Data-driven modeling of morphological dynamics and intracellular transport of organelles

Submitted in partial fulfillment of the requirements for
the degree of
Doctor of Philosophy
in
Department of Biomedical Engineering

Hao-Chih Lee
B.S., Life Science, National Taiwan University
M.S., Mathematics, National Taiwan University

Carnegie Mellon University
Pittsburgh, PA

September, 2015

Abstract

Data-driven modeling is essential to understanding complex cellular processes. In this thesis, we present a series of studies of analyzing morphological dynamics and intracellular transport of organelles using techniques of mathematical modeling, image processing and machine learning. We first characterized the morphology of organelles, focusing specifically on mitochondria. We developed a morphological data processing pipeline. Using this pipeline, we discovered a bi-modal distribution of mitochondrial sizes, with a stable mean value in each mode. We then developed a data-driven model to investigate how fusion/fission of mitochondria modulates their sizes. For further analysis of morphology of mitochondria as well as other cellular components, we developed a general purpose machine learning algorithm, which we refer to as shape component analysis (SCA). We used it for dimension reduction and classification of mitochondrial morphology and protein geometry.

In addition to studying the morphological dynamics of cellular organelles using data-driven modeling, we investigated the intracellular transport of organelles. We first proposed a probabilistic model for studying the relation between mitochondrial size and the velocity of their active transport. The proposed model not only explained the relation between mitochondrial size and velocity observed in experiments under normal conditions but also suggested a novel relation under changed conditions. Further analysis of the proposed model also suggested a way to evaluate the binding/unbinding rates of motors carrying the mitochondria. We further studied the global organization of organelle transport. We proposed an image processing framework to characterize the spatiotemporal dynamics of intracellular transport in terms of the spatial localization of stationary organelles and the spatiotemporal patterns of organelle movement, respectively. We used this framework to analyze time-lapse images of Lamp1 transport and found different global transport patterns. Overall, our studies produced both computational modeling methods and specific biological results for quantitative and systems-level understanding the complex behavior of intracellular organelles.

Acknowledgments

Firstly, I would like to express my gratitude to my advisor Ge Yang for the continuous support of my Ph.D. study and related research. I would like to thank the rest of my thesis committee Christopher Langmead, Jelena Kovačević and Robert Tilton for their insightful comments and encouragement, but also for the hard question which motivated me to widen my research from various perspectives.

I would also like to thank my fellows in the lab for helpful support and inspiring discussion. Many thanks to Yiyi Yu and Jackie Chen for kindly data sharing and technical support. This thesis would be an impossible mission without their help. A special thank goes to Huan-Kai Peng for his suggestion on presentation and endless trash talking.

I would also like to thank my parents, my brother and sister. They were always supporting me to pursue my dream with their best wishes.

The work was supported by US NSF grants MCB-1052660, DBI-1149494 and DBI-1052925, and a Ji-Dian Liang Graduate Research Fellowship and a Bertucci Graduate Research Fellowship from the Carnegie Institute of Technology, Carnegie Mellon University.

Contents

List of Tables

List of Figures

Chapter 1

Overview

1.1 Thesis statement

Biology depends critically on using experimental methods to study complex cellular processes. Given the complexity of such processes, a commonly used strategy is to perturb them in a well-controlled manner and qualitatively describe how they respond. Complementary to this strategy, quantitative analysis and mathematical modeling are increasingly used to study biological processes. Still, many challenges remain in resolving the complexity of biological processes even with quantitative approaches. This thesis research aims to overcome some of these challenges.

This thesis is centered around quantitative analysis of organelle dynamics. Organelles are specialized and membrane-enclosed units within a cell that serve diverse cellular functions and exhibit diverse dynamic behavior. Taking mitochondria as an example, they are known to show diverse behaviors in terms of their localization, morphology and biogenesis. We believe that organelles are excellent subject of study for quantitative approaches because 1) they are relatively simple in constitution and can be considered as individual physical units and 2) they demonstrate diverse biological

functions. In this regard, this thesis aims to achieve the following two goals

1. *To characterize dynamic behaviors of organelles using imaging-based methods.*

2. *To decipher the governing principles of organelle dynamics using mathematical modeling.*

1.2 Outline of the thesis

In this thesis, we present a series of work on characterizing and analyzing the morphological dynamics and active transport of organelles. Using mitochondria as an example, the intracellular dynamics of mitochondria is summarized as in Figure 1.1. Mitochondria are known to shown diverse dynamics. As an example, the morphology of mitochondria is constantly remodeled by the fusion and fission machinery. The localization of mitochondria is also changing over time by active transport of motor proteins. Furthermore, the distribution of mitochondria over the whole cell might be regulated by the cellular status. Maintenance of these mitochondrial dynamics are critical to the survival of cells while defects in either mitochondrial morphology or localization are related to neurodegenerative diseases. In this thesis, we aim to address the following questions:

1. How is the fission/fusion of mitochondria related to their morphology?

2. How is the mitochondrial morphology related to their motility?

3. Is the distribution of these organelles regulated at the whole cell scale?

To address these questions, we developed computational methods for studying organelle dynamics. The main results of this thesis are summarized as follows.

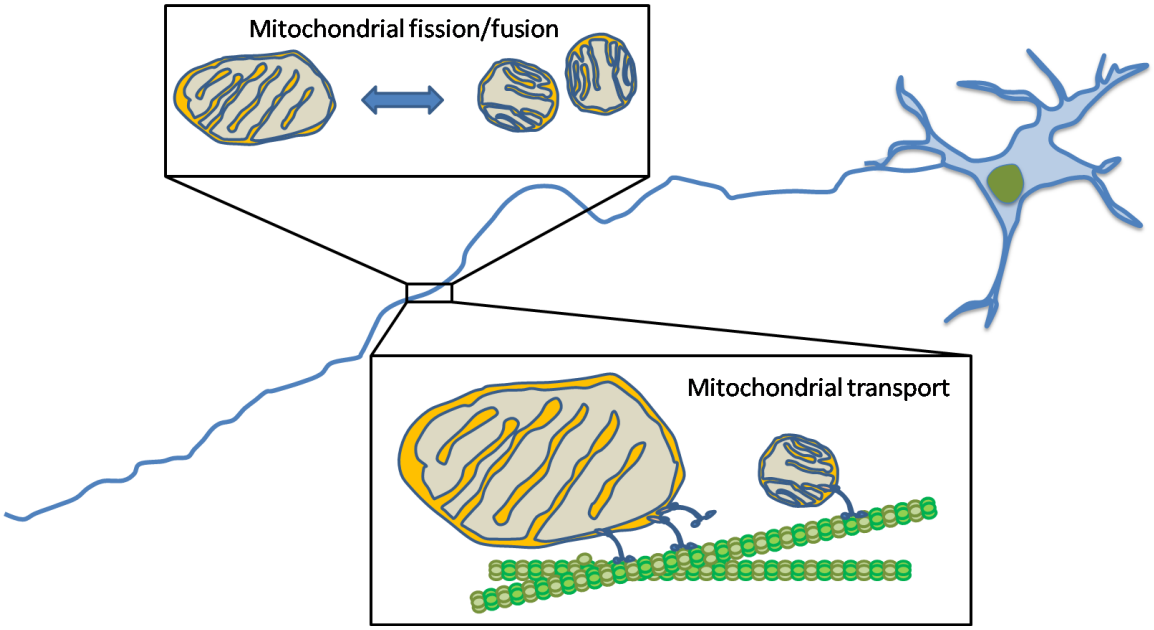

Figure 1.1: **Dynamic behavior of mitochondria in the axon.** Mitochondria are known to change their morphology through fusion/fission as well as their distribution through transport. Dynamics of mitochondria is critical to maintain their distribution at the whole cell scale.

1. In Chapter 2, we investigated the dynamics of mitochondrial morphology. We developed a data processing pipeline for quantifying mitochondrial morphology and motility. Using this pipeline, we developed a data-driven model to account for the fission/fusion mediated mitochondrial size distribution. The proposed model demonstrated can predict the relative occurrence of fission and fusion from the distribution of mitochondrial size.

2. In Chapter 3, we investigated the size-velocity relation of mitochondria within the axon. We first proposed a hierarchical model to account for the relation between mitochondrial size and velocity. The proposed model explains the size-velocity relation observed in in experiments and suggests a novel size-velocity relation given different spatial distribution of axonal microtubules. Furthermore, we demonstrate that it is possible to infer the binding behavior of motors carrying the mitochondria from the distribution of mitochondrial velocity.

3. In Chapter 4, we studies the global organization of Lamp1 transport. While organelles are individual units in cells, we wonder if they are regulated globally to meet an optimal distribution. Although to fully address this question is beyond the scope of this thesis, we, for the first time, designed a method to investigate the global dynamics of lysosomal transport. In chapter 4, we present an image processing framework to represent Lamp1 transport at the whole cell scale.

4. In Chapter 5, we developed methods for studying biological shapes such as mitochondrial morphology and protein geometry. We developed a technique for nonlinear dimension reduction of 2D and 3D biological shape representations on their Riemannian spaces. A key feature of this technique is that it approximately preserves distances between different shapes under reduced dimensions.

We demonstrated an application of this technique by combining it with nonlinear mean-shift clustering on the Riemannian spaces for unsupervised clustering of shapes of cellular organelles and proteins.

1.3 Summary of research contributions of the thesis

The research contributions of this thesis are summarized as follows:

1. In Chapter 2, we provided insights on how sizes of mitochondria are related to their fusion/fission dynamics. We modeled the fission/fusion dynamics based on queuing theory and conducted a rigorous analysis that revealed the dependency between mitochondrial fusion and fission. Based on our analysis, we also developed a way to predict the relative occurrence rate of mitochondrial fission and fusion. The developed model is general and can be adapted to investigate biological regulation on mitochondrial fission and fusion.

2. In chapter 3, we provided insights on how mitochondrial sizes are related to their transport velocities. We developed a model and conducted corresponding analysis. Our model and analysis revealed two kinds of size-dependent mitochondrial transport. Among these two types of dependency, one has been observed in the experimental data while the other remains to be tested in experiments.

3. In Chapter 4, we designed an image-based method for studying the organelle transport at the whole cell scale. Our method supports direct visualization of whole-cell scale organelle transport and is robust to image noise. Our method also revealed different transport patterns, which is crucial to understanding the

global control on organelle transport at the whole cell scale.

4. In Chapter 5, we developed a dimension reduction technique for analyzing biological morphology. One critical challenge of analyzing morphological data is in handling its intrinsic nonlinear structure originated from registration between objects. Our method is designed to preserve this nonlinear structure while reducing dimensionality. We demonstrated the superior performance of the proposed method in comparison to other existing methods and its application to analyzing biological morphology.

Chapter 2

Data-driven modeling of mitochondrial morphology

2.1 Introduction

Mitochondria are critical organelles in eukaryotic cells. Contrary to the traditional thinking of mitochondria as static powerhouses providing energy for cells, recent studies have revealed rich and dynamic behaviors of mitochondria, including their frequent movement and morphological changes [1]. Transported by molecular motors, mitochondria change their localization to meet changing energy needs within the cell. Also, mitochondrial morphology is remodeled constantly by fusion and fission processes.

The biological implication of fusion/fission process is center to the control on mitochondrial morphology and quality. Despite their fundamental roles, researches are still limited to either quantifying static mitochondrial morphology [2] or building models purely from theoretical perspective [3]. In this work, we first integrated par-

ticle tracking and image segmentation to retrieve the mitochondrial dynamics from image sequences. Based on exploratory analysis on collected data, we proposed a probabilistic model to see how fusion/fission is shaping mitochondrial morphology.

In this chapter, we quantitatively characterized mitochondrial morphology. The goal is twofold. First we aim to determine the major factors of mitochondrial morphology in the axon to facilitate further analysis and modeling on mitochondrial dynamics in the next chapter. For the other, we want to study how the fission/fusion mechanism is shaping the distribution of mitochondrial morphology.

The rest of this chapter is organized as follows: data collection, processing procedures and subsequent statistical analysis are first summarized in Section 2.2. An exploratory analysis on mitochondrial morphology is presented in Section 2.3. In Section 2.4, we proposed a probabilistic model of the fission/fusion process to investigate how the mitochondrial size is regulated by fission/fusion process. We concluded with a summary and a short discussion.

2.2 Methods

In this section, the overall workflow to retrieve morphological dynamics from images is presented. This approach consists of three steps, including 1) image collection and image segmentation, 2) mitochondrion tracking; and 3) morphology representation. The pipeline of data pre-processing is shown in Figure 2.1.

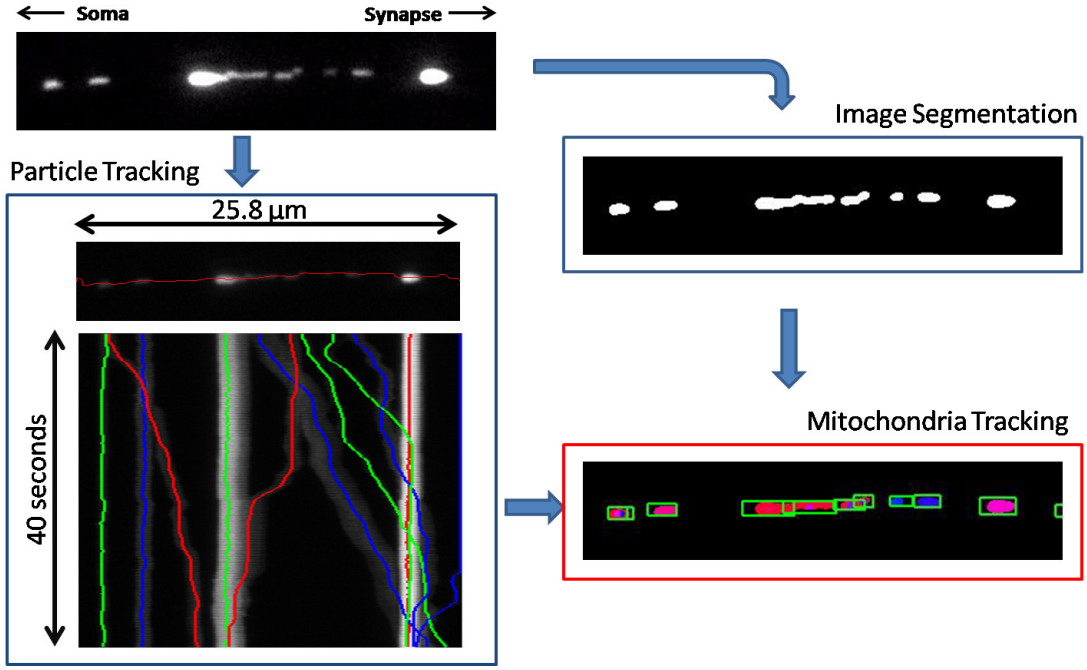

Figure 2.1: **Work flow of data processing**. Image sequences are processed by active-mask segmentation and particle tracking algorithms separately. Afterward, the processed results are combined to recover the morphology as well as motility dynamics of mitochondria.

2.2.1 Image collection and segmentation

Mitochondrial dynamics in segmental nerves of dissected Drosophila 3rd instar larvae is visualized by fluorescence live imaging. The imaging is performed on a Nikon Eclipse Ti-E inverted microscope with 5-frames-per-second sampling rate and 100x magnification. Sample drift is corrected by imageJ plug-in software. The effective pixel size is 64.5nm. Active-mask algorithm is used to segment mitochondria from background [4]. An example of collected image and segmented mask are shown in Figure 2.2.

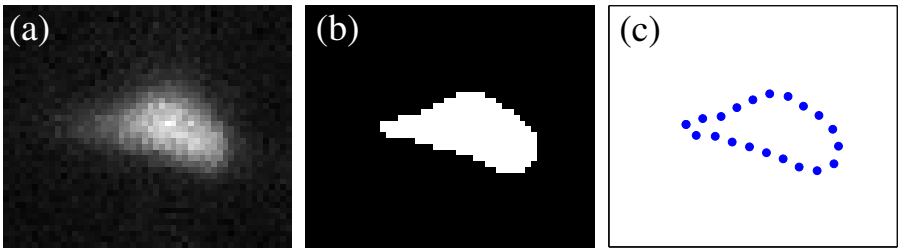

Figure 2.2: **Landmark representation of a mitochondrion.** (a) Raw image. (b) Segmented image. (c) Landmark representation. Only 20 landmarks are plotted for demonstration.

2.2.2 Tracking mitochondria

Following image segmentation, trajectory of each mitochondrion are reconstructed using software published in [5]. This software outputs trajectories of mitochondria as well as their motility measurements, such as types of motion (anterograde, retrograde, reversing and stationary) and velocity. Henceforth, a point in a trajectory is referred to as a *track point*.

2.2.3 Integration of mitochondrial localization and morphological dynamics

The outputs from image segmentation algorithm and tracking software are two sets of independent measurement about mitochondria morphology and their movements. Here we describe our method to recover the morphological dynamics by combining results from particle tracking software and the segmentation algorithm. One of the major challenge is analysis artifacts caused mainly by overlapped mitochondria (See Figure. 2.3). We propose a two-stage processing to minimize this artifact. The track points in tracking trajectories and connected components in segmented images are first identified. In the first stage, the number of track points within one connected

components is checked. If there is only one track point, this connected component will be identified as a member of the corresponding trajectory. If there is more than one track point, this region will be discarded to avoid artifacts in subsequent analysis. In the second stage, the average size of each mitochondrion is computed as the average size of bounding boxes of segmented regions. Then each connected components are discarded if their average bounding boxes are overlapped with others. An average bounding box for each trajectory is calculated as the third quartile among bounding boxes of all its corresponding non-overlapped regions from first stage. This is to avoid a problem of active-mask algorithm, which may split an overlapped region into two separate regions and lead to confusion in subsequent processing. Subsequently, morphological features, particularly area and aspect ratio of the best fitted ellipse, can be calculated from these connect components.

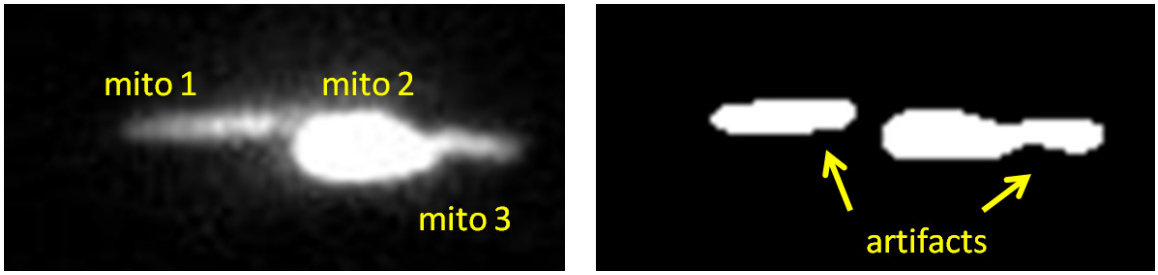

Figure 2.3: **Analysis artifacts due to overlapping mitochondria.** Upper panel shows the cropped image of three overlapping mitochondria. Lower panel shows the segmentation result having only two connected components.

2.2.4 Representing mitochondrial morphology by morphological features

Based on segmented binary regions, 9 commonly used features are computed to describe the mitochondrial morphology. These features are all invariant of rotations and can be classified into three categories: 2 of them are area-based features, another 4

features are related to length and the other 3 are ratios of chosen features and hence invariant of scaling. Table 2.1 summarized these morphological features.

morphological features		
area related	**length related**	**scale free**
area	perimeter	compactness
convex hull area	major length	aspect ratio
	minor length	skeleton-area ratio
	skeleton length	

Table 2.1: Descriptors for mitochondrial morphology

2.2.5 Landmark representation of mitochondrial morphology

To provide a comprehensive analysis of mitochondrial morphology, we use both morphological features and landmark representation to describe mitochondrial morphology. While morphological features, such as segmented area, are straightforward to analyze and interpret, the design of features requires prior knowledge on which feature should be chosen and hence may lead to biased description. To resolve this issue, we also adapt landmark representation to provide a compensatory description of the mitochondrial morphology. The morphology of a segmented connected region is represented by a set of control points positioned along its contour. This representation is called landmark representation. To be precise, supposed that a planar shape is represented by D control points along its contour (**Fig. 2.2**). This ordered sequence of control points is referred to as a *configuration*. A planar shape configuration is represented by a D-dimensional complex vector $\mathbf{Y}$, with $\mathbf{Y}_j = S_{1,j} + iS_{2,j}$, where $S_{k,j}$ records k-th coordinate of the j-th control point [6].

To compute the landmark representation, the contour is first extracted from the connected component and then fitted by a periodic cubic spline [7]. Then 200 equally spaced points are sampled from this cubic spline. The number of landmarks is chosen

empirically to ensure adequate descriptions of most contours. Since mitochondrial morphology may fluctuate over time, we calculated the mean shape among each trajectory for noise removal. Details of computing the mean shape are provided in the Chapter 5.

2.2.6 Nonlinear dimension reduction and mean-shift clustering

Dimension reduction is an essential tool for analyzing and understanding high dimensional data. A wide range of dimension reduction techniques have been developed [8]. However, for effective dimension reduction of biological shape representations, it is crucial to take into account their specific structures and properties. To give an example, biological shapes are often represented by points on high-dimensional Riemannian spaces [6]. Indeed, nonlinear Riemannian geometry of shape spaces is proposed as a tool of choice to depict geometric differences between shapes [9].

We developed a nonlinear dimension reduction technique on the shape spaces [10, 11]. We also applied our dimension reduction approach and nonlinear mean-shift clustering for unsupervised clustering of shapes of mitochondria and proteins. Experimental results confirmed that the proposed dimension reduction technique, when combined with mean-shift clustering, provided generally equivalent clustering performance but at significantly reduced processing time. This dimension reduction technique and unsupervised clustering are presented in Chapter 5 due to its high demanding in mathematics.

2.3 Investigating mitochondrial morphology

After processing, each mitochondrion is described by 1) a trajectory and corresponding motility measurements, 2) a vector consists of morphological features and 3) its mean shape represented by landmarks. These descriptors will be used in subsequent data analysis.

2.3.1 Size and aspect ratio describe mitochondrial morphology in the axon

To determine the major factors in mitochondria morphology, we collected morphology features from 980 wild-type mitochondria, and performed principal component analysis on these features. 9 morphology features are selected as described in previous section. Results of principal component analysis showed that 2 principal components (PCs) [12] are enough to explain 93% of the data. The biplot and the loading vectors of each descriptor are shown in Figure 2.4. It can be seen the major and minor length are most representative features among all designed features. However, the major feature l_M and minor length l_m are constrained by the inequality $l_M \geq l_m$. To get a parametrization of mitochondrial morphology without any constraints, we proposed to use their product ($l_M \cdot l_m$) and ratio (l_m/l_M) to equivalently describe the mitochondrial morphology. In fact, this two features represent the area and aspect ratio of the best fit ellipse of mitochondrial shape.

2.3.2 Big and round mitochondria generally do not move

The relation between morphology and motility is analyzed. Figure 2.5 shows the scatter plot of mitochondrial size versus aspect ratio. While the upper right half of the

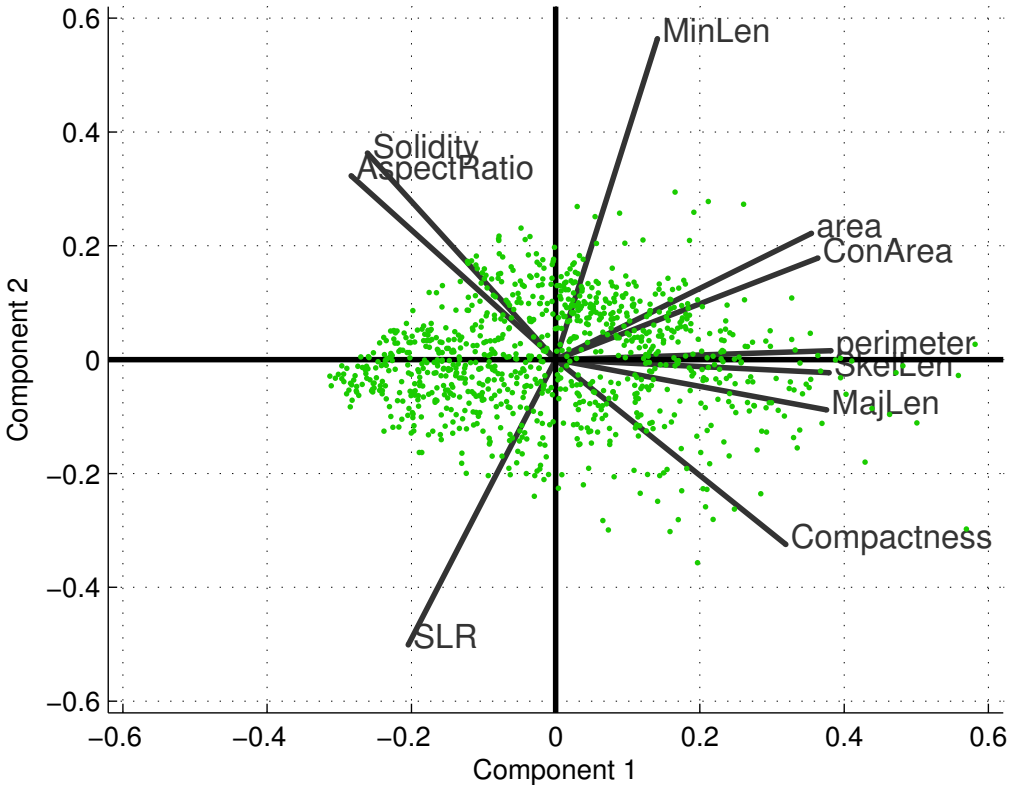

Figure 2.4: **Visualization of morphological descriptors**. Left: Morphological features visualized by principal component analysis. Right: Morphological features visualized by nonnegative matrix factorization. In both panels, data scattering is visualized in the subspace spanned by the first two optimal basis. Lines projecting from the origin show the loading vectors of the original morphological features.

graph is occupied by stationary mitochondria (blue dots), the moving (red triangles) and stationary mitochondria share the lower left of the graph. This suggests a general relation between morphology and motility, namely big and round mitochondria usually do not move. While the size-velocity relation among moving mitochondria is more intriguing upon closer examination, we will focus on stationary mitochondrial in this chapter. Analysis of moving mitochondria will be presented in the next chapter.

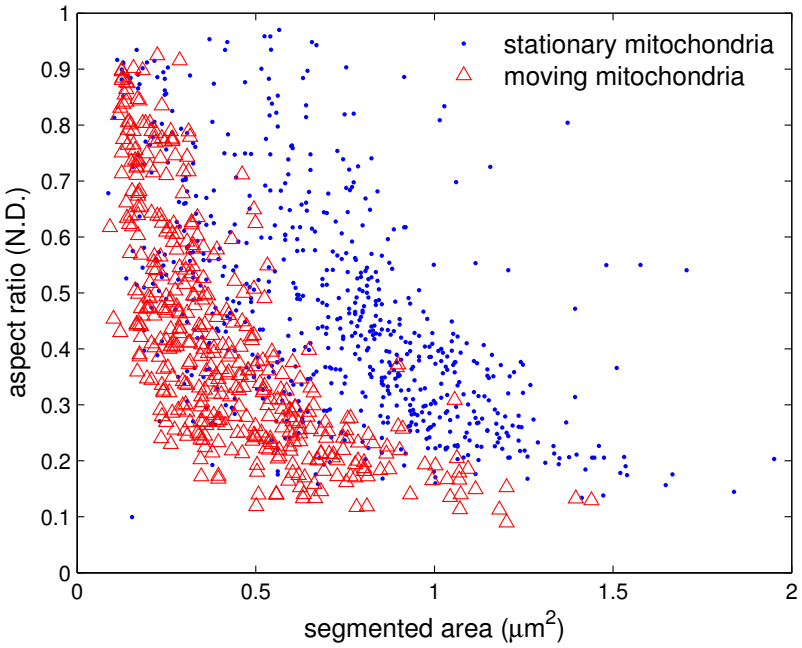

Figure 2.5: **Morphological difference in moving and stationary mitochondria.** x and y axes indicate mitochondrial size and aspect ratio respectively. Moving mitochondria are labeled in blue while stationary mitochondria are labeled in red.

2.3.3 Size of stationary mitochondria follows a bi-modal distribution

We also examined the changes in mitochondrial size distribution over space. Figure 2.6 shows histograms of stationary mitochondria's size. We observed that mitochondrial sizes follow a bi-modal distribution with changing portion in each mode over space. This observation is confirmed by fitting a Gaussian mixture model to mitochondrial size distribution. Figure 2.7 shows the mean and portion of each mode in fitted Gaussian mixture models. While the value of each modes remains stable, portions are changing over space. Larger mitochondria appear more frequently toward the synapse. On the contrary, the size of moving mitochondria still follows a bi-modal distribution with dominant numbers in the group of smaller size. This suggests the

mitochondrial size is regulated differently according to whether a mitochondrion is moving or not.

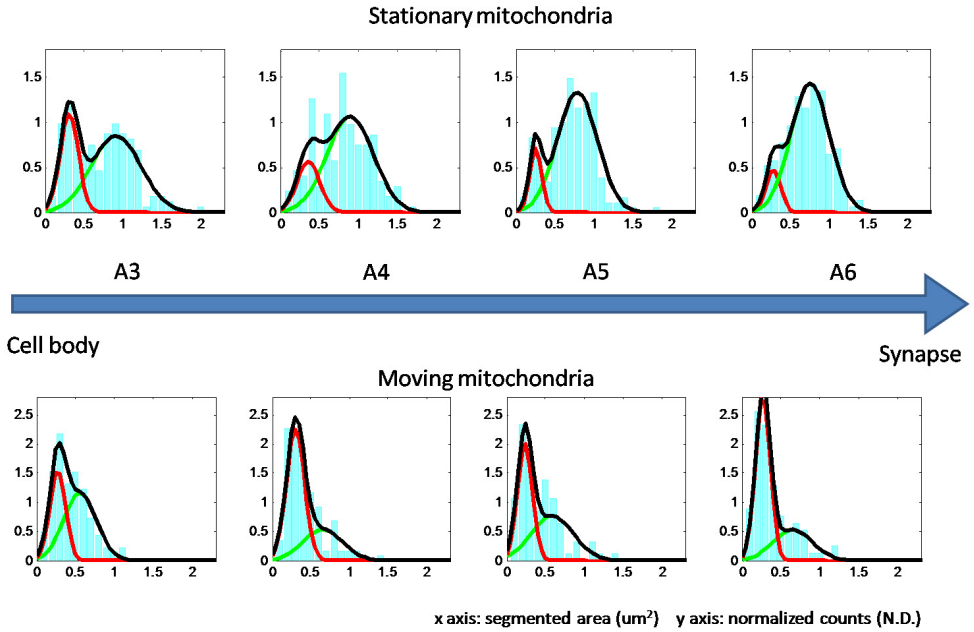

Figure 2.6: **Evolving mitochondrial size over space**. Size distribution of moving and stationary mitochondria are shown in the lower and upper panels respectively. Curves in each histogram depict the fitted Gaussian mixture model.

2.4 Inferring fission/fusion dynamics from mitochondrial size distribution

Mitochondrial size are known to be regulated by fission/fusion mechanism. A recent study observed that fission and fusion in the axon are mostly involved with stationary mitochondria [13]. Given this observation, we hypothesize that stationary mitochondria of larger size are end products of repetitive fission and fusion. Following we present a probabilistic model of mitochondrial fission and fusion to see if the observed distribution of mitochondrial size is solely shaped by fission/fussion

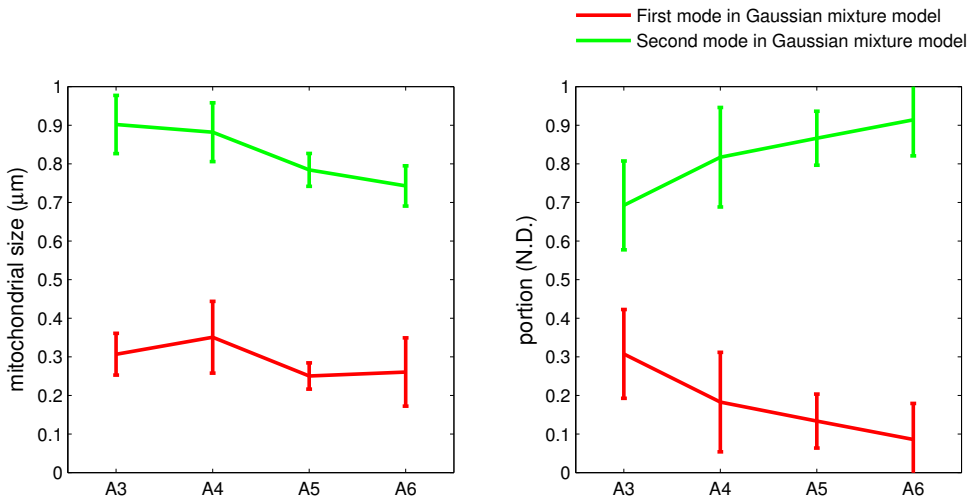

Figure 2.7: **Statistics of fitted Gaussian mixture models**. x-axis shows the labels of body segments, from proximal end (A3) to distal end (A6). y-axis represents mean (left panel) and portion (right panel) of each mode. Error bars show the 95% confidence intervals

mechanism.

2.4.1 Modeling mitochondrial size modulated by fission and fusion independently

Motivated by the data analysis in Section 2.3, we made the following assumption on mitochondrial fission/fusion process:

1. We assume a stationary mitochondrion is fixed in space with a time-varying mass $S(t)$.

2. We assume fusion to $S(t)$ occurs at rate λ_{fu}. At each fusion event, a mass M (followed the size distribution of moving mitochondria) is added to $S(t)$.

3. We assume fission of $S(t)$ occurs at rate λ_{fi}. We assume that, at each fission

event, the mass $S(t)$ is changed to $pS(t)$ for a constant $p \in (0,1)$.

Before proceeding further, we summarized notations that will be used in the following. We use X_n and Y_n to denote the time span between n-th and $n+1$-th fusion and fission events respectively. The cumulative waiting time up to n events is denoted by $T_{fu}(n)$ and $T_{fi}(n)$. i.e., $T_{fu}(n) = \sum_{i=1}^{n} X_i$. We also define $N_{fu}(\Delta t)$ and $N_{fi}(\Delta t)$ the counting processes of the occurrence of fusion (to $S(t)$) and fission (of $S(t)$) within any time span Δt. These two counting processes are defined by

$$N_{fi}(\Delta t) = \max\{n : \sum_{i=1}^{n} X_i \le \Delta t\},\ N_{fu}(\Delta t) = \max\{n : \sum_{i=1}^{n} Y_i \le \Delta t\}.$$

Note that assumptions (2) and (3) are equivalently defined by either 1) X_i and Y_i follow exponential distribution with mean $1/\lambda_{fu}$ and $1/\lambda_{fi}$ or 2) $N_{fu}(\Delta t)$ and $N_{fi}(\Delta t)$ are Poisson counting processes with rates λ_{fu} and λ_{fi} respectively.

Overall, we propose to model the fusion/fission modulated mitochondrial size $S(t)$ by

1. Initially $S(0) = S_0$.

2. Given cumulative waiting time $T_{fu}(n)$ and $T_{fi}(n)$ of fusion and fission, $S(t)$ is changed according to the rules

   - if $t = T_{fu}(n)$, the mass changes from $S(t)$ to $S(t) + M_n$. M_n are random variables with a common distribution f_M.
   - if $t = T_{fi}(n)$, the mass changes from $S(t)$ to $p \cdot S(t)$ with a fixed constant p.

To analyze how the mitochondrial size evolves over time, we first define the imbedded Markov chain Q_n [14]. The imbedded Markov chain Q_n describes the mitochondrial size right after n-th fission event, i.e.,

$$Q_n = S(T_{fi}(n)^+).$$

Given the independent arrivals of fusion events, the mass change from Q_n to Q_{n+1} is given by

$$Q_{n+1} = Q_n + \sum_{i=1}^{N_{fu}(Y_{n+1})} M_i,$$

while $\sum_{i=1}^{N_{fu}(Y_{n+1})} M_i$ is the overall mass added due to $N_{fu}(Y_{n+1})$ fusions happening within the time span Y_{n+1}. The mitochondrial size $S(t)$ can therefore be formulated by

$$S(t) = Q_{N_{fi}(t)} + \sum_{i=1}^{N_{fu}(C(t))} M_i$$

where $C(t)$ is the current lifetime of fission events up to time t defined by

$$C(t) = t - T_{fi}(N_{fi}(t)).$$

In the following we will drop the subscripts for random variables Y_n and M_i for notational simplicity since they are all independently and identically distributed.

2.4.2 Independent fission/fusion leads to oversized mitochondria

Independent fission/fusion with constant rates is typically assumed in modeling mitochondrial dynamics [3]. Using the model presented in section 2.4.1, the mean and

variance of mitochondrial size after repeating fusion and fission is computed as follows:

$$\begin{aligned} \mathbb{E}S_\infty &= \lim_{t\to\infty}\mathbb{E}S(t) &&= \frac{1}{1-p}\frac{\lambda_{fu}}{\lambda_{fi}}\mathbb{E}M \\ varS_\infty &= \lim_{t\to\infty} varS(t) &&= \frac{1}{1-p^2}\cdot\left((\mathbb{E}M)^2\cdot(\frac{\lambda_{fu}}{\lambda_{fi}})^2 + ((\mathbb{E}M)^2 + var(M_i))\cdot\frac{\lambda_{fu}}{\lambda_{fi}}\right) \end{aligned}$$

Detailed calculation can be found in the Appendix. These two equations can be solved for estimating the ratio of fusion rate to fission rate in terms of the mean and variance of size of stationary and moving mitochondria, i.e.,

$$\begin{aligned} \frac{\lambda_{fu}}{\lambda_{fi}} &= \frac{2\frac{varS_\infty}{\mathbb{E}M\mathbb{E}S_\infty} - (1+\frac{var(M)}{(\mathbb{E}M)^2})}{1+\frac{varS_\infty}{(\mathbb{E}S_\infty)^2}} \\ p &= \frac{\mathbb{E}S_\infty}{\eta\mathbb{E}M+\mathbb{E}S_\infty} \end{aligned}$$

However, As shown in table 2.2, an negative value of estimated $\frac{\lambda_{fu}}{\lambda_{fi}}$ is obtained when plugging in statistics estimated by Gaussian mixture model. In table 2.3, we reported the mean and variance of the first mode in moving mitochondria and of the second mode in stationary mitochondria. The predicted fission/fusion ratio are computed by plugging these quantities in to equations above. The negative predicted value suggests this simple model is incompatible with the collected data. A simulated size distribution is shown in Figure 2.8. It can be seen that the simulated distribution has a heavy tail compared to the observed distribution. This suggests independent fission/fusion process might lead to oversized mitochondria.

2.4.3 A general model for fission-regulated fusion

We hypothesize that the heavy tail observed in the previous model is due to the unlimited occurrence of fusion between two fission events. In practice, contradict to the independent fusion/fusion in the previous model, fission/fusion might be regulated to prevent unlimited occurrence of fusion. To address this issue, we propose a fission-

regulated fusion model to account for the regulation between fission and fusion. We reformulate the process $S(t)$ as

$$\begin{aligned} Q_{n+1} &= Q_n + \sum_{i=1}^{f(N_{fu}(Y))} M_i \\ S(t) &= Q_{N_{fi}(t)} + \sum_{i=1}^{f(N_{fu}(C(t)))} M_i. \end{aligned}$$

Here, the counting process $N_{fu}(Y)$ is interpreted as the number of possible fusion within a time span Y, while $f(n)$ is the number of realized fusion on top of n possible fusions. The form of f is designed to model the regulation on mitochondrial fusion. Two examples of f are given in the following:

1. $f(n) = n$. This simply gives the model with independent fission/fusion in the section 2.4.1.

2. $f(n) = \min(n,\ 1)$. This models the case that at most one fusion can occur between two fission events.

A subtle issue is that the definition of fusion rate in this model requires clarification. The rate λ_{fu} only represents the occurrence of possible fusion events, but not realized fusion events. Therefore, we define the fusion-to-fission ratio by

$$\eta = \mathbb{E} f(N_{fu}(Y)).$$

This quantity describes the average number of fusions occurring between two fission events. Note that in the case of $f(n) = n$, this ratio is exactly the ratio of fusion and fission rates $\lambda_{fu}/\lambda_{fi}$.

2.4.4 Regulated fusion model predicts the fusion-to-fission ratio

We analyzed the case that there is at most one fusion can occur between two fission events. The mean and variance of the stationary mitochondria's size are computed as

$$\begin{aligned} \mathbb{E}S_\infty &= \lim_{t\to\infty}\mathbb{E}S(t) &&= \tfrac{1}{1-p}\eta\cdot\mathbb{E}M_i \\ varS_\infty &= \lim_{n\to\infty} varS(t) &&= \tfrac{1}{1-p^2}\cdot((\mathbb{E}M_i)^2\cdot\eta(1-\eta)+var(M_i)\cdot\eta)\,. \end{aligned}$$

Given the statistics of stationary and moving mitochondria, we can inversely solve the fusion-to-fission ratio as

$$\begin{aligned} \eta &= \frac{1+\frac{var(M)}{(\mathbb{E}M)^2}-2\frac{varS_\infty}{\mathbb{E}M\mathbb{E}S_\infty}}{1-\frac{varS_\infty}{(\mathbb{E}S_\infty)^2}} \\ p &= 1-\eta\frac{\mathbb{E}M}{\mathbb{E}S_\infty}. \end{aligned}$$

The predicted ratio is computed by plugging in the estimated statistics of mitochondrial size from fitted Gaussian mixture models (table 2.2). This gives a value of 0.48, which is fairly close to the value (0.57) reported in the literature [13]. The simulated distribution of mitochondrial size using this ratio is shown in and figure 2.8.

Table 2.2: predicted fusion-to-fusion ratio

fusion-to-fusion ration		
model of independent fusion/fision	model of fission regulated fusion	reported value
-0.3280	0.4284	0.5745

Table 2.3: Statistics of mitochondrial size

moving mitochondria		stationary mitochondria	
mean	variance	mean	variance
0.2775	0.0116	0.8150	0.0881

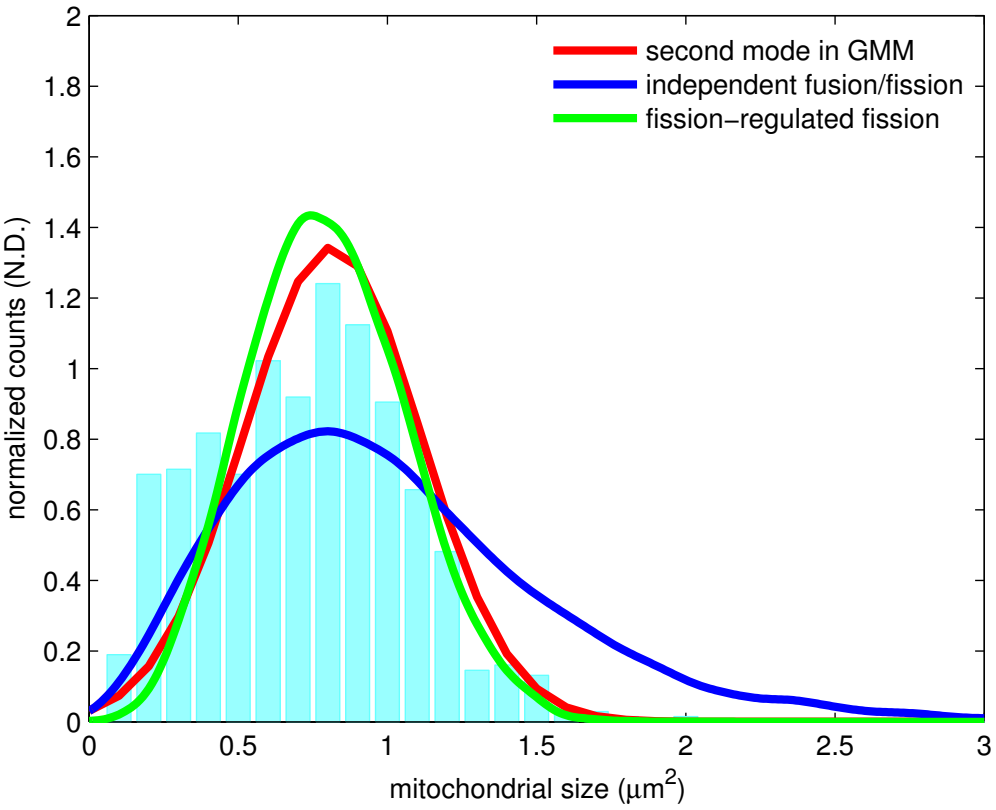

Figure 2.8: **Comparison of empirical and simulated distributions of mitochondrial size**. Blue bars: histogram of mitochondrial size. Red curve: the second mode fitted from Gaussian Mixture model. Blue and green curves: the mitochondrial size distributions simulated with independent fission/fusion and fission-regulated fusion respectively.

2.5 Conclusions

In this chapter we presented an in-depth analysis on mitochondrial morphology and its biological implication. We found that the mitochondrial morphology in the axon is relatively simple and can be summarized by their size and aspect ratio. We also proposed a probabilistic model to explain the bi-modal distribution of mitochondrial size. Combining these two, we demonstrated a way to infer the fission-to-fusion ratio based on observed mitochondrial size distribution. The estimated value is close to

the reported value in the literature.

Contradict to the commonly assumed independent fission/fusion process, our analysis suggests the independent fission/fusion leads to oversized mitochondria and cannot explain the data observed. The proposed model of fission-regulated fusion fixed this issue and provides a better prediction to the ratio of fission-to-fusion rate. However, whether other control mechanism can lead to the same prediction is still an open question. In this regard, our modeling approach can still be adapted to analyzing other possible control mechanism.

2.6 Appendix: Analysis of fission-regulated fusion model

Here we present mathematical details of analyzing the fission-regulated fusion model defined in the Section 2.4. The system is described by

$$\begin{aligned} Q_{n+1} &= Q_n + \textstyle\sum_{i=1}^{f(N_{fu}(Y))} M_i \\ S(t) &= Q_{N_{fi}(t)} + \textstyle\sum_{i=1}^{f(N_{fu}(C_t))} M_i, \end{aligned}$$

where M_i are i.i.d. to some distribution with known mean $\mathbb{E}M$ and variance $var(M)$. Y is the waiting time between two fission events and has exponential distribution with parameter $1/\lambda_{fi}$. $N_{fu}(t)$ and $N_{fi}(t)$ are Poisson counting processes with rate λ_{fu} and λ_{fi} respectively. $f(n)$ is a deterministic function mapping from nonnegative integers to nonnegative integers. C_t is the current life time.

We start with computing the mean and variance of the imbedded Markov chain, which defined as the mitochondrial mass right after a fission event.

Theorem 1.

$$\begin{aligned} \mathbb{E}Q_n &= (S_0 - \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(Y))) \cdot p^n + \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(Y)) \\ varQ_n &= (1-p^{2n})\tfrac{p^2}{1-p^2}\left((\mathbb{E}M_i)^2 \cdot varf(N_{fu}(Y)) + var(M_i)\cdot \mathbb{E}N_{fu}(Y)\right) \end{aligned}$$

Proof.

$$\begin{aligned} \mathbb{E}Q_{n+1} &= p\cdot \mathbb{E}Q_n + p\cdot \mathbb{E}M\mathbb{E}f(N_{fu}(Y)) \\ varQ_{n+1} &= p^2\cdot varQ_n + p^2\cdot var(\textstyle\sum_{i=1}^{N_{fu}(Y_{n+1})} M_i) \end{aligned}$$

Solving these linear difference equation and use the initial condition that $\mathbb{E}Q_0 = S_0$ and $varQ_0 = 0$, we got the results. □

Next we calculate the mean and variance of the imbedded Markov chain at an arbitrary time point.

Theorem 2.

$$\begin{aligned} \mathbb{E}Q_{N_{fi}(t)} &= (S_0 - \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(Y))) \cdot e^{-(1-p)\lambda_{fi}t} + \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(Y)) \\ varQ_{N_{fi}(t)} &= (S_0 - \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(C_t))) \cdot e^{-(1-p)\lambda_{fi}t} \end{aligned}$$

Proof. A direct calculation shows that

$$\mathbb{E}p^{N_{fi}(t)} = \sum_{i=0}^{\infty} p^n \frac{\lambda_{fi}^n t^n}{n!} e^{-\lambda_{fi}t} = e^{-(1-p)\lambda_{fi}t}.$$

Hence

$$\begin{aligned} \mathbb{E}Q_{N_{fi}(t)} &= (S_0 - \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fi}(t)))\mathbb{E}p^{N_{fi}(t)} + \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fi}(t)) \\ &= (S_0 - \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(Y))) \cdot e^{-(1-p)\lambda_{fi}t} + \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(Y)). \end{aligned}$$

Similarly,

$$\begin{aligned} varQ_{N_{fi}(t)} &= \mathbb{E}(var(Q_{N_{fi}(t)}|N_{fi}(t))) + var(\mathbb{E}(Q_{N_{fi}(t)}|N_{fi}(t))) \\ &= (1 - e^{(1-p^2)t/\lambda_{fi}})varQ_\infty + (S_0 - \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(C_t))) \cdot e^{-(1-p)t/\lambda_{fi}} \end{aligned}$$

□

We will need the following auxiliary theorems to compute the mean and variance of $S(t)$. The next theorem gives the distribution of the current life time of a mitochondrion.

Theorem 3. *Given Y follows exponential distribution with mean $1/\lambda$, then the current life C_t follows the distribution*

$$\mathbb{P}(C_t \le x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \le t \\ 1 & \text{otherwise} \end{cases}$$

Furthermore, $C_t \to Y$ in distribution as $t \to \infty$ and for any function $g(x) \ge 0$ with $\mathbb{E}g^2(Y) < \infty$,

$$\lim_{t\to\infty} \mathbb{E}g(C_t) = \mathbb{E}g(Y) \text{ and } \lim_{t\to\infty} varg(C_t) = varg(Y)$$

Proof. The distribution of C_t can be found in ([15], p.174). To compute the mean value,

$$\mathbb{E}g(C_t) = \int_0^t g(x)\lambda e^{-\lambda x}dx + g(t)e^{-\lambda t} \to \int_0^\infty g(x)\lambda e^{-\lambda x}dx = \mathbb{E}g(Y)$$

by dominated convergence theorem. Similarly we can show that

$$\lim_{t\to\infty} varg(C_t) = varg(Y).$$

□

Next we compute the joint distribution the number of fission events and the current life time of a mitochondrion.

Theorem 4. *Given Y follows exponential distribution with mean $1/\lambda$, the joint density of C_t and $N_{fi}(t)$ is given by*

$$f_{N(t),Ct}(n,x) = \begin{cases} \frac{\lambda^n t^n}{n!} e^{-\lambda t} \frac{n}{t} (\frac{t-x}{t})^{n-1} & \text{for } n \geq 1 \text{ and } x < t \\ e^{-\lambda t} & \text{for } n = 0 \text{ and } x = t \\ 0 & \text{otherwise} \end{cases}$$

Proof. We first compute the case for $n \geq 1$ and $x < t$.

$$\begin{aligned} \mathbb{P}(N(t) = n,\ C_t \leq x) &= \mathbb{P}(t - x < T_n < t,\ T_n + Y_{n+1} > t) \\ &= \int_{t-x}^{t} \int_{t-t_2}^{\infty} f_Y(t_1) f_{T_n}(t_2) dt_1 dt_2 \\ &= \int_{t-x}^{t} \int_{t-t_2}^{\infty} \lambda e^{-\lambda t_1} \frac{\lambda^n t_2^{n-1}}{(n-1)!} e^{-\lambda t_2} dt_1 dt_2 \\ &= \frac{\lambda^n t^n}{n!} e^{-\lambda t} (1 - (\frac{t-x}{t})^n). \end{aligned}$$

The density is obtained by taking derivative with respect to x. □

Theorem 5 shows that the mitochondrial mass right after a fission event is uncorrelated with the mass fused after the last fission event.

Theorem 5. *Assuming the function $g(y) \equiv \mathbb{E} f(N_{fu}(y))$ satisfies $0 \leq g(y) < \tilde{C}_1 t^{\alpha}$, for some constant α, then*

$$\lim_{t \to \infty} cov(Q_{N_{fi}(t)}, \sum_{i=1}^{f(N_{fu}(C_t))} M_i) = 0.$$

Proof. Writing

$$\begin{aligned}\mathbb{E}Q_n - \mathbb{E}Q_{N_{fi}(t)} &= (S_0 - \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(C_t))) \cdot (p^n - e^{-(1-p)\lambda_{fi}t}) \\ &\equiv \tilde{C}_2(p^n - e^{-(1-p)\lambda_{fi}t}),\end{aligned}$$

we have

$$\begin{aligned}&\mathbb{E}((Q_{N_{fi}(t)} - EQ)\textstyle\sum_{i=1}^{f(N_{fu}(C_t))} M_i) \\ &= \mathbb{E}\mathbb{E}((Q_{N_{fi}(t)} - \mathbb{E}Q_{N_{fi}(t)})\textstyle\sum_{i=1}^{f(N_{fu}(C_t))} M_i) \\ &= \tilde{C}_2\mathbb{E}M \cdot \mathbb{E}((p^{N_{fi}(t)} - e^{-(1-p)\lambda_{fi}t})g(C_t)|N_{fi}(t), C_t) \\ &= \tilde{C}_2\mathbb{E}M \\ &\quad \cdot \Big(\textstyle\sum_{n=1}^{\infty}\int_0^t (p^n - e^{-(1-p)\lambda_{fi}t})g(y)\frac{\lambda_{fi}^n t^n}{n!}e^{-\lambda_{fi}t}\frac{n}{t}(1-\frac{y}{t})^{n-1}dy + (1 - e^{-(1-p)\lambda_{fi}t})g(t)e^{-\lambda_{fi}t}\Big) \\ &= \tilde{C}_2\mathbb{E}M \cdot (\lambda_{fi}t\textstyle\sum_{n=0}^{\infty}\cdot\frac{\lambda_{fi}^n t^n}{n!}e^{-\lambda_{fi}t} \cdot (p^{n+1} - e^{-(1-p)\lambda_{fi}t}) \\ &\quad \cdot \textstyle\int_0^t \frac{g(y)}{t}(1-\frac{y}{t})^{n-1}dy + (1 - e^{-(1-p)\lambda_{fi}t})g(t)e^{-\lambda_{fi}t})\end{aligned}$$

From the assumption we have $0 \le \int_0^t \frac{g(y)}{t}(1-\frac{y}{t})^{n-1}dy \le \tilde{C}_3 t^{\alpha}$ for a constant α, and hence

$$\begin{aligned}&t\textstyle\sum_{n=0}^{\infty}\frac{\lambda_{fi}^n t^n}{n!}e^{-\lambda_{fi}t} \cdot p^{n+1} \cdot \int_0^t g(y)\frac{t}{n}(1-\frac{y}{t})^{n-1}dy \\ &\le t\textstyle\sum_{n=0}^{\infty}\frac{\lambda_{fi}^n t^n}{n!}e^{-\lambda_{fi}t} \cdot p^n \cdot \tilde{C}_3 t^{\alpha} \\ &= \tilde{C}_3 t^{\alpha+1}e^{-(1-p)\lambda_{fi}t} \to 0 \text{ as } t \to \infty\end{aligned}$$

and

$$\begin{aligned}&t\textstyle\sum_{n=0}^{\infty}\frac{\lambda_{fi}^n t^n}{n!}e^{-\lambda_{fi}t} \cdot e^{-(1-p)\lambda_{fi}t} \cdot \int_0^t g(y)\frac{t}{n}(1-\frac{y}{t})^{n-1}dy \\ &\le t\textstyle\sum_{n=0}^{\infty}\frac{\lambda_{fi}^n t^n}{n!}e^{-\lambda_{fi}t} \cdot e^{-(1-p)\lambda_{fi}t} \cdot \tilde{C}_3 t^{\alpha} \\ &= \tilde{C}_3 t^{\alpha+1}e^{-(1-p)\lambda_{fi}t} \to 0 \text{ as } t \to \infty.\end{aligned}$$

These shows

$$\lim_{t\to\infty}\mathbb{E}((Q_{N_{fi}(t)} - EQ)\sum_{i=1}^{f(N_{fu}(C_t))} M_i) = 0$$

and hence the covariance between $Q_{N_{fi}(t)}$ and $\sum_{i=1}^{f(N_{fu}(C_t))} M_i$ goes to 0 as time ap-

proaching to infinity. □

Theorem 6 gives the mean and variance of fission/fusion modulated mitochondrial mass.

Theorem 6. *Assuming the functions*

$$g_1(t) = \mathbb{E}\sum_{i=1}^{f(N_{fu}(t))} M_i = \mathbb{E}M \cdot \mathbb{E}f(N_{fu}(t))$$

$$g_2(t) = var\sum_{i=1}^{f(N_{fu}(t))} M_i = (\mathbb{E}M)^2 \cdot var(f(N_{fu}(t))) + varM \cdot \mathbb{E}f(N_{fu}(t))$$

satisfies $0 \le g_1(y) < \tilde{C}_1 t^\alpha$ *for some constant* α, $\mathbb{E}g_1^2(Y) < \infty$ *and* $\mathbb{E}g_2(Y) < \infty$ *then*

$$\begin{aligned}
\lim_{t\to\infty} \mathbb{E}S(t) &= \tfrac{1}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(Y)) \\
\lim_{t\to\infty} varS(t) &= \tfrac{1}{1-p^2}\left((\mathbb{E}M_i)^2 \cdot varf(N_{fu}(Y)) + var(M_i)\cdot \mathbb{E}N_{fu}(Y)\right)
\end{aligned}$$

Proof. From assumptions and theorem 2, we first compute

$$\mathbb{E}\sum_{i=1}^{N_{fu}(C_t)} M_i = \mathbb{E}g_1(C_t) \to \mathbb{E}g_1(Y) \text{ as } t \to \infty$$

$$\begin{aligned}
var\textstyle\sum_{i=1}^{N_{fu}(C_t)} M_i &= varg_1(C_t) + \mathbb{E}g_2(C_t) \\
&\to varg_1(Y) + \mathbb{E}g_2(Y) \text{ as } t \to \infty.
\end{aligned}$$

$\mathbb{E}S(t)$ is thus computed as

$$\begin{aligned}
\mathbb{E}S(t) &= \mathbb{E}Q_{N_{fi}(t)} + \mathbb{E}\textstyle\sum_{i=1}^{N_{fu}(C_t)} M_i \\
&= \tfrac{1}{1-p^2}\left((\mathbb{E}M_i)^2 \cdot varf(N_{fu}(C_t)) + var(M_i)\cdot \mathbb{E}N_{fu}(C_t)\right) \\
&\to \tfrac{1}{1-p^2}\left((\mathbb{E}M_i)^2 \cdot varf(N_{fu}(Y)) + var(M_i)\cdot \mathbb{E}N_{fu}(Y)\right) \text{ as } t \to \infty
\end{aligned}$$

$$
\begin{aligned}
varQ_{N_{fu}(t)} &= \mathbb{E}(var(Q_{N_{fi}(t)}|N_{fi}(t))) + var(\mathbb{E}(Q_{N_{fi}(t)}|N_{fi}(t))) \\
&= (1 - e^{(1-p^2)t/\lambda_{fi}})varQ_\infty + (S_0 - \tfrac{p}{1-p}\mathbb{E}M\mathbb{E}f(N_{fu}(C_t))) \cdot e^{-(1-p)t/\lambda_{fi}}
\end{aligned}
$$

$$
\begin{aligned}
varS(t) &= varQ_{N_{fi}(t)} + cov(Q_{N_{fi}(t)}, \textstyle\sum_{i=1}^{N_{fu}(C_t)} M_i) + var \textstyle\sum_{i=1}^{N_{fu}(C_t)} M_i \\
&\to \tfrac{1}{1-p^2}\left((\mathbb{E}M_i)^2 \cdot varf(N_{fu}(Y)) + var(M_i) \cdot \mathbb{E}N_{fu}(Y)\right)
\end{aligned}
$$

□

Following we analyzed two concrete examples based on these general theorems.

2.6.1 Model of independent fusion/fision

Given $f(n) = n$, we have

$$
\begin{aligned}
\mathbb{E}f(N_{fu}(Y)) &= \tfrac{\lambda_{fu}}{\lambda_{fi}} \\
varf(N_{fu}(Y)) &= \tfrac{\lambda_{fu}}{\lambda_{fi}} + (\tfrac{\lambda_{fu}}{\lambda_{fi}})^2.
\end{aligned}
$$

Plug these into theorem 4, we get

$$
\begin{aligned}
\lim_{t\to\infty} \mathbb{E}S(t) &= \tfrac{1}{1-p}\tfrac{\lambda_{fu}}{\lambda_{fi}} \cdot \mathbb{E}M_i \\
\lim_{n\to\infty} varS(t) &= \tfrac{1}{1-p^2} \cdot \left((\mathbb{E}M_i)^2 \cdot (\tfrac{\lambda_{fu}}{\lambda_{fi}})^2 + ((\mathbb{E}M_i)^2 + var(M_i)) \cdot \tfrac{\lambda_{fu}}{\lambda_{fi}}\right)
\end{aligned}
$$

Or conversely,

$$
\begin{aligned}
\eta &= \frac{2\frac{var(S)}{\mathbb{E}M\mathbb{E}S} - (1 + \frac{var(M)}{(EM)^2})}{1 + \frac{var(S)}{(\mathbb{E}(S))^2}} \\
p &= 1 - \eta \cdot \tfrac{\mathbb{E}S}{\mathbb{E}M}
\end{aligned}
$$

2.6.2 Model of fission-regulated fusion

Given $f(x) = \min(x, 1)$, the random variable $f(N_{fu}(Y))$ is actually a bernoulli random variable. Hence we have

$$\begin{aligned} \mathbb{E}f(N_{fu}(Y)) &= \mathbb{P}(N_{fu}(Y) > 0) \equiv \eta \\ varf(N_{fu}(Y)) &= \eta(1-\eta). \end{aligned}$$

Plug these into theorem 4, we get

$$\begin{aligned} \lim_{t\to\infty} \mathbb{E}S(t) &= \frac{1}{1-p} \cdot \eta \cdot \mathbb{E}M_i \\ \lim_{n\to\infty} varS(t) &= \frac{1}{1-p^2} \cdot ((\mathbb{E}M_i)^2 \cdot \eta(1-\eta) + var(M_i) \cdot \eta) \end{aligned}$$

Or conversely,

$$\begin{aligned} \eta &= \frac{1+\frac{var(M)}{(\mathbb{E}M)^2}-2\frac{var(S)}{\mathbb{E}M\mathbb{E}S}}{1-\frac{var(S)}{(\mathbb{E}S)^2}} \\ p &= 1-\eta\frac{\mathbb{E}M}{\mathbb{E}S} \end{aligned}$$

Bibliography

[1] H. M. McBride, M. Neuspiel, and S. Wasiak, "Mitochondria: more than just a powerhouse," *Current Biology*, vol. 16, no. 14, pp. R551–R560, 2006.

[2] J.-Y. Peng, C.-C. Lin, Y.-J. Chen, L.-S. Kao, Y.-C. Liu, C.-C. Chou, Y.-H. Huang, F.-R. Chang, Y.-C. Wu, Y.-S. Tsai, and C.-N. Hsu, "Automatic morphological subtyping reveals new roles of caspases in mitochondrial dynamics," *PLoS Computional Biololgy*, vol. 7, no. 10, 2011.

[3] P. Patel, O. Shirihai, and K. Huang, "Optimal dynamics for quality control in spatially distributed mitochondrial networks," *PLoS Computional Biololgy*, vol. 9, no. 7, p. e1003108, 2013.

[4] K.-C. Chen, Y. Yu, R. Li, H.-C. Lee, G. Yang, and J. Kovacevic, "Adaptive active-mask image segmentation for quantitative characterization of mitochondrial morphology," in *IEEE International Conference on Image Processing*, 2012, pp. 2033–2036.

[5] G. Reis, G. Yang, L. Szpankowski, C. Weaver, S. Shah, J. Robinson, T. Hays, G. Danuser, and L. Goldstein, "Molecular motor function in axonal transport in vivo probed by genetic and computational analysis in Drosophila," *Molecular Biology of the Cell*, 2012.

[6] I. Dryden and K. Mardia, *Statistical shape analysis*. John Wiley & Sons, 1998.

[7] N. Graham, "Smoothing with periodic cubic splines," *Bell System Technical Journal*, vol. 62, no. 1, pp. 101–110, 1983.

[8] I. Fodor, "A survey of dimension reduction techniques," *Technical Report UCRL-ID-148494*, 2002.

[9] G. Kendall, D. Barden, T. Carne, and H. Le, *Shape and shape theory*. John Wiley & Sons, 1999.

[10] H.-C. Lee and G. Yang, "Integrating dimension reduction with mean-shift clustering for biological shape classification," *IEEE International Symposium on Biomedical Imaging*, pp. 254–257, 2014.

[11] H.-C. Lee, T. Liao, Y. J. Zhang, and G. Yang, "Shape component analysis: structure-preserving dimension reduction on biological shape spaces," *submitted to Bioinformatics*.

[12] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.

[13] Y. Yu, H.-C. Lee, K.-C. Chen, J. Suhan, M. Qiu, and G. Yang, "Inner membrane fusion mediates spatial distribution of axonal mitochondria," *submitted to Scientific Reports*.

[14] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. Oxford university press, 2001.

[15] S. Karlin, *A First Course in Stochastic Processes*. Academic press, 1975.

Chapter 3

Probabilistic modeling of the size-velocity relation of mitochondrial transport

3.1 Introduction

Mitochondria are essential organelles of eukaryotic cells, serving a broad range of important functions that include energy production, metabolic regulation, signal transduction, and stress response [1, 2]. To fulfill changing needs at different subcellular locations, mitochondria undergo motor protein mediated active transport to reach their destinations [3]. In the meantime, to sustain their physiological functions, mitochondria undergo fusion and fission mediated content mixing, which also results in substantial concomitant changes to their morphology [4]. Indeed, fusion and fission play critical roles in defining the wide variety of morphology exhibited by mitochondria inside cells [4]. One of the fundamental questions regarding the dynamics of mitochondria is whether, and if so how, differences in their morphology result in dif-

ferences in their transport behavior. So far, this question remains largely unanswered.

Neurons provide a powerful model system to address this question because of their polarized structure and their critical dependence on mitochondrial dynamics for survival and function [5]. The long and thin axon, in particular, provides a simplified setting for high-resolution quantitative analysis of relations between mitochondrial morphology and transport. In this study, we used Drosophila third instar larvae as our model organism and tracked individual mitochondria within the axon of the larval neurons using high-resolution image analysis techniques. We found that these mitochondria exhibited a wide range of sizes and a wide range of velocities. We then focused specifically on analyzing the relation between their size and velocity.

To understand the mechanism underlying the mitochondrial size-velocity relation observed in our experiments, we developed a probabilistic model that takes into account the multiple variables and their interdependency in defining the mitochondrial size-velocity relation. To give an example of such variables and their interdependency, we note that, while larger mitochondria will experience stronger viscous drag force during transport, they can also recruit more molecular motors, which in turn can generate stronger actuation force. Our model can successfully explain the observed size-velocity relation. Furthermore, it identifies the ratio of mean velocity to its standard deviation as a measure of the binding behavior of motor proteins to the microtubule. Using this ratio, we gained new insights into how microtubule associated protein tau modulates cargo motility.

Many studies have been conducted to model the behavior of molecular motors [6] and intracellular transport [7]. In this study, we have focused on modeling and under-

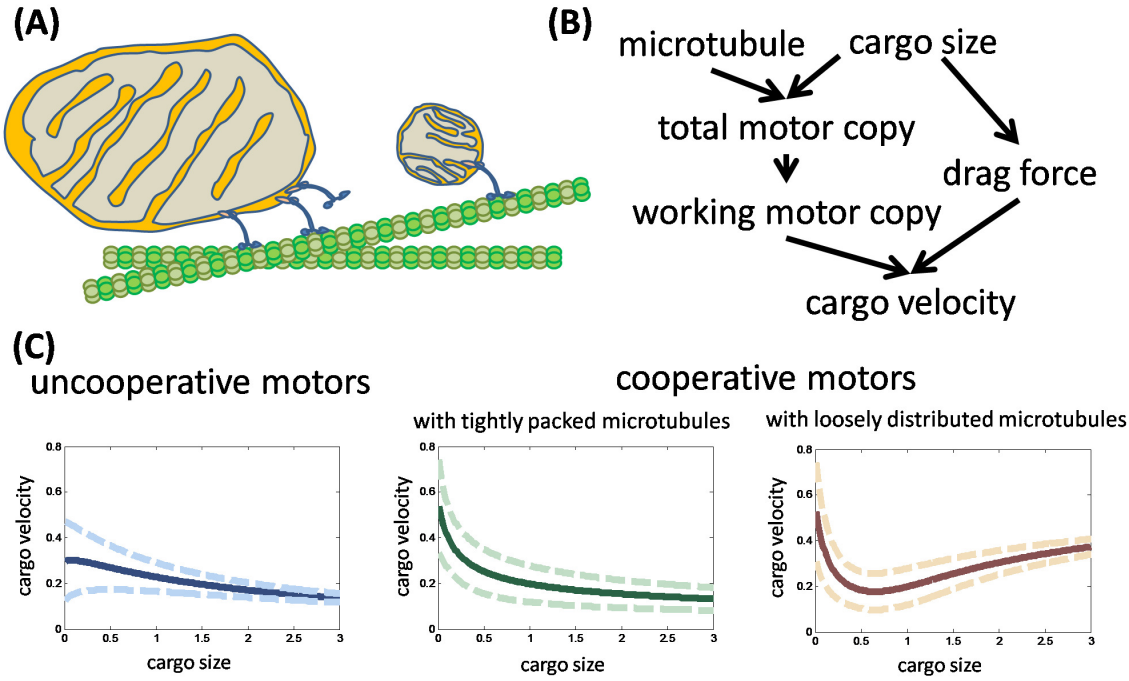

Figure 3.1: **Summary of the proposed model and size-velocity relations revealed in our model**. (A) Cargoes with various sizes are carried by motor proteins on a microtubule. (B) A hierarchy of factors considered in the proposed model. (C)size-dependent cargo velocity under different scenarios. Uncooperative motors has slow decreasing average velocity whose dynamic range is shrinking with size. Cooperative motors has also decreasing average velocity when moving on sparsely distributed microtubules. On the other hand, larger cargoes move faster on average when moving on densely distributed microtubules.

standing the relation between the morphology and transport of mitochondria within the axonal cytoskeleton. Studying the mitochondrial size-velocity relation provides new insights into how groups of motors function inside living cells. Furthermore, because of the large sizes of mitochondria, studying their size-velocity relation provide new insights into the mechanical properties of the network of cytoskeletal filaments within the axon.

3.2 Methods

3.2.1 Modeling

Individual mitochondria in the axon of Drosophila third instar larval neurons were imaged and tracked as mentioned in Chapter 2. The general configuration of transport of mitochondria by molecular motors walking on microtubules is illustrated in Fig. 3.1A. The key variables that define the mitochondrial size-velocity relation and their interdependency are shown in Fig. 3.1B. Anterograde and retrograde transport of mitochondria are driven by kinesin and dynein, respectively. Because the models developed in this section apply to both kinesin and dynein, we refer to them simply as "motor". The mathematical notations used in this paper are summarized in Table 1.

average velocity of a single mitochondrion

For a mitochondrion moving in a certain direction over a time span Δt, if the total number of steps taken by the motors pulling this mitochondrion within the time span is $S_{\Delta t}$, and the step size is d, the average velocity of this mitochondrion is defined by the following equation

$$V = \frac{d}{\Delta t} S_{\Delta t}. \tag{3.1}$$

In the following, we model two dependency relations. 1) the dependency of $S_{\Delta t}$ on the number of working motors M and 2) the dependency of M on the mitochondrion radius r, assuming a mitochondrion is a sphere.

stochastic stepping of a single motor

Movement of a mitochondrion is ultimately driven by the chemo-mechanical cycles of individual motors. We assume that, over a time span Δt, the number of steps taken by a single motor follows a Poisson distribution with a mean of $\Delta t/\tau$, where τ is the average duration of the chemo-mechanical cycle. This assumption is based on the experimental observation that the waiting time between consecutive steps of kinesin and dynein follows an exponential distribution [8, 9].

We simplify the complicated stepping of multiple motors as a Poisson process whose rate is depended on the number of motors m. This microscopic stepping rate can be determined from the macroscopic average velocity. Suppose the mitochondrion is moved by m motors at velocity v_m, the assumed Poisson distribution of motor stepping gives

$$v_m = \mathbb{E}(\frac{d}{\Delta t} S_{\Delta t} | M = m) = d/\tau(m). \tag{3.2}$$

The explicit form of v_m is derived in the following.

velocity of a mitochondrion with cooperative motors

We make the same mean field assumption as in [10] and assume that the external load on a mitochondrion is equally shared by its working motors. Furthermore, we assume a linear force-velocity relation for the motor [11]. Under these assumptions, the velocity of the mitochondrion pulled by m working motors under external load is given by the following equation

$$v_m(f) = v_{max} \max(1 - \frac{f}{m f_s},\ 0), \tag{3.3}$$

where v_{max} is the velocity when the external load is zero. We further assume that the shape of the mitochondrion can be approximated by a sphere with a radius r. Following the Stoke's law, by balancing the viscous drag force and the pulling force of cooperative motors, we have

$$f = 6\pi\eta r v_m = 6\pi\eta r v_{max} \max(1 - \frac{f}{m f_s}, 0). \tag{3.4}$$

Solving this equation, the average velocity of a mitochondrion pulled by m motors is given by

$$v_m = v_{max}(1 - \frac{C_0 r}{m f_s + C_0 r}) \tag{3.5}$$

with $C_0 = 6\pi\eta v_{max}$. This equation defines the size-velocity relation of a mitochondrion pulled by m cooperative motors.

velocity of a mitochondrion with non-cooperative motors

At the other extreme, for non-cooperative motors, we simply assume that the average velocity is independent of the number of pulling motors. That is, the average velocity is given by

$$v_m = \begin{cases} v_{max}(1 - \frac{C_0 r}{f_s + C_0 r}) & \text{if } m > 0 \\ 0 & \text{otherwise} \end{cases}$$

size-dependency of total motor copy number N

Total number of motors as a function of mitochondrial size Next we model the total number of motors N on a mitochondrion. We assume that spatial distribution of motors on a mitochondrion is random and follows a spatial Poisson distribution with a rate constant λ [12]. The probability of having n motors in a region of area A on

Parameter			values used	
notation	unit	explanation	Kinesin	Dynein
π_a	1/s	binding rate of a motor to a microtubule	5^*	1.6^*
ϵ_d	1/s	unbinding rate of a motor to a microtubule	1^*	0.27^*
v_{max}	$\mu m/s$	velocity of a single motor without loading	0.27	0.4
f_s	pN	stall force of a single motor	6^*	1.1^*
η	cP	cytoplasmic viscosity	200	200
C_1	N.D.	multiplicative constant in total numbers of working motor	3	2
α	N.D.	exponent in total numbers of working motor	0.3	0.3

Table 3.1: Summary of parameters used in this chapter. Values with * are taken from [10]

variables and mathematical operations	
notation	description
τ	average length of the chemomechanical cycle
d	step size of a motor protein
λ	motor density on a mitochondrion
$\mathbb{E}, \mathbb{P}, var$	expectation, probability, and variance of a random variable
r	radius of a spherical mitochondrion
$\mathbb{1}$	indicator function which take value of 1 when the statement is true and 0 otherwise
N	total number of motors that can pull a mitochondrion
M	number of motors that are exerting force on a mitochondrion
Δt	time span of observing a mitochondrion
V, V_c, V_n	mitochondrion velocity, subscripts indicate the mitochondrion is moved by cooperative and non-cooperative motors, respectively
$C(r)$	contact area between a mitochondrion and microtubules

Table 3.2: Summary of variables and mathematical operations used in this Chapter.

the mitochondrion surface is given by the following equation:

$$\mathbb{P}(N(A) = n) = \frac{(\lambda A)^n}{n!} e^{-\lambda A}. \tag{3.6}$$

The actual number of working motors that can exert force on the mitochondrion depends on the contact area between the mitochondrion and microtubules. We model this contact area as

$$C(r) = C_1' r^\alpha \tag{3.7}$$

where C_1' and α characterize the number and distribution of microtubules around the mitochondrion, respectively. In the Result section, we will show by simulation that *the number* of microtubules only change $C(r)$ up to a multiplicative constant C_1', while *the way* of how microtubules distributed impacts the exponent α. In [13], *the number* of microtubules is shown to increase the mitochondrion motility. However, we will demonstrate that the size-dependency of mitochondrion motility is critically determined by α, suggesting *the way* microtubules are distributed is more critical in regulating mitochondrion velocity.

By combining the model of the total number of molecular motors (equation (3.6)) and the model of the contact area (equation 3.7), we can compute the probability of having n motors that exert force on a mitochondria as

$$\mathbb{P}(N(C(r)) = n) = \frac{(C_1 r^\alpha)^n}{n!} e^{-C_1 r^\alpha}, \tag{3.8}$$

where $C_1 = \lambda C_1'$.

size-dependency of working motor copy number M

Previous works in [10] have described the steady state behavior of M given a fixed total number of motors $N = n$. With the Poisson-distributed total copy number N, the unconditional distribution of M can be calculated. In [10], the stationary distribution of $M|N = n$ is derived as $\mathbb{P}(M = m|N = n) = P_m = \binom{n}{m}\left(\frac{\pi_a}{\epsilon_d}\right)^m / \left(1 + \frac{\pi_a}{\epsilon_d}\right)^n$ for $n \geq m$ and 0 otherwise under the assumption that motors function independently. Together with the assumption of the random motor copy N, it can be shown that

$$\mathbb{P}(M = m) = \sum_{n=0}^{\infty} \mathbb{P}(M = m|N = n)\mathbb{P}(N = n) = \frac{z_r^m}{m!} e^{-z_r}$$

where $z_r = \frac{\pi_a/\epsilon_d}{1+\frac{\pi_a}{\epsilon_d}} C_1 r^{\alpha}$. This suggests the number of working motors M is also Poisson distributed with mean z_r.

3.2.2 Experimental methods

Drosophila handling

To image axonal mitochondria, we crossed male UAS-mito-GFP/CyO (Bloomington Drosophila Stock Center) with female pGAL4 SG26-1. UAS-mito-GFP/+; pGAL4 SG26-1/+ was used as a control. To image APP vesicles, we first crossed male UAS APP-YFP with female pGAL4 SG26-1. UAS APP-YFP/+; pGAL4 SG26-1/+ was used as a control. We further crossed male UAS APP-GFP/+; pGAL4 SG26-1/+ with female tau-RNAi, htau^{wt} and hTau^{R406w}, to image mitochondrial dynamics in axons harbording tau knockdown, human tau overexpression or human hTau^{R406w} mutation. We maintained Drosophila stocks at 25°C, and set up crossing at 29°C.The single neuron Gal4 driver pGAL4 SG26-149 was a gift from Dr. Lawrence Goldstein (University of California San Diego).

fly dissection and imaging

Third instar larvae were dissected in calcium free HL3 buffer (in nM; 128 NaCl, 1 ethylene glycol tetraacetic acid (EGTA), 4 MgCl2, 2 KCl, 5 HEPES, and 36 sucrose) to expose the segmental nerve as previously described [14]. Fly were dissected and imaged as previously described [14, 15]. Time-lapse movies were collected using a Nikon Eclipse Ti-E inverted microscope with a CoolSNAP HQ2 camera (Photometric) and a 100×/1.40 NA oil objective lens. The effective pixel size was 0.0645 μm. Mito-GFP was imaged using a FITC filter set. Mitochondrial time-lapse movie was collected at 5 frames per second rate for 1 min for motility and morphology analysis.

APP vesicle time-lapse movie was collected at 10 frames per second for 30 sec.

primary neuron culture and imaging

Mouse E18 Hippocampal tissue was purchased from Brainbitllc. Tissue was dissociated using 0.25% Trypsin with EDTA (Life Technology). Neurons were cultured on #1.5 cover glass in Neurobasal media supplemented with B27 and Glutamax (Life Technology). On 6 days in vitro (DIV), neurons were transfected with APP-YFP using lipofectamine 2000 (Life Technology). After 24~48 hours of transfection, neurons were used for live imaging.

trajectory analysis

Time-lapse movies of mitochondria, APP and synaptotagmin 1 vesicles were processed and traced using in house software [15]. The trajectories were analyzed as previously described [15].

3.2.3 Simulation

We conduct a simulation study to study how the contact area between microtubules and a cargo changes with cargo radius r. The simulation is set up as follows:

1. A circle with radius r centered at $(r, 0)$ is used to indicates a cargo's occupation in the cross section of an axon.

2. The locations of microtubules at the cross section are simulated according to a spatial Poisson process. microtubules are assumed to be projecting perpendicularly into the cross section.

3. Compute the overall segment length of microtubules that can be reached from the cargo. Given the size of a motor l_0 and a microtubule located l_1 away

from the center of the cargo, the segment length that is reachable to a motor is determined by $2\sqrt{(r+l_0)^2 - l_1^2}$ if $l_1 \leq r + l_0$ and 0 otherwise. The overall segment length is computed as the sum of individual segment length given simulated microtubule locations in step 2.

4. For each batch of simulations, repeat above steps 1000 times to determine the mean and variance of the overall segment length.

For the case of bundled microtubules, we simulated the spatial Poisson process with isotropic Gaussian density centered at the center of the axon. The bandwidth of the Gaussian density σ depicts the tightness of the microtubule. Higher value in σ gives wider spread of microtubles located around the center. Since microtubules shouldn't pass though the occluded region of the cargo, the density is set to be 0 in this occluded region. We calibrated the integral of the occluded Gaussian density to ensure the average number of microtubules remain unchanged under different cargo setting. The case of randomly distributed microtubules was simulated similarly except a uniform density is used instead of Gaussian density. Spatial Poisson processes are all simulated by thinning [12].

3.3 Results

To investigate how morphology influence the cargo motility, we first present predictions of the size-velocity dependency from the proposed model and then followed by the empirical data validation. Mathematical details can be found in the supplement material. Following we will use V_n and V_c to denote the size-dependent velocity of cargoes carried by non-cooperative and cooperative motors respectively.

3.3.1 Size-velocity relation of mitochondria with non-cooperative motors

In the case of non-cooperative motors, mean and variance of the cargo velocity are calculated as (Theorem 2 in the supplement)

$$\mathbb{E}(V_n|N>0) = \frac{d}{\tau}\mathbb{P}(M>0|N>0)$$

$$var(V_n|N>0) = \frac{d^2}{\tau\Delta t}\mathbb{P}(M>0|N>0) + (\frac{d}{\tau})^2\mathbb{P}(M>0|N>0)(1-\mathbb{P}(M>0|N>0))$$

respectively, where $\mathbb{P}(M=0|N>0)$ can be explicitly calculated as $\frac{e^{-z_r}-e^{-f(r)}}{1-e^{-f(r)}}$. To further investigate the dependency between cargo size and velocity, we conducted the asymptotic analysis to see changes in velocity when cargoes are extremely small or large. The analysis (Theorem 3 in the supplement) shows when the mean cargo velocity as well as the variance are eventually decreasing to zero for large cargo size. On the other hand, the mean and variance velocity for cargoes with insignificant size are determined by the attach/deattach rate of motor proteins. Particularly, it can be shown that the ratio of cargo velocity's mean and variance converges to the ratio of attach and detach rates of a single motor, i.e.,

$$\frac{\lim_{r\to 0}\mathbb{E}^2(V_n|N>0)}{\lim_{r\to 0} var(V_n|N>0)} = \frac{\pi_a}{\epsilon_d}(\frac{1}{1+(\frac{1+\pi_a/\epsilon_d}{\pi_a/\epsilon_d})\frac{dv_{opt}}{\Delta t}}). \tag{3.9}$$

3.3.2 Size-velocity relation of mitochondria with cooperative motors

Next we investigate how morphology influence the motility of cargoes carried by cooperative motors. Although the explicit form of the expectation $\mathbb{E}(V_c)$ and and

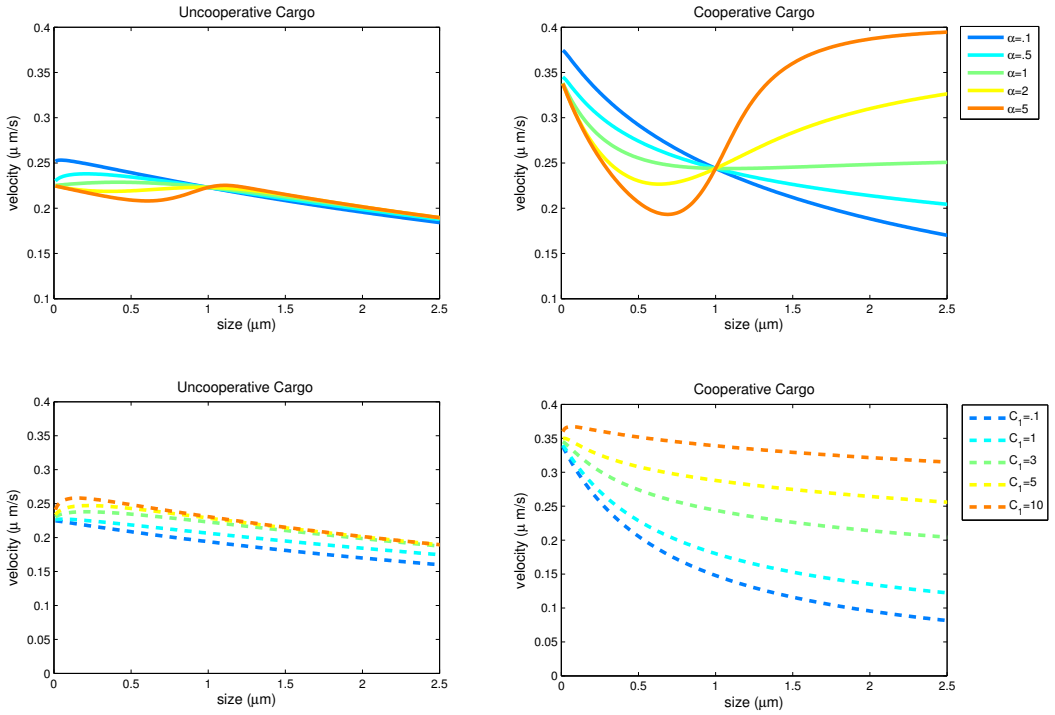

Figure 3.2: **Predicted size-velocity relations**. Each curve represents a size-velocity relation computed from the proposed model. The cases of non-cooperative motors are shown in the left panel while the case of cooperative motors are shown in the right panel. Parameters are reported in table 3.1 unless specified in the figure legend.

variance $var(V_c)$ are not available, we conduct an asymptotic analysis (Theorem 4 and 5 in the supplement) on the behavior of these two quantities of interest. Surprisingly, our analysis shows that the limiting behavior of $\mathbb{E}(V_c)$ is critically depended on z_r. We found that the mean velocity decays to zero, when the cargo size is significantly large, if $z_r = O(r^\alpha)$ with $\alpha < 1$ (Figure 3.2). However, the profile of mean velocity converges to v_{opt} if $z_r = O(r^\alpha)$ with $\alpha > 1$. For the case of cargoes with insignificant sizes, the ratio of variance and square of expectation is as it is in the case of non-cooperative cargoes

$$\frac{\lim_{r\to 0}\mathbb{E}^2(V_c|N>0)}{\lim_{r\to 0} var(V_c|N>0)} = \frac{\pi_a}{\epsilon_d}\left(\frac{1}{1+\left(\frac{1+\pi_a/\epsilon_d}{\pi_a/\epsilon_d}\right)\frac{dv_{opt}}{\Delta t}}\right). \tag{3.10}$$

3.3.3 In-silico study of the impact from microtubule distribution

To investigate how the contact length between microtubules and a cargo changes with cargo size, we conducted a simulation study considering the following two scenarios: 1) microtubules packed around the center of an axon and 2) microtubules scatter randomly in an axon. For the first scenario, locations of microtubules at the cross section of an axon is simulated by a spatial Poisson distribution with isotropic Gaussian density centered at the center of an axon (Fig. 3.3, right). The bandwidth of the Gaussian density depicts how tight microtubules are packed. The second scenario of randomly scattered microtubules is simulated by a spatial Poisson process with a uniform density (Fig. 3.3, left). Details can be found in Material and Method.

Our simulation results showed that the size-dependency of contact lengths has a convex profile for randomly scattered microtubules, while a concave profile is observed for packed microtubules (Fig. 3.3). For the case of packed microtubules, we varied

the bandwidth of the Gaussian density form 0.05 to 0.3 (μm), while the average number of microtubules simulated are kept to be 10. Uniform density are used to simulated 5, 10, 20 randomly scattered microtubules. Both profiles were not changed by the tightness of packed microtubules nor the number of randomly distributed microtubules. In section 1, we hypothesized that the total number of motors on a cargo is determined by the contact area between microtubules and a cargo and hence is proportional to r^{α} on average, where r is the cargo radius. These simulation results not only support this assumption but also suggest the choice of α in equation 3.8. The concave profile from packed microtubules is better approximated by setting $\alpha < 1$ while $\alpha > 1$ better depicts the convex profile from randomly scattered microtubules.

3.3.4 In-vivo observation of size-velocity dependency for large cargoes

To test the proposed model, we imaged the axonal transport of mitochondria in Drosophila. Mitochondria are chosen as the subject of study because of their various sizes. Typical, mitochondrial size (measured as segmented area from fluorescent images) ranges from $0.5\mu m^2$ to $2\mu m^2$. The collected images are processed by image segmentation and particle tracking algorithms to estimate the size and velocity of a mitochondrion. We also define the effective mitochondrial radius as the root mean square of major and minor axes's length computed form the best fit ellipse to the segmented area. Figure 3.4 shows the scatter plot of mitochondrial motility to the effective mitochondrial radius, where motility is further grouped by its direction (anterograde and retrograde).

On top of data scatter, we indicated the region that is less than 1 standard devi-

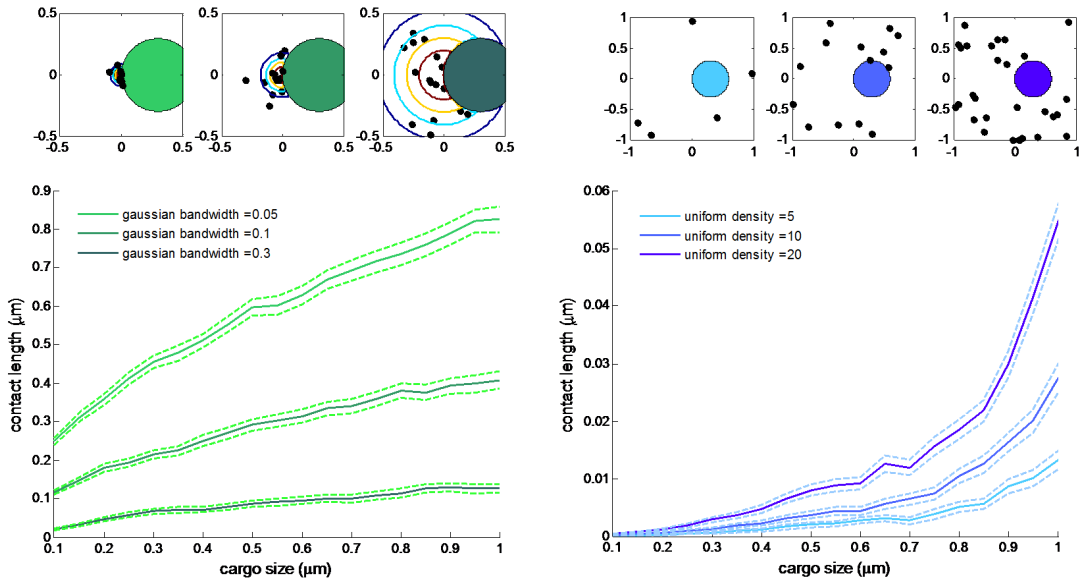

Figure 3.3: **The simulated size-dependency on contact length between microtubules and a cargo**. Left: contact length between randomly distributed microtubules and a cargo. Curves represent cases of 5, 10, 20 microtubules simulated per unit area as illustrated in the upper panels. Right: contact length between packed microtubules and a cargo. microtubules are simulated according to a Gaussian density whose bandwidth represents the tightness of packed microtubules. Three cases with bandwidths of 0.05, 0.1 and 0.3 are simulated. Simulated examples are shown on the upper panels where microtubule are shown in black and the cargo is shown in blue.

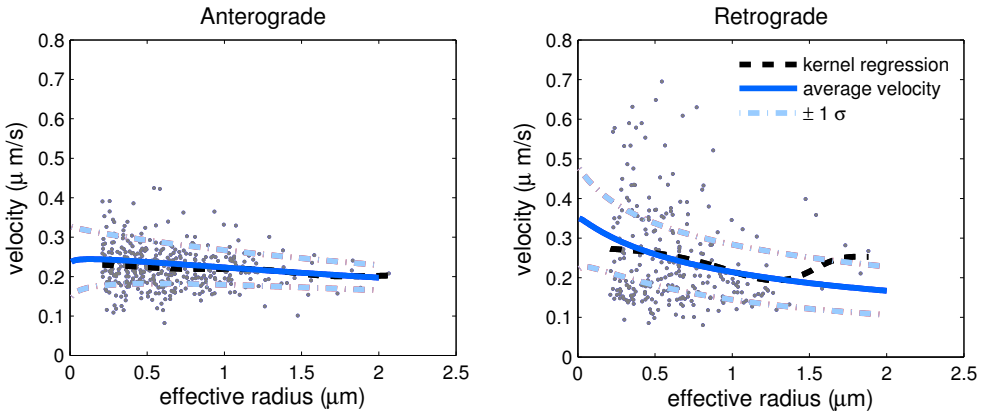

Figure 3.4: **Predicted tendency fits empirical data**. Each dot represents the size and velocity of an mitochondrion. The average velocity given a size is computed by kernel regression and depicted by the black dash curve. The blue solid curve is the predicted average velocity given size. Dashed lines indicates the regions within 1-standard deviation away the predicted average velocity.

ation away from the average velocity predicted by the proposed model. Parameters used are either obtained from literature or manually adjusted as summarized in Table 3.2. We used the model of non-cooperative and cooperative cargoes to plot the regions shown in panels of anterograde and retrograde motion respectively. The average velocity estimated by kernel smoothing is also shown in the figure 3.4. It can be seen that there is a good agreement between average velocity either predicted theoretically or accessed empirically.

3.3.5 Model predicts attach/detach rates of small vesicles

So far, we have shown that our model well predicts the relation of size and velocity of axonal moving mitochondria. We further looked at the behavior of cargoes with insignificant sizes (later called small vesicles). Our model predicted that the ratio of square mean $(\mathbb{E}(V)^2)$ to variance $(var(V))$ of small vesicles' velocities was equal to the ratio of attach to detach rate (π_a/ϵ_d) of motor proteins (equation 1 and 2).

To assess this prediction, we calculated the empirical ratio of average velocity to its standard deviation (E/SD ratio) and compared it with square root of π_a/ϵ_d ratio of kinesin-1, kinesin-3 and dynein motors estimated using literature reported values [10] (Table 3.1).

To assess these empirical values, we imaged the transport of APP and synaptotagmin1 in fly motor neuron or in mouse primary neurons. As previously suggested, the anterograde transport of APP and synaptotagmin1 vesicles are mainly driven by kinesin-1 motors and kinesin-3 motors respectively [16, 17, 18], while retrograde transport of both APP and synaptotagmin1 vesicles are driven by dynein motors [ref]. Mean, standard deviation, and the E/SD ratio of transport velocity are reported in table 2, in which the 99% confidence intervals of the E/SD ratios are estimated by bootstraping [ref]. We found the literature reported $(\pi_a/\epsilon_d)^{1/2}$ value of dynein motor was in the 99% confidence interval of the empirical E/SD values of retrograde transport of fly APP vesicles and synaptotagmin1 vesicles. Also, literature reported $(\pi_a/\epsilon_d)^{1/2}$ values of kinesin-1 and kinesin-3 are close to but statistically different from empirical E/SD values of APP and synaptotagmin1 vesicles (table 1). Overall, our model suggested that the binding behavior of motors π_a/ϵ_d can be estimated simply by evaluating the E/SD ratio. This prediction is particularly useful since the value of π_a/ϵ_d is generally unknown and experimentally inaccessible.

3.3.6 Model predicts changes of single-motor binding behaviors in neurodegeneration model

Next we proceed to investigate if the changes of E/SD ratio could reflects neurodegenerative deficits (ND) of axonal transport . Microtubule associate protein tau has long

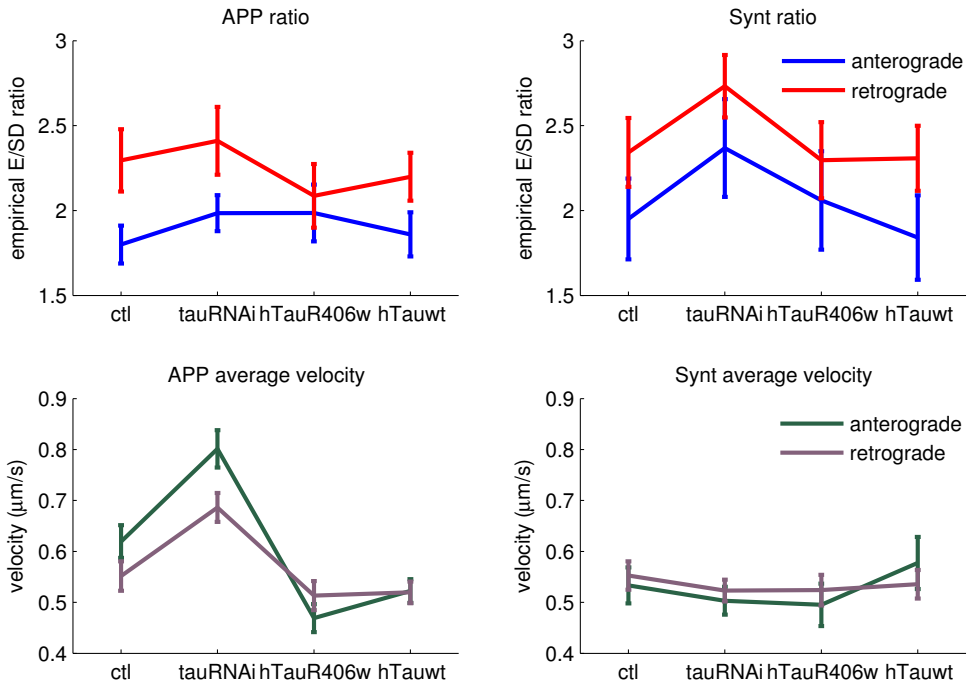

Figure 3.5: **E/SD Ratio and vesicle velocity under different expression level of tau proteins**. Error bar indicates 95 % confidence interval estimated by bootstrapping.

	Empirical values			Motor	Estimated (π_a/ϵ_d)
	E	SD	E/SD (95% C.I.)		
Fly APP anterograde vesicle	0.62	0.34	1.80 (1.75, 1.86)	Kinesin-1	2.24
Mouse APP anterograde vesicle	0.44	0.27	1.63 (1.24, 1.90)		
Fly APP retrograde vesicle	0.55	0.24	2.29 (2.20, 2.38)	Dynein	2.43
Fly Syt1 retrograde vesicle	0.55	0.24	2.34 (2.25, 2.43)		
Mouse APP retrograde vesicle	0.44	0.26	1.70 (1.17, 2.23)		
Fly Syt1 anterograde vesicle	0.53	0.27	1.97 (1.86, 2.09)	Kinesin-3	??

Table 3.3: E/SD ratios of APP and Sytn1 vesicle transport. E and SD represent the average and standard deviation of vesicle velocity respectively. C.I.: confidence interval

been suggested to play an important role in the pathology of NDs. Previous studies showed that tau expression and tau mutation impaired kinesin-1 [19, 20], kinesin-3 [21] and dynein [19] mediated transport. However, it is not clear that how tau protein mechanistically affects axonal transport driven by different motor proteins. In [22], tau is hypothesized as a "road block" for to modulate cargos' binding on a microtubule. Under this hypothesis, we expect to see 1) an increase in π_a/ϵ_d ratio under reduced tau expression and 2) a decrease in π_a/ϵ_d ratio under tau overexpression.

To see how π_a/ϵ_d ratio is changed in tau pathology models, we analyzed transport velocities of APP and synaptotagmin1 vesicles in axons harboring tau overexpression, tau mutation and tau knockdown. E/SD ratio is computed to estimate $(\pi_a/\epsilon_d)^{1/2}$ ratio and reported in Figure 3.5. We found E/SD ratios are all increased in tau knockdown for APP and synaptotagmin1 cargoes. In the cases of tau mutation or over expression, E/SD ratios either are slightly decreased or remain unchanged for all vesicles, with the exception of APP antegrograde vesicles.

For the vesicle motility, our data are consistent with previous studies [16, 14]. We found that the transport velocity of kinesin-1 cargo (i.e. anterograde transport of

APP vesicles) was the most sensitive to tau dose. Tau overexpression and mutation retarded anterograde transport velocity of APP vesicles by 15% and 24% respectively, while tau knockdown boosted anterograde transport velocity of APP vesicles by 30% (Figure 3.5). We found dynein and kinesin-3 driven transport were less affected by tau dose. Tau overexpression and mutation slightly reduced retrograde transport velocities of APP and synaptotagmin1 vesicles by 3~6%. Also tau knockdown, tau overexpression and mutation reduced or increased anterograde transport velocities of synaptotagmin1 vesicles by 6~7%.

3.4 Discussion

We presented a hierarchical model and its analysis on the dependency between cargo velocity and size. Our work generalized the model proposed by [10] by including factors such as cargo sizes, single motor stepping and microtubule distribution. Although we did not explore this direction, we would like to point out that the kinetics of ATP hydrolysis can be easily incorporated with our model.

Our model suggested that the motility of large cooperative cargoes depends critically on the way of how microtubules distribute around the cargoes. When the contact area between microtubules and cargoes grows sub-linearly with the increase of cargo size, larger cargoes would tend to move slower. Using axonal transport of mitochondria as an example, since the diameter of a microtubule (~20 nm) is around 25 folds smaller than typical radius of a mitochondrion (~500 nm), the contact area between these two should be determined by the dimension of a microtubule and hence grows sub-linearly with the cargo size. This sublinear dependency is also confirmed by our simulation. The decreasing tendency in cargo velocities of large mitochondria

observed in empirical data also supports our model prediction.

Our model also recovered the dynamics of cargoes with vanishing size. Our analysis suggested a way to evaluate the ratio of attach/detach rate at single molecule level from empirically accessible velocity measurements. The predicted value is close to the value in literature, despite the statistically significant difference in some cases. This model also predicted the changes of single-motor binding behaviors of APP and Synt1 cargoes under the mutation, overexpression and knockdown of microtubule-associate protein tau.

3.5 Appendix: Mathematical analysis of size-velocity relation

As it is set up in the main text, we use random variables as follows

1. N the total number of motors that is potentially can exert force on a cargo. N is a Poisson random variable with parameter $C_1 r^{\alpha}$.

2. M the number of motor that is exerting force on a cargo, i.e., the number of working motors. $M|N$ is a multi-nominal random variable, i.e., $\mathbb{P}(M = m|N = n) = P_m = \binom{n}{m}\left(\frac{\pi_a}{\epsilon_d}\right)^m / \left(1 + \frac{\pi_a}{\epsilon_d}\right)^n$ for $n \geq m$ and 0.

3. $S_{\Delta t}$ number of forward steps that motors move in Δt, $S_{\Delta t}|M = m$ is a Poisson random variable with parameter $\Delta t/\tau(m)$.

4. $V \equiv \frac{d}{\Delta t} S_{\Delta t}$ defines the velocity of a cargo within a timespan Δt

The cooperative strategy of motors is modeled by the dependency of $\tau(M)$ on M:

1. for non-cooperative motors, $1/\tau(M) = \begin{cases} \frac{v_{opt}}{d}(1 - \frac{C_0 r}{f_s + C_0 r}) & \text{if } M > 0 \\ 0 & \text{otherwise} \end{cases}$

2. for cooperative motors sharing forces evenly, $1/\tau(M) = \frac{v_{opt}}{d}(1 - \frac{C_0 r}{M f_s + C_0 r})$.

The goal here is provide mathematical details of our analysis on how the velocity changes with respect to cargo size r

3.5.1 Size-velocity relation for cargos carried by un-cooperative motor

First we compute the mean and variance of the cargo velocity.

Theorem 7. *Define the velocity process* $V \equiv \frac{d}{\Delta t} S_{\Delta t}$*, where* $S_{\Delta t}|M = m$ *is a Poisson random variable with parameter* $\Delta t/\tau(m)$*. Then*

$$\mathbb{E}(V|N > 0) = d\mathbb{E}(1/\tau(M)|N > 0), \text{ and}$$

$$var(V|N > 0) = \frac{d}{\Delta t}\mathbb{E}(V|N > 0) + d^2 var(1/\tau(M)|N > 0)$$

Proof. Let τ be a function of m denoted by $\tau(m)$, we can compute

$$\begin{aligned} \mathbb{E}(V|N > 0) &= \mathbb{E}(\mathbb{E}(V|M)|N > 0) \\ &= \mathbb{E}(\mathbb{E}(\tfrac{d}{\Delta t} S_{\Delta t}|M)|N > 0) \\ &= \mathbb{E}(d/\tau(M)|N > 0) \end{aligned}$$

and

$$\begin{aligned} var(V|N > 0) &= \mathbb{E}(var(V|M)|N > 0) + var(\mathbb{E}(V|M)|N > 0) \\ &= \mathbb{E}(\tfrac{d^2}{\Delta t \lambda(M)}|N > 0) + var(\tfrac{d}{\tau(M)}|N > 0) \end{aligned}$$

□

Note that the first term in variance is dependent on the time span of observing the motor. We shall ignore this term in the subsequent analysis since 1) we want to analyze the intrinsic noise of the cargo velocity and 2) this term is small under practical experimental setting. We can estimate $\mathbb{E}(\frac{d^2}{\Delta t \tau(M)}|N > 0) \leq \frac{d}{\Delta t} v_{opt}$. Plug in the values that $d = 8nm$, $v_{opt} = 0.4\mu m/s$ and $\Delta t = 3s$, we get this term is no more than $1.06 \times 10^{-3}(\mu m^2/s^2)$. AS will be shown latter, for small cargoes, the second term is roughly v_{opt}^2, which is of the value $0.16(\mu m^2/s^2)$.

In the case of non-cooperative motors, the average length of mechanochemical cycle is independent of the number of motors. With this assumption and theorem 1, we get:

Theorem 8. *If* $1/\tau(m) = 1/\tau$ *for* $m > 0$ *and 0 otherwise, then*

$$\mathbb{E}(V|N>0) = \frac{d}{\tau}\mathbb{P}(M>0|N>0)$$

$$var(V|N>0) = \frac{d^2}{\tau\Delta t}\mathbb{P}(M>0|N>0) + (\frac{d}{\tau})^2\mathbb{P}(M>0|N>0)(1-\mathbb{P}(M>0|N>0))$$

Following theorem shows the mean and variance of cargo velocity when the cargo size is significantly large.

Theorem 9. *If* $1/\tau(m) = \frac{v_{opt}}{d}(1-\frac{C_0 r}{f_s+C_0 r})$*, then*

$$\begin{aligned}
\lim_{r\to 0}\mathbb{E}(V|N>0) &= \frac{\pi_a/\epsilon_d}{1+\pi_a/\epsilon_d}v_{opt} \\
\lim_{r\to\infty}\mathbb{E}(V|N>0) &= 0 \\
\lim_{r\to 0} var(V|N>0) &= \frac{\pi_a/\epsilon_d}{(1+\pi_a/\epsilon_d)^2}v_{opt}^2 + \frac{d}{\Delta t}\frac{\pi_a/\epsilon_d}{1+\pi_a/\epsilon_d}v_{opt}.
\end{aligned}$$

3.5.2 Size-velocity relation for cargos carried by cooperative motor

For cooperative motors, we first compute the mean and variance of the cargo velocity when the cargo size is significantly small.

Theorem 10. *If* $1/\tau(m) = \frac{v_{opt}}{d}(1-\frac{C_0 r}{m f_s+C_0 r})$*, then*

$$\begin{aligned}
\lim_{r\to 0}\mathbb{E}(V|N>0) &= \frac{\pi_a/\epsilon_d}{1+\pi_a/\epsilon_d}v_{opt} \\
\lim_{r\to 0} var(V|N>0) &= \frac{\pi_a/\epsilon_d}{(1+\pi_a/\epsilon_d)^2}v_{opt}^2 + \frac{d}{\Delta t}\frac{\pi_a/\epsilon_d}{1+\pi_a/\epsilon_d}v_{opt}.
\end{aligned}$$

Proof. We shall show that $\frac{v_{opt}e^{-z_r}}{1-e^{-f(r)}}\sum_{m\geq 2}(\frac{mf_s}{mf_s+C_0r})(\frac{z_r^m}{m!}) \to 0$ as $r \to 0$ and hence

$$\begin{aligned}\lim_{r\to 0}\mathbb{E}(V|N>0) &= \frac{v_{opt}e^{-z_r}}{1-e^{-f(r)}}(\frac{f_sz_r}{f_s+C_0r}+\sum_{m\geq 2}(\frac{mf_s}{m+C_0r})(\frac{z_r^m}{m!})) \\ &= \lim_{r\to 0}\frac{v_{opt}e^{-z_r}}{1-e^{-f(r)}}\frac{f_sz_r}{f_s+v_{opt}C_0r} \\ &= \frac{\pi_a \diagup \epsilon_d}{1+\pi_a \diagup \epsilon_d}v_{opt}.\end{aligned}$$

The first limit is because

$$\begin{aligned}\sum_{m\geq 2}(\frac{mf_s}{mf_s+C_0r})(\frac{z_r^m}{m!}) &= \sum_{m\geq 2}(\frac{f_s}{mf_s+C_0r})(\frac{z_r^m}{(m-1)!}) \\ &= z_r\sum_{m\geq 1}(\frac{f_s}{(m+1)f_s+C_0r})(\frac{z_r^m}{m!}) \\ &\leq \frac{z_rf_s}{2f_s+C_0r}\sum_{m\geq 1}(\frac{z_r^m}{m!}) \\ &= \frac{z_rf_s}{2f_s+C_0r}(e^{z_r}-1) \to 0 \text{ as } r\to 0\end{aligned}$$

Similarity, since $0 \leq (\frac{mf_s}{mf_s+C_0r})^2 \leq (\frac{mf_s}{mf_s+C_0r})$, we have

$$\frac{v_{opt}^2e^{-z_r}}{1-e^{-f(r)}}\sum_{m\geq 2}(\frac{mf_s}{mf_s+C_0r})^2(\frac{z_r^m}{m!}) \to 0 \text{ as } r\to 0$$

and hence

$$\lim_{r\to 0}\mathbb{E}(V^2|N>0) = \lim_{r\to 0}\frac{e^{-z_r}z_r}{1-e^{-f(r)}}(\frac{v_{opt}f_s}{f_s+v_{opt}C_0r})^2 = \frac{\pi_a \diagup \epsilon_d}{1+\pi_a \diagup \epsilon_d}v_{opt}^2$$

overall

$$\begin{aligned}\lim_{r\to 0}var(V|N>0) &= \lim_{r\to 0}\mathbb{E}(V^2|N>0)-\mathbb{E}(V|N>0)^2 \\ &= \frac{\pi_a \diagup \epsilon_d}{1+\pi_a \diagup \epsilon_d}v_{opt}^2-(\frac{\pi_a \diagup \epsilon_d}{1+\pi_a \diagup \epsilon_d})^2v_{opt}^2 \\ &= \frac{\pi_a \diagup \epsilon_d}{(1+\pi_a \diagup \epsilon_d)^2}v_{opt}^2\end{aligned}$$

□

Now we compute the mean velocity for large cargoes carried by cooperative motors.

Theorem 11. *For cooperative motors with* $1/\tau(m) = \frac{v_{opt}}{d}(1 - \frac{C_0 r}{m f_s + C_0 r})$*, then*

$$\lim_{r\to\infty} \mathbb{E}(V|N>0) = \begin{cases} 0 & \text{if } \alpha < 1 \\ v_{opt} & \text{if } \alpha \geq 1 \end{cases}$$

Proof. First we compute the upper bound of $\mathbb{E}(V|N>0)$.

$$\begin{aligned}
\mathbb{E}(V|N>0) &= v_{opt} \textstyle\sum_m (1 - \frac{C_0 r}{m f_s + C_0 r})(\frac{z_r^m}{m!} e^{-z_r}) \\
&= v_{opt}(\textstyle\sum_m (\frac{m f_s}{m f_s + C_0 r})(\frac{z_r^m}{m!} \frac{e^{-z_r}}{1-e^{-f(r)}})) \\
&\leq v_{opt}(\frac{z_r}{C_0 r} \textstyle\sum_m (\frac{m f_s}{C_0 r})(\frac{z_r^m}{m!} \frac{e^{-z_r}}{1-e^{-f(r)}})) \\
&= v_{opt}(\frac{f_s}{C_0}) \frac{z_r}{r(1-e^{-f(r)})} \\
&= \mathcal{O}(r^{\alpha-1}) \to 0 \text{ if } \alpha < 1 \text{ as } r \to \infty
\end{aligned}$$

Next we compute the lower bound of $\mathbb{E}(V|N>0)$.

$$\begin{aligned}
\mathbb{E}(V|N>0) &= v_{opt} \textstyle\sum_m (1 - \frac{C_0 r}{m f_s + C_0 r})(\frac{z_r^m}{m!} \frac{e^{-z_r}}{1-e^{-f(r)}}) \\
&= v_{opt}(1 - \textstyle\sum_m (\frac{C_0 r}{m f_s + C_0 r})(\frac{z_r^m}{m!} \frac{e^{-z_r}}{1-e^{-f(r)}})) \\
&= v_{opt}(1 - \frac{C_0 r}{z_r} \textstyle\sum_m (\frac{1}{m f_s + C_0 r})(\frac{z_r^{m+1}}{m!} \frac{e^{-z_r}}{1-e^{-f(r)}})) \\
&= v_{opt}(1 - \frac{C_0 r}{z_r} \textstyle\sum_m (\frac{m+1}{m f_s + C_0 r})(\frac{z_r^{m+1}}{(m+1)!} \frac{e^{-z_r}}{1-e^{-f(r)}})) \\
&\geq v_{opt}(1 - \frac{C_0 r}{z_r (1-e^{-f(r)}) f_s}) \text{ for large } r \\
&\to v_{opt} \text{ if } \alpha > 1 \text{ as } r \to \infty
\end{aligned}$$

Clearly, $\mathbb{E}(V|N>0) \leq v_{opt}$. □

Bibliography

[1] H. M. McBride, M. Neuspiel, and S. Wasiak, "Mitochondria: more than just a powerhouse," *Current Biology*, vol. 16, no. 14, pp. R551–R560, 2006.

[2] J. Nunnari and A. Suomalainen, "Mitochondria: in sickness and in health," *Cell*, vol. 148, no. 6, pp. 1145–1159, 2012.

[3] Z.-H. Sheng, "Mitochondrial trafficking and anchoring in neurons: new insight and implications," *The Journal of Cell Biology*, vol. 204, no. 7, pp. 1087–1098, 2014.

[4] D. Chan, "Fusion and fission: interlinked processes critical for mitochondrial health," *Annual Review of Genetics*, vol. 46, pp. 265–287, 2012.

[5] H. Chen and D. Chan, "Critical dependence of neurons on mitochondrial dynamics," *Current Opinion in Cell Biology*, vol. 18, no. 4, pp. 453–459, 2006.

[6] A. Kolomeisky and M. Fisher, "Molecular motors: a theorist's perspective," *Annual Review of Physical Chemistry*, vol. 58, pp. 675–695, 2007.

[7] P. Bressloff and J. Newby, "Stochastic models of intracellular transport," *Reviews of Modern Physics*, vol. 85, no. 1, p. 135, 2013.

[8] A. Yildiz, M. Tomishige, R. Vale, and P. Selvin, "Kinesin walks hand-over-hand," *Science*, vol. 303, no. 5658, pp. 676–678, 2004.

[9] S. Reck-Peterson, A. Yildiz, A. Carter, A. Gennerich, N. Zhang, and R. Vale, "Single-molecule analysis of dynein processivity and stepping behavior," *Cell*, vol. 126, no. 2, pp. 335–348, 2006.

[10] S. Klumpp and R. Lipowsky, "Cooperative cargo transport by several molecular motors," *PNAS*, vol. 102, no. 48, pp. 17 284–17 289, 2005.

[11] K. Svoboda and S. Block, "Force and velocity measured for single kinesin molecules," *Cell*, vol. 77, no. 5, pp. 773–784, 1994.

[12] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan, *Statistical analysis and modelling of spatial point patterns*. John Wiley & Sons, 2008, vol. 70.

[13] J. Wortman, U. Shrestha, D. Barry, M. Garcia, S. Gross, and C. Yu, "Axonal transport: How high microtubule density can compensate for boundary effects in small-caliber axons," *Biophysical Journal*, vol. 106, no. 4, pp. 813–823, 2014.

[14] G. Reis, G. Yang, L. Szpankowski, C. Weaver, S. Shah, J. Robinson, T. Hays, G. Danuser, and L. Goldstein, "Molecular motor function in axonal transport in vivo probed by genetic and computational analysis in Drosophila," *Molecular Biology of the Cell*, 2012.

[15] Y. Yu, H.-C. Lee, K.-C. Chen, J. Suhan, M. Qiu, and G. Yang, "Inner membrane fusion mediates spatial distribution of axonal mitochondria," *submitted to Scientific Reports*.

[16] R. Barkus, O. Klyachko, D. Horiuchi, B. Dickson, and W. Saxton, "Identification of an axonal kinesin-3 motor for fast anterograde vesicle transport that facilitates

retrograde transport of neuropeptides," *Molecular Biology of the Cell*, vol. 19, no. 1, pp. 274–283, 2008.

[17] A. Kamal, G. Stokin, Z. Yang, C.-H. Xia, and L. Goldstein, "Axonal transport of amyloid precursor protein is mediated by direct binding to the kinesin light chain subunit of kinesin-I," *Neuron*, vol. 28, no. 2, pp. 449–459, 2000.

[18] O. Lazarov, G. Morfini, E. Lee, M. Farah, A. Szodorai, S. DeBoer, V. Koliatsos, S. Kins, V. Lee, P. Wong, and Others, "Axonal transport, amyloid precursor protein, kinesin-1, and the processing apparatus: revisited," *The Journal of Neuroscience*, vol. 25, no. 9, pp. 2386–2395, 2005.

[19] R. Dixit, J. Ross, Y. Goldman, and E. Holzbaur, "Differential regulation of dynein and kinesin motor proteins by tau," *Science*, vol. 319, no. 5866, pp. 1086–1089, 2008.

[20] K. Stamer, R. Vogel, E. Thies, E. Mandelkow, and E. Mandelkow, "Tau blocks traffic of organelles, neurofilaments, and APP vesicles in neurons and enhances oxidative stress," *The Journal of Cell Biology*, vol. 156, no. 6, pp. 1051–1063, 2002.

[21] N.-W. Tien, G.-H. Wu, C.-C. Hsu, C.-Y. Chang, and O. Wagner, "Tau/PTL-1 associates with kinesin-3 KIF1A/UNC-104 and affects the motor's motility characteristics in C. elegans neurons," *Neurobiology of Disease*, vol. 43, no. 2, pp. 495–506, 2011.

[22] P. Baas and L. Qiang, "Neuronal microtubules: when the MAP is the roadblock," *Trends in Cell Biology*, vol. 15, no. 4, pp. 183–187, 2005.

Chapter 4

Characterize spatiotemporal dynamics of intracellular transport at the whole-cell scale

4.1 Introduction

Intracellular transport is the process of distributing and collecting cargoes to meet the changing structural and functional needs at different sites inside a cell. Because the intracellular environment is highly dynamic and heterogeneous, a fundamental question regarding intracellular transport is how it is controlled spatially and temporally to deliver cargoes to the right place at the right time. So far, the molecular machinery of intracellular transport has largely been identified [1]. However, the spatial and temporal control mechanism of intracellular transport remains poorly understood, especially at the whole-cell scale. Addressing this question requires computational methods for quantitative characterization of whole cell-scale spatiotemporal dynamics of intracellular transport. However, such methods are currently lacking.

To address this problem, we developed image-based computational methods for characterizing spatiotemporal dynamics of intracellular transport. As a first step, we developed a computational method to decompose signals of stationary and moving cargoes into separate image channels. Specifically, we chose the transport of lysosome associated membrane protein 1 positive (Lamp1) cargoes as our model system. Because of the highly dynamic nature of Lamp1 transport, a main challenge is to infer its steady-state localization. [2]. Our method addresses this challenge by providing a way to partition the spatiotemporal dynamics of Lamp1 transport into two parts: the localization of stationary cargoes in space and the pattern of cargo movement in space and time.

4.2 Characterizing global spatiotemporal dynamics of intracellular transport

Tracking movement of individual cargoes is often the first step to characterize intracellular transport. Currently, the common practice is to first locate individual cargoes through particle detection and then follow their movement using single particle tracking techniques. This approach, however, has several limitations. First, cargoes are often labeled by fluorescent proteins. Due to factors such as variation of expression level and diffusion of fluorescent proteins, the images often have strong background and low signal-to-noise ratio. This is known to cause significant degradation of performance in single particle tracking [3]. Second, when the spatial density of cargoes is high, to resolve movement of individual particles is often not reliable. Furthermore, complete trajectories of individual particles may not be essential for identifying global patterns of cargo movement.

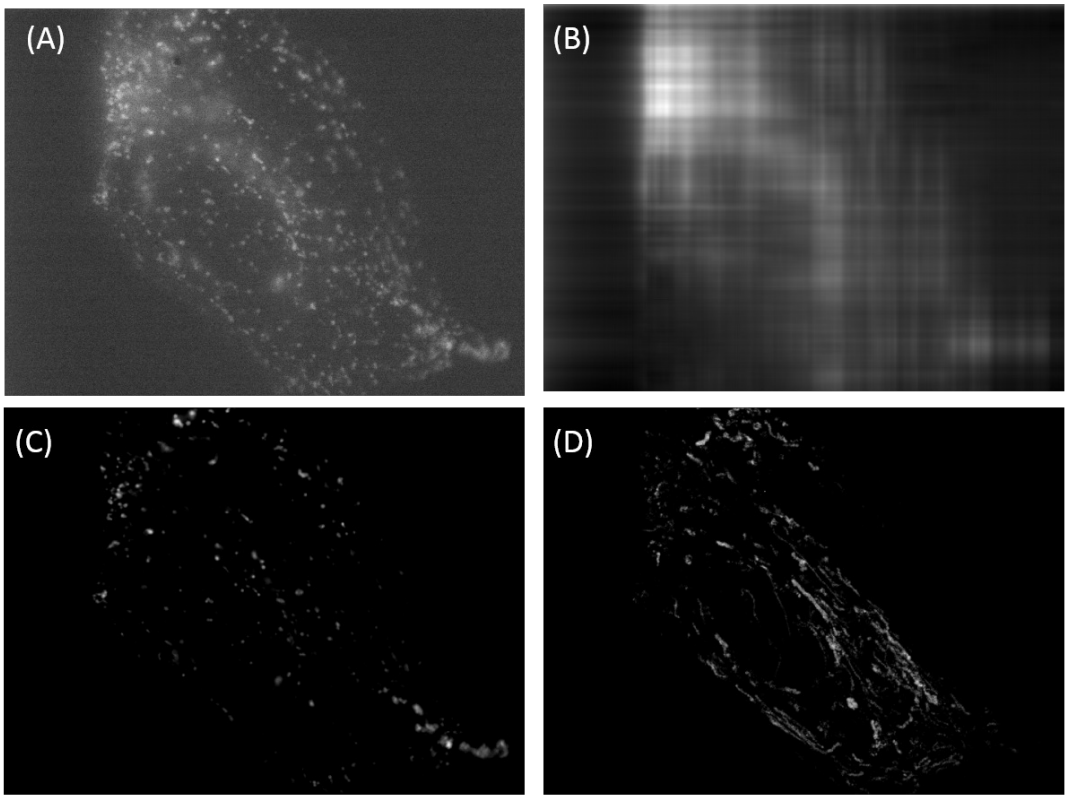

Figure 4.1: **Decomposition of an image sequence of the Lamp1 transport**. (A) The first frame of the image sequence. (B) Decomposed background fluorescence (C) Decomposed signals from stationary Lamp1 cargoes. (D) A 2D flow map generated from decomposed signals of moving Lamp1 cargoes.

Here we propose a signal decomposition method to characterize whole-cell scale cargo transport without single particle tracking. The main idea is to separate signals from background fluorescence, stationary and moving cargoes using a signal decomposition approach. The separation of moving cargoes from stationary cargoes makes it possible for us to characterize global organization of these two types of cargoes separately. The signal decomposition method for separation of signal of background fluorescence in a static image was first developed in [4] and then extended to dynamic time-lapse images [5].

4.2.1 Image formation

We consider the image model as follows:

$$I_t = B_t + g * (S_t + M_t) + \epsilon_t,$$

where I_t, B_t, S_t and M_t are m-by-n matrices representing the given image, background fluorescence signal, the foreground fluorescence signal from stationary and moving cargoes, respectively. ϵ_t denotes a zero mean random noise. The subscript t stands for time t and g is the point spread function of the microscope. Throughout this paper, we use symbols (for example, I) without subscript t to refer to 3-fold tensors representing stacks of matrices [6]. Our goal is to separate B, S, and M into three separate channels based on their different spatiotemporal characteristics. Typically background fluorescence is diffusive in space and slow-varying in time [4]. Thus, we assume the background fluorescence B_t is structured and thus low-rank both in space and time. Signal from stationary cargoes is slow-varying in time and sparsely distributed in space. Thus, we assume the foreground fluorescent signal from stationary cargoes S_t is structured and thus low-rank in time and sparse in space. Signal M_t from moving cargoes is also assumed to be sparse in space. In the following we propose a two-stage approach to separate signals form stationary/moving cargoes and the background fluorescence based on these characteristics. The characteristics of each signal is summarized in Table 4.1

The main idea of the proposed approach is to successively separate signals in to a structured part and a sparse part. In the first stage, we separate the combination of B_t and S_t, called A_t apart from M_t. As can be seen in Table 4.1, A_t is stable in time and M_t is only sparse. At the second stage, we further decompose the combination A_t

Table 4.1: Characteristics of signals

	Signal Source	Space	Time	Sparsity
B_t	Background fluorescence	structured	stable	no
S_t	Stationary cargo	not structured	stable	yes
M_t	Moving cargo	not structured	not stable	yes

into B_t and S_t. And this is done similarly as in the first stage since B_t is structured in space while M_t is only sparse in space.

4.2.2 Separation of signals of moving cargoes

To separate the signal of moving particles, we formulate it as the following optimization:

$$[A^*, M^*] = \text{argmin}_{A,\,M} \frac{1}{2}\|I - (A + g * M)\|_2^2 + \lambda_1 \|A_{(3)}\|_* + \lambda_2 \|M\|_1, \qquad (4.1)$$

subject to $M \geq 0$. $(\cdot)_{(3)} : R^{m\times n\times k} \to R^{k\times(nm)}$ is the unfolding operator along time dimension [6], i.e., the operation to stack vectors of the intensity profile over time at a fixed pixel into a matrix. The nuclear norm $\|(\cdot)\|_*$, defined by the sum of singular values of a matrix, is a convex relaxation of matrix rank. Using this penalty, we expect to confine A to have less complicated behaviors in time and therefore contains only signals from background fluorescence and stationary cargoes. Also, the l_1 norm $\|(\cdot)\|_1$, defined by the sum of absolute values of elements in a matrix, is a sparsity measurement, which would enforce signals of sparsely distributed moving cargoes into M^*. This optimization is the relaxed optimization problem of robust principal component analysis [7] with a modified data error term to account for the point spread function.

We solved this optimization problem by the forward-backward algorithm [8]. The update rules of forward-backward algorithm are derived as follows: Given initial

values A_0 and M_0, iteratively update

$$A^+ = (\mathcal{S}^*_{h\lambda_2}((A - h(A + g * M - I))_{(3)}))^{(3)},$$

$$M^+ = \mathcal{S}^{1+}_{h\lambda_3}(M - h\hat{g} * (A + g * M - I)),$$

where $(\hat{g})_{i,j} = g_{-i,-j}$ and the operation $(\cdot)^{(3)} : R^{k\times(nm)} \to R^{m\times n\times k}$ is the inverse operator of the unfolding operator $(\cdot)_{(3)}$ defined above. $\mathcal{S}^*_{h\lambda_2}$ and $\mathcal{S}^{1+}_{h\lambda_3}$ are soft thresholding operators on the singular values and values of elements respectively as defined in the appendix.

4.2.3 Separation of signals of stationary cargoes

In the second stage, the signals S_t and B_t are further decomposed from their combination A_t by solving the optimization problem

$$[B^*_t, S^*_t] = \text{argmin}_{B_t,\, S_t} \frac{1}{2}\|A^*_t - (B_t + g * S_t)\|^2_2 + \lambda_3\|B_t\|_* + \lambda_4\|S_t\|_1, \qquad (4.2)$$

subject to $S_t \geq 0$. A^* is obtained from the previous step and symbols with subscript t denotes t^{th} frame in the corresponding image sequence. We would like to point out that this algorithm can be stand-alone for removing the background fluorescence in a single image with sparse point features [4]. A typical example is shown in (**Fig. 1A**). We note that this optimization problem can be solved similarly by the forward-backward algorithm [8, 4]. The update rules of FB are as follows [4]:

$$B^+_t = \mathcal{S}^*_{h\lambda_3}(B_t - h(B_t + g * S_t - A^*_t)),$$

$$S_t^+ = \mathcal{S}_{h\lambda_4}^{1+}(S_t - h\hat{g} * (B_t + g * S_t - A_t^*)),$$

where $(\hat{g})_{i,j} = g_{-i,-j}$. Detailed derivation of the updating rule and convergence analysis can be found in the Appendix. Specifically, the gradient of this objective function is 2−Lipschitz continuous provided that $\sum_{i,j} g_{i,j} = 1$ and $g_{i,j} \geq 0$. In this case, we can choose a constant step size for convenient implementation and rapid convergence. We choose $t = 1.9/2$ throughout this paper.

4.2.4 Representing global structure and flow of cargoes

Fig. 1 shows an example of three separated signals in an image sequence. With our decomposition method, a visualization of two-dimensional cargo transport is conveniently provided by taking maximum over the temporal dimension of the 3-order tensor M^* (**Fig. 1D**). We define the *2D-flow map* as

$$\hat{M}_{i,j} = \max_k M^*_{i,j,k}.$$

The 2D flow map visualizes the spatiotemporal dynamics of a two-dimensional transport process as curvilinear features of objects in an image. As illustrated in **Fig. 1D**, 2D flow map provide a direct representation of intracellular transport at the whole-cell scale and is suitable for studying its global dynamics. Furthermore, representing the spatiotemporal dynamics of intracellular transport in a flow map allows the usage of image processing techniques, such as multi-resolution analysis, to study global organization of intracellular transport.

Similarly, the localization of stationary organelles can be visualized as an image

$$\hat{S}_{i,j} = \max_k S^*_{i,j,k}.$$

Overall, we propose to characterize the spatiotemporal dynamics of intracellular transport at the whole-cell scale by characterizing the steady-state distributions of stationary cargoes in conjunction with the dynamic patterns of cargo movement. We refer these two fine organization as the *structure* and *flow* of intracellular transport.

4.3 Validation and applications

4.3.1 Validation by simulation

syndetic image data

To test our algorithm, we generated a synthetic image with a known background according to

$$B^{\text{true}}_{i,j,k} = A \times (2 - \frac{1}{2}[(X_{i,j} - Y_{i,j})^2 + \frac{1}{c^2}(X_{i,j} + Y_{i,j})^2])$$

where $X_{i,j} = -1 + 0.01 \times i$ and $Y_{i,j} = -1 + 0.01 \times j$ for $0 \leq i,\ j \leq N$. N represents the width and height of the simulation image and is chosen to be 201 in our case. The profile of B^{true} is chosen to have elliptic level sets with aspect ratio $1/c$ and intensity values within $[A,\ 2A]$. Specifically, we choose $A = 1000$ and $c = 2$. The foreground F^{true} consists of 100 point features uniformly distributed in the first image plane with intensity level 500. Among them, 50 point features are selected to perform random walk with fixed drift while the other 50 remain stationary. 10 frames of synthetic images are then generated as the sum of background and foreground signals plus an additive Gaussian noise with different noise levels. The variance of the Gaussian noise ranges from 0, 100 to 200 so that the corresponding signal ratio are ∞, 5 and 2.5, respectively.

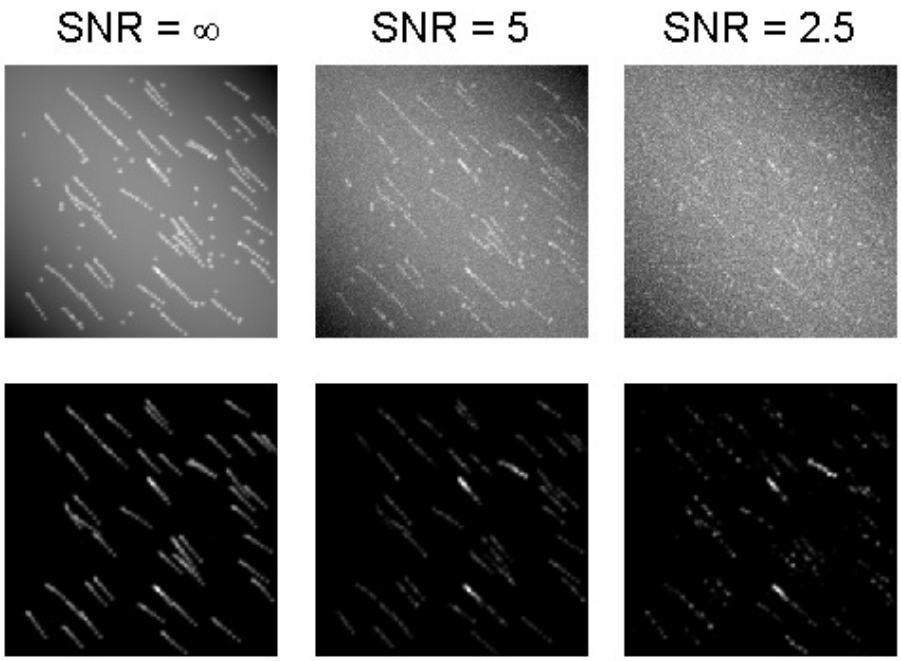

Figure 4.2: **Reconstructed 2D flow maps from synthetic image data**. Upper row shows the maximum projection of simulated image sequences under different noise level. Reconstructed 2D-flow maps are shown in the lower row.

validation on M

To evaluate the performance of our algorithm, we report the F1-score to see if the reconstructed 2D-flow map contain information of trajectories in ground truth images. Specifically, we define the F1-score of a matrix $M \in R^{m \times n}$ as

$$\mathrm{F1}(M) = 2 \times \frac{\mathrm{precision}(M) \times \mathrm{recall}(M)}{\mathrm{precision}(M) + \mathrm{recall}\ (M)},$$

while

$$\mathrm{precision}(M) = \frac{\#\{M > 0 \cap M_{\mathrm{true}} > 0\}}{\#\{M > 0\}} \text{ and}$$

$$\mathrm{recall}(M) = \frac{\#\{M > 0 \cap M_{\mathrm{true}} > 0\}}{\#\{M_{\mathrm{true}} > 0\}}.$$

Table 4.2: F1 score

SNR= ∞

$\lambda_1 \setminus \lambda_2$	10	20	40
1000	88.5 ± 0.6	90.1 ± 0.2	84.9 ± 0.5
2000	61.5 ± 1.3	88.3 ± 0.6	90.1 ± 0.4
3000	50.4 ± 1.2	86.1 ± 0.8	88.3 ± 0.5
5000	49.6 ± 1.2	85.4 ± 1.0	87.5 ± 0.6

SNR= 5

$\lambda_1 \setminus \lambda_2$	10	20	40
1000	85.8 ± 0.7	1.3 ± 0.5	0 ± 0.0
2000	56.8 ± 1.0	85.2 ± 0.8	8.7 ± 2.5
3000	48.6 ± 0.8	88.1 ± 0.7	85.0 ± 0.8
5000	48.6 ± 0.8	48.7 ± 0.7	81.1 ± 1.3

SNR= 2.5

$\lambda_1 \setminus \lambda_2$	10	20	40
1000	15.8 ± 1.6	0 ± 0.0	0 ± 0.0
2000	56.5 ± 0.6	11.1 ± 1.5	0 ± 0.0
3000	49.4 ± 0.8	79.2 ± 1.2	8.4 ± 1.7
5000	49.4 ± 0.8	49.5 ± 0.8	75.7 ± 0.7

The performance of our method is tested with various parameters λ_1 and λ_2. In our simulation, we found that the optimal result determined by F1-score agrees with visual inspection quite well. Examples of reconstructed 2D-flow maps determined by F1-score are shown in figure 4.2. Detailed results are summarized in Table 4.2. Due to space limitation, only cases with medium (SNR = 5) and low (SNR = 2.5) signal-to-noise ratio are reported. Note that the signal-to-noise ratio of the last simulated example is lower than 4 [3], which is known to cause substantial performance degradation of single particle tracking. Our algorithm can significantly enhance image quality and can be used as pre-processing to enhance the performance of single particle tracking.

validation on background removal

To evaluate the performance of the second stage of our algorithm, we use two error metrics for reconstructed foreground and background, respectively. For the reconstructed foreground, we use the normalized root mean square error with respect to the intensity level of point features. Specifically, we define the normalized root mean square error of a matrix $F \in R^{m \times n}$ as

$$nRMSE(F) = \sqrt{\frac{\sum_{i,j}(F_{i,j} - F^{\text{true}}_{i,j})^2}{m \times n \times \sigma^2}}.$$

For the reconstructed background, we report the normalized sum-of-square error, i.e.,

$$E^2(B) = \|B - B^{\text{true}}\|_F \big/ \|B^{\text{true}}\|_F.$$

We test our algorithm with various values of λ_1 while keeping $\lambda_2 = 20$. Results are summarized in Table 4.3. In [9], a background estimation method based on filtering spiky signals from first singular vectors of each image frame is proposed (TADN). In comparison, our algorithm gives better results.

4.3.2 Application to real image data

background removal

We first tested the performance of our algorithm on removing the background signal of diffusive nature from real images. Figure 4.3 (A) and (B) show results of processing images with point features and diffusive background signals. These images are single frames from image sequences of stochastic optical reconstruction microscopy (STORM). The computation typically takes a tens of seconds to be completed on a

Table 4.3: Reconstruction Errors

$E^2(B)$ (‰)

$\sigma \setminus \lambda_1$	100	200	400	1000	TADN
0	6.1 ± .24	2.5 ± .11	4.6 ± .06	26 ± 0.10	70 ± .007
5	6.4 ± .25	2.5 ± .09	4.6 ± .05	26 ± 0.10	70 ± .007
10	7.3 ± .22	2.6 ± .08	4.8 ± .06	26 ± 0.11	70 ± .007
50	26 ± .11	13 ± .11	14 ± .21	57 ± 0.71	70 ± .012

$nRMSE(F)$ (‰)

$\sigma \setminus \lambda_1$	100	200	400	1000	TADN
0	28 ± 8.9	19 ± .54	18 ± .57	72 ± 0.36	252 ± .02
5	29 ± 9.6	19 ± .57	19 ± .57	73 ± 0.34	253 ± .05
10	29 ± 9.0	19 ± .55	20 ± .54	74 ± 0.38	254 ± .10
50	30 ± 8.9	48 ± .54	84 ± .65	253 ± 2.9	272 ± .46

desktop workstation (2× Intel Xeon E5503 2.00 GHz and 8G RAM). Diffusive background is often a major technical barrier in STORM that often leads to significant performance deterioration in subsequent image processing. So far, overcoming this barrier relies on laborious optimization of experimental protocols to reduce diffusive fluorophores. Our algorithm provides a convenient computational solution to remove such kind of background signals from a single image and hence may substantially lower the experiment barrier of STORM imaging. Figure 4.3 (C) shows another processed image of a *Xenopus* egg extract spindle under wide-field fluorescence speckle microscopy.

global patterns of transport

To visualize the transport of Lamp1 cargo, BSC-1 cells were transfected with Lamp1-green fluorescent protein (Invitrogen) and visualized by fluorescence live imaging. Images were collected on a Nikon Ti-E inverted microscope at 5 frames per second, under a NA of 1.41 and a 100x magnification.

Figure 4.4 shows three 2D-flow maps processed from collected images with manual optimization of parameters. Surprisingly, these selected examples revealed different global flow patterns of Lamp1 cargoes. For example, figure 4.4 (B) shows the cargo flow is moving in and out of the cell center, while figure 4.4 (C) shows the cargo flow is polarized along the cell periphery. We are currently investigating physiological implications of these patterns.

4.4 Conclusions

In this study we developed image-based computational methods for characterizing the spatiotemporal dynamics of intracellular transport at the whole-cell scale. Key

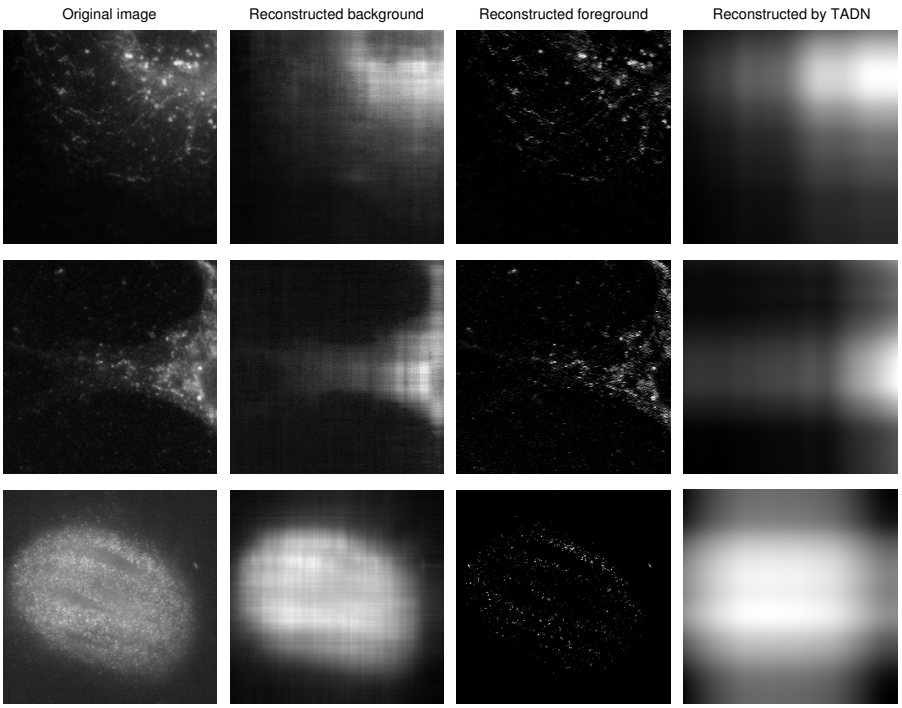

Figure 4.3: Background removal for single-particle images. First column: original images. Second column: separated background. Third column: images after background removal. Fourth column: background estimated by TADN. Note that original image in panel (C) is larger than other two figures. Here, we set λ_1 = 1000 and λ_2 = 20 for panels (A) and (B) and λ_1 = 300 and λ_2 = 5 for panel (C). Contrast has been adjusted for each image.

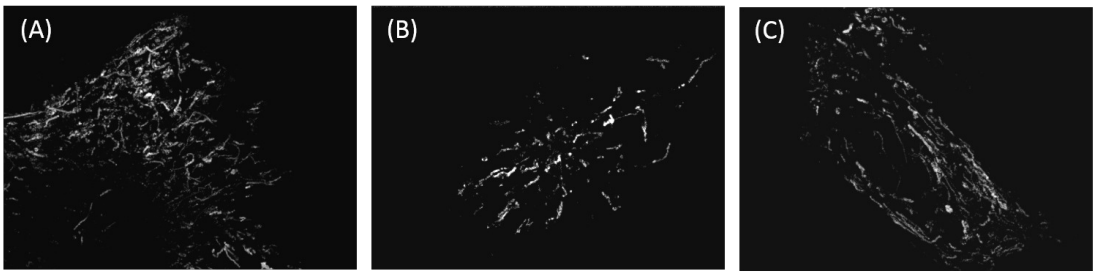

Figure 4.4: Examples of different 2D flow patterns. (A) unorganized flow (B) radial flow (C) polarized flow.

to these methods is a signal decomposition algorithm, which allows us to separate the background fluorescence signal, the stationary cargo signal, and the moving cargo signal into three different channels based on their different spatiotemporal properties. We have demonstrated our method's applicability to recover the movement of cargoes even when the signal-to-noise is low. Furthermore, using this method, we found different global patterns in intracellular transport of Lamp1 cargoes. The 2D flow map generated can be combined with different curvilinear feature detection techniques for further quantitative characterization of the spatiotemporal dynamics of intracellular transport. Although the focus of this study is on intracellular transport, the signal decomposition method we developed is general and applicable to other cellular processes that are composed of signals of different spatiotemporal properties.

4.5 Appendix

4.5.1 Forward-backward algorithms

We present an overview of the forward-backward algorithm (FB) [8] used for solving proposed optimization problems. FB is designed to solve optimization problems whose objective function $f(x)$ is separable into two parts, i.e., $f(x) = f_1(x) + f_2(x)$. Here, x is an element in a Hilbert space and $f_1(x)$ can be any differentiable function and $f_2(x)$ is a non-differentiable penalty term. Given a function f, the proximal operator with a step size t is defined by

$$prox_{tf_2}(x) = \text{argmin}_z \frac{1}{2}\|z - x\|^2 + tf_2(z).$$

The essence of FB is the analytically derived proximal operator. Examples of the proximal operator are

1. $f_2(x) = \lambda\|x\|_1$ for $x \in R^n$. The proximal operator in this case is the soft thresholding on components of the vectors x [10], i.e.,

$$[prox_{tf_2}(x)]_i = [S^1_{t\lambda}(x)]_i = \text{sgn}(x_i)\max(|x_i| - t\lambda,\ 0).$$

2. $f_2(x) = \lambda\|x\|_1 + \chi_+(x)$ for $x \in R^n$, where $\chi_+(x)$ is the characteristic function of the half space $\{x \geq 0\}$. The proximal operator in this case is the modified soft thresholding on components of the vectors x, i.e.,

$$[prox_{tf_2}(x)]_i = [S^{1+}_{t\lambda}(x)]_i = \max(x_i - t\lambda,\ 0).$$

3. $f_2(X) = \lambda\|X\|_*$ for $X \in R^{m\times n}$. The proximal operator in this case is the soft thresholding on singular values of the matrix X [11], i.e.,

$$prox_{tf_2}(X) = S^*_{t\lambda}(X) = U\Sigma_{t\lambda}V^T,$$

where $U\Sigma V^T$ is the singular value decomposition of X and $[\Sigma_{t\lambda}]_{i,j} = \max([\Sigma]_{i,j} - t\lambda, 0)$, assuming that all of the singular values are chosen to be non-negative.

Iterative update in FB follows

$$x_{k+1} = prox_{tf_2}(x_k - t\nabla f_1(x_k)), \tag{4.3}$$

where t is a step size constant. FB is guaranteed to converge with rate $\mathcal{O}(1/k)$ provided that ∇f_1 is L-Lipschitz continuous and $t < 2/L$.

Following we derive the updating rule of the optimization problem 4.2

$$\begin{aligned} &\min_{B_t,\, S_t} \quad \tfrac{1}{2}\|A_t^* - (B_t + g * S_t)\|_2^2 + \lambda_3\|B_t\|_* + \lambda_4\|S_t\|_1, \\ &\text{subject to} \quad S_t \geq 0 \end{aligned}$$

and analyze the convergence of FB. The results can be adapted for the optimization problem 4.1 with minor changes.

4.5.2 Derivation of the updating rule

To derive the updating rule of the forward-backward algorithm, we compute the proximal operator of $g(B, S) = \lambda_3\|B\|_* + \lambda_4\|vec(S)\|_1$ with $S \geq 0$. First we reformulate the problem. Let $Z = [Z_1\ Z_2]$ for each $Z_i \in R^{m\times n}$. The objective function can be rewritten as

$$f(Z) = \frac{1}{2}\|Z_1 + g * Z_2 - A\|_F^2 + \lambda_3\|Z_1\|_* + \lambda_4\|vec(Z_2)\|_1 + \chi_+(Z_2)$$

where χ_+ is the characteristic function of the nonnegative real line. We first show that the proximal operator of $\lambda_1\|Z_1\|_* + \lambda_2\|vec(Z_2)\|_1 + \chi_+(Z_2)$ is

$$prox_t(Z) = [S^*_{t\lambda_3}(Z_1)\ S^{1+}_{t\lambda_4}(Z_2)].$$

This is achieved by observing that the objective function is separable in Z_1 and Z_2, i.e., denoting $Z = [Z_1\ Z_2]$ and $X = [X_1\ X_2]$,

$$\begin{aligned} &\tfrac{1}{2}\|Z - X\|_F^2 + \lambda_3\|Z_1\|_* + \lambda_4\|vec(Z_2)\|_1 + \chi_+(Z_2) \\ &= \left[\tfrac{1}{2}\|Z_1 - X_1\|_F^2 + \lambda_3\|Z_1\|_*\right] + \left[\tfrac{1}{2}\|Z_2 - X_2\|_F^2 + \lambda_4\|vec(Z_2)\|_1 + \chi_+(Z_2)\right] \end{aligned}$$

Hence, proximal operator can be obtained by solving two subproblems

$$\begin{aligned} &\text{argmin}_{Z_1 \in R^{m\times n}} \quad \tfrac{1}{2}\|Z_1 - X_1\|_F^2 + \lambda_3 \|Z_1\|_* \\ &\text{argmin}_{Z_2 \in R^{m\times n}} \quad \tfrac{1}{2}\|Z_2 - X_2\|_F^2 + \lambda_4 \|vec(Z_2)\|_1 + \chi_+(Z_2) \end{aligned}$$

whose exact solutions are (see examples 2 and 3 in **Section 2.3**)

$$[Z_1^* \; Z_2^*] = [S^*_{t\lambda_3}(X_1) \; S^{1+}_{t\lambda_4}(X_2)].$$

Also note that the derivative of the smooth part is given by

$$\begin{aligned} \tfrac{\partial f_1}{\partial Z_1} &= Z_1 + g * Z_2 - A, \\ \tfrac{\partial f_1}{\partial Z_2} &= \hat{g} * (Z_1 + g * Z_2 - A), \end{aligned}$$

where $(\hat{g})_{i,j} = g_{-i,-j}$. The updating rule can be immediately obtained by substituting these results into equation (4.3).

Given our objective function, we derived the updating rules of FB as follows: Given initial values B^0 and S^0, iteratively update

$$B^{k+1} = S^*_{t\lambda_3}(B^k - t(B^k + g * S^k - A)), \tag{4.4}$$

$$S^{k+1} = S^{1+}_{t\lambda_4}(S^k - t\hat{g} * (B^k + g * S^k - A)), \tag{4.5}$$

where $(\hat{g})_{i,j} = g_{-i,-j}$.

4.5.3 Lipschitz continuous gradient of f_1

Here we show that the smooth part of the objective function has a Lipschitz continuous gradient with Lipschitz constant 2 provided that $\sum_{i,j} g_{i,j} = 1$ and $g_{i,j} \geq 0$.

Denoting $Z = [Z_1\ Z_2]$, $\tilde{Z} = [\tilde{Z}_1\ \tilde{Z}_2]$ and $\Delta Z_i = Z_i - \tilde{Z}_i$, we have

$$\|\nabla f_1(Z) - \nabla f_1(\tilde{Z})\|_F^2 = \|\Delta Z_1 + g * \Delta Z_2\|_F^2 + \|\hat{g} * (\Delta Z_1 + g * \Delta Z_2)\|_F^2$$

Using Young's inequality and the triangle inequality of Frobenius norm, we have

$$\begin{aligned} \|\hat{g} * (\Delta Z_1 + g * \Delta Z_2)\|_F^2 &\leq \|\Delta Z_1 + g * \Delta Z_2\|_F^2 \\ &\leq (\|\Delta Z_1\|_F + \|\Delta Z_2\|_F)^2 \end{aligned}$$

These conclude with

$$\|\nabla f_1(Z) - \nabla f_1(\tilde{Z})\|_F^2 \leq 2(\|\Delta Z_1\|_F + \|\Delta Z_2\|_F)^2 \leq 4\|Z - \tilde{Z}\|_F^2,$$

which determines the Lipschitz constant to be smaller than 2.

4.5.4 Variants

FB can be further accelerated up to quadratic convergence [10]. We implemented the accelerated version of our algorithm as follows: Given initial values $\tilde{B} = B^0$, $\tilde{S} = S^0$ and $t_1 = 1$, iteratively update

$$B^k = S^*_{t\lambda_1}(\tilde{B} - t(\tilde{B} + g * \tilde{S}_k - I)),$$

$$S^k = S^{1+}_{t\lambda_2}(\tilde{S} - t(\hat{g} * (\tilde{B} + g * \tilde{S} - I))).$$

$$t_{k+1} = \frac{1 + \sqrt{1 + t_k^2}}{2},\ \alpha_k = \frac{t_k - 1}{t_{k+1}}$$

$$\tilde{B} = B^k + \alpha_k(B^k - B^{k-1})$$

$$\tilde{S} = S^k + \alpha_k(S^k - S^{k-1})$$

Note that the proposed objective function has a Lipschitz continuous gradient and therefore backtracking is not required when accelerating the algorithm [10].

In the special case when point spread function is not considered (i.e. $g = 1$), the block coordinate descent algorithm (BCD) [12] can also be used to find the minimum of our objective function. BCD alternatively minimizes a few disjoint sets of variables, the matrices B and S in our case, and is guaranteed to converge to the global minimum of a convex problem. Specifically, the BCD is stated as follows: initialize B^0 and S^0, for each B^k, S^k, sequentially update

$$B^{k+1} = S^{*}_{\lambda_1}(A - S^k),$$

$$S^{k+1} = S^{1+}_{\lambda_2}(A - B^{k+1}),$$

In practice, we observed a faster convergence in BCD than in FB (data not shown).

Bibliography

[1] S. Takamori *et al*, "Molecular anatomy of a trafficking organelle," *Cell*, vol. 127, no. 4, pp. 831–846, 2006.

[2] J. Falcón-Pérez, R. Nazarian, C. Sabatti, and E. Dell'Angelica, "Distribution and dynamics of Lamp1-containing endocytic organelles in fibroblasts deficient in BLOC-3," *Journal of Cell Science*, vol. 118, no. 22, pp. 5243–5255, 2005.

[3] N. Chenouard, I. Smal, F. De Chaumont, M. Maška, I. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter *et al.*, "Objective comparison of particle tracking methods," *Natature Methods*, vol. 11, no. 3, pp. 281–289, 2014.

[4] H.-C. Lee and G. Yang, "Computational removal of background fluorescence for biological fluorescence microscopy," *IEEE International Symposium on Biomedical Imaging*, pp. 205–208, 2014.

[5] H.-C. Lee and G. Yang, "An image-based computational method for characterizing whole-cell scale spatiotemporal dynamics of intracellular transport," *IEEE International Symposium on Biomedical Imaging*, pp. 699–702, 2015.

[6] T. Kolda and B. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[7] E. Cand, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–37, 2011.

[8] P. Combettes and V. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling and Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[9] H.-C. Lee, B.-L. Lin, W.-H. Chang, and I.-P. Tu, "Toward automated denoising of single molecular Forster resonance energy transfer data," *Journal of Biomedical Optics*, vol. 17, no. 1, pp. 11 007–1–11 007–10, 2012.

[10] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[11] R. Mazumder, T. Hastie, and R. Tibshirani, "Spectral regularization algorithms for learning large incomplete matrices," *Journal of Machine Learning Research*, vol. 11, pp. 2287–2322, 2010.

[12] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *Journal of Optimization Theory and Applications*, vol. 109, no. 3, pp. 475–494, 2001.

Chapter 5

Shape component analysis: structure-preserving dimension reduction on biological shape spaces

5.1 Introduction

Geometrical shapes are a fundamental property of biological structures. Quantitative shape analysis is required by many biological studies across diverse scales. For example, at the molecular scale, quantitative analysis of shapes of proteins is essential for understanding their functions and interactions ([1]). As another example, at the cellular scale, quantitative analysis of shapes of cells is essential for understanding their morphogenesis and migration ([2]). Recently, high-throughput and systems-level biological studies have started to produce large volume of biological shape data from structural analysis ([1, 3]) or image analysis ([4, 5, 6]). The biological shape data

produced often have high dimensionality, which poses a significant challenge for their analysis and understanding.

Dimension reduction is an essential tool for analyzing and understanding high dimensional data. A wide range of dimension reduction techniques have been developed ([7]). However, for effective dimension reduction of biological shape representations, it is crucial to take into account their specific structures and properties. To give an example, biological shapes are often represented by points on high-dimensional Riemannian spaces ([8, 9]). Differences between distinct shapes are best represented by their Riemannian distances rather than Euclidean distances, which are commonly used in dimension reduction. This can be seen from the example in Fig. 5.1, which shows that Riemannian distances better differentiate between shapes. Indeed, nonlinear Riemannian geometry of shape spaces is proposed as a tool of choice to depict geometric differences between shapes ([10]).

In this study, we described 2D shapes using their landmark representations ([8, 9]). 3D shapes are described using spherical harmonic representation ([11]) in case of the lack of well-defined landmarks. We developed a technique for nonlinear dimension reduction of these two shape representations on their Riemannian spaces. A key feature of our technique is that the nonlinear distances between shapes are preserved under reduced dimensionality. We demonstrated an application of our dimension reduction approach by combining in with nonlinear mean-shift clustering on Riemannian spaces ([12])) for unsupervised clustering of shapes of mitochondria and proteins. Experimental results confirmed that the proposed dimension reduction technique, when combined with mean-shift clustering, provided generally equivalent clustering performance but at reduced processing time. The proposed dimension reduction approach is general and can also be combined with other shape analysis techniques.

The paper is structured as follows. We first outline the theory of shape space

and then describe our dimension reduction technique and its integration with unsupervised mean-shift clustering on shape spaces. The overall shape analysis work flow is summarized in Fig. 5.2. We present experimental results on a variety of shape datasets, first on a 2D generic shape dataset, then on a 2D mitochondrial shape dataset and a 3D protein shape dataset.

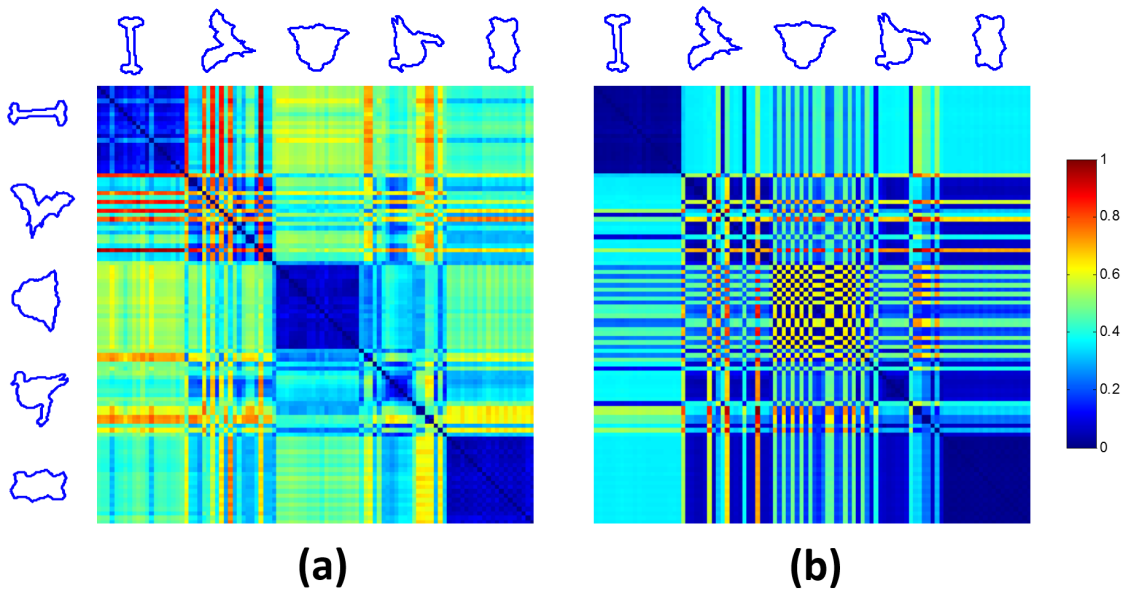

Figure 5.1: **Pairwise distances between different shape clusters.** Five groups of shapes were taken from the MPEG-7 dataset ([13]). An example of each shape cluster is shown on the top and left of the distance map. (a) Riemannian distance of shape manifold; (b) Euclidean distance with elliptic Fourier descriptors ([14]). Color codes represent the normalized value of each distance as shown in the color bar. Note that the difference in performance is not resulted from which the descriptor in use since the Euclidean distance of the elliptic Fourier descriptors is equivalent to the Euclidean distance between two landmark shapes after matching their orientations.

5.2 Methods

5.2.1 Representation of 2D biological shapes

In this study we describe 2D biological shapes using their landmark representation ([9]). It should be pointed out that landmark representation can also be used to describe higher dimensional shapes as long as landmark points are specified.

A N dimensional shape is usually represented by a sequence of its D control points. In the case of a planar object, its 2D shape can be represented by landmarks along its contour. This ordered sequence of control points is referred to as a *configuration*. A configuration can be mathematically represented by a N-by-D real matrix $\mathbf{Y}$, where $\mathbf{Y}_{k,j}$ records k-th coordinate of the j-th control point ([9]). A *pre-shape* $\mathbf{Z}$ is the configuration that is independent of translation and scaling. Typically it is determined by setting $\mathbf{Y}$ to be zero-mean in centroid and unit length in size. The space of pre-shape can be defined as

$$\{\mathbf{Z} \in \mathbb{R}^{N \times D} : \sum_{j=1}^{D} \mathbf{Z}_{k,j} = 0 \ \forall k, \ \|\mathbf{Z}\|_F^2 = 1\}$$

where the Frobenius norm $\|(\cdot)\|_F$ is the square root of the sum of the absolute squares of elements in a matrix. This space is a high-dimensional hyper-sphere in the Euclidean space. Note that the dimension of a configuration ($N \cdot D$-2) are different from the dimension of the physical world (N).

5.2.2 Riemannian geometry of the shape space

A shape is an invariance of a given configuration under translation, rotation, and scaling. Within this context, a N-dimensional shape space Σ_N^D is defined as the collection of equivalent configurations under translation, scaling, and rotation. Accordingly, the

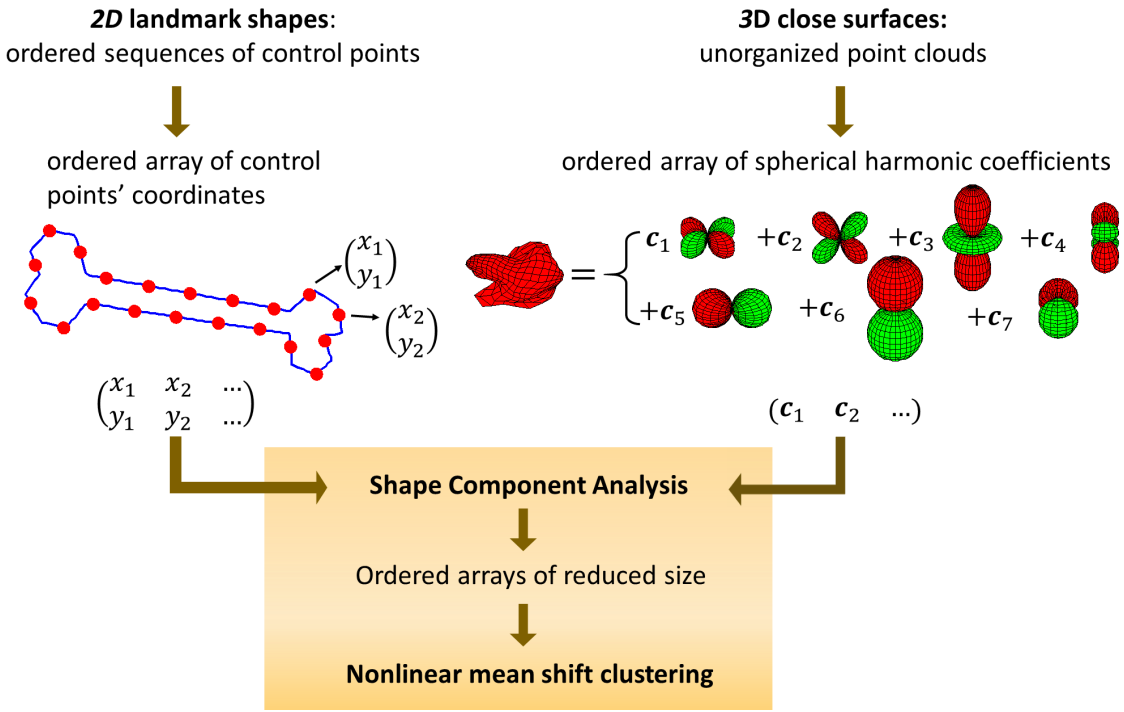

Figure 5.2: **Overall work flow of the proposed dimension reduction method, with application in mean-shift clustering.** Both 2D and 3D shapes are represented as ordered arrays of coefficients to be analyzed by Shape component analysis and nonlinear mean-shift clustering.

Riemannian geometry of the N-dimensional shape space can be explicitly formulated. Here we list several related properties that will be used in the rest of the paper. For ease of implementation, we represent the shape space as a submanifold embedded in $\mathbb{R}^{N\times D}$ with a lifted tangent space.

To define the Riemannian geometry of the shape space, we need to specify: 1) a mathematically defined set containing all possible shapes and 2) the pairwise distance between two shapes. In the theory of landmark representation, a shape is defined as the equivalent class of a pre-shape $\mathbf{Z}$ up to rotations ([9]). Mathematically, this equivalent class can be represented by

$$[\mathbf{Z}] = \{\mathbf{OZ} : \mathbf{O} \in SO(N)\}, \tag{5.1}$$

where $SO(N)$ is the group of N-dimension rotation matrices. The N dimensional shape space Σ_N^D is a Riemannian manifold containing all possible shapes $[\mathbf{Z}]$ and equipped with a distance $\rho(\cdot,\cdot)$. This Riemannian distance is given by

$$\rho([\mathbf{Z}_0],[\mathbf{Z}_1]) = \cos^{-1}(\max_{\mathbf{O}\in SO(N)} \langle \mathbf{O}\mathbf{Z}_0,\ \mathbf{Z}_1\rangle). \tag{5.2}$$

Note that this distance is equivalent to the l_2 distance ρ_F after two shapes are rotationally aligned, that is,

$$\rho_F([\mathbf{Z}_0],[\mathbf{Z}_1]) = \min_{O\in SO(N)} \|O\mathbf{Z}_0 - \mathbf{Z}_1\|_F. \tag{5.3}$$

From the computational perspective, the structure of the tangent space is crucial for developing efficient algorithms. Here we introduce three concepts: the tangent space of a shape space, the Exponential map and its inverse Log map. Roughly speaking, the tangent space contains information of a function's gradient on a Riemannian manifold, while the Exponential and Log maps are used to exploit the direction of a gradient for searching extrema of the function. The lifted tangent space at $[\mathbf{Z}]$ is represented by

$$H_{\mathbf{Z}}(\Sigma_N^D) = \{\mathbf{X}\in\mathbb{R}^{N\times D} | \langle \mathbf{Z},\ \mathbf{X}\rangle = 0, \mathbf{X}\mathbf{Z}^T = \mathbf{Z}\mathbf{X}^T\}. \tag{5.4}$$

Given an equivalent class $[\mathbf{Z}]$ and a tangent vector $\mathbf{v}\in H_{\mathbf{Z}}(\Sigma_N^D)$, the exponential map on the shape space is given by

$$\mathbf{Exp}_{[\mathbf{Z}]}(\mathbf{v}) = [\mathbf{Z}\cos(\|\mathbf{v}\|_F) + \frac{\mathbf{v}}{\|\mathbf{v}\|_F}\sin(\|\mathbf{v}\|_F)]. \tag{5.5}$$

Conversely, when two pre-shapes $\mathbf{Z}_0$ and $\mathbf{Z}_1$ and a rotation $\mathbf{R}\in SO(N)$ satisfy $\mathbf{R}\mathbf{Z}_1\mathbf{Z}_0^T$

being symmetric and positive definite, we obtain the (horizontally lifted) inverse exponential map, which is

$$\mathbf{Log}_{\mathbf{Z}_0}(\mathbf{Z}_1) = \frac{s_0}{\sin(s_0)}(\mathbf{R}^T\mathbf{Z}_1 - \mathbf{Z}_0\cos(s_0)), \tag{5.6}$$

where $s_0 = \rho(\mathbf{Z}_0, \mathbf{Z}_1)$.

Lastly, we note that the Riemannian distance $\rho(\cdot,\cdot)$ and the Log map can be computed using polar decomposition ([15, 10]) as follows:

1. Given two preshapes $\mathbf{Z}_0$ and $\mathbf{Z}_1$, compute the singular value decomposition of their covariance matrix

   $$\mathbf{Z}_0\mathbf{Z}_1^T = \mathbf{USV}^T.$$

   Here, $\mathbf{S}$ is a diagonal matrix formed by singular values and U and V are two orthonormal matrices.

2. Compute the distance $\rho(\mathbf{Z}_0, \mathbf{Z}_1)$ as the arc-cosine of sum of absolute value of singular values, i.e., $s = \cos^{-1}\sum_i(|\mathbf{S}_{ii}|)$.

3. Compute the Log map as

   $$\mathbf{Log}_{\mathbf{Z}_0}(\mathbf{Z}_1) = \frac{s}{\sin(s)}(\mathbf{RZ}_1 - \mathbf{Z}_0\cos(s))$$

   where $\mathbf{R} = \det(\mathbf{VU}^T)\mathbf{VU}^T$.

For 2D shapes, explicit formulas for computing distance and Log map have been derived by representing a $2 \times N$ matrix as a $1 \times N$ complex vector ([9]). This allows a fast implementation without calculating the singular value decomposition.

5.2.3 Dimension reduction on the shape space

In this section we define a mapping that projects a high-dimension shape space into another low-dimension shape space while the pairwise distance is preserved when restricted in the subset defined by Eq. 5.11. This allows us to adapt, for example, the mean-shift algorithm on the shape space with substantially lower dimensionality.

Low dimensional shape space and projection error

The essence of principal component analysis is to find the low-dimensional subspace that minimizes projection error in the original Euclidean space. Analogous to this concept, we first define embedded low-dimensional shape spaces and their corresponding projection errors. Given a D-by-r matrix $\mathbf{R}$ satisfying $\mathbf{R}^T\mathbf{R} = \mathbf{I}_D$, we define the embedded r-dimensional shape space induced by $\mathbf{R}$ as

$$\{[\mathbf{M}\mathbf{R}^T] : \mathbf{M} \in \mathbb{R}^{N\times r},\ \|\mathbf{M}\|_F = 1\} \tag{5.7}$$

which is the image of a continuous function

$$f_R : \Sigma_N^r \to \Sigma_N^D,\ \ f_R([\mathbf{M}]) = [\mathbf{M}\mathbf{R}^T]. \tag{5.8}$$

For every shape $[\mathbf{Z}] \in \Sigma_N^D$, we propose to use the cosine of the Riemannian distance as the similarity measurement to assess the performance of representing $[\mathbf{Z}]$ in the embedded shape space. That is, we define the similarity measurement as

$$\begin{aligned} E([\mathbf{Z}],\,[\mathbf{R}]) &\equiv \max_{\mathbf{M}\in R^{n\times r},\,\|\mathbf{M}\|_F=1} \cos(\rho([\mathbf{Z}],[\mathbf{M}\mathbf{R}^T])) \\ &= \max_{\mathbf{M}\in\mathbf{R}^{n\times r},\,\|\mathbf{M}\|_F=1,\,O\in SO(N)} \langle \mathbf{O}\mathbf{Z},\,\mathbf{M}\mathbf{R}^T\rangle \\ &= \|\mathbf{Z}\mathbf{R}\|_F. \end{aligned} \tag{5.9}$$

This similarity measurement is continuous in $[\mathbf{Z}]$ and invariant under left multiplication of $SO(N)$ and hence its well-definiteness is guaranteed by universal property of the quotient ([16]). Note that this similarity measure also reflects the extent of information of a data point that is preserved under the projection. And the set of zero similarity scores is an analogy to the set that is perpendicular to a Euclidean subspace space. Thus, an embedded submanifold's accuracy to represent a given data set can be easily checked by the similarity scores of each data point.

Structure-preserving dimension reduction

Next we describe the desired mapping. To provide a strong link to the implementation, we consider the shape space Σ_N^D and its tangent space embedded in $\mathbb{R}^{N\times D}$ as described in section 2.2. The following statement is the cornerstone of our proposed method.

Theorem 12. *Given a D-by-r matrix* $\mathbf{R}$ *satisfying* $\mathbf{R}^T\mathbf{R} = \mathbf{I}_D$*, the mapping*

$$\begin{aligned} T_r: \quad & \Sigma_N^D \bigcap \{E(\cdot,\, \mathbf{R}) > 0\} \mapsto \Sigma_N^r \\ & T_r([\mathbf{Z}]) = [\mathbf{ZR}/\|\mathbf{ZR}\|_F] \end{aligned} \tag{5.10}$$

is well-defined. Furthermore, distance is preserved under T_r *in the set*

$$\{[\mathbf{MR}^T] : \mathbf{M} \in \mathbb{R}^{N\times r},\ \|\mathbf{M}\|_F = 1\}. \tag{5.11}$$

Proof. The well-definiteness is followed by verifying that the mapping is continuous and satisfies the universal property of the quotient. The preserving of pairwise distance is ensured since the inner product is preserved in this subset, i.e.,

$$\langle \mathbf{M}_1\mathbf{R}^T,\, \mathbf{M}_2\mathbf{R}^T\rangle = tr(\mathbf{M}_1\mathbf{R}^T\mathbf{R}\mathbf{M}_2^T) = \langle \mathbf{M}_1,\, \mathbf{M}_2\rangle. \tag{5.12}$$

□

Next we address the question of determining the optimal basis matrix from a given data set.

Determining the optimal basis matrix

Given pre-shapes $\mathbf{X}_i$, we propose to determine the optimal basis matrix $\mathbf{R}$ by solving the following optimization problem:

$$\max_{\mathbf{R}^T\mathbf{R}=\mathbf{I}_k} \sum_i E(\mathbf{X}_i, \mathbf{R})^2 = \max_{\mathbf{R}^T\mathbf{R}=\mathbf{I}_k} \sum_i \|\mathbf{X}_i\mathbf{R}\|_F^2. \tag{5.13}$$

Note that the optimal solution is the first k eigenvectors of the covariance matrix $\sum_i \mathbf{X}_i^T\mathbf{X}_i$. The solution to this optimization formulation is similar to the one of two-dimensional principal component analysis ([17]), though we derived this optimization problem from a different perspective. We henceforth refer to this method as shape component analysis (SCA). The overall procedure is summarized in Algorithm 1.

Algorithm 1 shape component analysis

1: Given pre-shapes $\mathbf{z}_j$, $j = 1, \ldots, n$
2: specify the reduced dimension r
3: $\mathbf{C} = \sum_i \mathbf{z}_i^T\mathbf{z}_i$
4: Compute the first r eigenvectors of $\mathbf{C}$
5: List the first r eigenvectors in columns of a matrix $\mathbf{R}$
6: $\tilde{\mathbf{z}}_j \leftarrow \mathbf{z}_j\mathbf{R}/\|\mathbf{z}_j\mathbf{R}\|_F$
7: Output $\tilde{\mathbf{z}}_j$ for mean-shift clustering

5.2.4 Spherical harmonic representation of 3-dimensional biological shapes

To use landmark representation to describe shapes, ordered sequences of control points are required to compare different shapes. However, 3-dimensional surfaces are often recorded as unstructured point clouds and hence may not have well-defined landmark points. To solve this problem, we adapted spherical harmonics representation (SHR) to transform point clouds of 3-dimensional surfaces into ordered sequence of coefficients in the frequency domain. This approach allows a systematic treatment of SHR similar to the landmark representation.

Spherical harmonics $\{Y_m^l(\theta,\phi)\}$ is an orthonormal system of complex functions for decomposing square-integrable functions into a series of coefficients indexed by integers l and m. Given a closed surface $\mathbf{x}(\theta,\phi) = (x(\theta,\phi), y(\theta,\phi), z(\theta,\phi))^T$ parameterized in polar coordinate, $\mathbf{x}(\theta,\phi)$ can be represented as the linear combination of spherical harmonics

$$\mathbf{x}(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \mathbf{c}_m^l Y_m^l(\theta,\phi), \tag{5.14}$$

where $\mathbf{c}_m^l = (c_{m,l}^x,\ c_{m,l}^y,\ c_{m,l}^z)^T$. Each coefficient is calculated by the integration with a specific spherical harmonics, for example,

$$c_{m,l}^x = \int_{S^2} x(\theta,\ \phi) Y_m^{l*}(\theta,\ \phi) d\Omega. \tag{5.15}$$

In practice, the infinite series is truncated by limiting l to $0 \le l \le L$. Detailed explanation of SHR can be found in [11], [18] and references therein.

Next we would like to point out that the landmark representation and spherical harmonics representation share the same Riemannian geometry. As mentioned in Section 2.2, a Riemannian geometry is specified by the set containing all elements and

the pairwise distance between any two elements. We shall define the SHR "shapes" and show that they 1) form a structurally equivalent set and 2) have the pairwise distance as landmark shapes do. As in the landmark representation, a truncated series of spherical harmonic coefficients can be arranged in an ordered array of coefficients. Such an ordered array is represented as a 3-by-$2L^2$ matrix whose elements record the real and imaginary parts of spherical harmonics coefficients. To define the SHR "pre-shape", we propose to

1. Set the coefficient $\mathbf{c}_0^0$ to be zero to make SHR invariant of translation. This operation will set a closed surface to be centered at the origin since

$$\mathbf{c}_0^0 = (\int_{S^2} x(\theta,\phi)d\Omega,\ \int_{S^2} y(\theta,\phi)d\Omega,\ \int_{S^2} z(\theta,\phi)d\Omega)^T$$

   is the centroid of the surface.

2. Normalize $\tilde{\mathbf{C}}$ to unit norm to make SHR invariant of scaling.

We denoted this calibrated matrix as $\tilde{\mathbf{C}}$ to represent SHR pre-shape. Clearly, SHR pre-shapes reside on a high-dimensional hyper-sphere as the landmark pre-shapes. Furthermore, The SHR shape can be defined as the equivalent set of a SHR pre-shape up to rotations, i.e.,

$$[\tilde{\mathbf{C}}] = \{\mathbf{O}\tilde{\mathbf{C}},\ \mathbf{O} \in SO(3)\}. \tag{5.16}$$

Note that this set is also the equivalent set of a closed surface since rotating the surface is equivalent to applying a rotation matrix to its coefficients array. This can be seen from

$$\mathbf{O}\mathbf{x}(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left(\mathbf{O}\mathbf{c}_m^l\right)Y_m^l(\theta,\phi). \tag{5.17}$$

So far we have shown that the space of SHR shape can be represented as the collection of equivalent sets in hyper-sphere.

The pairwise distance of SHR shapes can be defined as the metric between landmark shapes. The key idea behind the Riemannian distance is to capture the minimum L_2 distance between two rotationally aligned objects. For example, given two closed surfaces, we would compute $\min_{\mathbf{O}\in SO(3)} \int_{S^2} \|\mathbf{O}\mathbf{x_1} - \mathbf{x_2}\|_F d\Omega$ for comparing how different these two surfaces are. Because of the orthnormality of the spherical harmonics, this quantity can be written in terms of SHR coefficient matrices, i.e.,

$$\min_{\mathbf{O}\in SO(3)} \int_{S^2} \|\mathbf{O}\mathbf{x_1} - \mathbf{x_2}\|_F d\Omega = \min_{\mathbf{O}\in SO(3)} \|\mathbf{O}\mathbf{C_1} - \mathbf{C_2}\|_F. \tag{5.18}$$

Note that this is the same distance used in landmark representation in equation (5.3). The direct implication here is that the SHR shapes and landmark shapes share the same pairwise distance.

Overall, we have shown that the SHR shape and landmark shapes share the same defining set and pairwise distance and hence the same Riemannian geometry. Since our dimensional reduction technique in the previous section and clustering algorithm presented in next section are designed according to the Riemannian geometry of shape spaces, SHR coefficient arrays can be used as inputs to our algorithm once they are transformed properly. The overall procedure of transforming SHR coefficient arrays to their shape representations is summarized in Algorithm 2.

Algorithm 2 transforming SHR into SHR Shape

1: Given SHR $\mathbf{C}$
2: $\mathbf{C}_0^0 \leftarrow 0$
3: $\tilde{\mathbf{C}} \leftarrow [\text{real}(\mathbf{C})\ \text{imag}(\mathbf{C})]$
4: $\tilde{\mathbf{C}} \leftarrow \tilde{\mathbf{C}}/\|\tilde{\mathbf{C}}\|_F$
5: Output $\tilde{\mathbf{C}}$ as a preshape for other analysis

5.2.5 Mean-shift clustering on a Riemannian manifold

The mean-shift clustering algorithm is a peak-finding algorithm that searches for representatives supported by the majority of data ([19]). A mean-shift clustering algorithm defines modes as local maxima of the underlying probability density distribution estimated by a kernel density estimation and iteratively searches for local maximums using a gradient-based update rule. In [12], mean-shift clustering is generalized on a Riemannian manifold by replacing the L_2 norm in the kernel estimation with the Riemannian distance. The update rules are described as follows: given $\mathbf{x}_0$ as a data point, iteratively update

$$\begin{aligned} \mathbf{m}_h(\mathbf{x}_k) &= \frac{\sum_i g(d^2(\mathbf{x}_k,\mathbf{z}_i)/h)\mathbf{Log}_{\mathbf{z}_i}(\mathbf{x}_k)}{\sum_i g(d^2(\mathbf{x}_k,\mathbf{z}_i)/h)} \\ \mathbf{x}_{k+1} &= \mathbf{Exp}_{\mathbf{x}_k}\left(\mathbf{m}_h\left(\mathbf{x}_k\right)\right) \end{aligned}, \tag{5.19}$$

where $\mathbf{z}_i$ are pre-shapes given, h is the bandwidth of the kernel function $g(x) = e^{-x/2}$. Convergence of this algorithm is guaranteed. We implemented mean-shift clustering on the shape space following the procedure shown in Algorithm 3. We note that with the proposed SCA, mean-shift clustering can be performed on the shape space of substantially lower dimensionality without any changes in procedure.

5.3 Results

We tested our dimension reduction algorithm on three different shape datasets. All computation was performed on a desktop workstation (2×Intel Xeon E5503 2.00GHz and 8G RAM).

Algorithm 3 mean-shift clustering on the shape space

Given pre-shapes $\mathbf{z}_i,\ i = 1,\ \ldots,\ n$
for $i = 1,\ \ldots,\ n$ **do**
    $\mathbf{x} \leftarrow \mathbf{z}_i$
    while $\|\mathbf{m}_h(\mathbf{x})\| < \epsilon$ **do**
        for $j = 1,\ \ldots,\ n$ **do**
            $[U_j,\ S_j,\ V_j] = svd(\mathbf{z}_j\mathbf{x}^T)$
            $R_j = det(V_jU_j^T)V_jU_j^T$
            $\rho_j = \cos^{-1}(tr(S_j))$
            $T_j = \frac{s_0}{\sin(s_0)}(R\mathbf{z}_j - \mathbf{x}\cos(s_0))$
        end for
        $\mathbf{m}_h(\mathbf{x}) \leftarrow \frac{\sum_i g(d^2(\mathbf{x},\mathbf{z}_i)/h)\mathbf{T}_i}{\sum_i g(d^2(\mathbf{x},\mathbf{z}_i)/h)}$
        $s \leftarrow \|\mathbf{m}_h(\mathbf{x})\|$
        $\mathbf{x} \leftarrow \frac{\sin(s)}{s}\mathbf{m}_h(\mathbf{x}) + \cos(s)\mathbf{x}$
    end while
    Retain $\mathbf{x}$ as a local mode
end for
Return distinct local modes

5.3.1 Application I: 2D generic shapes

Data preparation

We first tested our dimension reduction algorithm on a generic 2D shape dataset. Specifically, we selected 5 categories of shapes from the MPEG-7 dataset ([13]), with 20 examples in each category. The first example of each category is shown in Fig. 5.3. The boundary of each example was fitted by a cubic spline. Then a total of 200 semi-landmark points was sampled from the fitted spline. In this way, each shape is represented by a 2×200 matrix.

Results: geometric approximation

We tested whether our dimension reduction technique can reliably preserve distances between different shapes under reduced dimensions. We first calculated the pairwise Riemmanian distances between different shapes and used them as the ground truth. We then calculated Riemannian distance between shapes under different levels of dimension reduction (Fig. 5.3) and the normalized mean square error. For comparison, we also calculated the pairwise Euclidean distance on the tangent space of the mean shape. This allowed us to compare our dimension reduction technique to the principal geodesic analysis (PGA) technique ([20]), which is based on Euclidean distance on the tangent space.

The level of distortion under reduced dimension was calculated as the normalized error of the distance matrix M compared to the ground truth distance matrix $\mathbf{M}_{\text{truth}}$, i.e.,

$$\text{Distortion}(\mathbf{M}, \mathbf{M}_{\text{truth}}) = \frac{\|\mathbf{M} - \mathbf{M}_{\text{truth}}\|_F}{\|\mathbf{M}_{\text{truth}}\|_F},$$

where $\|\cdot\|_F$ is the Frobenius norm. In addition, we reported the reconstruction accuracy of the proposed algorithm, which is calculated by the normalized root mean

square error (nRMSE) of each reconstructed shape compared to the original shape. The results were summarized in Table 5.1.

It can be seen that there was a non-vanishing distortion in pairwise distances when PGA and the Euclidean distance were used. In contrast, SCA consistently reduced the approximation error when increasing dimension of projection submanifold. In terms of reconstruction errors, SCA is comparable with PGA under larger reduced dimensions.

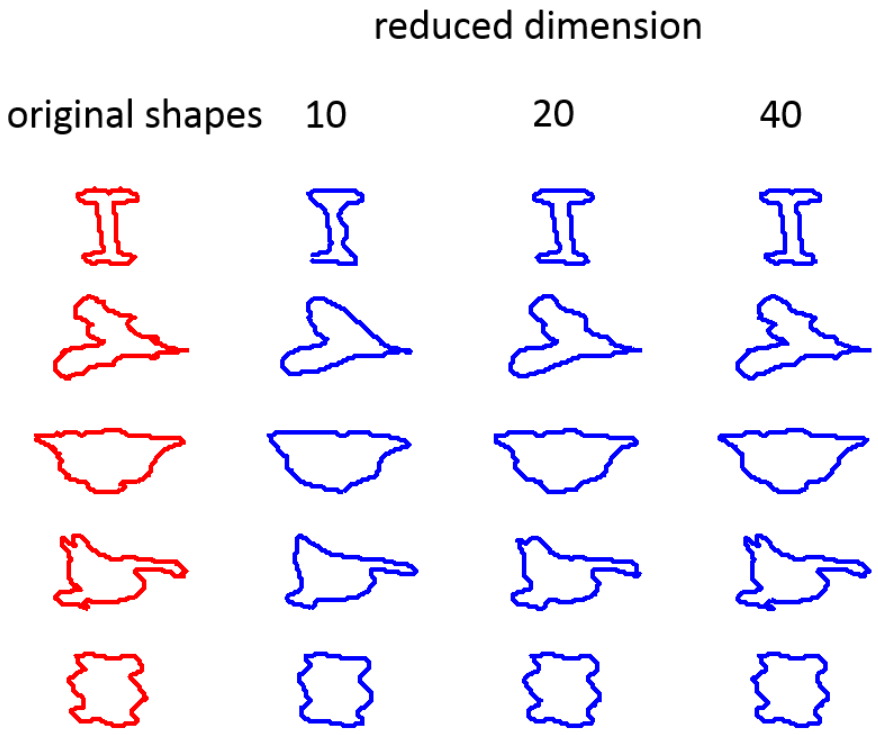

Figure 5.3: **Reconstructed shapes with various reduced dimension.** Dimension reduction is done with SCA. Examples from different shape classes are shown in rows. Reconstructed shapes with different reduced dimension are shown in columns.

	PGA			SCA		
dimension	10	20	40	10	20	40
nRMSE	5.9	3.0	1.2	11.1	5.3	2.3
distortion (%)	2.3	2.1	2.2	5.3	1.5	0.4

Table 5.1: Reconstruction errors of 2D generic shapes by PGA and SCA.

Results: shape clustering

We further examined the performance of integrated dimension reduction and mean-shift clustering. We compared clustering results of five different methods:

1. Riemannian mean shift clustering (RMS).

2. RMS with shape componenet analysis (RMSSCA).

3. Mean shift clustering on the tangent space, using Euclidean distance (tMS). By testing this method, we examined the difference between using Riemannian and Euclidean distance in shape clustering.

4. Mean shift clustering with elliptic Fourier descriptors (MSFD) ([14]) as inputs. This method was chosen to show the performance of using rotation-invariant descriptors in shape clustering.

5. Mean shift clustering with features obtained from Laplacian eigenmap ([21]) (MSLAP). This method was chosen to show the performance of a common strategy of "flattening the nonlinear manifold." We implemented Laplacian eigenmap with 5-nearest neighbors to project the Riemanniam manifold into a 10 dimensional Euclidean space, where the mean-shift clustering was performed to find clusters.

	2D shapes				
	RMS	RMSSCA	tMS	MSFD	MSLAP
# clusters	7	7	7	8	8
Purity	0.8	0.8	0.68	0.64	0.54
NMI	0.73	0.73	0.62	0.53	0.48
AR	0.64	0.64	0.47	0.37	0.26
RunTime(s)	2.1	0.6	0.3	0.1	0.06

Table 5.2: Performance of different algorithms on 2D shape clustering. 20 examples from each 5 shape clusters were used for testing in the 2D case.

For all algorithms, clusters were first identified and then used to classify each data point using one-nearest-neighbor classification. The final clustering results were evaluated by 3 performance metrics ([22]), including

1. Purity = $\frac{1}{N}\sum_k \max_j n_{j,k}$.

2. Normalized information

$$\text{NMI} = -2\frac{\sum_j \sum_k \frac{n_{j,k}}{N}\log(\frac{n_{j,k}}{a_j b_k/N})}{(\sum_j \frac{a_j}{N}\log(a_i/N) + \sum_k \frac{b_k}{N}\log(b_k/N))}.$$

3. Adjusted Rand index (AR)

$$\text{AR} = \frac{\sum_{j,k}\binom{n_{j,k}}{2} - \sum_j \binom{a_j}{2}\sum_k \binom{b_k}{2}}{\frac{1}{2}(\sum_j \binom{a_j}{2} - \sum_k \binom{b_k}{2}) - (\sum_j \binom{a_j}{2}\sum_k \binom{b_k}{2})/\binom{n}{2}}.$$

Here, $a_j = \#\{g_j\}$, $b_k = \#\{c_k\}$, $n_{k,j} = \#\{c_k \cap g_j\}$, and c_j and g_j are the j-th set of members form clustering and ground truth labeling, respectively.

In some cases, these performance metrics could be biased by the number of clusters. To avoid this problem, we adjusted parameters so that each algorithm generated 7-8 clusters. Performance metrics are summarized in the left half of Table 5.2. The results showed that algorithms using Riemannian distance such as RMS and RMSSCA performed better than algorithms using Euclidean distance such as tMS and MSFD. MSLAP performed the worst among all the algorithms. This confirmed the merit of using nonlinear similarity in clustering. Although RMSSCA provided similar clustering performance as RMS, its computational time was significantly reduced, by a factor of 3~6.

5.3.2 Application II: 2D mitochondrial shapes

We further tested our SCA algorithm on a mitochondrial shape dataset. Mitochondria within segmental nerves of dissected Drosophila 3^{rd} instar larvae were visualized by fluorescence live imaging. Images were collected on a Nikon Ti-E inverted microscope at 5 frames per second, under a NA of 1.41 and a 100× magnification. Because of the constraint imposed by the axon geometry, shapes of axonal mitochondria can be represented in 2D rather than in 3D. An adaptive active-mask algorithm was used to segment mitochondria ([23]). Individual mitochondria were tracked as in [24]. The boundary of each segmented mitochondrion was fitted by a cubic spline. Then a configuration of 200 points was sampled from the fitted spline. A total of 4000 configurations were collected from 800 mitochondria with 5 repeats sampled at different time points. The shape classes identified by the proposed algorithm are shown in Fig. 5.4. The original mean-shift algorithm detected 8 diverse shape clusters. Mean shift with dimension reduction provided similar clustering results, but the run time was significantly reduced. While the original mean-shift algorithm took 20756 seconds to finish the computation, mean-shift with dimension reduction took only 1390 seconds in total (480 seconds in dimensional reduction and 910 seconds in mean-shift clustering), a reduction of the computing time by a factor of ~15.

5.3.3 Application III: 3D protein surfaces

Experiment design

We also tested our algorithm on 3-dimensional protein shape dataset. We selected three series of protein structures from Data base of Macromolecular Movements (http://www.molmovdb.org/), including insulin, S100-Ca^{2+} sensor and small G-protein Arf6. These three proteins were selected for their distinct geometry (See Fig. 5.5).

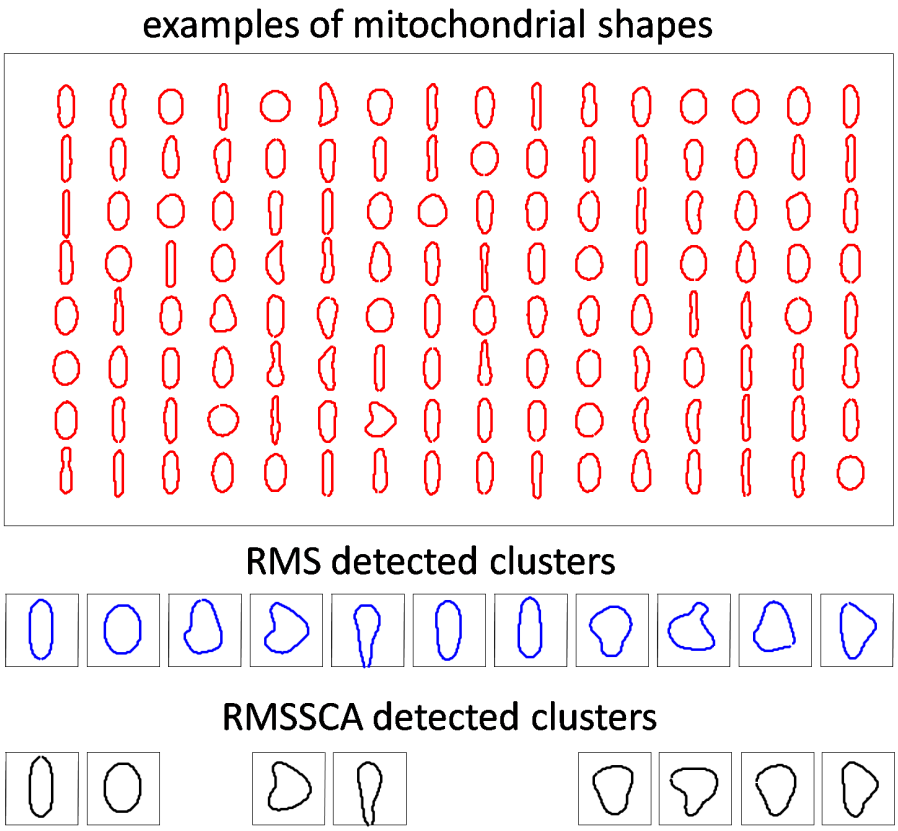

Figure 5.4: **Clustering of mitochondrial shapes**. Red: examples of mitochondrial shapes (selected randomly from 4000 shapes). Black and blue: shape clusters identified by RMS and adaptive RMSSCA, respectively.

Insulin tends to have an ellipsoidal shape, Arf6 is more spherical overall, whereas Ca^{2+} sensor is of Y-shaped with a blunt end. Snapshots of the motion sequences of these three proteins were used as examples of each protein class. Overall, 20 examples were collected used from each protein class.

Data preprocessing

Surfaces of the selected proteins were generated using Gaussian kernel functions as in ([25, 26]). A multi-level summation of Gaussian kernel functions was employed to generate implicit models from atomic resolution data of the selected proteins. An unique

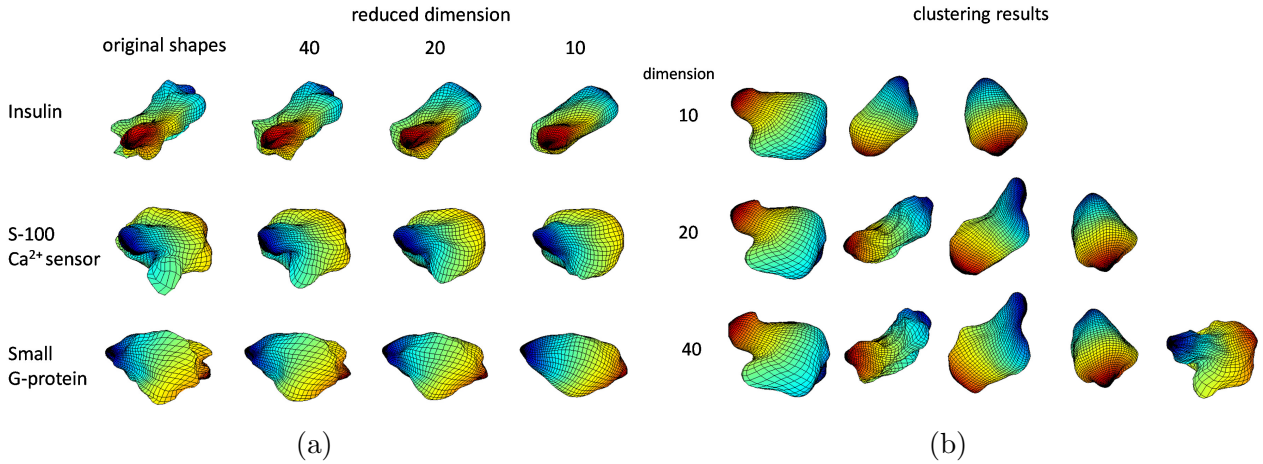

Figure 5.5: **Dimension reduction and clustering of protein shapes.** (a) Reconstructed protein surfaces with different reduced dimensions. (b) Clusters of molecule surfaces determined by RMSSCA with different reduced dimensions.

	SCA		
reduced dimension	40	80	160
Surface area distortion	2.43 ± 0.9	0.48 ± 0.3	0.02 ± 0.02
nRMSE in SH coefficients	5.04 ± 0.68	2.7 ± 0.26	0.54 ± 0.26
nRMSE in coordinates	3.98 ± 0.63	1.59 ± 0.19	0.19 ± 0.09

Table 5.3: Reconstruction errors of 3D protein surfaces. Normalized Errors in surface area, normalized RMSE in spherical coefficients and normalized RMSE in coordinates of control points on the molecule surface are reported. All numbers are reported in percentage.

strength of this method is that it allows local resolution control on protein surfaces. Parameters were chosen manually to ensure a genus-zero surface. After the control points on the molecule surfaces were generated, spherical harmonics parametrization were computed using SPHARM-MAT software ([27]). We computed spherical harmonics up to order 31 to ensure a $0.2\mathring{A}$ accuracy in approximating original molecule surfaces. In this way, each molecular surface was represented by a matrix of dimension 3×2048.

Results: dimension reduction

First we tested the performance of our algorithm in approximating protein surfaces with low dimensionality. SCA was performed on the set of 60 protein molecule surfaces and approximations of 20, 40, 80 and 160 dimensions were generated. To evaluate the level of surface distortion, we calculated three metrics: normalized RMSE in spherical harmonic coefficients and in coordinates of control points, and normalized error of surface area. Results are summarized in Table 5.3. In this case, SCA was able to approximate original surfaces with 99% accuracy using 3×160 real matrices. Compared to the original 3×2048 coefficient arrays, dimension was reduced by a factor of ~12.

	3D shapes				
	RMS	RMSSCA	MSSH	MSSH+PCA	MSLAP
# clusters	5	5	5	4	6
Purity	0.73	0.75	0.70	0.70	0.57
NMI	0.489	0.494	0.39	0.39	0.24
AR	0.426	0.421	0.34	0.34	0.06
RunTime(s)	33.1	5.6	0.08	0.06	0.13

Table 5.4: Performance of different algorithms on 3D shape clustering. 20 examples from each 3 shape clusters were used for testing in the 3D case.

Results: shape clustering

Next we tested performance of the proposed algorithm in clustering molecule surfaces. For comparison, we also implemented clustering methods including:

1. Riemannian mean shift clustering without dimension reduction with (RMSSCA) and without dimension reduction (RMS).

2. Mean shift clustering with 31 rotation invariant spherical harmonic features ([11]) as inputs (MSSH). Principal component analysis was used to select a 6-dimensional subspace that account for 99% variance of data for mean-shift clustering (MSSH + PCA).

3. Mean shift clustering with features obtained from Laplacian eigenmap ([21]) (MSLAP). Laplacian eigenmap was implemented as in the case of 2D shape clustering for examining the performance of a common strategy of "flattening the nonlinear manifold."

As in Section 3.1, metrics such as purity, mutual information and adjusted Rand index were used to evaluate clustering results. The results are summarized in Table 5.4. Overall, Riemannian mean shift clustering provided the best performance. RMSSCA provided similar performance with as RMS, but RMSSCA's computational time was

reduced by a factor of 5 compared to RMS. We note that among these algorithms, only RMS and RMSSCA provide backward projection to the original space and thus their clustering results can be visualized. In Figure 5.5, clustering results via RMSSCA with various reduced dimension are shown. The three basic structures (ellipsoidal, spherical and Y-shape) were detected via RMSSCA with dimensionality of 10, while different levels of details of the molecule surface were captured by RMSSCA with higher dimensions.

5.4 Conclusions

In this study, we proposed a technique for dimension reduction on the Riemannian manifold of 2D and 3D biological shapes. A key advantage of this technique is that it preserves the distances between different shapes over the manifold. We showed that although spherical harmonic representation of 3D shapes differs from the landmark representation of 2D shapes, they share the same Riemannian geometry and thus can be processed through the same dimension reduction technique. We verified the proposed algorithm on datasets of 2D and 3D shapes. In particular, we demonstrated an application of the algorithm by combining it with nonlinear mean-shift clustering for unsupervised classification of biological shapes. The proposed dimension reduction approach is general and provides a tool for analyzing and understanding large sets of high-dimensional shape data. It can be integrated with shape analysis techniques other than mean-shift clustering.

SCA is closely related to principal geodesic analysis which performs principal component analysis on the tangent plan of the mean shape. We have demonstrated two merits of our approach, that is, SCA better preserves pairwise distance and retains

the Riemannian geometry of the reduced space. However, our current approach can only handle genus-zero surfaces and not yet allow adaptive control on local features.

Bibliography

[1] J.-M. Chandonia and S. Brenner, "The impact of structural genomics: expectations and outcomes," *Science*, vol. 311, no. 5759, pp. 347–351, 2006.

[2] K. Keren, Z. Pincus, G. Allen, E. Barnhart, G. Marriott, A. Mogilner, and J. Theriot, "Mechanism of shape determination in motile cells," *Nature*, vol. 453, no. 7194, pp. 475–480, 2008.

[3] E. Oueslati, I. Messaoudi, Z. Lachiri, and N. Ellouze, "A new way to visualize dna's base succession: the caenorhabditis elegans chromosome landscapes," *Medical & biological engineering & computing*, pp. 1–12, 2015.

[4] T. Saito, M. Ohtani, H. Sawai, F. Sano, A. Saka, D. Watanabe, M. Yukawa, Y. Ohya, and S. Morishita, "Scmd: Saccharomyces cerevisiae morphological database," *Nucleic acids research*, vol. 32, no. suppl 1, pp. D319–D322, 2004.

[5] M. D'Ambrosio and R. Vale, "A whole genome RNAi screen of drosophila s2 cell spreading performed using automated computational image analysis," *The Journal of Cell Biology*, vol. 191, no. 3, pp. 471–478, 2010.

[6] E. Sumiya, H. Shimogawa, H. Sasaki, M. Tsutsumi, K. Yoshita, M. Ojika, K. Suenaga, and M. Uesugi, "Cell-morphology profiling of a natural product library

identifies bisebromoamide and miuraenamide a as actin filament stabilizers," *ACS Chemical Biology*, vol. 6, no. 5, pp. 425–431, 2011.

[7] I. Fodor, "A survey of dimension reduction techniques," *Technical Report UCRL-ID-148494*, 2002.

[8] J. Kent, "The complex bingham distribution and shape analysis," *Journal of the Royal Statistical Society Series B*, pp. 285–299, 1994.

[9] I. Dryden and K. Mardia, *Statistical shape analysis*. John Wiley & Sons, 1998.

[10] G. Kendall, D. Barden, T. Carne, and H. Le, *Shape and shape theory*. John Wiley & Sons, 1999.

[11] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, "Rotation invariant spherical harmonic representation of 3D shape descriptors," vol. 6, pp. 156–165, 2003.

[12] R. Subbarao and P. Meer, "Nonlinear mean shift over Riemannian manifolds," *International Journal of Computer Vision*, vol. 84, no. 1, pp. 1–20, 2009.

[13] M. Bober, "MPEG-7 visual shape descriptors," *IEEE Circuits and Systems for Video Technology*, vol. 11, pp. 716–719, 2001.

[14] F. Kuhl and C. Giardina, "Elliptic Fourier features of a closed contour," *Computer Graphics and Image Processing*, vol. 18, no. 3, pp. 236–258, 1982.

[15] G. Golub and C. van Van Loan, *Matrix Computations*. The Johns Hopkins University Press, 1996.

[16] T. Hungerford, *Algebra*. Springer, 1980.

[17] J. Yang, D. Zhang, A. Frangi, and J. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.

[18] R. Morris, R. Najmanovich, A. Kahraman, and J. Thornton, "Real spherical harmonic expansion coefficients as 3D shape descriptors for protein binding pocket and ligand comparisons," *Bioinformatics*, vol. 21, no. 10, pp. 2347–2355, 2005.

[19] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[20] T. Fletcher, C. Lu, S. Pizer, and S. Joshi, "Principal geodesic analysis for the study of nonlinear statistics of shape," *IEEE Trans on Medical Imaging*, vol. 23, no. 8, pp. 995–1005, 2004.

[21] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[22] N. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *The Journal of Machine Learning Research*, vol. 11, pp. 2837–2854, 2010.

[23] K.-C. Chen, Y. Yu, R. Li, H.-C. Lee, G. Yang, and J. Kovacevic, "Adaptive active-mask image segmentation for quantitative characterization of mitochondrial morphology," in *IEEE International Conference on Image Processing*, 2012, pp. 2033–2036.

[24] M. Qiu, H.-C. Lee, and G. Yang, "Nanometer Resolution Tracking and Modeling of Bidirectional Axonal Cargo Transport," in *IEEE International Symposium on Biomedical Imaging*, 2012, pp. 992–995.

[25] Y. Zhang, G. Xu, and C. Bajaj, "Quality meshing of implicit solvation models of biomolecular structures," *Computer Aided Geometric Design*, vol. 23, no. 6, pp. 510–530, 2006.

[26] T. Liao, Y. Zhang, P. Kekenes-Huskey, Y. Cheng, A. Michailova, A. McCulloch, M. Holst, and A. McCammon, "Multi-core CPU or GPU-accelerated multiscale modeling for biomolecular complexes," *Molecular based mathematical biology*, vol. 1, pp. 164–179, 2013.

[27] C. Brechbühler *et al*, "Parametrization of closed surfaces for 3D shape description," *Computer Vision and Image Understanding*, vol. 61, no. 2, pp. 154–170, 1995.

Chapter 6

Conclusions and future work

6.1 Conclusions

In this thesis we characterized, analyzed, and modeled the morphology and active transport of mitochondria. Our conclusions are summarized as follows:

1. **Mitochondrial fusion and fission are not independent**. The proposed model in Chapter 2 shows that independent fusion/fission would lead to oversized mitochondria, which contradicts experimental observation. The proposed fission regulated fusion model resolves this issue and gives a realistic estimation on the relative occurrence of fission/fusion events.

2. **The mitochondrial size and velocity are dependent**. The proposed model in Chapter 3 reveals the subtle dependency between mitochondrial size and velocity. The model suggests that larger organelles tend to move slower given that the microtubules are tightly packed. The proposed model also shows that the size-velocity relation is critically dependent on the spatial distribution of microtubules around the organelle.

We also developed a suite of general tools to computationally characterize and model the organelle dynamics, including

1. **A computational tool for representing the spatiotemporal dynamics of organelle transport at the whole-cell scale**. We presented an image based method to separate signals of stationary organelles and moving organelles. Based on this method, we further define image representations of localization of stationary organelles and flow of moving organelles. Our results show that the lysosomal transport might be regulated at the whole-cell scale.

2. **A dimension reduction technique that can be deployed for large scale analysis of biological morphology**. We proposed a dimension reduction technique that preserves the nonlinear structure of shape manifold and demonstrate its compatibility with nonlinear mean-shift clustering to speed up the unsupervised clustering on biological morphology.

Overall, we hope the proposed approaches would lay the foundation for a computational framework for automated systems-level analysis of organelle dynamics.

6.2 Future work

6.2.1 Using organelles as sensors for probing cellular environment

Probing the changing cellular environment is fundamental to understanding how cells are functioning. In Chapters 2 and 3, we demonstrated that it is possible to estimate parameters of cellular processes by measuring physical properties of organelles. To achieve this, one needs to first mathematically formulate the governing rules on

organelle dynamics, and subsequently uses these mathematical rules and measured organelle's properties to estimate the target parameters. We believe this idea would provide a good direction to achieve a non-invasive estimation on cellular environments.

6.2.2 Analyzing and modeling the logistics of intracellular transport at the whole cellular scale

In Chapter 4, we proposed an image representation on the global structure and flow of organelle transport. However, subsequent analysis is still required to extract its biological implication. We would expect further analysis to answer questions such as:

1. Are there any transport paths with high traffic volume?
2. How many global transport patterns can be observed?

In Chapter 3, we proposed a model describing the individual behavior of organelle transport. A natural extension is to consider a new model that takes account of the intracellular transport at the whole cell scale.

1. How could cells direct materials to multiple destinations?
2. How do cells achieve this goal in an energy-economical way?

We believe these questions are fundamental to advancing our understanding of intracellular transport.