single time series, we expect these results to be more robust than if we were to use a single experiment. We did a 2d projection of the 6 joint variable system using UMAP and found peaks in the KDE, clustering around their neighbourhood(fig 3.3). From this process, we computed a "fight ethogram"(fig 3.4) .

There are some discernible patterns to be discussed, and some clusters are more explored in the time period correspondent to the fight bouts, while others are more explored in the period both fish happen to be at the bottom of the tank, however these observations by themselves are not conclusive.

With reference in the previous ethogram, we compute the average value of the joint variables in each cluster (fig 3.5) . And we can make some observations from the results. The clusters that in the previous figure were more common during the fight section have a higher mean tail beating rate than the global average, which is a main feature of aggressive states. The clusters where the distance is higher than the global average correspond to freezing states, where the pair is in general in lower activity, something also reflected in the mean speed.

Using the labelling process described in fig 2.3, we labelled the cluster as belonging to different behavioral states as a function of their average properties (fig 3.6). The first observation to make is that none of these clusters exhibit a specific behavior like circling or display, but they are part of a compound behavioral states that includes these specific behaviors, or have average properties similar to the ones exhibited by the pair when performing displays or during circling events. A better way to refer to these clusters were these behaviors are present are clusters that

exhibit display-like or freeze-like behavior. The small clusters are joined in larger structures of similar average properties. If we consider our labelling process as adequate, then this result suggests our compound behavioral states have finer distinct structures, and that there is a finer possible description of behavior from the clusters recovered. There is close proximity between display clusters and aggressive clusters, while these are separate from freezing clusters by the aforementioned display clusters. This structure suggests a particular dynamic for the fight, one of escalation and deescalation that is not abrupt. It suggests transition from a symmetric aggressive interaction to a freezing interaction that is fundamentally asymmetric is transient rather than instantaneous.

A look through the transition matrix (fig 3.7) shows us that the clusters are more highly connected to their neighbouring clusters in the behavioral map we recovered. It also shows transitions between larger sets of clusters that are to be expected from the graphical interpretation. There are values closer to zero for transitions between behavioral clusters that are further apart in the map (i.e freezing states to aggressive states), in accordance to our possible interpretation of the fighting dynamics. From the transition matrix we could do spectral analysis and explore the properties of the fish-pair system.

The first observation we made from the plot (fig 3.8) is that we do not have a Markovian process, which is fairly unremarkable, given we're exploring animal behavior. Another observation is that the second, third and fourth largest eigenvalues have a behavior that is distinctly non-Markovian, with the 2nd eigenvalue decaying at a far slower rate relative to the other eigenvalues up until the order of  $10^2$  transitions. This suggests these eigenvalues capture long-term behavioral patterns.

Looking at the eigenvectors (fig 3.9), they seem to describe activity between different sets of clusters, suggests these eigenvalues are associated with the dynamics of the fight.

By projecting the value of eigenvectors over the clusters they correspond (fig 3.10), we have a better picture of the dynamics associated with it. The clusters with negative and positive coefficient represent different dynamical states, and the eigenvalues describe the dynamics between these different states. As these large dynamic states roughly correspond to the labelled compound behavioral states, we can interpret the eigenvalues as describing the transition between the behavioral states which occur at a long timescale.

From all of the results we have from the fight, we can build a picture of how to interpret the dynamics for a fish pair. The fight can be interpreted as a stochastic process with the pair



transitioning across different stereotyped behavioral states. The transition between these states has dynamic modes of different timescales, these timescales being captured by their associated eigenvalues.

A possible interpretation to what the pipeline does at a physical level, it that the dynamic modes are representation of metastable regions of the phase space, close to stable fixed points that are captured by our model and projected onto a lower dimension. The variables we used for the embedding process can be seen as observables of the system.

Another possible interpretation is that the 2d histogram recovered is equivalent to a lower dimensional projection of a potential energy manifold, where the dynamic modes are associated with energy minima, and the transition probabilities are functions of the potential barriers between the different minima.

All of these pictures are equivalent in some fashion and give us a framework to interpret fish interactions in a physical language.

### **3.2.1 For future projects**

Some things we may want to explore are different variables, or to develop an accurate reasoning for the choice of these variables, other representation of joint variables and even the addition of single fish variables may change the dynamics recovered.

The dimensionality reduction process is also something that requires testing. Perhaps a toy higher dimensional system may be projected down onto lower dimensions and the dynamics of that system may be explored to see what is lost and preserved in the process.

While we are happy with the results obtained, there are several things we think could improve our results, while still working with the same set of variables. A key problem would be using the default Euclidean metric on the joint variables, where some non-Euclidean variables (the alignments) were included. We used a transform of these variables for better implementation, but a possible change in the future would be to explore non-Euclidean metrics, which would involve the transformation of our Euclidean variables. We hope to see in the future, how using different metrics would change our resulting projection and the interpretations we derived from them.

It will be of interest to explore how the eigenvalues are associated with different scales for a fight. Our results show we can have a finer description of the fight, as such we may explore the possibility of building an hierarchical structure of the fight based on the clusters recovered.

Something to explore would be to test the robustness of this model. This was built from the combination of several different experiments, so it will be interesting to see how the picture changes with the inclusion of more experiments, perhaps leading to a statistically significant model of the fight.

Our physical interpretation of our variables as being observables of a system, leave a couple of questions in the air. Whether or not properties of a phase space can be reproduced by observables is something we've yet to determine, so it is a priority in further to clarify it. Perhaps a possible approach would be to consider the case where observables are functions of phase space variables. After having a sufficiently robust model, we can start exploring its properties, and making predictions. As zebrafishes are used in genetic studies, it will be interesting to explore changes to the system when using mutants in a fight.

# Chapter 4

## Conclusion

We began by introducing the zebrafish fight as an example of an interesting example of social interactions between animals, in specific, contests. We addressed the challenges that exist when attempting to model these small number interactions. We presented some unsupervised techniques that can be used for a more principled model development, and ultimately, used a system of easily interpretable joint variables. We used a set of compound behavioral states for an initial analysis of the kinematic properties reflected in our variables. We embedded jointly-defined variables in a lower dimensional representation, and from the resulting map, we built a symbolic representation of the fight as a function of the density peaks in the map, we analyse the kinematic properties of the resulting clusters and identify them with a compound behavioral state.

The model we developed may capture distinct behavioral states not perceptible to human observation and labelling, and a larger structure of the dynamics of the fight by looking at the joint variables. It is also possible this picture will change with a different representation of variables, or with a more detail posture of the fish (namely including the fins).

It may be of interest to develop more on the built model of the interaction, specially in how the different eigenvalues are associated with long timescale behaviors. And how looking at different temporal resolutions will change the picture of the fight. Testing the robustness and dependency of the model to different parameters is fundamental in case we want to use it for any further analysis. If this observation is consistent across the different experiments to be done, then it brings an interesting framework for the analysis of a fight, namely how this fight dynamic changes with different mutations or environmental conditions.

As for what this project represents for the larger question of a physical model for behavior, it

doesn't even scratch the surface. That would be a fairly presumptuous assertion. What it does is bring a certain framework from which we can answer this larger question, and possible questions that come from it. Exploring social behaviors as a joint property has certain implications in the nature of behavior, so it would also be interesting to explore it from another perspective.

The end of this project comes with certainly more questions than answers, and yet I see this as a positive outcome. After all, what is the pursuit of science, if not a series of questions to be answered?

## 4.1 Appendix - Recovering Almost Invariant Sets in Dynamical Systems

It is relatively complicated to describe certain dynamical systems, especially chaotic ones. However, some representations allow us to infer properties of the system and make predictions, like almost invariant sets. Prior to understand almost invariant sets, it is perhaps useful to describe invariant sets first.

An invariant set is a region in the phase, where if a point in that region is selected and tracked over time, it will remain in the region.

Mathematically, it can be expressed:

Consider an operator  $T : U \rightarrow U$ . A subspace  $V \subset U$  is invariant if  $\forall x \in V, T(x) \in V$ . [70]

Almost invariant sets are defined as regions of relative stability in the phase space, where trajectories spend a lot of time before transitioning to another region. While it may be hard to describe the trajectory of chaotic systems, it is possible to describe their dynamics in function of these almost invariant sets [71].

The eigenmode projection recovers the invariant/almost invariant sets present in the Dynamical System, being possible to partition a phase space into almost invariant sets. The eigenvalue may be associated with a dynamic switching pattern between the two invariant sets.

$$\tau_{switch} = -\frac{\tau}{\log |\lambda|} \approx -\frac{1}{\log |\lambda_0|}$$

with  $\tau_{switch}$  being the average time between transitions, and  $\lambda_0$  being the eigenvalue at  $\tau = 0$ .

Our process of building a transition matrix from dynamic modes, is effectively a coarse-grained process for the estimation of almost invariant sets.

To understand what the UMAP process does to a typical dynamical system, we will apply the pipeline to a couple of well known systems with the purpose of having a clear understanding of the process.

We will use dynamical systems with two almost-invariant sets for simplicity.

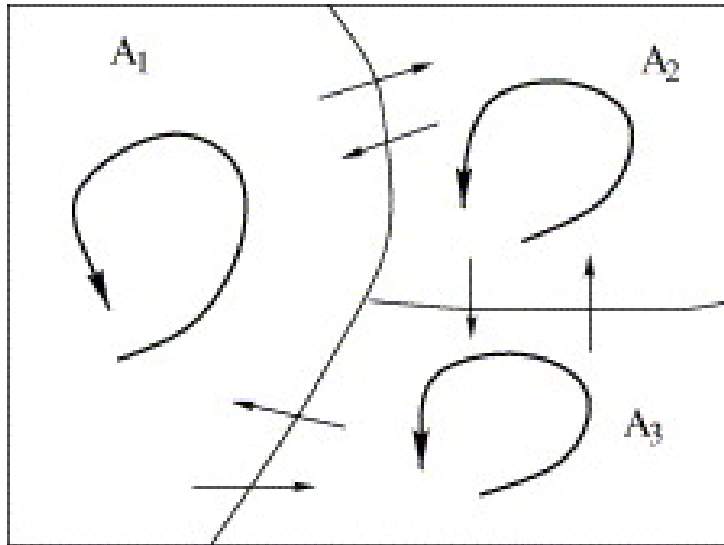


Figure 4.1: Representation of three almost-invariant sets (from [71])

### 4.1.1 Lorenz system

The Lorenz system is a classic and well known dynamical system that exhibits chaotic behavior [72], that has more than one almost invariant set. so it is the obvious first choice when it comes down to exploring dynamical properties. It is defined by:

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = x(\rho - z) - y \\ \dot{z} = xy - \beta z \end{cases}$$

We simulate the system with the following parameters:  $\sigma = 10$ ,  $\rho = 8/3$  and  $\beta = 28$ , with a  $\Delta t = 0.005$  s and  $T = 200$  s.

By having this many samples, these are the results we recovered:

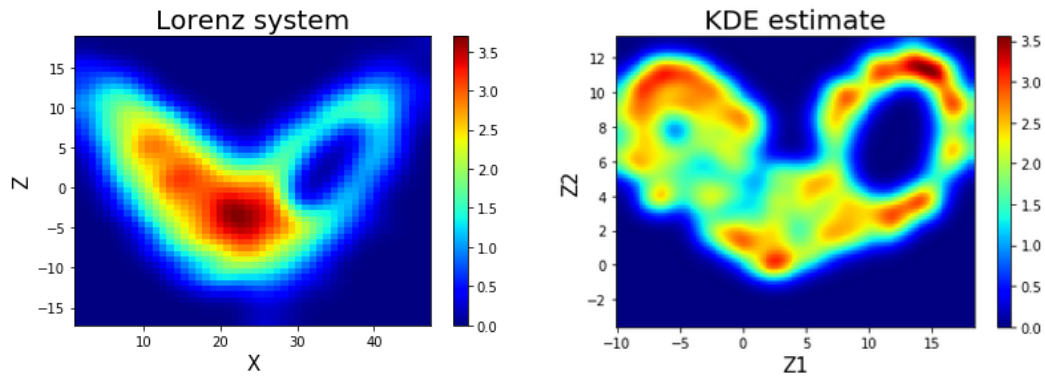


Figure 4.2: KDE of the XZ projection for the Lorenz System, and the recovered 2 dimensional projection

It is possible to see a resemblance between the projected UMAP result and the XZ projection of the Lorenz attractor. This suggests some of the overall topological structure of the attractor is preserved.

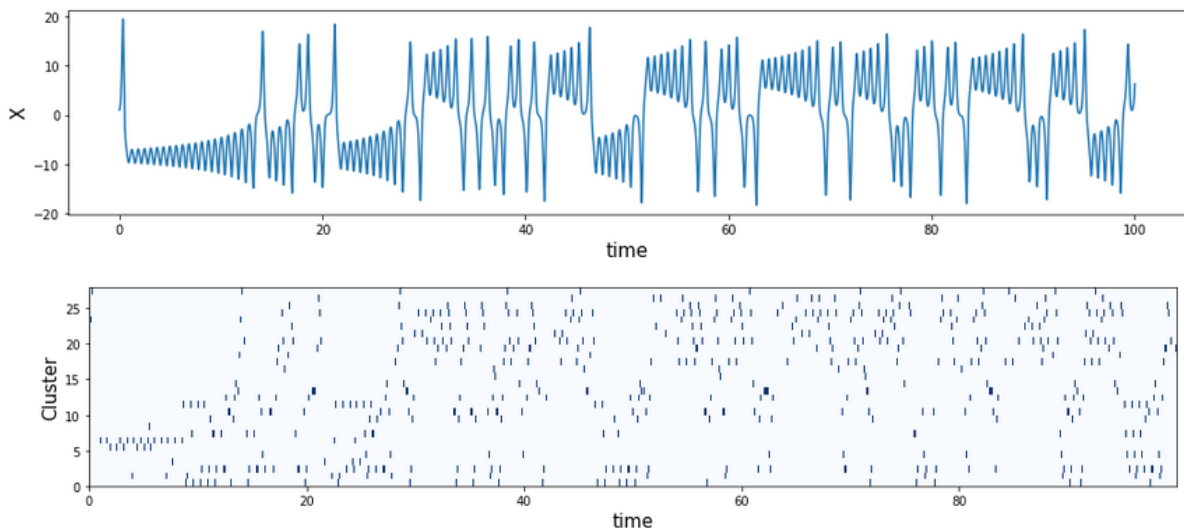


Figure 4.3: Time series of the X coordinate in the Lorenz system; "Ethogram" of the UMAP-embedded Lorenz system

The activity patterns across the different clusters is also reminiscent of the dynamical patterns in the X variable.

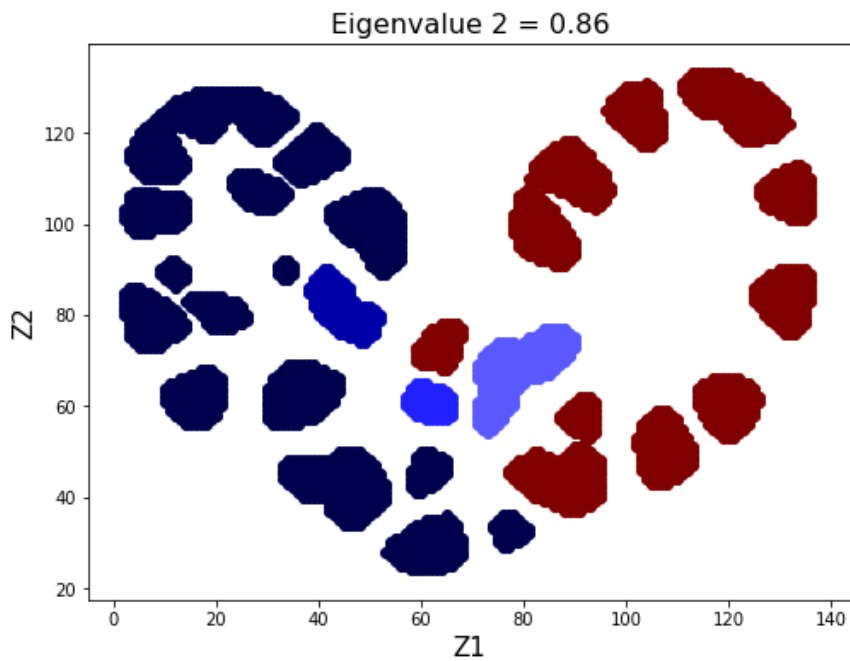


Figure 4.4: Eigenmode projection of the Lorenz System for the second largest eigenvalue

From the second eigenvalue, we are able to recover representations of the almost invariant sets present in the Lorenz system.

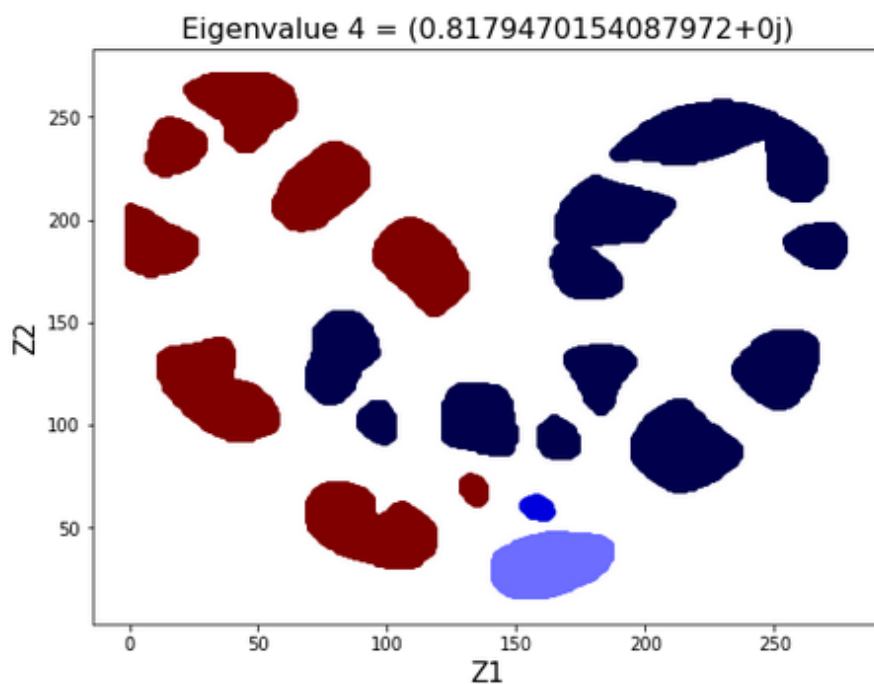


Figure 4.5: Eigenmode projection of the Lorenz System for the fourth largest eigenvalue

A second run gives us the fourth eigenvalue associated with it instead

Note: It seems the biggest factor determining this is the process from getting the clusters to making the transition matrix, another common thread is how these invariant set eigenmodes are



associated with real eigenvalues, this may be a key towards determining the properties of this method and improving it).

### 4.1.2 Chua's circuit system

Chua's circuit is an electronic circuit that exhibits chaotic behavior, and has two almost invariant sets [73]. It is described by the following equation:

$$\begin{cases} \dot{x} = \alpha(y - x - g(x)) \\ \dot{y} = x - y + z \\ \dot{z} = -\beta y \end{cases}, g(x) = \gamma x + \frac{1}{2}(\delta - \gamma)(|x + 1| - |x - 1|)$$

For the Chua system we use parameters as  $\alpha = 15, \beta = 25.58, \gamma = -5/7$  and  $\delta = -8/7$ , with a  $\Delta t = 0.01$  s and  $T = 200$  s.

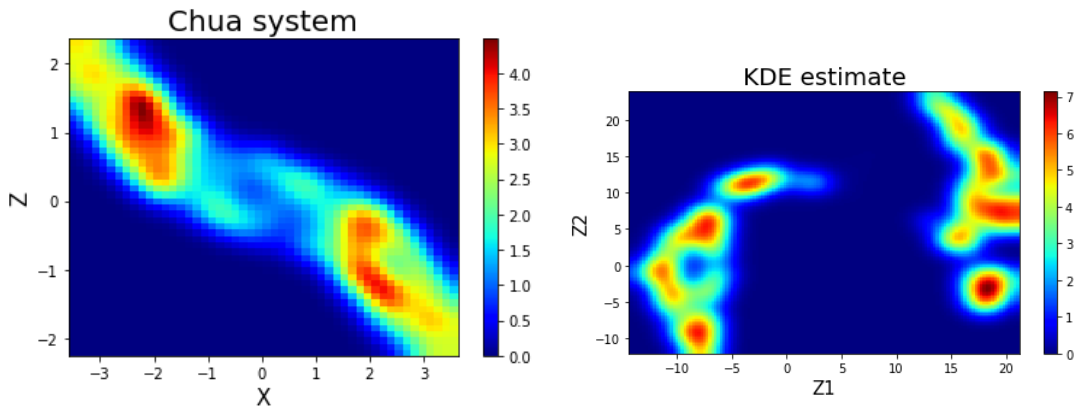


Figure 4.6: KDE of the XZ projection for the Chua's System, and the recovered 2 dimensional projection

Different from the first case, there is not a clear similarity between the embedding and the projection, however the two density clusters are clearly distinct, which is an interesting feature associated with the topological equivalency of the representation.

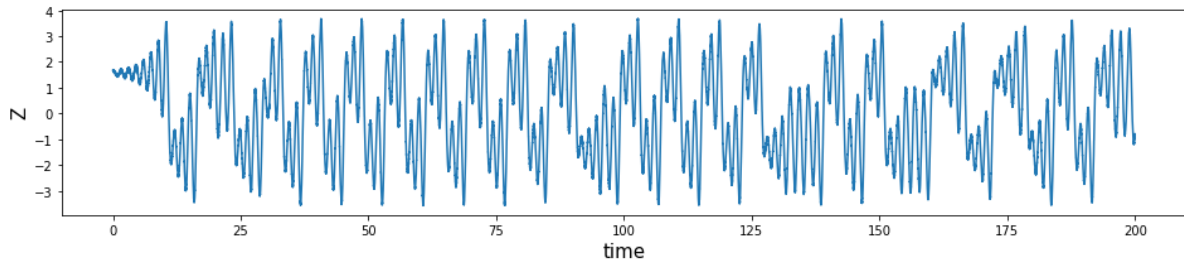


Figure 4.7: Time series of the Z coordinate in the Chua's system

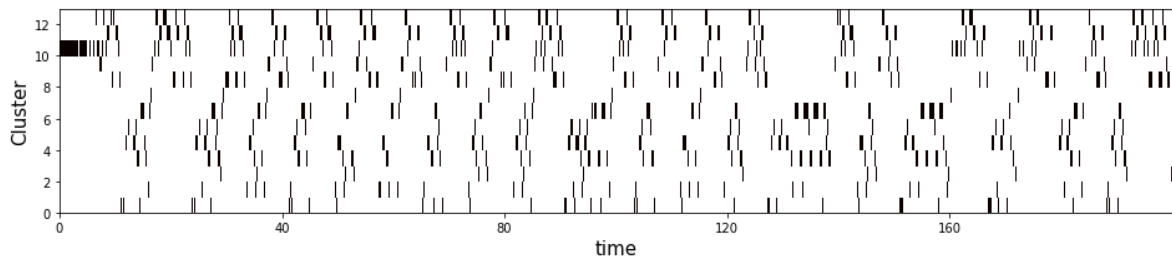


Figure 4.8: "Ethogram" of the UMAP-embedded Chua's system

The activity patterns across the different clusters is also reminiscent of the dynamical patterns in the Z variable, similar to the previous cases relative to the Lorenz System.

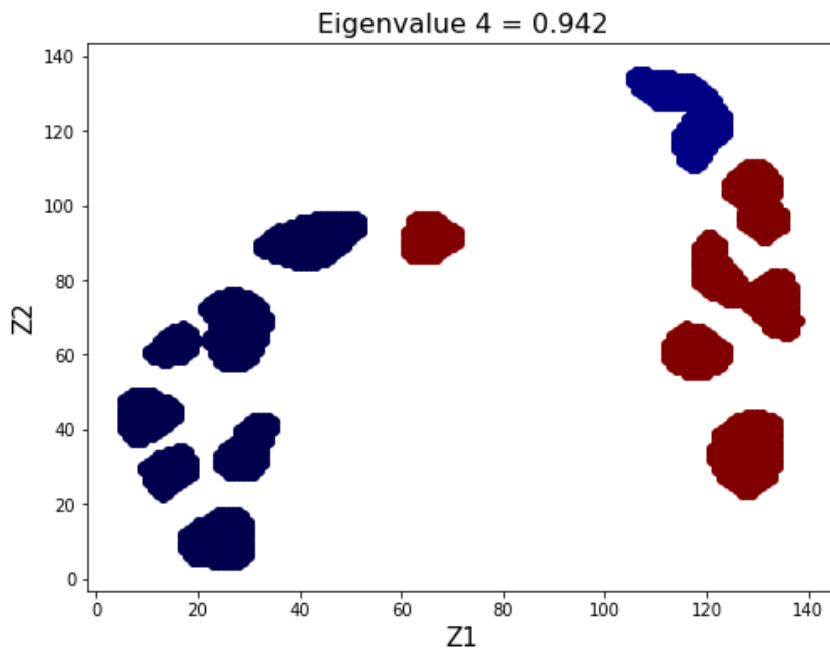


Figure 4.9: Eigenmode projection of the Chua's System for the fourth largest eigenvalue

In this case, the fourth eigenvalue seems to be the one associated with the eigenmode distinguishing different invariant modes.

### 4.1.3 Discussion

The UMAP process consists essentially of embedding data to a uniform manifold and projecting said manifold to a lower dimension [63]. So, a possible explanation for these results might be that the almost invariant sets form a uniform manifold, and therefore the properties preserved by the manifold projection would be equivalent to the dynamic properties of the almost invariant sets. If we can prove a particular invariant set is smooth and bounded, it can be embedded in a uniform manifold. Something we might have to consider is whether this is something similar to a coarse-grained view of the phase space, as we embedded phase space variables, explaining the observed invariant sets. There is a problem with consistency when it comes to detecting these invariant sets from the eigenmodes. The eigenmodes with which they are associated are highly inconsistent, and may be due to the construction of the transition matrix, as it does not seem to be fairly consistent across different runs. This may be a problem for systems for which we do not know what the description is, as the one explored in the experiment. That can be addressed by choosing a random seed, which was implemented in the main project.

Another thing we need to change is perhaps the fact we used the phase space variables directly in this case while in general, we may only have access to observable variables. If we can show we are able to recover invariant sets from UMAP projections, we can then argue that the behavioral map for the zebrafish fight we recovered not only has behavioral relevancy but also physical relevancy, as we can show that the invariant set of the social dynamic system is recovered through the observable variables of the system.

# Bibliography

- [1] A. Brown and B. de Bivort. Ethology as a physical science. *Nature Physics*, 14(7), 2018.
- [2] T. Vicsek and A. Zafeiris. Collective motion. *Physics Reports*, 517(3), 2012.
- [3] SJ. Helms, W.M Rozemuller, AC Costa, L Avery, GJ. Stephens, and TS. Shimizu. Modelling the ballistic-to-diffusive transition in nematode motility reveals variation in exploratory behaviour across species. *Journal of the Royal Society Interface*, 16(157):1–46, 2019.
- [4] G.J. Berman. Measuring behavior across scales. *BMC Biology*, 16(23), 2018.
- [5] J. Altmann. Observational study of behavior: Sampling methods. *Behaviour*, 49(3/4):227–267, 1974.
- [6] J. Kain, C. Stokes, Q. Gaudry, X. Song, J. Foley, R. Wilson, and B. De Bivort. Leg-tracking and automated behavioural classification in Drosophila. *Nature Communications*, 4(1), 2013.
- [7] C. Mendes, I. Bartos, Z. Márka, T. Akay, S. Márka, and R. Mann. Quantification of gait parameters in freely walking rodents. *BMC Biology*, 13(1), 2015.
- [8] J. Taylor, L. Bordeaux, T. Cashman, B. Corish, C. Keskin, E. Soto, D. Sweeney, J. Valentin, B. Luff, A. Topalian, E. Wood, S. Khamis, P. Kohli, T. Sharp, S. Izadi, R. Banks, A. Fitzgibbon, and J. Shotton. Efficient and precise interactive hand tracking through joint, continuous optimization of pose and correspondences. *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2016*, 35, July 2016.
- [9] S. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh. Convolutional pose machines. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4724–4732, 2016.

- [10] K. Branson, A. Robie, J. Bender, P. Perona, and M. Dickinson. High-throughput ethomics in large groups of drosophila. *Nature Methods*, 6:451–457, 2009.
- [11] P. Gao and S. Ganguli. On simplicity and complexity in the brave new world of large-scale neuroscience. *Current Opinion in Neurobiology*, 32:148–155, 2015.
- [12] S. Panda, J. B. Hogenesch, and S.A. Kay. Circadian rhythms from flies to human. *Nature*, 417:329 – 335, 2002.
- [13] R.J. Full and D.E. Koditschek. Templates and anchors: neuromechanical hypotheses of legged locomotion on land. *Journal of Experimental Biology*, 202(23):3325–3332, 1999.
- [14] G.J. Stephens, L. C. Osborne, and W. Bialek. Searching for simplicity in the analysis of neurons and behavior. *Proceedings of the National Academy of Sciences*, 108(Supplement 3):15565–15571, 2011.
- [15] AY Katsov, L Freifeld, M Horowitz, S Kuehn, and Clandinin TR. Dynamic structure of locomotor behavior in walking fruit flies. *eLife*, 6, 2017.
- [16] G.J. Berman, D.M. Choi, W. Bialek, and J.W. Shaevitz. Mapping the stereotyped behaviour of freely moving fruit flies. *Journal of the Royal Society Interface*, 11(99), 2014.
- [17] A.G. Barto and S. Mahadevan. Recent advances in hierarchical reinforcement learning. *Discrete Event Dynamic Systems*, 13(4):341–379, Oct 2003.
- [18] R. Dawkins. Hierarchical organisation: A candidate principle for ethology. *Growing Points in Ethology*, 1976.
- [19] A Gomez-Marin, GJ Stephens, and AE Brown. Hierarchical compression of *Caenorhabditis elegans* locomotion reveals phenotypic differences in the organization of behaviour. *Journal of the Royal Society Interface*, 13(121), 2016.
- [20] JC Marques, S Lackner, R Félix, and MB Orger. Structure of the zebrafish locomotor repertoire revealed with unsupervised behavioral clustering. *Current Biology*, 28(2):181 – 195.e5, 2018.
- [21] G.J. Stephens, B. Johnson-Kerner, W. Bialek, and W.S. Ryu. Dimensionality and dynamics in the behavior of *C. elegans*. *PLoS Computational Biology*, 4(4), 2008.

- [22] G.J. Stephens, M.B. De Mesquita, W.S. Ryu, and W. Bialek. Emergence of long timescales and stereotyped behaviors in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, 108(18):7286–7289, 2011.
- [23] AB Wiltschko, MJ Johnson, G Iurilli, RE Peterson, JM Katon, SL Pashkovski, VE Abraira, RP Adams, and SR Datta. Mapping sub-second structure in mouse behavior. *Neuron*, 88(6):1121–1135, 2015.
- [24] A. E.X. Brown, E. I. Yemini, L. J. Grundy, T. Jucikas, and W. R. Schafer. A dictionary of behavioral motifs reveals clusters of genes affecting *Caenorhabditis elegans* locomotion. *Proceedings of the National Academy of Sciences of the United States of America*, 110(2):791–796, 2013.
- [25] J.G. Todd, J.S. Kain, and B.L. De Bivort. Systematic exploration of unsupervised methods for mapping behavior. *Physical Biology*, 14(1), 2017.
- [26] U. Klibaite, G.J. Berman, J. Cande, D.L. Stern, and J.W. Shaevitz. An unsupervised method for quantifying the behavior of paired animals. *Physical Biology*, 14(1):015006, feb 2017.
- [27] SA Lee, G Vallortigara, M Flore, ES Spelke, and VA Sovrano. Navigation by environmental geometry: The use of zebrafish as a model. *Journal of Experimental Biology*, 216(19):3693–3699, 2013.
- [28] SA Lee, A Ferrari, G Vallortigara, and VA Sovrano. Boundary primacy in spatial mapping: Evidence from zebrafish (*Danio rerio*). *Behavioural Processes*, 119:116–122, 2015.
- [29] P. Goupillaud, A. Grossmann, and J. Morlet. Cycle-octave and related transforms in seismic signal analysis. *Geoexploration*, 23(1):85–102, 1984.
- [30] L. van der Maaten and G. Hinton. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.
- [31] The C. Elegans Sequencing Consortium. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science*, 282(5396):2012–2018, 1998.
- [32] JG White, E Southgate, JN Thomson, and S Brenner. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond. B*, 314, 1986.

- [33] R Gonzalez and R Woods. *Digital Image Processing*. Prentice Hall Press, 2002.
- [34] R Gonzalez and R Woods. *Classification and Regression Trees*. Taylor & Francis, 1984.
- [35] JH Baek, P Cosman, Z Feng, J Silver, and WR Schafer. Using machine vision to analyze and classify *Caenorhabditis elegans* behavioral phenotypes quantitatively. *Journal of Neuroscience Methods*, 118(1):9–21, 2002.
- [36] YC Fung. *A First Course in Continuum Mechanics*. Prentice Hall Press, 1993.
- [37] J Sznitman, PK Purohit, P Krajacic, T Lamitina, and PE Arratia. Material properties of *Caenorhabditis elegans* swimming at low reynolds number. *Biophysical Journal*, 98(4):617–626, 2010.
- [38] Y Rabets, M Backholm, K Dalnoki-Veress, and WS. Ryu. Direct measurements of drag forces in *c. elegans* crawling locomotion. *Biophysical Journal*, 107(8):1980–1987, 2014.
- [39] EE. Keaveny and AE Brown. Predicting path from undulations for *C. Elegans* using linear and nonlinear resistive force theory. *Physical Biology*, 14(2), 2017.
- [40] N. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library, 2007.
- [41] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 1984.
- [42] D. Struik. *Lectures on Classical Differential Geometry*. Courier Corporation, 1961.
- [43] E. Racca and A. Porporato. Langevin equations from time series. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 71(2):18–20, 2005.
- [44] OD Broekmans, JB Rodgers, WS Ryu, and GJ Stephens. Resolving coiled shapes reveals new reorientation behaviors in *C. elegans*. *eLife*, 5(September):1–17, 2016.
- [45] SW. Linderman, MJ. Johnson, AC. Miller, RP. Adams, DM. Blei, and L Paninski. Bayesian learning and inference in recurrent switching linear dynamical systems. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS 2017*, 54, 2017.

- [46] J.H. Ward. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301):236–244, 1963.
- [47] MA. Muñoz. Colloquium: Criticality and dynamical scaling in living systems. *Reviews of Modern Physics*, 90(3):31001, 2018.
- [48] AC. Costa, T Ahamed, and GJ. Stephens. Adaptive, locally linear models of complex dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 116(5):1501–1510, 2019.
- [49] M Enquist and O Leimar. Evolution of fighting behaviour: Decision rules and assessment of relative strength. *Journal of Theoretical Biology*, 102:387–410, 1983.
- [50] RH Payne and M Pagel. Escalation and time costs in displays of endurance. *Journal of Theoretical Biology*, 183:185–193, 1996.
- [51] Payne R. Gradually escalating fights and displays: the cumulative assessment model. *Animal Behaviour*, 56(3):651 – 662, 1998.
- [52] D Stuart-Fox. Testing game theory models: fighting ability and decision rules in chameleon contests. *Proceedings of the Royal Society B: Biological Sciences*, 273:1555 – 1561, 2006.
- [53] M Koops and J Grant. Weight asymmetry and sequential assessment in convict cichlid contests. *Canadian Journal of Zoology*, 71:475–479, 1993.
- [54] A Kalluef, M Gebhart, and AM Stewart. Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond. *Zebrafish*, 10(1):70–85, 2013.
- [55] R Oliveira, J Silva, and J Simões. Fighting Zebrafish: Characterization of Aggressive Behavior and Winner Loser Effects. *Zebrafish*, 8(2):73–81, 2011.
- [56] MS Abreu, AC Giacomini, R Genario, B Santos, LG da Rosa, KA Demin, EA Wappler-Guzzetta, and A Kalueff. Neuropharmacology, pharmacogenetics and pharmacogenomics of aggression: The zebrafish model. *Pharmacological research*, 141:602–608, 2019.
- [57] F Romero-Ferrero, M Bergomi, R Hinz, F Heras, and GG de Polavieja. idtracker.ai: tracking all individuals in small or large collectives of unmarked animals. *Nature Methods*, 16:179–182, 2019.



- [58] A Laan, M Iglesias-Julios, and GG de Polavieja. Zebrafish aggression on the sub-second time scale: evidence for mutual motor coordination and multi-functional attack manoeuvres. *Royal Society Open Science*, 5, 2018.
- [59] Z Cao, G Hidalgo Martinez, T Simon, S Wei, and YA Sheikh. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.
- [60] TD Pereira, DE Aldarondo, L Willmore, M Kislin, SS Wang, M Murthy, and JW Shaevitz. Fast animal pose estimation using deep neural networks. *Nature Methods*, 16:117–125, 2019.
- [61] AL Menke, JM Spitsbergen, APM Wolterbeek, and RA Woutersen. Normal anatomy and histology of the adult zebrafish. *Toxicologic Pathology*, 39(5):759–775, 2011.
- [62] S. Willard. *General Topology*. Dover Publications, 2004.
- [63] LMcInnes, J Healy, and J Melville. Umap: Uniform manifold approximation and projection for dimension reduction, 2020.
- [64] A. Van den Bos. Complex gradient and hessian. *IEE Proceedings, Visual, Image and Signal Processing*, 141(6):380 – 382, 1995.
- [65] J. Stewart. *Multivariate Calculus: Concepts and Contexts*. Brooks/Cole, 2005.
- [66] C Fiorio and J Gustedt. Two linear time union-find strategies for image processing. *Theoretical Computer Science*, 154:165 – 181, 1996.
- [67] S. Resnick. *Adventures in Stochastic Processes*. Birkhauser, 1992.
- [68] D Aldous and JA Fill. Reversible markov chains and random walks on graphs, 2002. Unfinished monograph, recompiled 2014, available at <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- [69] T. Ahamed, AC. Costa, and GJ. Stephens. Capturing the continuous complexity of behaviour in *Caenorhabditis elegans*. *Nature Physics*, 17(2):275–283, 2021.
- [70] J. Meiss. *Differential Dynamical Systems*. SIAM, 2007.

- [71] Gary Froyland. Statistically optimal almost-invariant sets. *Physica D: Nonlinear Phenomena*, 200(3):205–219, 2005.
- [72] EN Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, 1963.
- [73] M. Dellnitz and O. Junge. Almost invariant sets in chua’s circuit. *International Journal of Bifurcation and Chaos*, 07(11):2475–2485, 1997.