

# Quantifying the Golgi

Gert-Jan Both

Supervised by:  
P. Sens  
C. Storm

Technical university of Eindhoven  
January-November 2018

# Abstract

Title: Quantifying the Golgi

Abstract: The Golgi apparatus is a key component in intracellular trafficking, maturing and directing proteins essential to the cell. Despite years of research, a model coupling Golgi size and function to the cells' transport properties is lacking. In this thesis we develop such a model, describing the Golgi as an active droplet. New experimental data sheds more insight in the spatial organization of the trafficking and I have also devised two new methods relying on image gradients and neural networks to analyze this data and confront it with our model.

# Contents

## ABSTRACT

i

### 1 INTRODUCTION

1.1	The secretory pathway: biology 101 for physicists . . . . .
1.2	This thesis . . . . .

### 2 MODEL FITTING

2.1	The concept . . . . .
2.2	Step 1 - Smoothing and denoising . . . . .
2.3	Step 2 - Derivatives . . . . .
2.4	Step 3 - Segmentation . . . . .
2.5	Step 4 - Fitting . . . . .

### 3 DATA ANALYSIS

3.1	Experimental data . . . . .
3.2	Initial analysis . . . . .
3.3	Analysis of LS-fit . . . . .
3.4	Conclusion . . . . .

### 4 PHYSICS INFORMED NEURAL NETWORKS

4.1	Neural Networks . . . . .
4.2	Physics Informed Neural Networks . . . . .
4.3	Conclusion . . . . .

<b>5</b>	<b>GOLGI AS A PHASE SEPARATED DROPLET</b>
<b>5.1</b>	Phase separation . . . . .
<b>5.2</b>	Effective droplet . . . . .
<b>5.3</b>	Golgi as an active droplet . . . . .

<b>6</b>	<b>GOLGI MODEL</b>
<b>6.1</b>	Two-component model . . . . .
<b>6.2</b>	Conclusion . . . . .

**7 CONCLUSION**

**APPENDIX I: SOME EXTRA STUFF**

**REFERENCES**

# 1

## Introduction

The cell uses thousands of proteins and lipids to function and many of these are produced in the Endoplasmic Reticulum (ER), an organelle found in eukaryotic cells. Exiting the ER, the cargo is transported throughout and outside the cell, completing what is called the *secretory pathway*. Key component in this intracellular trafficking is the Golgi apparatus, an organelle responsible for biochemically maturing proteins and lipids exiting the ER and directing them to the right location. Intense research over the last years has identified key players in the secretory pathway (1, 2) but an integrated model coupling Golgi size and function to the intracellular transport is lacking (3, 4). In this thesis, we seek exactly such a model.

### 1.1 THE SECRETORY PATHWAY: BIOLOGY 101 FOR PHYSICISTS

Proteins produced in the ER exit the organelle at specific places known as ER Exit Sites - ERES. At the ERES, cargo is packaged into a lipid bilayer and this package, known as a vesicle, buds off into the cytoplasm<sup>5</sup>. As ERES are located throughout the cell, the vesicles need to be transported to the Golgi apparatus. In general, intracellular transport has two modes: diffusive and active<sup>6</sup>. In the active mode, molecular motors pull the vesicles across microtubules. Microtubules (MTs) are tubular polymers which form a network throughout the cell (in fact, they are part of the cytoskeleton) and act as the backbone for intracellular transport. They nucleate and organise around objects known as MicroTubular Organisation Centers (MTOCs). The primary MTOC is the centrosome, an organelle located next to the nucleus, but strong evidence exists that the Golgi apparatus acts a nucleation center too<sup>7,8</sup>.

MTs are polarized, having two distinct ((+) and (-)) ends, and use different molecular motors (dynein and kinesin respectively) for transport towards each end<sup>9</sup>. Research has shown that cargo constantly switches between these two transport directions in a random fashion, making active transport in the cell a stochastic process which can be described by a tug-of-war model<sup>10</sup>. It's also possible for the cargo to completely detach from the MT. The cargo will then move through the cytoplasm in a diffusive way, until it reattaches to a microtubule.

The appearance of the Golgi is strongly dependent on the cell type. In plants for example, the Golgi is distributed throughout the cell in separate but fully functional subunits<sup>11</sup> known as stacks, whereas in mammals all these stacks are localized in a single organelle, the *Golgi Ribbon*<sup>7</sup>. Each Golgi stack consists of a number of stacked compartments with a disk-like shape called cisternae. These are membrane enclosed objects with enzymes responsible for biochemically altering proteins. The golgi has distinct entry and exit faces known respectively as the cis and trans face, with the cisternae being labeled analogously. At the cis-face vesicles fuse with the golgi, releasing their cargo into a compartment, while their lipid bilayer becomes part of the golgi membrane. At the trans-face, the cargo is encapsulated again in a lipid bilayer made from the golgi membrane and is transported to its destination. The Golgi thus acts as a sort of post-office of the cell, receiving cargo, repackaging and sending it to the right destination<sup>12</sup>.

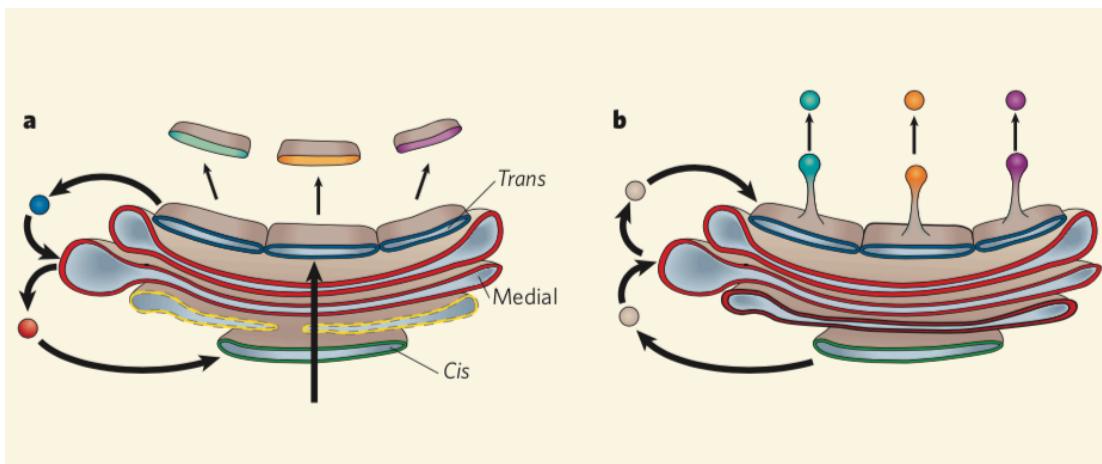


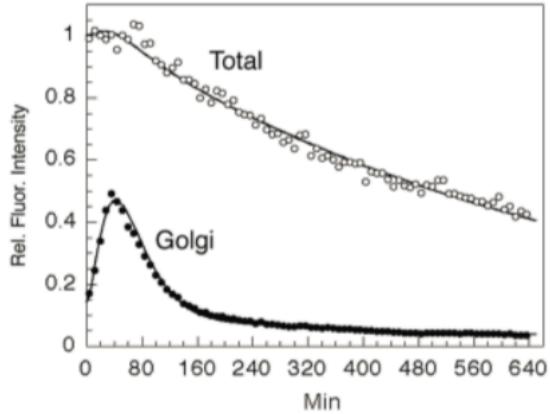
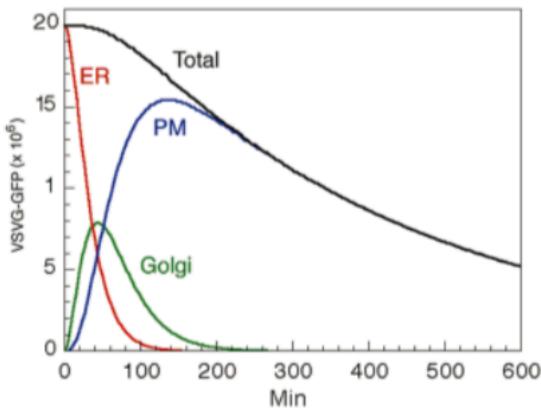
Figure 1.1: Left panel: the cisternal maturation model. Right panel: the vesicle-shuttle model. Image taken from 13

Exactly how this sorting and maturing inside the Golgi happens is debated<sup>14</sup>. The two main competing models are known respectively as the *cisternal maturation* and *vesicle shuttle*. In figure fig. 1.1 we show the structure of the Golgi and and a schematic view of each model.

In the cisternal maturation model (see the left panel of fig. 1.1) , incoming vesicles form a cis-compartment which then matures as whole to a medial and finally trans-compartment. Vesicles are budded off and any excess membrane is trafficked back to the cis face. In the vesicle shuttle model, the cisternae are static entities with a defined task and cargo is moved from one compartment to the next by vesicles. This is shown in the right panel. After budding off from the trans face, the cargo is transported to either some location in the cell or secreted.

### 1.1.1 QUANTITATIVE MODELS OF THE GOLGI

Although the Golgi has been intensively studied by biologists for many years, very few quantitative work has appeared. Our research only turned up a single attempt by Hirschberg et al 15 in which they use a kinetic model to describe the trafficking of a virus (VSVG) from the ER through the Golgi to the plasma membrane. We reprint their main result in figure fig. 1.2.

**B****C**

**Figure 1.2:** Left panel: First order rate model fitted to experimental data by Hirschberg et al. Right panel: Inferred concentration in ER, Golgi and PM using the fitted parameters from the left panel and their model. Image taken from 15

Hirschberg et al model the secretory pathway by dividing it into three reservoirs: each one representing respectively the ER, Golgi and the PM. The compartments are connected by a first-order rate equation, i.e.  $d\varphi_{golgi}/dt = k_{ER \rightarrow Golgi} \varphi_{ER}$  and they show (see figure fig. 1.2) that such an equation is sufficient to model both total concentration in the cell as well as the concentration in the Golgi. Although this model fits the experimental data well, it's a

phenomenological model, reducing many processes to a single rate parameter  $k$  and neglecting any spatial dependence. It is thus unable to provide any information on how the intracellular transport of the vesicles couples to the function and size of the Golgi.

## 1.2 THIS THESIS

In this thesis we seek to build a model which couples intracellular transport to the Golgi function and size. We hypothesize that we can describe the Golgi as an *active, phase separated droplet*. We also confront our model with experimental data, obtained by the group of Frank Perez at Institut Curie. This group developed a new technique called RUSH<sup>16</sup> which allows precise timing of the release of proteins from the ER and study the intracellular transport using fluorescence microscopy. In the next sections, we justify the description of the Golgi as an active phase-separated droplet and how we intend to perform the data analysis of the experimental data.

### 1.2.1 GOLGI AS AN ACTIVE PHASE-SEPARATED DROPLET

Many biological processes require a relatively high concentration of some protein or lipid to occur. This can be reached by physically separating this protein inside a membrane, but the cell has several membraneless organelles. Such an organelle thus needs a different process to reach such concentrations. The prime candidate is liquid-liquid phase separation and indeed it is able to correctly describe several phenomena such as P-granules<sup>17</sup> and centrosome growth<sup>18</sup>.

The strongest clue for describing the Golgi comes from its biogenesis: upon removal, the Golgi is able to form *de novo*. Ronchi et al<sup>8</sup> study golgi biogenesis in detail and found several phases of growth. First, vesicles are released from the ER until after some time stack-like structures are formed, including a cisternae-like topology. In the second phase, the stacks are transported by the microtubules to the location of the would-be golgi, where the stacks are fused into a single golgi ribbon. The first phase resembles phase transition: once the solution reaches some critical density, phase separation occurs and a dense droplet is formed. One crucial detail we've skipped over is that the golgi isn't a membraneless organelle. The cargo entering and exiting the Golgi isn't membraneless either; while these membranes ensure phase separation of the cargo, if we interpret the vesicles as the liquid to separate, phase separation occurs if the vesicles combine into a single dense droplet - the Golgi. A droplet arising from such a description would be a passive droplet however, forever growing. Contrarily, the golgi constantly takes up vesicles with immature proteins while vesicles with mature proteins bud off. Including a term to mimic this process makes

the droplet *active*. We can thus model a membrane-delimited organelle using tools normally used for membraneless ones.

### 1.2.2 BIOLOGICAL IMAGE ANALYSIS

Image analysis is a lively subject in cell biology, with many different methods and techniques being developed constantly. The last years image analysis is being developed with the specific goal of quantification in mind<sup>19</sup>, specifically as an insight into intracellular transport (20, 21, 22). To analyze intracellular transport or flow in a cell, single particle tracking (SPT) is a powerful method. SPT uses small fluorescent beads or particles suspended in a flow which can be tracked by a computer. The obtained trajectories can then be analyzed to obtain information about the transport. The RUSH method doesn't make use of such beads or particles, so that these techniques cannot be applied. Alternative methods mainly utilize some sort of correlation spectroscopy (23, 24, 25). One can derive a general relation between the correlation of the noise and transport properties such as the diffusion constant. These techniques require a signal roughly constant in time, which unfortunately is not a given for the RUSH experimental data. Thus, none of the techniques we found are applicable to the RUSH data.

All these methods are a specific solution to a general problem: how do we fit a model to a spatiotemporal dataset? Since most of the times a model is written in the form a partial differential equation (i.e.  $df/dt = \alpha df/dx + \beta d^2f/dx^2 + \dots$ ), while the data is  $f(x, t)$ , this is not a trivial problem. In this case, we're actually asking what set of parameters in our model (such as a diffusion coefficient) best describes the temporal evolution of a dataset. We have developed and evaluated two different methods for doing this. Our first method attacks the problem rather directly, by calculating spatial and temporal derivatives directly from the data using something known as image gradients. Our second alternative method is based on a recently invented method based on neural networks<sup>26</sup>. We'll show that by encoding physics into the neural network, we're not only able to infer the optimal parameters, but even an optimal parameter *field*.

### 1.2.3 STRUCTURE AND MAIN QUESTIONS

The rest of this thesis is divided into two parts. In the first part we show the two model fitting methods we have developed and apply them to the RUSH experimental data. The second part discusses the model we have developed for the Golgi. In a chapter-by-chapter breakdown, we have the following:

- Part I - Model fitting and data analysis
  - Chapter 2 introduces the framework we have developed for model fitting spatiotemporal data using image gradients.
  - Chapter 3 applies the method developed in chapter 2 to experimental data.
  - Chapter 4 shows an alternative method for model fitting based on neural networks.
- Part II - Model for Golgi
  - Chapter 5 introduces the Cahn-Hilliard equation, which describes phase separation, an approximation of it known as effective droplet theory and describes our model.
  - Chapter 6 contains the predictions the model developed in chapter 5 and investigates the biological implications.
- Chapter 7- is the concluding chapter and summarizes all the findings from the previous chapters.

# 2

## Model fitting

In this chapter we introduce the method we have developed for fitting a model in the form of a PDE to spatiotemporal data. We start with a section describing the general idea and subsequent sections elaborate on each step. The method principally works for any type of data and model, but was developed originally to analyze data from the RUSH experiments. We have chosen to illustrate the effects of each step with RUSH experimental data instead of synthetic data.

### 2.1 THE CONCEPT

Assume we have access to experimental data of some process  $f(x, t)$ . Parallelly, we have also developed a model describing this process, but it is in the form of a PDE:

$$\partial_t f(x, t) = \lambda_1 \nabla^2 f(x, t) + \lambda_2 \nabla f(x, t) + \lambda_3 f(x, t) + \lambda_4 \quad (2.1)$$

We now wish to investigate if this model fits the data  $f(x, t)$  and find the optimal value of coefficients  $\lambda_i$ . To do so, we consider each spatial term in  $f(x, t)$  in eq. 2.1 as some variable  $x_i$  but  $\partial_t f$  as  $y$ , so that we can rewrite it as:

$$y = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \lambda_4$$

If we thus can find the variables  $x_i$  and  $y$ , we can perform some fitting procedure such as least squares to obtain the coefficients  $\lambda_i$ . In other words, if we can calculate the spatial and

temporal derivatives of our data, we can fit the model. Although the concept seems trivial, its implementation is all but. Data is rarely noiseless and obtaining accurate derivatives from noisy data is notoriously hard, but forms the heart of our method. It's also possible to have distinct models in different areas of the data, so that we need to segment the data. Furthermore, the coefficients  $\lambda_i$  might not be constant but could be space- and time- dependent. The process of fitting the data thus has several steps:

1. Denoising and smoothing
2. Calculating derivatives
3. Segmenting
4. Fitting

In the next sections, we describe each step separately. Note that the method we present here has been developed empirically: there's no theoretical background as to why this particular combination should work optimally. Instead, it's been developed while analyzing the data, adapting each step on the go. However, we have found several parallel to another method from machine vision known as optical flow<sup>27</sup>.

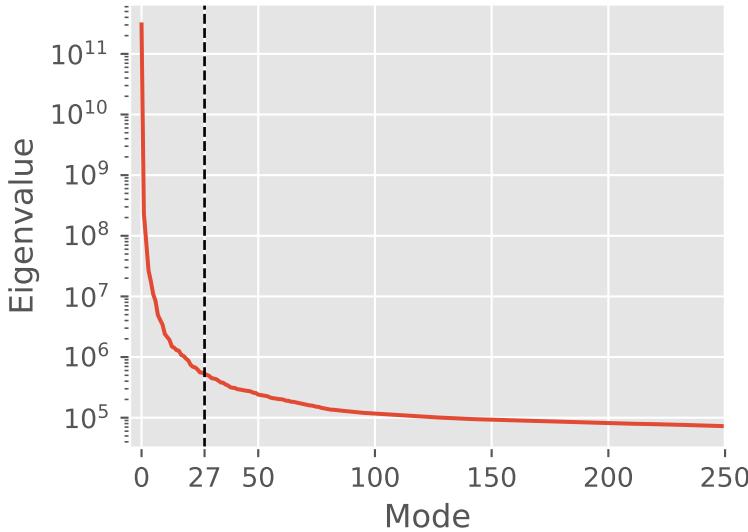
## 2.2 STEP 1 - SMOOTHING AND DENOISING

The first step in our pipeline is to denoise and smooth the data. The smoothing is necessary for accurately calculating the derivatives. Denoising still is a very active area of research (especially in life sciences) and dozens of different methods exist<sup>30</sup>. For example, one could Fourier transform the signal and use a high pass filter, but this would also get rid of small and sharp features. After evaluating several methods, we have settled on the so-called ‘WavInPOD’ method, introduced by 31 in 2016. In 32 they show that this methods outperforms several other advanced methods. WavInPOD combines Proper Orthogonal Decomposition (POD) with Wavelet filtering (Wav). Both subjects are vast (especially Wavelet transform) and we're only interested in the result of the technique, so we only present a short introduction here, adapted from 31.

POD is a so-called dimensionality reduction technique and is very closely related to the more well-known Principal Component Analysis (PCA) in statistics. In physics it's often used to analyze turbulent flows<sup>33</sup>. In POD we wish to describe a function as a sum over its variables:

$$f(x, t) = \sum_n^r \alpha_n(x) \varphi_n(t)$$

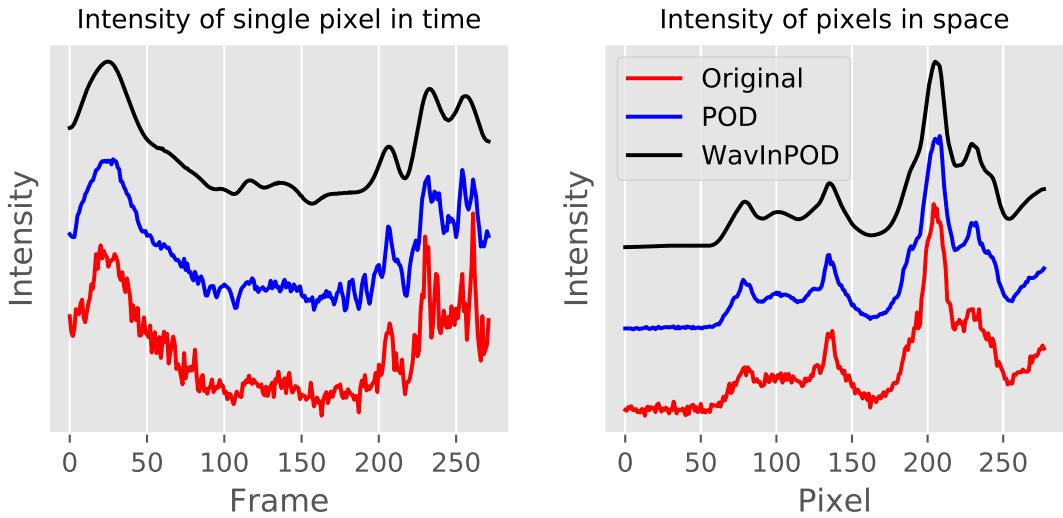
where  $\alpha_n$  and  $\varphi_n$  are called respectively the spatial and temporal modes. Associated with each mode  $n$  is an energy-like quantity  $E_n$ . Modes with a higher ‘energy’  $E_n$  contribute more to the signal  $f$  than modes with a lower energy and we can thus approximate the signal by selecting the  $k$  modes with the highest energy. A typical logio energy spectrum has a ‘knee’ in the values, as shown in figure fig. ???. Modes with an energy below the knee are noise, and modes above signal.



The wavelet transform is very similar to the Fourier transform, but uses wavelets as its basis. A fourier transform gives the frequency domain with infinite precision, but tells nothing about the locality of the frequencies (i.e when each frequency is present in a signal). By using a wavelet (a wave whose amplitude is only non-zero for a finite time), we sacrifice precision in the frequency domain but gain information on the locality instead. Performing a wavelet transform transforms the signal into the sum of an approximation and its details and we can filter this analogous to a fourier filter. Due to its locality however, noise is filtered out, by sharpness is retained.

WavinPOD combines these two techniques by applying wavelet filtering to the POD modes. In detail, one first decomposes the problem with a POD transformation. The energy spectrum of this transformation is shown in figure fig. ?? and we select a cutoff of 27. All retained modes are wavelet filtered and are then retransformed to give the denoised and smoothed signal. In figure fig. 2.1 we show the results of the smoothing in the time and spatial domain. In the left panel we show the signal of a single pixel in time, while we plot a line of pixels in a single frame in the right panel. The red lines denote the original (unfiltered) signal, the blue line the effect of just applying a POD and the black one the

result of the WavInPOD technique. Note that the effect of the wavelet filtering is to smooth the signal significantly and in comparing the original data to the filtered data that we've retained the sharpness of the features whilst obtaining a smooth signal.



**Figure 2.1:** Effect of POD with a cutoff of 27 and wavelet filtering with a level 3 db4 wavelet. Left panel shows the result in the time domain, right panel in the spatial domain. Lines have been offset for clarity.

### 2.3 STEP 2 - DERIVATIVES

After having denoised the images, we calculate the spatial and temporal derivatives. Obtaining correct numerical derivatives is hard and becomes much more so in the presence of noise<sup>35</sup>. Next to a finite-difference scheme, one can for example (locally) fit a polynomial and take its derivative or use a so-called tikhonov-regularizer<sup>36</sup>. The computational cost of these methods is high however and they don't scale well to dimensions higher than one. For our spatial derivatives these methods are thus not available. In fact, obtaining the gradient of a 2D discrete grid has another subtlety which we need to address.

Naively, one could obtain the gradient of a 2D grid by taking the derivative using a finite difference scheme with respect to the first and second axis. If there are features on the scale of the discretization ( $\sim$  few pixels), such an operation will lead to artifacts and underestimate the gradient. These issues have long been known and several techniques have been developed to accurately calculate the gradient of an ‘image’. The most-used image-gradient technique is the so-called Sobel operator. It belongs to a set of operations known as *kernel*

*operators.* Kernel operators are expressed as a matrix and by convolving this matrix with the matrix on which the operation is to be performed, we obtain the result of the operator. We show this for the Sobel operator.

Consider a basic central finite difference scheme:

$$\frac{df}{dx} \approx \frac{f(x_{i+1}) - f(x_{i-1})}{2h}$$

where  $h$  is defined as  $x_{i+1} - x_i$ . In terms of a kernel operator, we rewrite this as:

$$\frac{1}{2} \cdot [1 \ 0 \ -1]$$

where we have set  $h = 1$ , as the distance between pixels is one by definition. By convolving this matrix with the matrix  $A$  we obtain the derivative of  $A$ :

$$\partial_x A \approx A * \frac{1}{2} [1 \ 0 \ -1]$$

Finite difference

0	0	0
-1		1
0	0	0

3x3 Sobel

-1	0	1
-2		2
-1	0	1

As stated, this operation is inaccurate and introduces artifacts. To improve this, we wish to include the pixels on the diagonal of the pixel we're performing the operation on as well (see figure fig. ??). The distance between the diagonal pixels and the center pixel is not 1 but  $\sqrt{2}$  and the diagonal gradient also needs to be decomposed into  $\hat{x}$  and  $\hat{y}$ , introducing another factor  $\sqrt{2}$ . The kernel thus obtained is the classic  $3 \times 3$  Sobel filter:

$$G_x = \frac{1}{8} \cdot \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

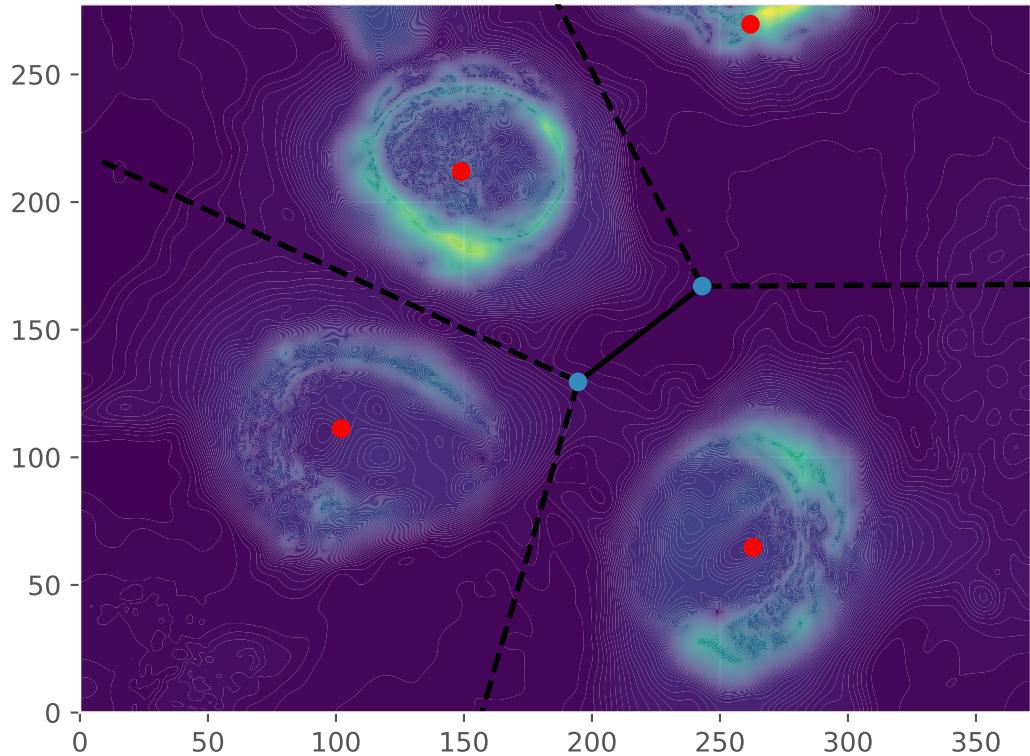
Increasing the size of the Sobel filter increases its accuracy and we've implemented a  $5 \times 5$  operator. Implementing the derivative operation as a kernel method is also beneficial from a computational standpoint, as convolutional operations are very efficient. The Sobel filter is usually applied to an image and hence is often said to calculate the image-gradient, but due to its separability is possible to scale this method to an arbitrary number of dimensions.

## 2.4 STEP 3 - SEGMENTATION

In the case of the RUSH data, obtained images and movies often contain multiple cells. Each of these cells can be further segmented into two more areas of interest: the cytoplasm, which is where we want to fit our model and the Golgi apparatus. We wish to make a mask which allows us to separate the cells from the background and themselves and divide each cell into cytoplasm or Golgi. Figure fig. ?? shows four typical frames in the MANII transport cycle. Note that no sharp edges can be observed, especially once the MANII localizes in the Golgi. No bright field images were available as well, together making use of techniques such as described in [37] unavailable. We have thus developed two methods which allow us to segment the image and the cells, based on Voronoi diagrams and the intensity.

### 2.4.1 VORONOI DIAGRAM

Consider again the frame on the left of figure fig. ???. Note that in early frames such as this one, the cargo (i.e. fluorescence) is spread circumnuclear. Applying a simple intensity based segmentation gives us a number of separate areas, which *very roughly* correspond to a cell. We can then pinpoint each cells' respective center. Given  $n$  points, Voronoi tessellation divides the frame into  $n$  areas, where point  $i$  is the closest point for each position in area  $A_i$ . The hidden assumption here is thus that each pixel belongs to the cell center it's closest too. Although this is a very big assumption, in practice we've found this to be reasonable. Furthermore, one can add 'empty' points to make the diagram match observations. Assuming small movements of the cell, this isn't an issue either for this technique, as we are assigning an area to each cell instead of very precisely bounding it. This also allows us to calculate the Voronoi diagram in the early frames and apply the segmentation to the entire movie. The result of this segmentation for MANII is shown in figure fig. 2.2. Each cell centre is denoted by a dot, while the lines denote the border between each voronoi cell.



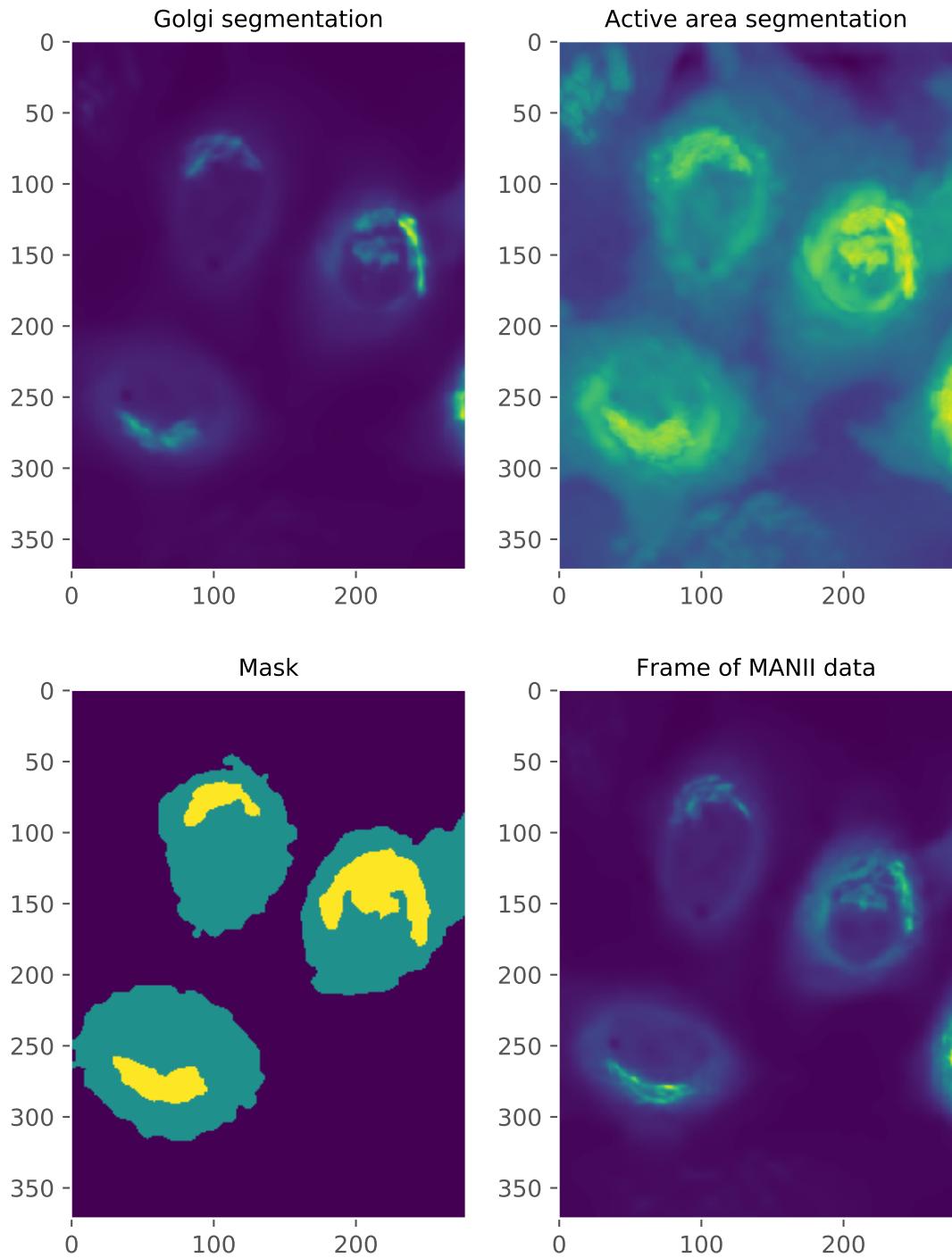
**Figure 2.2:** The obtained mask. Red dots are cell centers, dashed lines infinite edges and solid lines finite edges.

#### 2.4.2 INTENSITY

The Voronoi technique works very well for an area-based approach, but for analyzing our fitting data we would like a more precise mask - although we still don't require pixel-level accuracy. From the movies, the Golgi is clearly visible and we can separate the cytoplasm from the background, with a big 'gray' area inbetween. We thus turn to an intensity based approach. We have developed the following approach for localizing the Golgi:

1. Normalize the intensity  $I$  between 0 and 1.
2. Sum all the frames in time:  $\sum_n I(x, y, t_n)$ . A typical result is shown in figure fig. 2.3 .
3. Threshold the image to obtain the mask. This is either done automatically through an Otsu threshold or by manually adjusting the threshold until desired result.
4. The mask is postprocessed by filling any potential holes inside the mask.

This procedure was unable to reliably separate the background from the cytoplasm. Noting that while the cytoplasm might not have the intensity as the golgi, its time derivative should be higher than the rest of the areas. We replace step two by  $\log_{10} (\sum_n I(x, y, t_n) \cdot \partial_t I(x, y, t_n))$ , where the time derivative has been normalized between 0 and 1. Figure fig. 2.3 shows our final results. The upper two panels show the images obtained after performing the summing operation for the Golgi and cytoplasm respectively, while the lower left panel shows the final mask obtained after thresholding these two images. For comparison, we plot frame to compare the mask to.

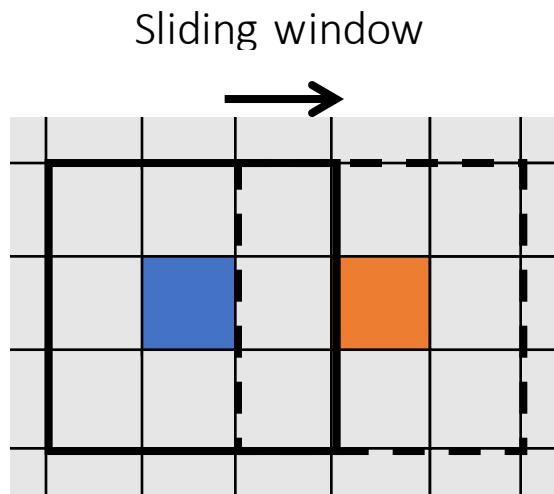


**Figure 2.3:** Four panels showing the different stages of making the mask. From segmenting the upper two panels we determine the golgi and active area, leading to the mask in the lower left. Compare the to the lower right.

## 2.5 STEP 4 - FITTING

The final step in our method is to fit the our model to the data. By calculating the derivatives, we have reduced our PDE-model to a generic model, which allows us to use virtually any fitting method. For simplicity, we use least-squares, but one could use a Bayesian method to obtain not only the fit, but also the uncertainty.

Equation eq. 2.1 assumes a model with constant coefficients. In reality, coefficients will be spatially and even temporally varying. To circumvent this issue, we assume the coefficients can be assumed to be locally constant. We thus assume that for a small area we can fit the model using constant coefficients. We perform this operation for every datapoint in a sliding-window manner, as shown in figure fig. 2.4, thus ending up with a coefficient field.



**Figure 2.4:** Schematic overview of the sliding window technique. The solid black line encompasses an area around its blue coloured central pixel and the fit output is assigned to that pixel. We then move the window (dashed black line) and perform the fit for the orange coloured pixel.

In the next chapter we apply this method to the RUSH experimental data.

# 3

## Data analysis

In this chapter we apply the method developed in the previous chapter to experimental data obtained using the RUSH technique. Our first section introduces the RUSH technique and discusses our model for the intracellular transport. In the following section we discuss the experimental data, investigate the fluorescence curves of several areas of interest, and gain more insight into the data by studying its time derivative. We then present a linear least squares fit and show that this can lead to unphysical results. We end with a short section of both recommendations for the experimentalists as well as a number of ways to improve the method.

### 3.1 EXPERIMENTAL DATA

The transport of vesicles from the ERES to the Golgi is both diffusive and directive and a technique known as RUSH (Retention Using Selective Hooks) has recently been developed<sup>16</sup> in the team of Frank Perez at Institut Curie to study this trafficking. RUSH allows precise timing of the release of proteins from the ER and can be used to follow the secretory pathway from the ER to the Golgi and even post-golgi using fluorescent live-cell imaging. Several other methods have been developed (???-golgi\_1997, ???-golgi\_2018), but lack the non-toxicity, timing and versatility of the RUSH technique.

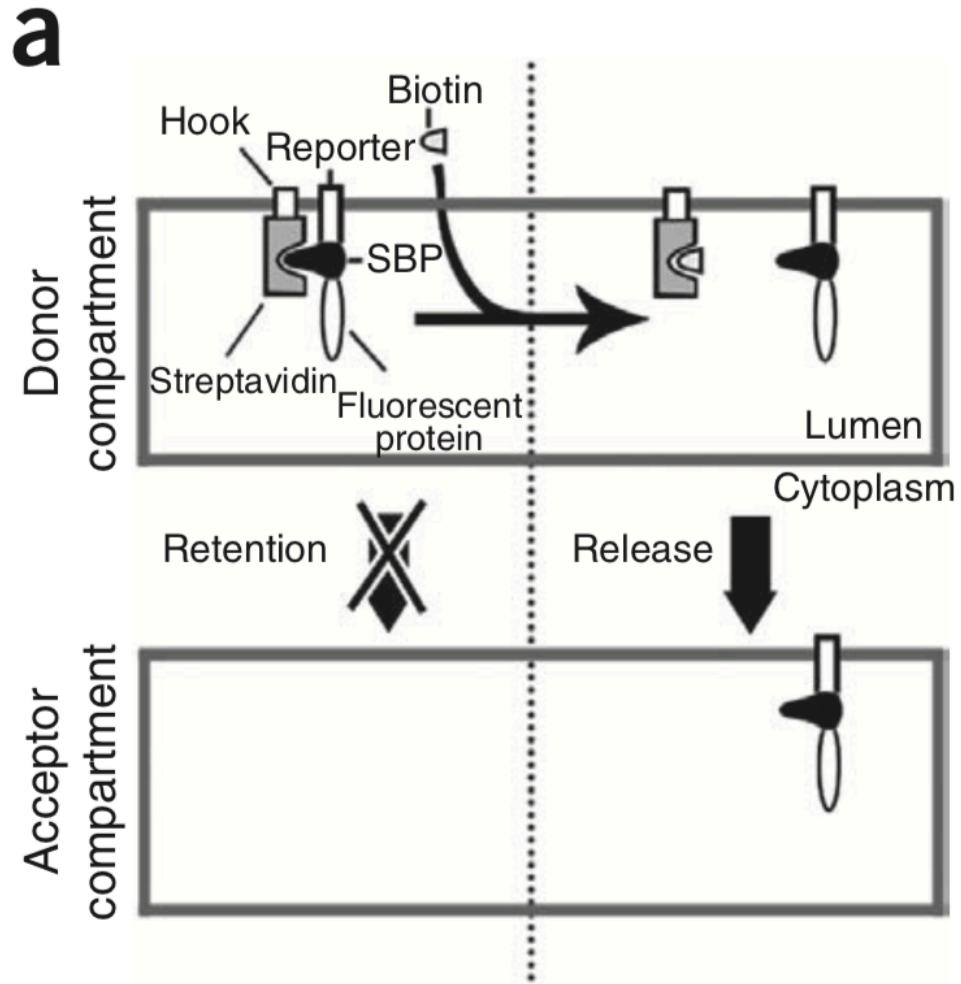
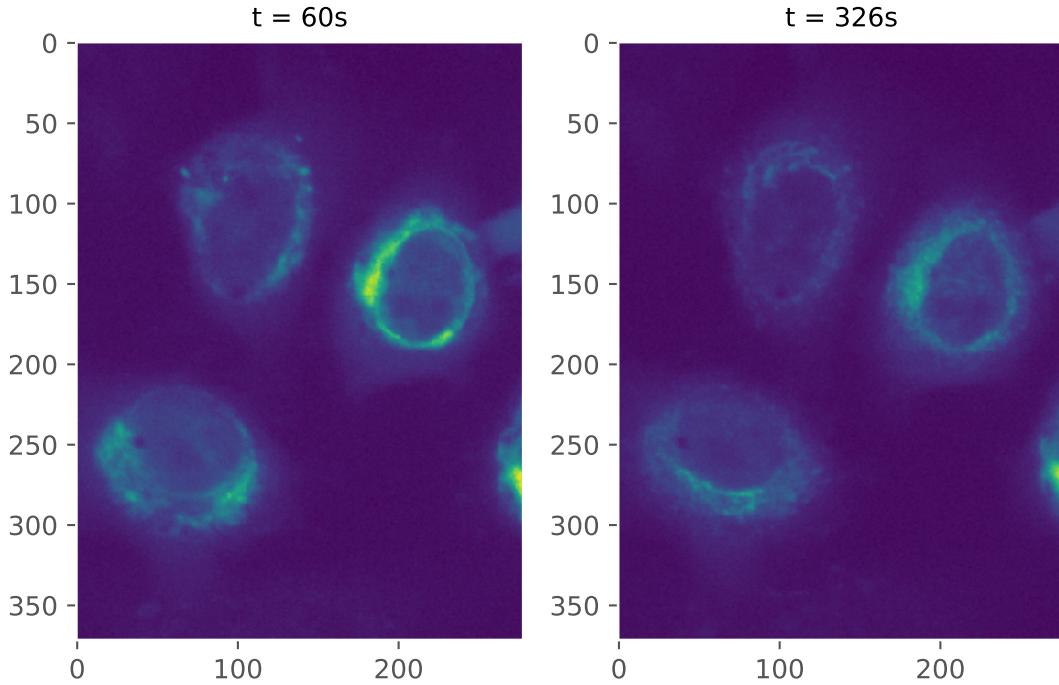


Figure 3.1: Schematic overview of the RUSH system. Image taken from 16

Figure fig. 3.1 shows the principle of the RUSH system. Inside the ER, a core streptavidin is fused to it using a hook protein. Another protein known as a streptavidin-binding-protein (SBP) binds to streptavidin, but connected to the SBP are also the protein to be transported ('reporter') and a fluorescent protein. Upon the addition of biotin, the SBP is released from the streptavidin as the biotin binds to it. The SBP-reporter-fluorescent complex then exits the ER and can be followed the entire secretory pathway with fluorescence microscopy.

The RUSH technique can be used for many different proteins, but in this thesis we mainly focus on the  $\alpha$ -mannosidase-II, generally referred to as ManII. The ManII protein resides in the Golgi apparatus and thus upon reaching it will remain there. This means that the data we obtain will only contain transport *towards* the golgi, greatly simplifying the analysis as we won't have to post and pre-golgi traffic. Figure fig. 3.2 shows two frames in a typical RUSH experiment of ManII trafficking. The left panel shows an image obtained just after the addition of the biotion, so that most of the cargo is still retained in the ER. A later frame is shown on the right: we can observe the localization of fluorescence in the Golgi, while there's still fluorescence in the rest of the cells.



**Figure 3.2:** Two frames of the ManII transport images using the RUSH technique.

### 3.1.1 MODEL

Vesicles exiting the ERES are transported towards the ER over the microtubules. This is a stochastic process with the proteins detaching from and (re-) attaching to the microtubules randomly, while the vesicles move diffusely once detached. Several models have been developed to describe such intracellular transport processes (6, 38), many in the light of virus trafficking (39, 40, 41). In general, these models assume a two population model, with one

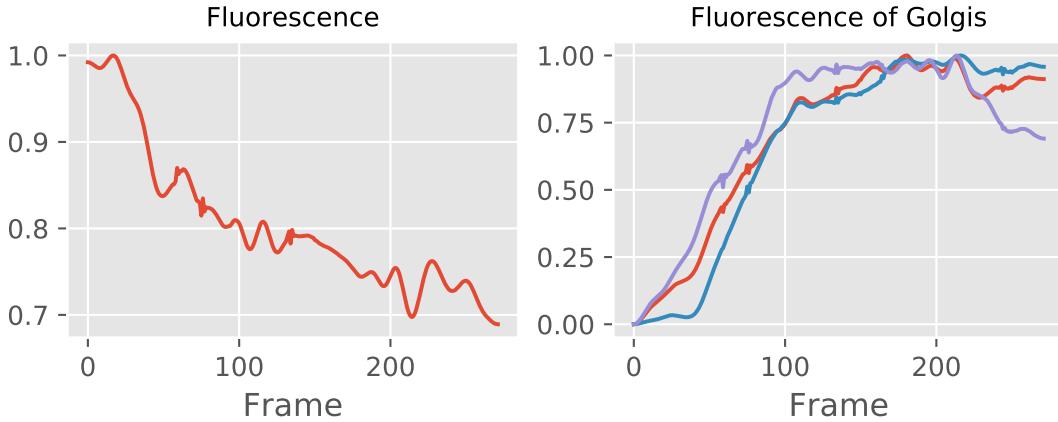
population being cargo attached to a microtubule and another cargo freely diffusing in the cytoplasm. If one assumes that the timescale for attaching and detaching from the microtubules is much smaller than the transport timescale, the two populations can be assumed to be in equilibrium. In this assumption, known as a quasi-steady-state reduction, the two population model reduces to a Fokker-Planck equation. As the Fokker-Planck equation is functionally equivalent to an advection-diffusion equation, we hypothesize that we can model protein transport using an advection-diffusion equation:

$$\partial_t c = D \nabla^2 c - v \nabla c \quad (3.1)$$

where  $c$  is the concentration of the cargo,  $D$  a diffusion coefficient and  $v$  an advection velocity. Equation eq. 3.1 is thus the model we fit our data to. Note that the fluorescence images obtained from the RUSH experiment return an intensity  $I$  and not a concentration  $c$ , and hence we make the assumption  $c \propto I$ .

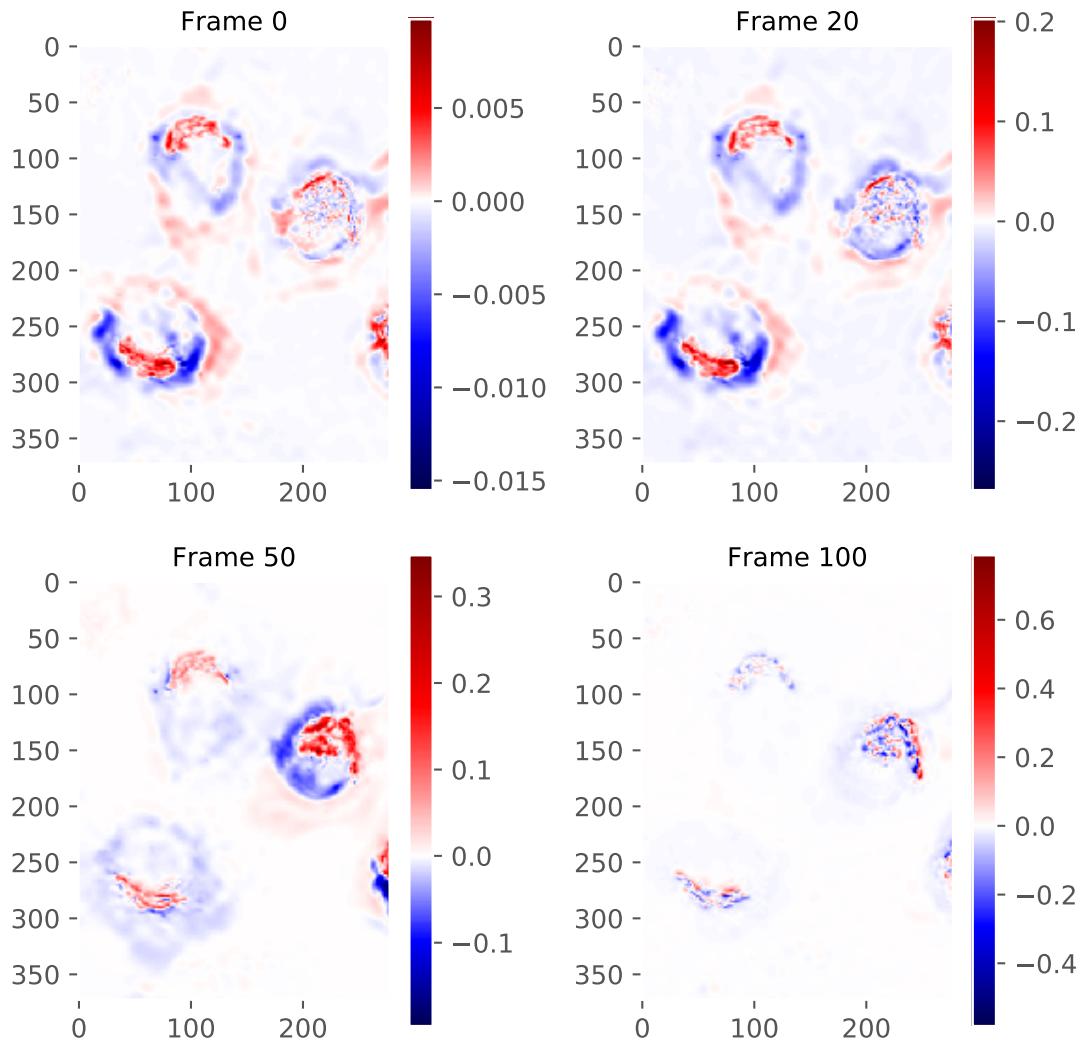
### 3.2 INITIAL ANALYSIS

We first study the evolution of the fluorescence in two ways, plotting both the mean fluorescence for each frame and the mean fluorescence in the golgi region. To get rid of the background in our statistics, we normalize the fluorescence between before computing the mean. The left panel of frame #fig:fluorescence shows the average fluorescence of each frame and shows a significant drop of almost 30% between the initial and final frame. We observe a strong initial drop and a slower decay after roughly frame 100, which can probably be attributed to photobleaching. The transition between the two decays roughly matches the saturation of the fluorescence in the Golgi (see the right panel), casting strong doubts on our assumption that  $c \propto I$ . Compensating for this is possible (see 15), but requires significant effort on the experimentalists' part and if the difference between two subsequent frames is small the effect on our fit is negligible.



In the right panel we show the fluorescence in the Golgi ROI for each of the three cells. Specifically, we plot the mean intensity in each ROI, normalized on the maximum intensity and compensated for the loss of fluorescence as shown in the left panel. Interestingly, we observe that all three curves roughly follow a similar pattern. Although one of the cells (blue line) seems to have some sort of delay, the fluorescence seems to increase in a linear way up to frame 100, when the fluorescence saturates. The purple cell shows a significant drop at frame 200, but since the ManII protein remains in the Golgi, this is not caused by any type of intracellular transport and thus not of interest to us. The linear increase and common pattern suggests that the transport properties are not concentration dependent at these concentrations.

To perform our fit in the next section, we have also determined the time derivative of the dataset,  $\partial_t I$ . We plot this for four different frames in figure fig. 3.3 .

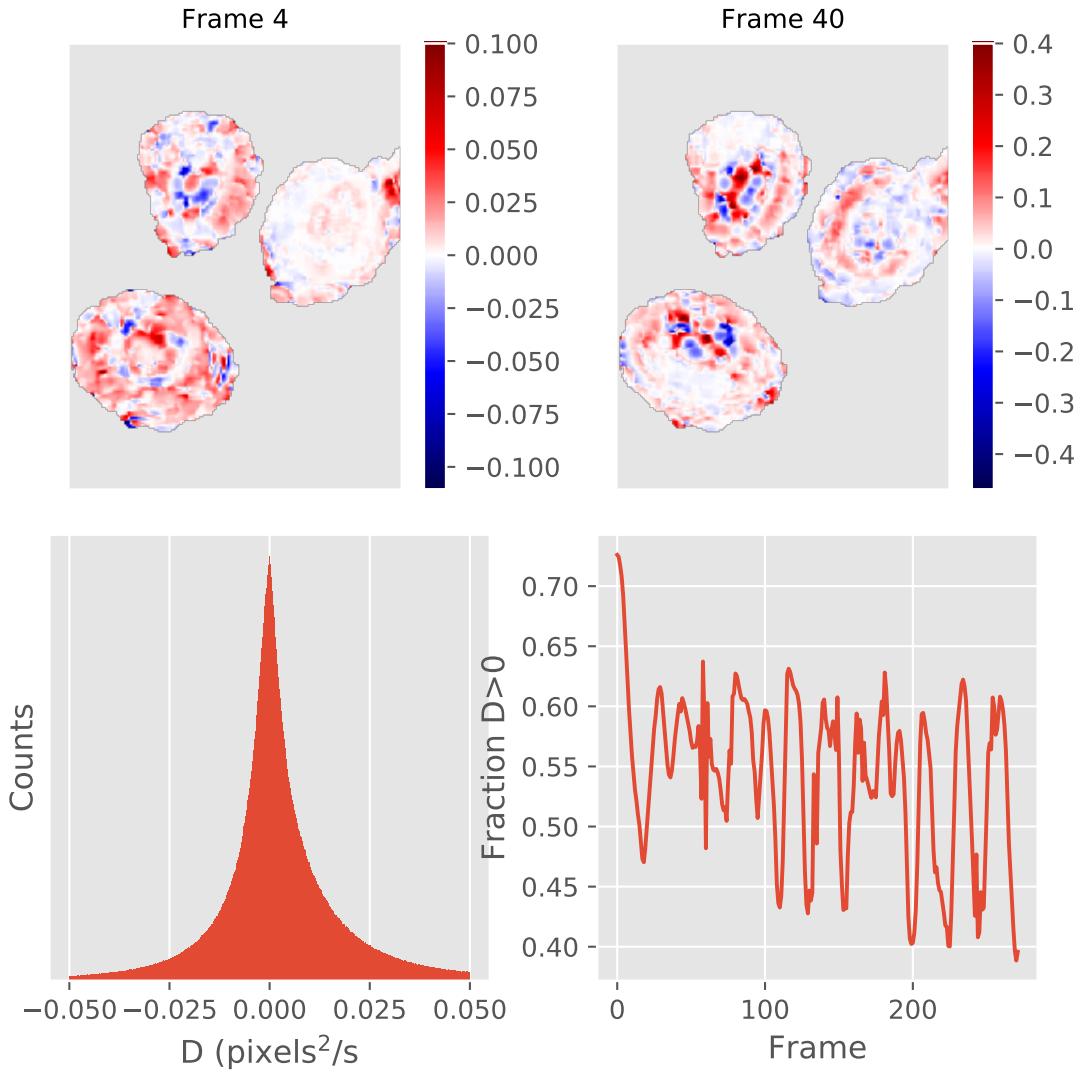


**Figure 3.3:** The determined time derivative four different frames of the ManII RUSH experiments.

Areas where the derivative is positive (thus were the concentration increases) are coloured red, while areas where the concentration decreases are coloured blue. As expected, the Golgi shows up in each cell as a bright red object. Note however that we also observe red areas towards the edges of the cell. As the concentration close to the Golgi decreases, the red area moves outwards and slowly takes over the blue area. This could be caused by ERES acting as a in that area, but given that ERES are located throughout the cell it seems more likely such a pattern would be caused by diffusion.

### 3.3 ANALYSIS OF LS-FIT

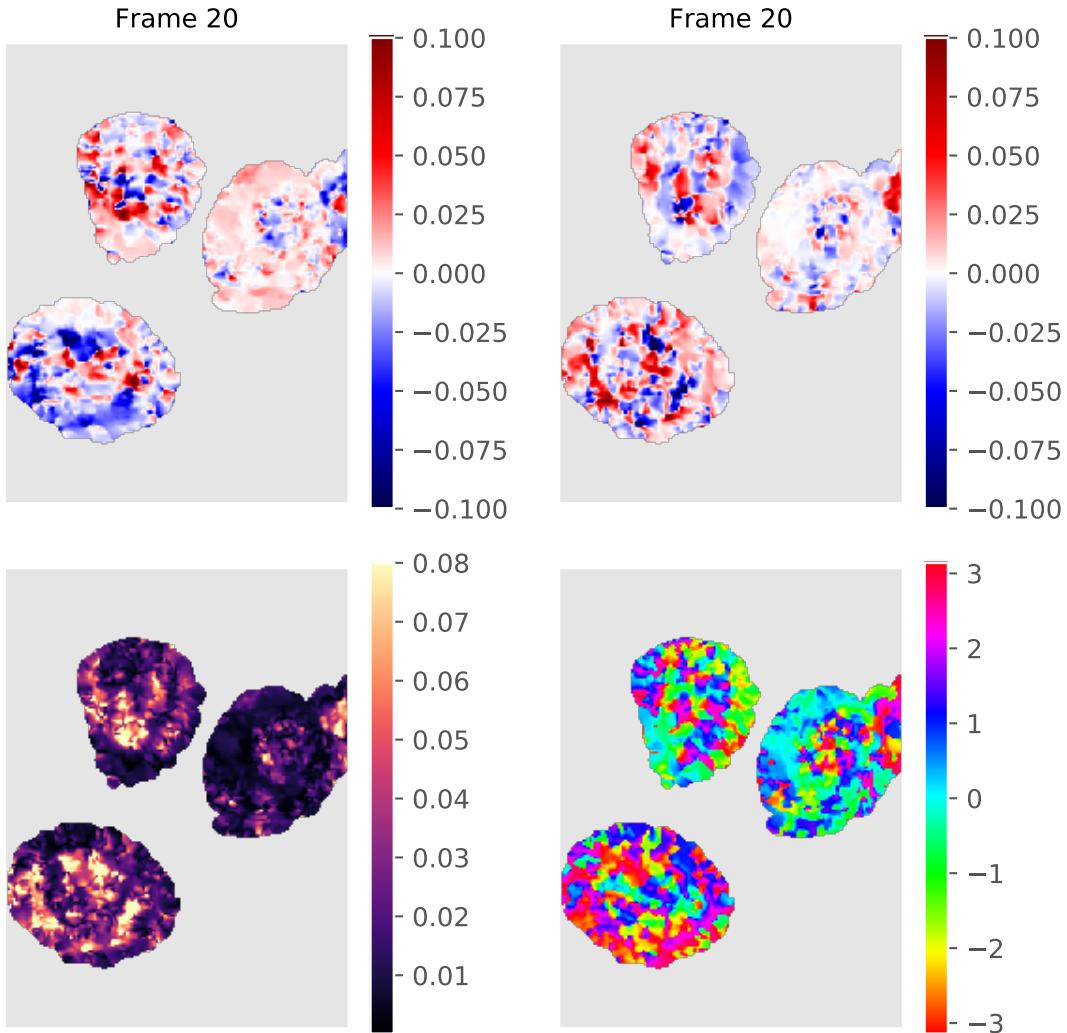
In this section we analyse results of the least squares fit. We've used a  $7 \times 7$  window in the spatial domain to perform the sliding window operation, fitting each frame of the movie independently. We analyse the diffusion, advection and the error of our fit. These fits are movies and we're hence unable to print them - please find our these movies at our GitHub. In figure fig. 3.4 we show two typical inferred diffusion fields in the upper row, just after addition of the biotin (frame 4) and halfway to the complete saturation of the Golgi (frame 40).



**Figure 3.4:** Analysis of the inferred diffusion field. The upper row shows the inferred field at two frames, while the lower row shows the distribution of values and the fraction of physical values as a function of time.

Another problem is that we observe many areas with a negative -unphysical- diffusion coefficient. In the lower left panel we plot the distribution of values. Analysis shows that roughly 40% of the inferred field has a negative diffusion coefficient. In the lower right panel we have calculated this fraction as a function of time. It shows that, save for a few initial frames, this fraction is not (strongly) time-dependent. Results are slightly skewed though, since many coefficients are negative but extremely close to zero (e.g  $-10^{-4}$ ). Neg-

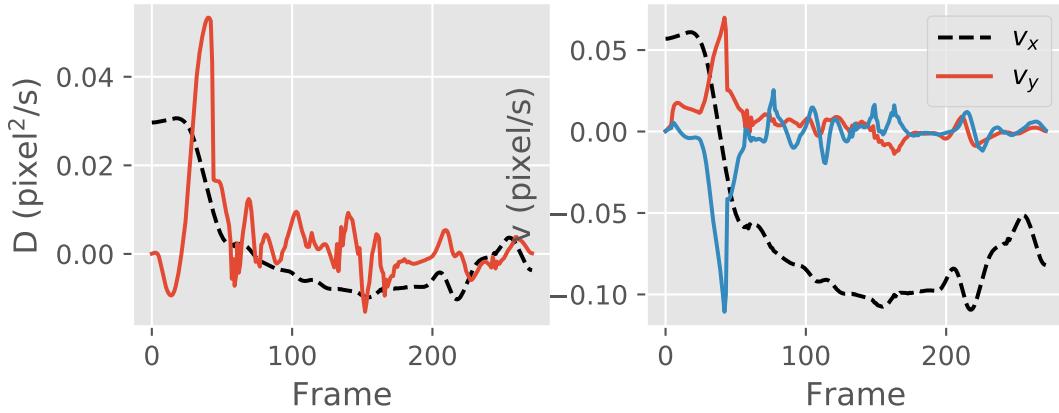
tive diffusion coefficients correspond to clustering, but could also be the result of an incorrect fit. We investigate this in depth after studying the advection profiles, which we show in figure fig. 3.5 . In the four panels we show the inferred velocity in the  $\hat{x}$  and  $\hat{y}$  direction, the magnitude of the velocity and the angle.



**Figure 3.5:** Analysis of the velocity fields. The upper rows show respectively  $v_x$  and  $v_y$ , while the lower row shows the magnitude and angle.

Similar to the diffusion, we observe patterns both in time and space bigger than our fitting window, meaning that the fit isn't completely random. On the other hand, we are not able

to discern any specific flow from the figures in fig. 3.5. To gain more insight into our fit, we analyze a single pixel in time. Figure fig. 3.6 shows the diffusion and advection velocities as a function of time. We've plotted a scaled and translated signal of that pixel in a black dashed line. This pixel is initially constant and then decreases to noise level. Note that initially, while the signal is constant, the diffusion constant is negative. Once the signal starts decreasing, or, in other words, cargo starts flowing from that pixel, we see a physical diffusion constant and non-zero velocities. Once the signal returns to around noise-level at pixel 50, the inferred velocities and diffusion constant seem to become random around 0. In other words, our method seems to work when cargo is flowing, but struggling when the signal is either constant or at noise level. We observe similar behaviour in other pixels, so we contribute (most of) the unphysical diffusion values to constant and noise-level signal. Performing the fit with the diffusion constant constrained to be larger than zero lead to the negative coefficients in the free fit to being set to zero.



**Figure 3.6:** Diffusion and advection velocities of a single pixel in time. We've plotted (scaled and translated) signal as a black dashed line to find any correlation.

This doesn't completely explain the magnitude of the coefficients though, which is significantly lower than expected. One possibility is the ‘mixing’ of the transport fluorescence with the fluorescence of the ER. After the addition of biotin, the fluorescent cargo gets released, but still has a finite residence time in the ER. Since the obtained images are projected over an axis, changes in fluorescence we observe can be both due to intracellular transport as well as processes inside the ER. If these processes have different timescales, this can strongly affect the inferred coefficients.

Another possibility is that we've assumed that a concentration of fluorescent particles leads

to some sort of ‘mean-field’ fluorescence which we can describe by an advection-diffusion equation. Once the size of the particles becomes on the order of the pixel size, this assumption breaks down. In the case of the ManII trafficking, the pixels are roughly two to three times the size of a vesicle, meaning that we are at the limits of the ‘mean-field’ assumption.

### 3.4 CONCLUSION

We’ve applied the method developed in the previous chapter to the RUSH trafficking data of the ManII protein. Although our fit shows patterns in the coefficient fields larger than our fitting window, pointing at some underlying pattern, we’re unable to make any conclusions about our model. It’s been shown that the model seems to perform well when the data is transient, i.e. when cargo is actually flowing. However, for most of the time the data is either constant or at noise level, preventing us from making any conclusions.

Improvements to our method fall roughly into two categories. The first category concerns improvements to the calculation of the derivatives. We’ve implemented a rather basic  $5 \times 5$  Sobel filter, but implementing more advanced methods which would result in more accurate derivatives would probably make the biggest improvement. The second category would be improve the fitting procedure. The most obvious candidate is implementing some sort of Bayesian method which would return not just the most probable coefficient, but the entire probability distribution.

# 4

## Physics Informed Neural Networks

In the previous chapters we showed the difficulties in fitting a model in the form of a partial differential equation to spatio-temporal data. The method we developed was a classical numerical approach, separating the problem into several substeps such as denoising, smoothing and numerical differentiating. In the last few years machine learning has been slowly making its way into physics. Very recently, a technique generally referred to as Physics Informed Neural Networks (PINNs) have shown great promise as both tools for simulation and model fitting (42, 43, 44, 26, 45). In this chapter, I will evaluate the use of this technique to fit the model to the RUSH data. I've divided the chapter into three parts:

- Neural Networks - This part will cover the basics of neural networks: their inner workings, how to train them and other general features.
- Physics Informed Neural networks - In this second part we introduce the concept behind PINNs, use it to solve a toy problem and apply it to our RUSH data.
- Conclusion - Finally we summarize the results and observations from the previous sections.

### 4.1 NEURAL NETWORKS

Artificial Neural Networks (ANNs) are networks inspired by biological neural networks. Contrary to other ways of computing, ANNs are not specifically programmed for a task -

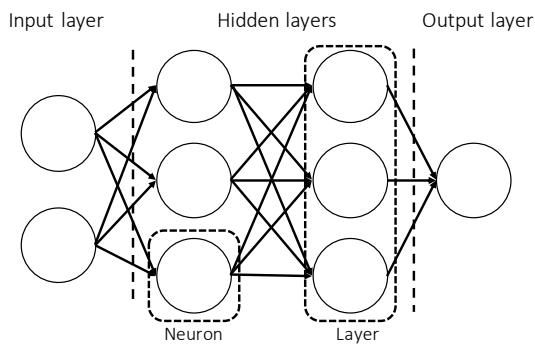
instead, ANNs are *trained* using a set of data. Research on artificial neural networks started in the '40s but never gained any critical mass, as no efficient training algorithm was known. Once an efficient training algorithm was found in 1975 by Werbos, interest resurged but it wasn't until the late '90s that deep learning started gaining widespread traction. The use of GPU's allowed ANNs to be efficiently trained and widely deployed at reasonable cost.

The advancements in machine learning in general and especially neural networks in the last ten years have yielded a wealth of techniques and approaches. In supervised learning, the network is given pre-labeled data so that it is trained by learning the mapping from the given inputs to the given outputs. Other types such as unsupervised learning, where the network needs to learn to discriminate between unlabeled data, and reinforcement learning don't have any obvious use for PINNs yet and I've thus chosen to omit them. In the next sections, I'll present the mathematics of an ANN and show how they are trained using the so-called *backpropagation* algorithm.

#### 4.1.1 ARCHITECTURE

*An excellent introduction is given by Michael Nielsen in his freely available book “Neural networks and deep learning.” The following section has been strongly inspired by his presentation.*

At the basis of each neural network lies the neuron. It transforms several inputs non-linearly into an output and we can use several neurons in parallel to create a *layer*. In turn, we several layers in series make up a network. The layers in the middle of the network are known as *hidden layers*, as shown in figure fig. ??



**Figure 4.1:** Schematic view of a neural network.

In the schematic shown in fig. ??, each neuron is connected to every neuron of the previ-

ous and next layer. This is known as a *fully connected* layer. Using only this type of layers, we've created a feed-forward network and it has been proven that a single hidden layer with enough neurons is a *universal function approximator*, i.e. a neural network can represent any continuous function using enough neurons.

As stated, a neuron takes several inputs and transforms them into an output. This is a two step process, where in the first step the neuron multiplies the input vector  $x$  with a weight vector  $w$  and adds a bias  $b$ :

$$z = wx + b \quad (4.1)$$

$z$  is called the weighted input and is transformed in the second step by the neuron *activation function*  $\sigma$ . This in turn gives the output of the neuron  $a$ , also known as the activation:

$$a = \sigma(z) = \sigma(wx + b) \quad (4.2)$$

The role of the activation function is to introduce non-linearity into the system. The classical and often used activation function is the *tanh*, as it is bounded between +1 and -1. Since we're working with multiple layers, it is useful to rewrite function eq. 4.2 in terms of the activation  $a^l$  of layer  $l$ :

$$a^l = \sigma(z^l) = \sigma(w^l a^{l-1} + b^l)$$

where  $w^l$  and  $b^l$  are respectively the weight matrix and bias of layer  $l$ .

#### 4.1.2 TRAINING

In supervised learning the task of training a machine means adjusting the weights and biases until the neural network predictions match the desired outputs. We thus need some sort of metric to define this 'distance' between prediction and desired output. Training the network than means minimizing the metric with respect to the weights and biases of the network. This metric is known as the cost function  $\mathcal{L}$  and the most used form is a mean squared error:

$$\mathcal{L} = \frac{1}{2n} \sum_i |y_i - a_i^L|^2 \quad (4.3)$$

where  $n$  is the number of samples,  $y_i$  the desired output of sample  $i$  and  $a_i^L$  the activation of the last function - the prediction of the network. Minimizing this is not trivial, as the problem can have many local minima. A solution can be found however using gradient descent techniques.

Gradient descent techniques are based on the fact that given an initial position, the fastest way to reach the minimum from that position is by following the steepest gradient. Thus, given a function  $f(x)$  to minimize w.r.t to  $x$ , we guess an initial position  $x_n$  and iteratively change until it converges:

$$x_{n+1} = x_n - \gamma \nabla f(x_n)$$

where  $\gamma$  is known as the learning rate. If a global minimum exists, this technique will converge on it. More advanced versions of this technique exist which are able to deal with local minima as well, since convexity of the cost function is not at all guaranteed.

Making use of gradient descent requires knowledge of the derivatives of the cost function w.r.t to the variables to be optimized. In the case of neural networks, we thus need to know the derivative w.r.t to each weight and bias. A naive finite difference scheme would quickly grow computationally untractable for even shallow networks. A solution to this problem was found by Werbos in the form of the backpropagation algorithm. Despite many years of ongoing research, it is still the go-to algorithm for each neural network implementation.

## BACK PROPAGATION AND AUTOMATIC DIFFERENTIATION

As we wish to minimize the cost function w.r.t. to each weight  $w$  and bias  $b$  using gradient descent, we need to find the derivative of the cost function w.r.t to each. Our argument simplifies if we move away from vector notation and introduce  $w_{jk}^l$ , the weight of the  $j$ -th neuron in layer  $l-1$  to neuron  $k$  in layer  $l$  and  $b_j^l$ , the bias of the neuron  $j$  in the  $l$ -th layer. We introduce the error of neuron  $j$  in layer  $l$  as:

$$\delta_j^l = \frac{\partial C}{\partial z_j^l}$$

We can rewrite this using the chain rule as:

$$\delta_j^l = \sum_k \frac{\partial C}{\partial a_{jk}^l} \frac{\partial a_{jk}^l}{\partial z_j^l}$$

However, the second term is always zero except when  $j = k$ , so the summation can be dropped. Remembering eq. 4.2, we note that  $\partial a_{jk}^l / \partial z_j^l = \sigma'(z_j^l)$ . For the last layer  $l = L$ , the first term turns into the derivative of the cost function, finally giving us:

$$\delta_j^L = |a_j^L - y_j| \sigma'(z_j^L) \quad (4.4)$$

Equation eq. 4.4 relates the error in the output layer to its inputs. This in turn is a function of all the previous inputs and errors and we thus need to find an expression relating the error in layer  $l$  with the error in layer  $l+1$ . Since we have an expression for the error in the last layer, we propagate the error going down the layers, hence the name *backpropagation*. Again using the chain rule gives:

$$\delta_j^l = \sum_k \frac{\partial C}{\partial z_{jk}^{l+1}} \frac{\partial z_{jk}^{l+1}}{\partial z_j^l} = \sum_k \delta_k^{l+1} \frac{\partial z_{jk}^{l+1}}{\partial z_j^l}$$

Using equation eq. 4.1, we obtain after substitution:

$$\delta_j^l = \sum_k \delta_k^{l+1} w_{kj}^{l+1} \sigma'(z_j^l) \quad (4.5)$$

Using equations eq. 4.4 and eq. 4.5, we can calculate the error in  $C$  due to each neuron. Finally, we need to relate the error in each error to  $\partial C / \partial w_{jk}^l$  and  $\partial C / \partial b_j^l$ . Making use yet again gives us the last two backpropagation relations:

$$\frac{\partial C}{\partial b_j^l} \frac{\partial b_j^l}{\partial z_j^l} = \frac{\partial C}{\partial b_j^l} = \delta_j^l \quad (4.6)$$

$$\sum_k \frac{\partial C}{\partial w_{jk}^l} \frac{\partial w_{jk}^l}{\partial z_j^l} = \delta_j^l \rightarrow \frac{\partial C}{\partial w_{jk}^l} = a_j^{l-1} \delta_j^l \quad (4.7)$$

Now that we now that all back propagation equations, we state the algorithm. It consists of four steps:

- i. Complete a forward pass, i.e., calculate the expected outcomes with the current weights and biases.

2. Calculate the error using eq. 4.4 and do a backward pass to obtain the error in each neuron using eq. 4.5. This can be used to calculate the gradients using eq. 4.6 and eq. 4.7
3. Adjust the weights and biases using the chosen optimizer (e.g. gradient descent)
4. Return to step 1 until the optimization problem converges.

Mathematically, back propagation is a special case of a technique known as automatic differentiation. Automatic differentiation is a third type of differentiation, next to numeric and symbolic. It allows for machine precision calculation of derivatives by writing it as a chain of simple operations combined with the chain rule, similar to backpropagation. Note that:

$$\delta_j^\circ = \frac{\partial C}{\partial x_j} \frac{\partial x_j}{\partial z_j^\circ}$$

so that:

$$\frac{\partial C}{\partial x_j} = a_j^\circ \delta_j^\circ$$

Thus neural networks also give us access to high precision derivatives with regard to each coordinate.

## 4.2 PHYSICS INFORMED NEURAL NETWORKS

On the face of things, the goal of physics and neural networks are opposite: whereas physics tries to build an understanding of things using models to make predictions, neural networks learn a *modelless* mapping to make predictions. Recent advancements however have merged the two approaches together in a concept known as Physics Informed Neural Networks (45, 26). In this approach, we encode physical laws into the network, so that the network respects the physics. This can be used to both numerically solve equations or fit a model to spatiotemporal data. Even more so, it should allow us to infer coefficient fields.

### 4.2.1 THE CONCEPT

Consider a set of  $1D+1$  spatiotemporal data, consisting of some property  $u(x, t)$  at coordinates  $(x, t)$ . The neural network can be learned the underlying physics by minimizing the cost function:

$$\mathcal{L} = \frac{1}{2n} \sum_i |u_i - d_i^L|^2$$

The process of learning requires a lot of data and is prone to overfitting. Now assume that we know that  $u(x, t)$  is governed by some process which is written as a partial differential equation:

$$\partial_t u = f(1, u, u_x, u_{xx}, u^2, \dots)$$

where  $f$  is a function of  $u$  or its spatial derivatives. Rewriting it as:

$$g = o = -\partial_t u + f(1, u, u_x, u_{xx}, u^2, \dots) \quad (4.8)$$

we see that in order to satisfy the PDE,  $g \rightarrow o$ . The idea of PINNs is to add this function  $g$  to the costfunction of the neural network:

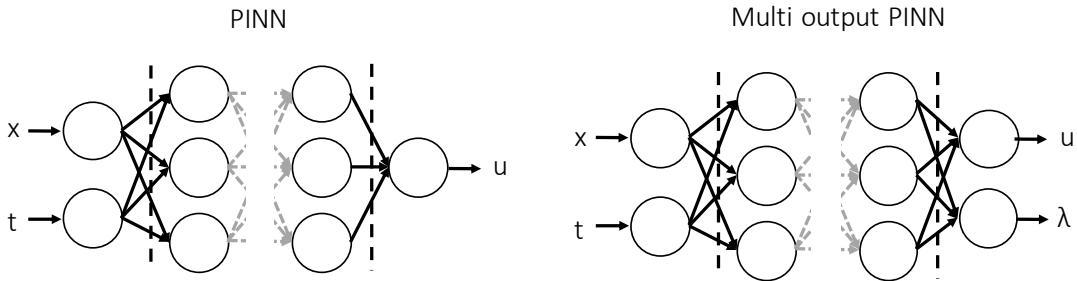
$$\mathcal{L} = \frac{1}{2n} \sum_i |u_i - d_i^L|^2 + \lambda \sum_i |g_i|^2 = MSE + \lambda PI$$

where  $\lambda$  sets the effective strength of the two terms. Observe that the cost function is higher if the PDE is not satisfied. Minimizing the costfunction will thus mean minimizing  $g$  and hence satisfying the PDE. We effectively penalize solutions not satisfying the physics we put in equation eq. 4.8; the added term acts a ‘physics-regularizer’. Concretely, the adding of physics constrains the solution space, preventing overfitting and making the neural network much more data efficient. The most useful feature however is that we don’t need a vast set of training data to train the network, as we solve the problem *by* training the network.

We can also remove the mean squared error term from the cost function and add initial and boundary conditions, similar to the PI term. If we now train the network, it will learn the solution to the given PDE whilst respecting the given boundary and initial conditions. This alternative means of numerically solving a model doesn’t need advanced meshing of the problem domain or carefully constructed (unstable) discretization schemes, as it requires the physics to be fulfilled at every point in the spatiotemporal domain. A useful analogy here is calculating the trajectory of a launched object. A classical numerical solver would take small steps in time, updating the position and speed of the object each step. A PINN however uses a completely different approach. Given the initial (random) state of the neural network, it calculates a first trajectory and keeps adjusting the weights of the network until the cost is minimized, i.e. until we obtain a solution satisfying the included physics

and initial and boundary conditions. A classical numerical approach tries once using a correct and methodical approach, whereas a PINN tries many times until the result satisfies its constraints.

We can also use this framework to fit models to spatiotemporal data by letting the coefficient of each term be a variable to be minimized as well. More concretely, where before the cost was a function of the weights and biases,  $\mathcal{L} = f(w^l, b^l)$ , we now let it be a function of the coefficients  $\lambda$  of the PDE as well:  $\mathcal{L} = f(w^l, b^l, \lambda_1, \lambda_2, \dots)$ . This is shown for several PDEs such as the Burgers, Schrodinger or Navier-Stokes equation in the papers of M. Raissi(45, 26). In the case of the Navier-Stokes equation, it's shown that it's also possible to infer the pressure field, which appears as a separate term. This is achieved by adding another output neuron to the PINN (shown in figure fig. 4.2), so that it predicts both the pressure and the flow. In theory it should also be possible to infer spatially and temporally varying *coefficient* fields. We investigate this claim in the next section.



**Figure 4.2:** Left panel: a normal single output PINN. Right panel: a multi-output PINN. The network now also predicts the coefficients values at each data point.

#### 4.2.2 PINNs IN PRACTICE

We now wish to evaluate the use of PINNs to analyze the RUSH data. Using a diffusive process as a toy problem, we first show how PINNs are able to accurately determine the diffusion constant, even in the presence of noise. Next, we prove that PINNs are indeed capable of inferring coefficient fields and finish by analyzing some parts of the RUSH data.

In our toy problem we have an initial concentration profile:

$$c(x, 0) = e^{-\frac{(x-0.5)^2}{0.02}}$$

diffusing in a 1D box according to:

$$\frac{\partial c(x, t)}{\partial t} = \nabla \cdot [D(x) \nabla c(x, t)]$$

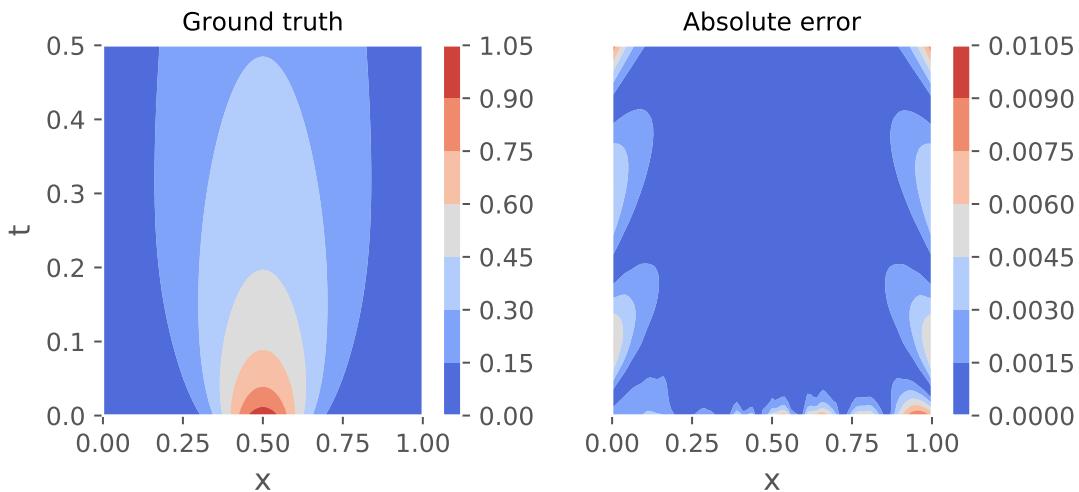
on the spatial domain  $[0, 1]$  with perfectly absorbing boundaries at the edges of the domain:

$$c(0, t) = c(1, t) = 0$$

If  $D(x) = D$ , this problem has an analytical solution through a Greens function. If the diffusion coefficient is spatially dependent though, the problem needs to be solved numerically. The code used to generate our data can be found in the appendix. Although this toy problem is simple and in 1D, our results easily generalize to higher dimensions and complexity at the cost of higher computational cost.

#### 4.2.2.1 CONSTANT DIFFUSION COEFFICIENT

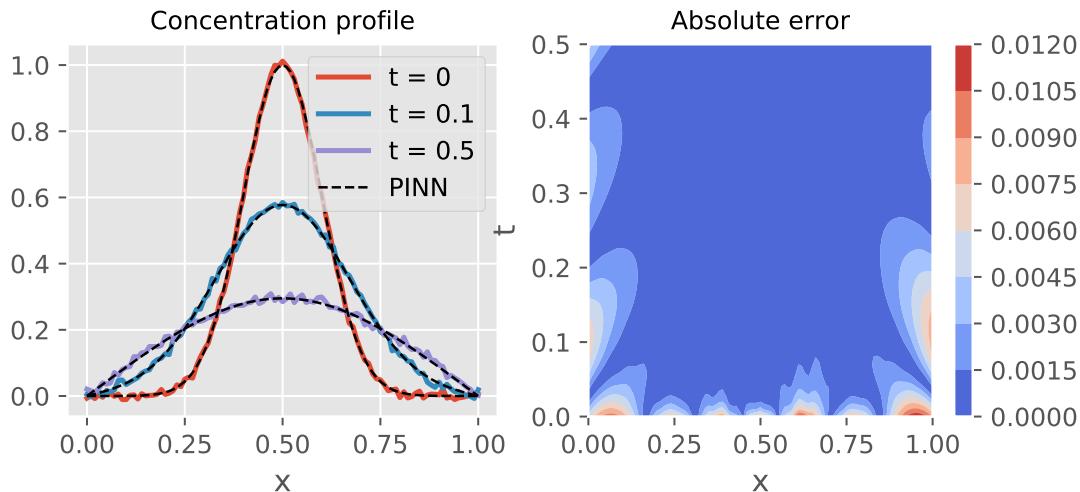
We now consider the mentioned problem with a diffusion coefficient of  $D(x) = D_0 = 0.1$  and simulate it between  $t = 0$  and  $t = 0.5$ . Using a spatial and temporal resolution of  $0.01$ , we have a datagrid of  $101$  by  $51$ , so that our total dataset consists of  $5151$  samples. The neural network consists of 6 hidden layers of  $20$  neurons each and  $\lambda = 1$ . Figure fig. 4.3 shows the ground truth for the problem and the absolute error of the neural network.



**Figure 4.3:** Left panel: Simulated ground truth of the problem. Right panel: The absolute error of neural network. Note that most of the error is located at areas with low concentration, i.e. signal.

The predicted diffusion coefficient is  $D_{pred} = 0.100026$ , giving an error of 0.026%. In [45], the authors obtain similar accuracies for significantly more complex problems such as the Schrodinger equation, which means that our accurate inference is not just due to the simplicity of the problem. Furthermore, Raissi et al. show that the result is robust w.r.t the architecture of the network. From the absolute error we observe that the error seems to be higher in areas with low concentration. This is a feature we've consistently observed: in areas with low ‘signal’, the neural network seems to struggle.

As good as these results are, the input data is noiseless and thus of limited practical interest. We now show that PINNs perform equally well with noisy data by adding 5% white noise to the data and performing the same procedure. The network is now doing two tasks in parallel: it’s both denoising the data and performing a model fit. In the left panel of figure fig. 4.4 we show the concentration profile at times  $t = 0, 0.1$  and  $0.5$ , with the prediction of the PINN superimposed in black dashed lines at each time. On the right panel we show again the absolute error from the ground truth. Observe the similarities with the noiseless case: most of the error localizes in areas with low concentration.

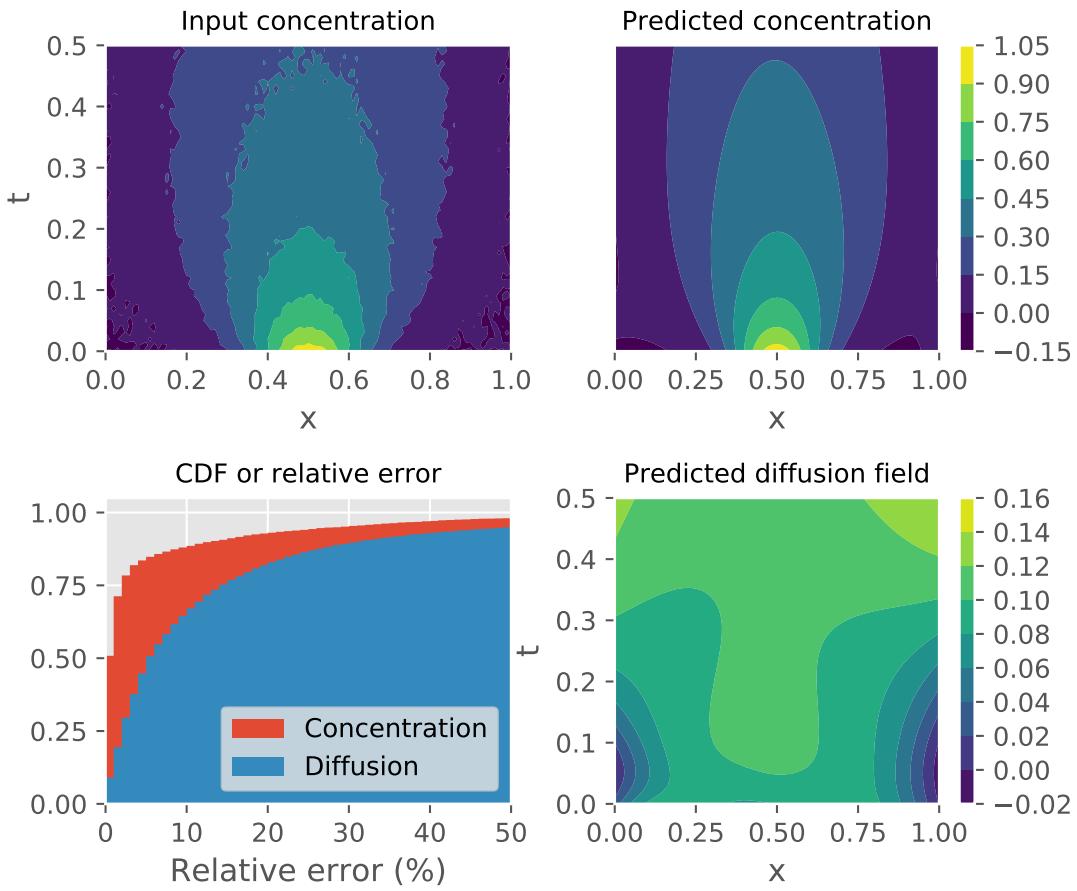


**Figure 4.4:** Left panel: The original noisy concentration profile with the neural network inferred denoised version super imposed. Right panel: The absolute error of neural network with respect to the ground truth. Note that most of the error is located at areas with low concentration.

The inferred diffusion constant is  $D_0 = 0.10052$ , giving an error of 0.52%. Although the error is slightly higher than in the noiseless version, it’s extremely impressive that we obtain the diffusion constant to this precision.

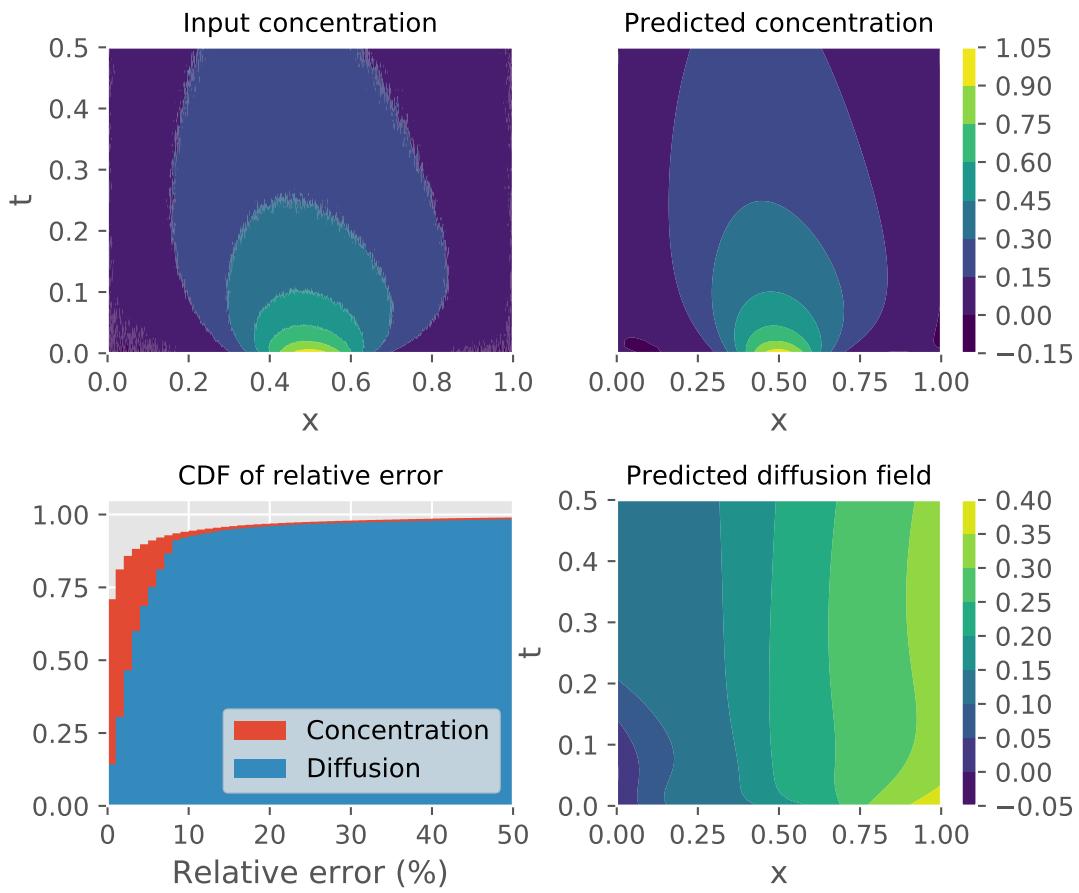
#### 4.2.2.2 VARYING D

As stated, it should be possible to infer coefficient fields by using a two output neural network. One output predicts the concentration while the other predicts the diffusion coefficient. Such a network is indeed capable of generating the right coefficient field as shown in figure fig. 4.5 . Here the network has been trained on the constant diffusion coefficient data we used before including 5% white noise, so that we should observe a diffusion field constant at  $D(x, t) = D_0 = 0.1$ . In the upper left we show the data on which the network is trained, with the upper right panel the predicted concentration profile, which shows a very good match. In the lower right panel we show the inferred diffusion field. We observe a good match in the middle of the plot, but the neural network again struggles in areas with low concentration, such as the lower left and right area. A more quantitative analysis of the predicted diffusion and concentration is presented in the lower left corner. Here we plot the Cumulative Distribution Function (CDF) of the absolute relative error. Note that the PINN predicts the concentration very well, but struggles more with the diffusion coefficient. This is expected, as the mean squared error of the cost function is quite explicit in its use of the concentration, whereas the diffusion coefficient is determined self-consistently in the PI part. We also observed similar but distinctive results in different runs, owing to the non-convexity of the problem. Overall the result is still remarkable, given that we've inferred a diffusion field from just concentration data with 5% noise.

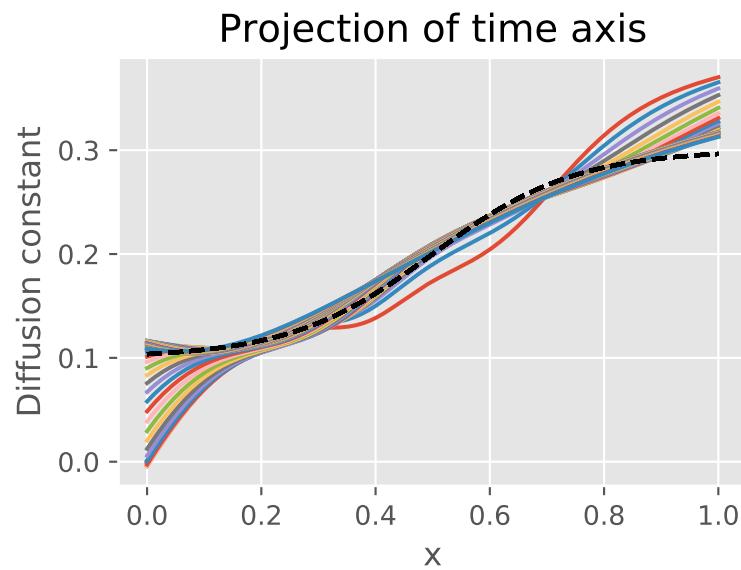


**Figure 4.5:** We show the training data and predicted concentration profile in the upper left and right panels. The lower right panel shows the inferred diffusion field while the lower left panel shows the CDF of the relative error of the diffusion and concentration.

In figure fig. 4.6 we show a similar analysis for a non-constant diffusion field. We've simulated data with a diffusion in the form of a  $\tanh(x)$ . Remarkably, the neural network is able to infer the field with reasonable accuracy, although it required a more sizeable dataset of 50000 points. Figure fig. 4.7 studies the inferred diffusion profiles in depth by projecting them along the time axis. Here we observe the strongest dissonance close to the edges, which, again, is where the concentration is low. Nonetheless, we've proven that a neural network is able to infer a coefficient field with reasonable accuracy from noisy data!



**Figure 4.6:** We show the training data and predicted concentration profile in the upper left and right panels. The lower right panel shows the inferred diffusion field while the lower left panel shows the CDF of the relative error of the diffusion and concentration.



**Figure 4.7:** Projection of the inferred diffusion profile along the time axis.

#### 4.2.2.3 REAL CELL

### 4.3 CONCLUSION

#### 4.3.1 WEAK POINTS AND HOW TO IMPROVE

# 5

## Golgi as a phase separated droplet

In this second part of the thesis we develop a model linking the Golgi function and size to the properties of the intracellular transport. We start with a general section on phase separation, followed by a section where we introduce an approximation. This approximation, known as the effective-droplet approximation, makes phase separation analytically tractable. We then introduce our model and its biological justification. The results of our model will be presented in the next section.

### 5.1 PHASE SEPARATION

Consider a mixture of two molecules, type A and B, with underling interaction strengths  $\chi_{ij}$ . Depending on the strength and sign of these interactions, the system is either in a mixed state with a constant concentration of A and B, or in a phase-separated state. Landau showed that instead of a complete statistical description, phase separation could be modeled by a system with a double well free energy function<sup>46</sup>. We can define an *order parameter*  $c = N_A/N_B$  which describes the state of the system and define a free energy density function  $f(c)$  with minima at  $c_o^-$  and  $c_o^+$ :

$$f(c) = \frac{b}{2(\Delta c)^2} (c - c_o^-)^2 (c - c_o^+)^2$$

where  $b$  characterizes the strength of molecular interactions and  $\Delta c = |c_o^- - c_o^+|$ . Once the system phase separates, the system will have two areas of concentration  $c_o^+$  and  $c_o^-$  with

a boundary inbetween. Associated with this boundary is a surface tension, so that the full free energy of the system becomes

$$F(c) = \int dV(f(c) + \frac{1}{2}\kappa(\nabla c)^2)$$

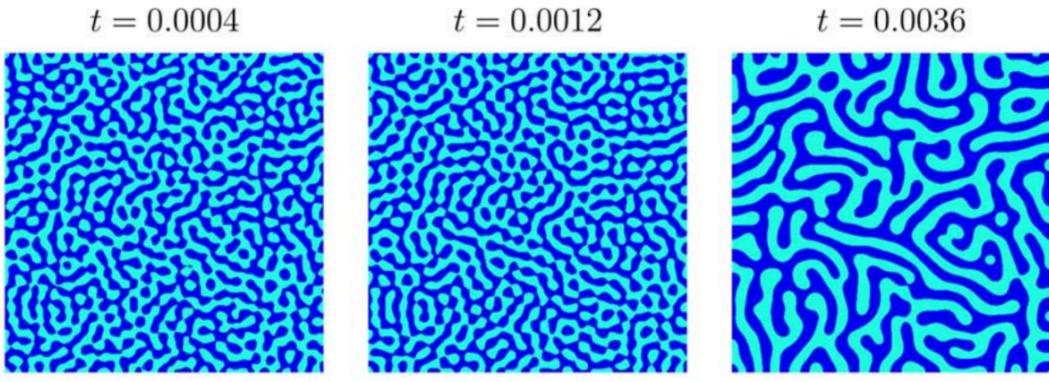
To find the equilibrium concentration profile, we minimize this free energy:

$$\frac{\delta F}{\delta c} = f'(c) - k\nabla^2 c = \mu(x) = 0 \quad (5.1)$$

where  $\delta F/\delta c$  is a functional derivative, as we minimize with respect to the concentration *profile*. Solving such an equation is generally not possible due to the third order terms of the free energy density function, but in 1D equation eq. 5.1 has been solved to yield:

$$c(x) = \frac{c_o^- + c_o^+}{2} + \frac{c_o^+ - c_o^-}{2} \tanh\left(\frac{x}{w}\right) \quad (5.2)$$

where  $w = \sqrt{\kappa/b}$  is the width of the boundary. When we quench a system from a mixed state into a phase separated state (a process known as spinodal decomposition), the actual concentration profile is a far cry from equation eq. 5.2. Inside the system, maze-like domains form and keep growing until a single dense and dilute phase are left. This is shown in figure fig. 5.1.



**Figure 5.1:** Cahn-Hilliard domains

In this process, the dynamics need to be taken explicitly into account. In the case of liquid-liquid phase separation, the order parameter  $c$  is conserved, as a molecule of type A cannot

change into type B. This means that the order parameter can only exchange locally, so that:

$$\partial_t c = -\nabla \cdot j$$

where  $j$  is a flux. We can relate the flux to the chemical potential:

$$j = -m\nabla\mu$$

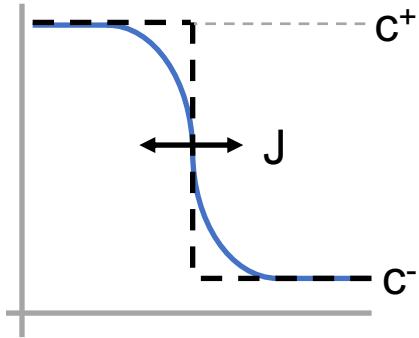
where  $m$  is a coefficient characterizing the mobility. Equation eq. 5.1 also gives us an expression for the chemical potential, so that we finally obtain the *Cahn-Hilliard equation*:

$$\frac{\partial c}{\partial t} = m\nabla^2[f(c) - k\nabla^2c]$$

It is this equation which governs the behaviour observed in figure fig. 5.1. Due to its non-linearity and fourth order derivatives simply solving the Cahn-Hilliard is usually forsaken in favour of deriving a scaling relation, which relates the domain growth speed  $dR/dt$  to the domain size  $R$  or some other system parameters. Another option is to study the system in the so-called effective droplet approximation, as we do in the next section.

## 5.2 EFFECTIVE DROPLET

Consider again equation eq. 5.2. If  $w \ll 1$ , we can approximate the system by describing it as two bulk phases, separated by an interface. If we apply correct boundary conditions at the interface, we can essentially split the system into two separate problems for the dense and dilute phase and match them at the interface. We thus model the phase separated system as a droplet which exchanges material with its environment through its interface, as shown in figure fig. ??



**Figure 5.2:** Model of an effective droplet. Blue line is full Cahn-Hilliard, black dashed line effective droplet.

By assuming the interface to be at thermodynamic equilibrium, the growth of the droplet is described in terms of the fluxes across the interface. Consider again equation eq. 5.1. As we can neglect the interfacial term, we have  $\mu = f(c)$ . In each bulk phase, we linearize the chemical potential around its equilibrium density ( $c_o^-$  or  $c_o^+$ ) to yield:

$$\frac{\partial c}{\partial t} = D \nabla^2 c \quad (5.3)$$

where we've absorbed all the coefficients into a single coefficient  $D$ . Observe that we have obtained a diffusion equation. Since we now have a linear equation in  $c$ , we can analytically solve this. Moreover, it is possible to add additional effects such as decay, production or advection and still keep an analytically solvable model, as we show in the next section and chapter.

By introducing the interface and breaking the problem into two separate parts, we need two new boundary conditions to solve equations such as eq. 5.3. By assuming the interface to be at thermodynamic equilibrium, we will derive a set of boundary conditions independent of the position, size or kinetic parameters of the droplet. Consider such a phase-separated system with an infinitely thin interface. The total free energy of the system can then be written as:

$$F = V_1 f(\varphi_1) + V_2 f(\varphi_2)$$

where  $V_i$  and  $\varphi_i$  are respectively the volume and concentration of phase  $i$  and  $f(\varphi_i)$  is the free energy density. Assuming incompressibility ( $V_1 + V_2 = V$ ) and conservation of particles ( $V_1 \varphi_1 + V_2 \varphi_2 = V \varphi$ ) constrains the system to two free variables, so that minimizing the free energy with respect to  $\varphi_1$  and  $V_1$  gives us two conditions:

$$f(\varphi_1) = f(\varphi_2)$$

$$0 = f(\varphi_1) + f(\varphi_2) + (\varphi_2 - \varphi_1)f'(\varphi_2)$$

Since  $f(\varphi) = \mu(\varphi)$ , the first condition states that both phases must have the same chemical potential, while the second one states that both phases must have equal pressure. The obvious solution to these equations is a completely mixed state with  $\varphi_1 = \varphi_2$ . A non-trivial phase-separated solution exists as well, where  $\varphi_1$  and  $\varphi_2$  are the two minima of the free energy density function  $f(\varphi)$ . Note that this is valid for droplets in 1D. In higher dimensions, the curvature of the droplet will affect the boundary conditions due to the Laplace pressure, but one can show that this leads to an extra term which scales with  $1/R^{47}$ .

Having defined boundary conditions, equations such as eq. 5.3 can be solved. Although one could solve these equations fully time-dependent using Green's functions, we assume a quasi-steady state so that  $dc/dt = 0$ . This will give us the concentration profiles in and outside the droplet, droplet growth however is determined by the fluxes across the interface. We show this in the next section.

### 5.2.1 FLUXES, ACTIVITY AND INTERFACES

Given a concentration profile  $c(x)$ , a diffusive flux can be calculated by applying Ficks' second law:

$$J(x) = -D \frac{\partial c}{\partial x}$$

Using this expression, the flux at the interface on the inside and outside of the droplet,  $J_{in}$  and  $J_{out}$ , can be calculated. Note that in and out respectively refer to inside and outside of the droplet rather than the direction of the flux; our boundary conditions fix the concentration at the interface but not the (direction of the) fluxes. If these fluxes are not balanced, there exists a net flux across the interface which leads to either growth or decay of the droplet. As the droplet changes size, the interface moves with a speed  $v_n$ . We derive an expression for  $v_n$  in terms of the fluxes across it. To move the interface a distance  $\Delta x$ , a net material gain of  $\Delta x \Delta c$  is needed. This net gain is given by the net flux in a time  $\Delta t$ , so that:

$$\Delta x \Delta c = (J_{in} - J_{out}) \Delta t$$

which can be rewritten as:

$$\frac{\Delta x}{\Delta t} = v_n = \frac{J_{in} - J_{out}}{\Delta c} \quad (5.4)$$

In the passive case (and assuming quasi-steady state), the concentration both inside and outside the droplet would be described by a solution to laplace's equation (i.e.  $\nabla^2 c = 0$ ), leading to a flat concentration profile  $c_{out}(x) = c_o^-$ ,  $c_{in}(x) = c_o^+$  and hence  $J = 0$  everywhere; the system is at thermodynamical equilibrium with  $\mu = 0$  and the droplet doesn't change size as  $v_n = 0$ . Placing such a droplet in a supersaturated environment where  $c_{out}(x) > c_o^-$  would lead to a lead to a non-zero  $J_{out}$ , resulting in the droplet growing to infinity by a process known as Ostwald ripening. As we show in the next section, active droplet suppress the Ostwald ripening<sup>48</sup>, leading to a droplet with a finite radius.

We now make the droplet *active* by adding a chemical reaction in the droplet which decays the droplet material A into some other other material B. Assuming material B diffuses very fast and is thus always in equilibrium, we ignore material B and describe solely A. Adding a decay term to equation eq. 5.3 gives:

$$D\nabla^2 c - kc = 0 \quad (5.5)$$

where  $k$  is a decay constant. Solving equation eq. 5.5 will always give a convex solution and hence a finite  $J_{in}$ . A typical concentration profile for an active droplet is shown in figure fig. 5.3.

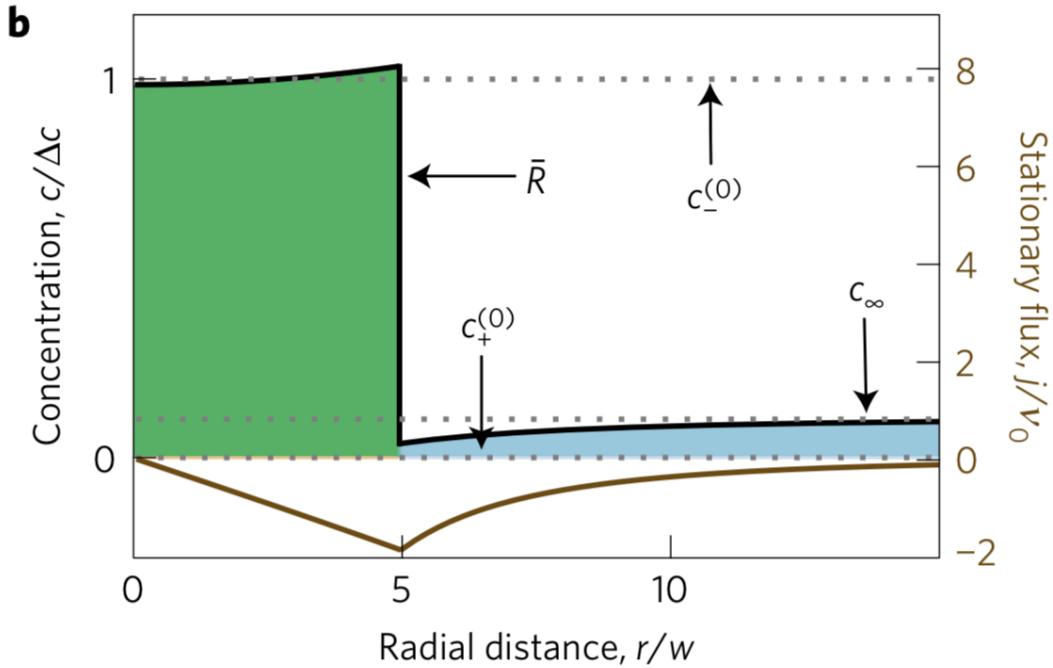


Figure 5.3: Typical concentration profile of active droplet. Taken from 49

The ‘decay flux’  $J_{in}$  will need to be balanced by a flux into the droplet  $J_{out}$  for a stable droplet to exist. This is a key property of active systems: while the system is at steady state (i.e.  $v_n = 0$ ), it is not at thermodynamical equilibrium, as  $\mu \neq 0$  and  $J \neq 0$ . Another fascinating property of active droplets is that they can propel themselves. To see this, consider a droplet of radius  $R$  at position  $x_o$  with two interfaces moving respectively at  $v_l$  and  $v_r$ . In a time  $dt$ , the droplet moves to a new position  $x_o + dx$  and will have a new radius  $R + dR$ :

$$x_o - R + v_l dt = x_o + dx - (R + dR)$$

$$x_o + R + v_r dt = x_o + dx + (R + dR)$$

Solving this set of equations for  $dx$  and  $dR$  gives:

$$\frac{dR}{dt} = \frac{1}{2}(v_r - v_l) \quad (5.6)$$

$$\frac{dx_o}{dt} = \frac{1}{2}(v_l + v_r) \quad (5.7)$$

Combining these equations with equation eq. 5.4 finally relates the growth and movement of the droplet to the fluxes across the interface:

$$\frac{dR}{dt} = \frac{I}{2\Delta c} [(J_{in}^{x=R} - J_{in}^{x=-R}) + (J_{out}^{x=-R} - J_{out}^{x=R})]$$

$$\frac{dx_o}{dt} = \frac{I}{2\Delta c} [(J_{in}^{x=-R} + J_{in}^{x=R}) - (J_{out}^{x=-R} + J_{out}^{x=R})]$$

Studying equations eq. 5.6 and eq. 5.7 shows that movement only happens if the fluxes are asymmetrical; the droplet center is displaced because one side of the droplet grows faster than the other. This imbalance is caused by a concentration gradient and thus *droplets will move up the gradient*, as the flux on the high concentration side will be higher than on the low side.

These equations allow us to find steady states both with respect to the size and the position of the droplet. In the next section we adapt the equations to model the Golgi and couple it to the intracellular transport through the calculations of the fluxes  $J_{out}$ . We do so in the next section.

### 5.3 GOLGI AS AN ACTIVE DROPLET

In the introduction we justified using a phase-separation approach to describe the Golgi. In this section we develop our model for the Golgi from biological considerations, but having established the mathematical background of phase separation, we parallelly present the mathematical description.

We can recognize four different populations in our system: immature cargo -heading to the Golgi-, mature cargo, -originating from the Golgi and which is produced from immature cargo in the golgi-, the Golgi itself and the cytoplasm, which acts as the solute. We start by reducing this set of populations to a system described by a single concentration  $c$ . Assuming maturation from the Golgi as a oneway process, i.e. immature cargo turns into mature cargo but not otherwise, and no interaction between the mature and immature cargo during intracellular transport, we can ignore the mature cargo. Modeling the solvent implicitly, the immature cargo in the cytoplasm is then represented by a dilute phase in some concentration  $c$ , while the Golgi is described by a dense phase in the same concentration.

Upon adding the drug nocadazole to mammalian cells, the microtubules are depolymerized and the Golgi ribbon breaks up into separate stacks<sup>50</sup>. These stacks are fully functional<sup>7</sup> and move away from their perinuclear location to collocate with an ERES. If we model not

the complete Golgi but a single stack, we can reduce our problem to 1D, where a droplet can move from one side of the system, representing the Golgi ribbon, to the other side, representing the ERES. As each stack is fully functional, we make no simplifications with respect to the function of the Golgi. Although many complex models of the maturation exist, we model it as a simple decay-like term:

$$\begin{aligned}\frac{\partial c}{\partial t} &= -\nabla J - kc \\ J &= -D\nabla c\end{aligned}$$

where  $k$  is a maturation constant. We now turn our attention to the intracellular transport. In our model fitting chapter we presented an argument that we could model the intracellular transport as an advection-diffusion equation. Evidence exists of vesicles refusing with the ER<sup>8</sup>, so we add an additional decay term, so that the concentration outside the droplet is described is described by:

$$D\partial_x^2 c(x) - v\partial_x c(x) - \alpha c(x) = 0 \quad (5.8)$$

with  $v$  an advection velocity and  $\alpha$  some decay constant. Solving this equation will lead to a concentration profile with a gradient and we thus model our golgi as an active droplet growing in a concentration gradient, inspired by 51. A study of the biogenesis of the golgi<sup>8</sup> shows that stacks are transported to the ribbon over the microtubules, so we thus add the advection also to the dense phase:

$$D\partial_x^2 c(x) - v\partial_x c(x) - kc(x) = 0 \quad (5.9)$$

Note that the dense and dilute phase description are thus almost similar, save for a different decay (maturation) constant. One could pick different diffusion constants and advection speeds, but for simplicity and without loss of generality we pick the same. As our free energy function has minima at  $c_o^+$  and  $c_o^-$ , our boundary conditions at the interface are:

$$c(x_o \pm R) = \begin{cases} c_o^+, & \text{inside} \\ c_o^-, & \text{outside} \end{cases}$$

As stated, we model our system in 1D, with one boundary representing the ERES and the other boundary as the location of Golgi Ribbon. We place the ERES on the left side of the system and thus model this boundary as source:

$$(-D\partial_x c + vc)|_{x=o} = J_{in}$$

whereas the right boundary is merely the edge of the system and we thus model it as a zero-flux boundary:

$$(-D\partial_x c + vc)|_{x=L} = 0$$

We solve this set of equations in the next chapter.

# 6

## Golgi model

In the previous chapter we introduced the Cahn-Hilliard equation. Its non-linearity makes it extremely hard to study and in this chapter we present an approximation known as effective droplet theory. In this chapter we introduce the effective droplet approximation, discuss our models for the Golgi apparatus and show the results for these models. We've divided the chapter into three sections:

- Effective droplet theory - In this section we introduce effective droplet theory.
- Golgi as an effective droplet- Here we introduce our model and show
- Two-component model- This third section contains a possible extension of the effective droplet model to two-components.

We end the chapter with a short summary of our conclusions.

### 6.0.1 SOLUTION

The general solution of equations eq. 5.8 and eq. 5.9 is given by:

$$c(x) = C_1 e^{-\frac{x}{l^-}} + C_2 e^{\frac{x}{l^+}}$$

where we have defined  $l^\pm$  as:

$$l^\pm = \frac{2D}{\sqrt{4kD + v^2} \pm v} \quad (6.1)$$

where  $k$  should be replaced by  $a$  outside the droplet. Note that  $l^\pm$  defines the lengthscale of the problem. Without advection it simplifies to  $\sqrt{D/k}$  and both lengthscales becomes similar. It is here we see that the advection combined with decay leads to some sort of ‘symmetry-breaking’ in the droplet and that the effect of advection is more than just translating the droplet. We study this is in the next section. Applying the boundary conditions and calculating the flux at position  $x_o + R$  and  $x_o - R$  gives the following fluxes outside:

$$J_{out}^{x=-R} = J_{in} \frac{(1 + \frac{l_-}{l_+}) e^{\frac{-(x_o - R)}{l_-}}}{Pe_- + Pe_+ \frac{l_-}{l_+} e^{\frac{-x_1}{l_-}}} + \frac{c_o^{out} D}{l_+} \frac{Pe_+ (1 - e^{\frac{-x_1}{l_-}})}{1 + \frac{l_-}{l_+} \frac{Pe_+}{Pe_-} e^{\frac{-x_1}{l_-}}} \quad (6.2)$$

$$J_{out}^{x=R} = -c_o^{out} D \frac{Pe_- Pe_+ (1 - e^{\frac{-x_2 + L}{l_-}})}{l_+ Pe_- + e^{\frac{-x_2 + L}{l_-}} l_- Pe_+} \quad (6.3)$$

where we have introduced the coordinates  $x_1$  and  $x_2$ , which are defined respectively as  $x_o \pm R$  and correspond to the left and right interface. We have also defined the Peclet-like numbers  $Pe^\pm$ :

$$Pe^\pm = 1 \mp \frac{vl^\pm}{D}$$

and a new combined lengthscale  $l = \frac{l^+ l^-}{l^+ + l^-} = 1/l^+ + 1/l^-$ . The flux on the left of the droplet consists of two terms, with the first one accounting for the influx and the second one for the interface with the droplet. The flux on the right is clearly similar, but lacks a second term since we’ve set a zero-flux boundary at  $x = L$ .

We now turn to the fluxes on the inside of the droplet. In this case however, the fluxes on the left and right side of the droplet are not of particular importance; considering equations eq. 5.6 and eq. 5.7, we’re interested in their sum and difference. Introducing  $J_{Rad} = J_{in}^{x=R} - J_{in}^{x=-R}$  and  $J_{Pos} = J_{in}^{x=R} + J_{in}^{x=-R}$ , we obtain:

$$J_{rad} = \frac{-2c_o^+ D \sinh \frac{R}{l_-} \sinh \frac{R}{l_+}}{l} \frac{\sinh \frac{R}{l_-} \sinh \frac{R}{l_+}}{\sinh \frac{R}{l}}$$

$$J_{pos} = 2c_o^{in} D \left[ \frac{Pe_-}{l_-} \frac{\sinh \frac{R}{l_+} \cosh \frac{R}{l_-}}{\sinh \frac{R}{l}} - \frac{Pe_+}{l_+} \frac{\sinh \frac{R}{l_-} \cosh \frac{R}{l_+}}{\sinh \frac{R}{l}} \right]$$

The equation for the radius can also be interpreted as the maturation flux. It contains an important point with respect to the validity of effective droplet theory. Consider the limit of  $R \rightarrow \infty$ :

$$\lim_{R \rightarrow \infty} = -2c_o^+ \sqrt{kD} \quad (6.4)$$

where for simplicity we have set the advection speed  $v$  to zero. For an active droplet one would expect the maturation flux to scale with the radius. However, equation eq. 6.4 shows that the maturation flux saturates. The answer to this conundrum is found by studying the concentration profile in the droplet. For  $R \gg l$ , the concentration in the center of the drop goes to zero, which is unphysical and an artifact of the effective droplet model. Thus effective droplet theory is only valid in the region  $R \ll l$ . When advection is added back into the mix this point is slightly more nuanced, as advection changes the concentration profile. We show this in the next section.

The equation for the positional flux hints at the effects of the advection and maturation. In a droplet without decay, the concentration profile would be flat and the positional flux would be proportional to  $c_o^+ v$ . Equation tells a different story. We investigate this in depth in the next section.

#### 6.0.2 NO DECAY IN THE DILUTE PHASE: THE EFFECT OF ADVECTION

We first study the effect of advection on the droplet by setting the decay in the dilute phase to zero. In this case the flux on the outside of the droplet becomes position-independent:

$$J_{in}^{x=-R} = J_{in}$$

$$J_{out}^{x=R} = 0$$

This is expected, as the cargo can only exit the system by decaying inside the droplet. The equations for the flux in the droplet remain unchanged because they are independent of the decay outside. Developing the internal droplet fluxes for  $R \ll l^\pm$  gives:

$$J_{rad} \approx -2c_o^+ kR$$

$$J_{pos} \approx$$

Combining these equations gives us expressions for the growth rate and speed of the droplet:

$$\frac{dR}{dt} =$$

$$\frac{dx_o}{dt} =$$

Armed with these equations we can study the effect of advection. In figure .. we show the minimum concentration,  $dx_o/dt$  and  $dR/dt$  as a function of  $R$  and  $v$ .

In the left panel we show the minimum concentration as a function of  $R$  and  $v$ . The white line shows the concentration corresponding to  $l = 0.5$  at  $v = 0$  and we indeed see that advection increases the minimum concentration. In the middle panel we show the increase in radius  $dR/dt$  as a function of again  $R$  and  $v$ . Red colors mean positive, blue negative and white zero, so we've plotted with x's the line  $dR/dt = 0$ , which corresponds to a stable droplet. On the right we show  $dx_o/dt$ ; where the  $dx_o/dt = 0$  line crosses with the  $dR/dt = 0$  line there's a droplet with stable radius and position.

To show this more clearly, we plot the droplet radius and minimum concentration as a function of the advection velocity in figure . Observe that advection decreases the droplet radius by almost 10%, while simultaneously also increasing the minimum concentration. The advection is thus ‘compacting’ the droplet. By what process is this happening? To see this, consider an active droplet without advection. The concentration profile is symmetric and convex; the flux on both sides due to diffusion is inwards and thus in opposite directions. If we now turn on advection, this symmetry is broken. On the left side of the droplet, the diffusion and advection point in the same direction but on the other side they are aligned opposite. This causes an asymmetry in the concentration and a difference in the flux to the centre: more goes to it from the left than is taken away by the right, hence compacting the droplet.

Can advection stabilize a droplet? Zwicker et al show that an active droplet such as ours can divide through a shape instability. Our model is 1D however and not easily scalable to higher D's because of the advection. Simulations would be needed to state anything on this.

We can make a few conclusions about the situation. For a very specific set of parameters, the droplet can both be stable in radius and position. There's no position dependence so not that interesting. Also observe that there are only two situations: either the droplet is moving to the left infinitely, or moving to the right infinitely, as there's no stable position on either side. Which situation happens is controlled by the gradient of the concentration

versus the advection. note that we need a non-zero convection to move the droplet from one mode to the other. The explanation is simple: the convection needs to overcome the gradient to reverse droplet movement. At the inflection point,  $dx_o/dt = 0$ .

We now make some very speculative biological connection to the observation of Golgi properties. A strong marker is that once the microtubules are depolymerized, the stacks move away from the perinuclear area and collocate with the ER. This is what we see when we turn off the advection, the droplet moves up the gradient to the source. A similar thing is seen with biogenesis, the stacks are made around the ER but need microtubules to be transported to the center.

We have a model which describes some features but its unstable. Can we make some position in our system inherently stable? We try in the next section by adding decay to the dilute phase.

### 6.0.3 DECAY IN THE DILUTE PHASE

We now wish to find a model in which the position has a stable position. So not just a set of parameters where  $dx_o/dt = 0$ , but a position  $x_o$  where the droplet is stable. The flux on the inside is independent of the position of the droplet, so to get a stable position we need a position dependent flux on the outside. To this end, we introduce a decay term in the dilute phase. We then end up with the full equations ... . We first numerically solve these equations before attempting an approximation. The result is shown in figure ... and shows a ‘phase-diagram’ of our active droplet model. In the left panel we show the minimum concentration in the droplet, while the middle and right panel shows the numerical value of  $dx_o/dt$  and  $dR/dt$ , all as a function of the CoM position  $x_o$  and radius  $R$ . From the right two plots we’ve extracted the lines  $dx_o/dt = 0$  and  $dR/dt = 0$ , which we’ve superimposed on all three panels. Where these lines cross a stable state exists. We first discuss each line independently and then discuss the stable points.

First consider the line  $dR/dt = 0$ , a droplet with stable radius. We observe that left of the line  $dR/dt > 0$ , with the opposite on the right. This means that each radius is stable: a perturbation w.r.t to the radius will not propagate. The  $dx_o/dt = 0$  is more interesting. It has the appearance of a ‘finger’ and has a stable and unstable branch. The arm below the point is stable, where the arm above the point is not. This has important implications for the crossings: the lower intersection is stable w.r.t. to all perturbations, while the upper intersection is unstable. Perturbations above the  $dx_o/dt = 0$  line will propagate, while perturbations below will move the droplet to the lower intersection.

This is all fine and dandy, but there’s one problem: the lower intersection, the stable point,

is unphysical. We characterize the droplet in terms of its center of mass  $x_o$  and radius  $R$ , meaning that the left interface is at a position  $x_o - R$ . The red line in the plots shows the line  $x_o = R$  and everything below this line is unphysical, as it means the left interface is at a position  $x_o - R <$ , past the system edge. Why does this happen? First, in a dynamical description this wouldn't happen, as the left interface boundary condition prevents moving past. In a static description this can however since we specify the edge of the system as a flux boundary condition. Inspection of the concentration (see figure ) learns that the droplet moves on top of the source: the flux inside the droplet at a point is similar to the boundary condition. In 2D this wouldnt be a problem but in our 1D description it is as the BC is specified as a flux and not as a source. That means we're left with just an unstable point. Why is this point unstable? Investigating the fluxes shows that despite the decay the left interface is an order of magnitude bigger than the right interface, thus behaviour is mainly determined by the left interface. Once the left interface starts moving due to a perturbation, the rest follows.

## 6.1 TWO-COMPONENT MODEL

Two components jeeej

## 6.2 CONCLUSION

# 7

## Conclusion

So what can we say about our project? We tried a lot but nothing worked ha-ha.

## Appendix I: Some extra stuff

Add appendix I here. Vivamus hendrerit rhoncus interdum. Sed ullamcorper et augue at porta. Suspendisse facilisis imperdiet urna, eu pellentesque purus suscipit in. Integer dignissim mattis ex aliquam blandit. Curabitur lobortis quam varius turpis ultrices egestas.

# References

1. Emr, S. *et al.* Journeys through the Golgi—taking stock in a new era. *The Journal of Cell Biology* 187, 449–453 (2009).
2. Tang, D. & Wang, Y. Cell cycle regulation of Golgi membrane dynamics. *Trends in Cell Biology* 23, 296–304 (2013).
3. Rothman, J. E. The Future of Golgi Research. *Molecular Biology of the Cell* 21, 3776–3780 (2010).
4. Gosavi, P. & Gleeson, P. A. The Function of the Golgi Ribbon Structure - An Enduring Mystery Unfolds! *BioEssays* 39, 1700063 (2017).
5. Budnik, A. & Stephens, D. J. ER exit sites - Localization and control of COPII vesicle formation. *FEBS Letters* 583, 3796–3803 (2009).
6. Bressloff, P. C. & Newby, J. M. Stochastic models of intracellular transport. *Reviews of Modern Physics* 85, 135–196 (2013).
7. Wei, J.-H. & Seemann, J. Unraveling the Golgi Ribbon. *Traffic* 11, 1391–1400 (2010).
8. Ronchi, P., Tischer, C., Acehan, D. & Pepperkok, R. Positive feedback between Golgi membranes, microtubules and ER exit sites directs de novo biogenesis of the Golgi. *Journal of Cell Science* 127, 4620–4633 (2014).
9. Newby, J. M. & Bressloff, P. C. Quasi-steady State Reduction of Molecular Motor-Based Models of Directed Intermittent Search. *Bulletin of Mathematical Biology* 72, 1840–1866 (2010).
10. Newby, J. & Bressloff, P. C. Random intermittent search and the tug-of-war model of motor-driven transport. *Journal of Statistical Mechanics: Theory and Experiment* 2010, Po4014 (2010).
11. Staehelin, L. A. & Kang, B.-H. Nanoscale Architecture of Endoplasmic Reticulum Export Sites and of Golgi Membranes as Determined by Electron Tomography. *PLANT PHYSIOLOGY* 147, 1454–1468 (2008).
12. Alberts. *Molecular biology of the Cell*.
13. Griffiths, H. Rubisco is said to be both the most important enzyme on Earth and surprisingly inefficient. Yet an understanding of the reaction by which it fixes CO<sub>2</sub> suggests that evolution has made the best of a bad job. 2
14. Glick, B. S. & Luini, A. Models for Golgi Traffic: A Critical Assessment. *Cold Spring Harbor Perspectives in Biology* 3, a005215–a005215 (2011).
15. Hirschberg, K. *et al.* Kinetic Analysis of Secretory Protein Traffic and Characterization of Golgi to Plasma Membrane Transport Intermediates in Living Cells. *The Journal of Cell Biology* 143, 1485–1503 (1998).
16. Boncompain, G. *et al.* Synchronization of secretory protein traffic in populations of cells. *Nature Methods* 9, 493–498 (2012).

17. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-Liquid Phase Separation in Biology. *Annual Review of Cell and Developmental Biology* 30, 39–58 (2014).
18. Zwicker, D., Decker, M., Jaensch, S., Hyman, A. A. & Jülicher, F. Centrosomes are autocatalytic droplets of pericentriolar material organized by centrioles. *Proceedings of the National Academy of Sciences* 111, E2636–E2645 (2014).
19. Sbalzarini, I. F. Seeing Is Believing: Quantifying Is Convincing: Computational Image Analysis in Biology. in *Focus on Bio-Image Informatics* (eds. De Vos, W. H., Munck, S. & Timmermans, J.-P.) 219, 1–39 (Springer International Publishing, 2016).
20. Chen, K.-C., Qiu, M., Kovacevic, J. & Yang, G. Computational Image Modeling for Characterization and Analysis of Intracellular Cargo Transport. in *Computational Modeling of Objects Presented in Images. Fundamentals, Methods, and Applications* (eds. Hutchison, D. et al.) 8641, 292–303 (Springer International Publishing, 2014).
21. Lee, H.-C. & Yang, G. An image-based computational method for characterizing whole-cell scale spatiotemporal dynamics of intracellular transport. in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)* 699–702 (IEEE, 2015). doi:[10.1109/ISBI.2015.7163969](https://doi.org/10.1109/ISBI.2015.7163969)
22. Yang, G. Bioimage informatics for understanding spatiotemporal dynamics of cellular processes: Spatiotemporal dynamics of cellular processes. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 5, 367–380 (2013).
23. Hebert, B., Costantino, S. & Wiseman, P. W. Spatiotemporal Image Correlation Spectroscopy (STICS) Theory, Verification, and Application to Protein Velocity Mapping in Living CHO Cells. *Biophysical Journal* 88, 3601–3614 (2005).
24. Kisley, L. et al. Characterization of Porous Materials by Fluorescence Correlation Spectroscopy Super-resolution Optical Fluctuation Imaging. *ACS Nano* 9, 9158–9166 (2015).
25. Semrau, S. & Schmidt, T. Particle Image Correlation Spectroscopy (PICS): Retrieving Nanometer-Scale Correlations from High-Density Single-Molecule Position Data. *Biophysical Journal* 92, 613–621 (2007).
26. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations. *arXiv:1711.10566 [cs, math, stat]* (2017).
27. Barron, J. L., Fleet, D. J. & Beauchemin, S. S. Performance of Optical Flow Techniques. 60 (1994).
28. Dong, G., Baskin, T. I. & Palaniappan, K. Motion Flow Estimation from Image Sequences with Applications to Biological Growth and Motility. in *2006 International Conference on Image Processing* 1245–1248 (IEEE, 2006). doi:[10.1109/ICIP.2006.312551](https://doi.org/10.1109/ICIP.2006.312551)
29. Vig, D. K., Hamby, A. E. & Wolgemuth, C. W. On the Quantification of Cellular Velocity Fields. *Biophysical Journal* 100, 1469–1475 (2016).
30. Garcia, D. Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics & Data Analysis* 54, 1167–1178 (2010).
31. Zimoń, M., Reese, J. & Emerson, D. A novel coupling of noise reduction algorithms for particle flow simulations. *Journal of Computational Physics* 321, 169–190 (2016).
32. Zimoń, M. et al. An evaluation of noise reduction algorithms for particle-based fluid simulations in multi-scale applications. *Journal of Computational Physics* 325, 380–394 (2016).

33. Grinberg, L., Yakhot, A. & Karniadakis, G. E. Analyzing Transient Turbulence in a Stenosed Carotid Artery by Proper Orthogonal Decomposition. *Annals of Biomedical Engineering* 37, 2200–2217 (2009).
34. Grinberg, L. Proper orthogonal decomposition of atomistic flow simulations. *Journal of Computational Physics* 231, 5542–5556 (2012).
35. Bruno, O. & Hoch, D. Numerical Differentiation of Approximated Functions with Limited Order-of-Accuracy Deterioration. *SIAM Journal on Numerical Analysis* 50, 1581–1603 (2012).
36. Knowles, I. & Renka, R. J. METHODS FOR NUMERICAL DIFFERENTIATION OF NOISY DATA. 12
37. Rizk, A. *et al.* Segmentation and quantification of subcellular structures in fluorescence microscopy images using Squash. *Nature Protocols* 9, 586–596 (2014).
38. Holcman, D. Modeling DNA and Virus Trafficking in the Cell Cytoplasm. *Journal of Statistical Physics* 127, 471–494 (2007).
39. Lagache, T. & Holcman, D. Effective Motion of a Virus Trafficking Inside a Biological Cell. *SIAM Journal on Applied Mathematics* 68, 1146–1167 (2008).
40. Dinh, A.-T., Theofanous, T. & Mitragotri, S. A Model for Intracellular Trafficking of Adenoviral Vectors. *Biophysical Journal* 89, 1574–1588 (2005).
41. Brandenburg, B. & Zhuang, X. Virus trafficking – learning from single-virus tracking. *Nature Reviews Microbiology* 5, 197–208 (2007).
42. Karpatne, A., Watkins, W., Read, J. & Kumar, V. Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling. *arXiv:1710.11431 [physics, stat]* (2017).
43. Sharma, R., Farimani, A. B., Gomes, J., Eastman, P. & Pande, V. Weakly-Supervised Deep Learning of Heat Transport via Physics Informed Loss. *arXiv:1807.11374 [cs, stat]* (2018).
44. Pun, G. P. P., Batra, R., Ramprasad, R. & Mishin, Y. Physically-informed artificial neural networks for atomistic modeling of materials. *arXiv:1808.01696 [cond-mat]* (2018).
45. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations. *arXiv:1711.10561 [cs, math, stat]* (2017).
46. Bray, A. J. Theory of phase-ordering kinetics. *Advances in Physics* 51, 481–587 (2002).
47. Weber, C. A., Zwicker, D., Jülicher, F. & Lee, C. F. Physics of Active Emulsions. *arXiv:1806.09552 [cond-mat]* (2018).
48. Zwicker, D., Hyman, A. A. & Jülicher, F. Suppression of Ostwald ripening in active emulsions. *Physical Review E* 92, (2015).
49. Zwicker, D., Seyboldt, R., Weber, C. A., Hyman, A. A. & Jülicher, F. Growth and division of active droplets provides a model for protocells. *Nature Physics* 13, 408–413 (2017).
50. Sengupta, D. & Linstedt, A. D. Control of Organelle Size: The Golgi Complex. *Annual Review of Cell and Developmental Biology* 27, 57–77 (2011).
51. Weber, C. A., Lee, C. F. & Jülicher, F. Droplet ripening in concentration gradients. *New Journal of Physics* 19, 053021 (2017).