Modeling Intracellular Transport during Messenger RNA Localization in *Xenopus* Oocytes

by

Maria-Veronica Ciocanel

B.S., Duke University, Durham, NC, 2012

M.Sc., Brown University, Providence, RI, 2013

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Division of Applied Mathematics at Brown University

PROVIDENCE, RHODE ISLAND

May 2017

Abstract of " Modeling Intracellular Transport during Messenger RNA Localization in *Xenopus* Oocytes " by Maria-Veronica Ciocanel, Ph.D., Brown University, May 2017

Many organisms need to establish spatial orientation during early development. In egg cells (oocytes) of the frog *Xenopus laevis*, spatial differentiation is achieved by localization of messenger RNA (mRNA), as these molecules move from the nucleus to the periphery of the egg cell during egg formation. Our goal is to understand how the long-term dynamics of mRNA molecules varies across the oocyte and how localization is regulated in space and time given parameters estimated using fluorescence recovery after photobleaching (FRAP) data. Although a large number of analytical and numerical models have been developed to extract binding and diffusion rates from FRAP recovery curves, active transport of molecules is typically not included in the existing models. We introduced a validated numerical method for estimating diffusion, binding/unbinding rates, and active transport velocities using FRAP data that captures intracellular dynamics through partial differential equation models. Given knowledge of these parameters, the effective velocity and diffusion of particles at large times are derived for linear and nonlinear PDE models of active transport using dynamical systems and stochastic methods. In combination with FRAP parameter estimates and predicted run times and lengths of particles, these asymptotic quantities quantify dynamical properties of localizing and non-localizing mRNA. Our results confirm the hypothesis of distinct transport dynamics in different regions of the egg cell and suggest that bidirectional transport of mRNA may influence the timescale of RNA localization in *Xenopus* oocytes. In addition, the parameter estimates inform numerical simulations of mRNA localization on model microtubule structures, which suggest that an anchoring mechanism at the cell periphery may be essential in reproducing localization patterns.

This dissertation by Maria-Veronica Ciocanel is accepted in its present form by the Division of Applied Mathematics as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date_____ _____

Björn Sandstede, Ph.D., Advisor

Recommended to the Graduate Council

Date_____ _____

Martin Maxey, Ph.D., Reader

Date_____ _____

Scott McKinley, Ph.D., Reader

Approved by the Graduate Council

Date_____ _____

Andrew Campbell, Dean of the Graduate School

Vitae

Maria-Veronica Ciocanel

b. February 4th, 1989

Ploiesti, Romania

Education

- Brown University, Providence, RI, 2011-2016

  Ph.D. in Applied Mathematics, expected May 2017

  M.Sc. in Applied Mathematics, 2013

- Duke University, Durham, NC, 2008-2012

  B.S. in Mathematics, Graduation with Distinction

  B.A. in French Studies

Honors and Awards

- 2016, AWM Poster Competition at SIAM Annual Meeting, Honorable Mention.
- 2016, SIAM Student Chapter Certificate of Recognition.

- 2016, SIAM Travel Award for Conference on Nonlinear Waves and Coherent Structures.
- 2016, AWM Travel Award for SIAM Conference on the Life Sciences.
- 2015, Reginald D. Archambault Award for Teaching Excellence in Summer Session Education.
- 2015, Red Sock Award (best poster presentation in dynamical systems) at SIAM Conference on Applications of Dynamical Systems, Snowbird, UT.
- 2015, NSF Travel Award for Opening Workshop of the Newton Institute program on Stochastic Dynamical Systems in Biology, Cambridge, UK.

Teaching Experience

- Teaching Assistant, GirlsGetMath ICERM: Summer Math Camp for High School Girls, ICERM, August 2016.
- Instructor, Methods of Applied Math II, Brown University, Summer 2015.
- Teaching assistant, Methods of Applied Math I, Brown University, Spring 2014.
- Teaching assistant, Statistical Inference I, Brown University, Fall 2013.
- Instructor, Beginning Math, English for Action, Providence, Fall 2014 – Fall 2016.

Service and Professional Activities

- Sheridan Center for Teaching and Learning

  - Sheridan Center Certificate II, Spring 2017
  - Sheridan Center Applied Math Departmental Liaison, 2014 – 2016
  - Sheridan Center Certificate III, 2015 – 2016
  - New TA orientation workshop leader, Fall 2014
  - Sheridan Center Certificate IV, 2014 – 2015
  - Sheridan Center Certificate I, 2013 – 2014

- Organizer of minisymposium "Modeling of Mechanisms of Intracellular Transport", SIAM Conference on the Life Sciences, Boston, MA. July 2016.
- Co-organizer of Workshop on Agent-Based Modeling, Division of Applied Mathematics at Brown University, Providence, RI. March 2015.
- Participant, MRC workshop on "Mathematics in Physiology and Medicine", Snowbird, UT. June 19-25, 2016.
- Participant, Mathematical Problems in Industry Workshop, Duke University. June 13-17, 2016.
- Participant, IMA workshop on " Mathematical Modeling in Industry Workshop for Graduate Students", Minneapolis, MN. August 5-14, 2015.
- Participant, NIMBioS Research Collaboration Workshop for Women in Mathematical Biology, University of Tennessee. June 22-25, 2015.
- Founder of the Brown Mathematical Contest for Modeling, Fall 2015.
- Other service: Treasurer and events organizer for Rose Whelan Society for Women in Math (2013 – 2016), Vice President of AWM Chapter (2013 – 2014), Vice President of SIAM Chapter (2015 – 2016).

Publications

- Analysis of Active Transport by Fluorescence Recovery after Photobleaching. MV Ciocanel, J Kreiling, J Gagnon, K Mowry, and B Sandstede. Accepted (Biophysical Journal), 2017.
- Modeling Autoregulation of the Afferent Arteriole of the Rat Kidney. MV Ciocanel, T Stepien, A Edwards and A Layton. In press (Research Collaborations in Mathematical Biology, Association for Women in Mathematics Springer Series), 2017.
- A Multicellular Model of the Renal Myogenic Response. MV Ciocanel, T Stepien, I Sgouralis, and A Layton. Submitted, 2016.
- Influenza spread on context-specific social networks. J Abrams, A Schwartz, MV Ciocanel, A Volkening, and B Sandstede. Submitted, 2016.
- Using *in vivo* imaging to measure RNA mobility in *Xenopus laevis* oocytes. E Powrie, MV Ciocanel, J Kreiling, J Gagnon, B Sandstede and K Mowry. Methods 98 (2016), pp. 60-65.

Presentations

- Contributed talk. WINRS New England Mathematical Meeting, Brown University, Providence, RI. March 4, 2017.
- Contributed talk. Joint Mathematics Meetings, AMS Session for Contributed Papers on Mathematical Biology, Atlanta, GA. January, 2017.
- Invited talk, Probability and Stochastics Seminar, Tulane University, New Orleans, LA. October 19, 2016.

- Invited talk, Bates College, Lewiston, ME. September 22, 2016.
- Invited talk, WPI, Worcester, MA. September 22, 2016.
- Minisymposium talk, SIAM Life Sciences, Boston, MA. July 11-14, 2016.
- Brown/BU PDE seminar, Boston University, Boston, MA. November 17, 2015.
- Seminar talk, Brown Applied Math graduate student seminar, Brown University, Providence, RI. October 17, 2016, and October 6, 2015.
- Contributed talk, Applied Math Days at Rensselaer Polytechnic Institute, Troy, NY. April, 2016 and 2015.
- Poster presentation. SIAM Nonlinear Waves and Coherence Structures, Philadelphia, PA. August 8-11, 2016.
- Poster presentation. AWM Workshop at SIAM Annual Meeting, Boston, MA. July 11-15, 2016.
- Poster presentation. Conference on Analysis of PDEs using Dynamical Systems Techniques at Boston University, Boston, MA, USA. June 1-3, 2016.
- Poster presentation. KUMU Conference on PDEs, Dynamical Systems and Applications at University of Missouri, Columbia, MO. April 23-24, 2016.
- Poster presentation. Opening Workshop on Stochastic Dynamical Systems in Biology, Newton Institute for Mathematical Sciences, Cambridge, UK. January 18-22, 2016.
- Poster presentation. SIAM Applications of Dynamical Systems, Snowbird, UT. May 17-21, 2015.

Professional Associations

- American Mathematical Society (AMS)
- Association for Women in Mathematics (AWM)
- Society for Industrial and Applied Mathematics (SIAM)
- Society for Mathematical Biology (SMB)

Dedication

To my parents, who supported my academic journey no matter how far it took me.

Acknowledgements

I would like to first thank my advisor Björn Sandstede, who has been incredibly supportive throughout my graduate school experience. I feel grateful to have had Björn as a teacher, as an instructor during my teaching assistantship, and as a research advisor who has taught me so much in every meeting. I am thankful for his excitement for our research collaboration and for his support for my professional and personal development. It is hard to imagine a better mentor, friend and colleague than Björn, and I feel honored and grateful to call him my PhD advisor.

I would also like to thank Professor Martin Maxey and Professor Scott McKinley for taking the time to be on my defense committee. Thank you to Professor Kimberly Mowry and her lab, for the research problem as well as for the helpful advice, collaboration, and discussions. I would also like to thank Scott for the productive discussions on my research.

Thank you to the members of the research group: Al, Blake, Chao, John, Kelly, Liz, Laura, and Paul, with whom I had productive conversations about my thesis work and beyond, and to Al, Amanda, Christian, Eirini, Ivana, and Paul for their friendship and encouragement throughout graduate school.

Finally, I would like to thank my parents (Emilia and Nelson) and Stefan for being there for me through the ups and downs. I cannot imagine going through graduate school without their love and support.

Contents

List of Figures

Chapter One

Introduction

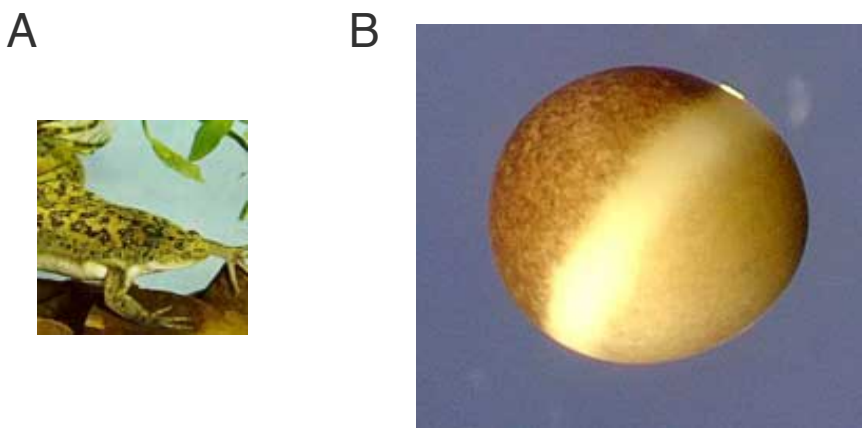

Figure 1.1: A: *Xenopus laevis* frog. B: Egg cell in the final stage of oogenesis, with one half (animal pole) colored brown and the other half (vegetal pole) colored yellow from the accumulation of egg yolk. Images from the Mowry lab.

1.1 Messenger RNA localization

Many organisms need to establish spatial orientation and patterning during early development. In particular, the egg cells (oocytes) of many organisms are asymmetrical and are organized along certain spatial axes. This spatial differentiation is key in cell division, where different material accumulates in different regions and the spatial axes of the oocyte control embryo formation [8]. After fertilization of the oocyte, the embryo consists of one cell which eventually develops into an organism with multiple cells. The large fertilized egg cell is divided through cell division into smaller cells [9] whose development in different locations in the newly formed embryo determines whether they form the skin and nervous system, the gut, or the muscle, blood and bone [10, 11].

This spatial differentiation in developing organisms can drive expression of genes that is confined to certain spatial regions and determines the healthy body plan of the embryo. This is often achieved through asymmetric accumulations of proteins and messenger RNAs (mRNAs) in oocytes [11, 12] (see Figure 1.1). In particular, active transport of mRNA is a widely used mechanism for generating and maintaining

a spatial orientation that is necessary for early development. In *Xenopus laevis* frog oocytes (see Figure 1.1), spatial differentiation is achieved by localization of messenger RNA (mRNA, red dots in Figure 1.2), which consists of an accumulation of these molecules at the periphery of the cell (denoted vegetal cortex or vegetal pole). mRNA particles are moved by molecular motor proteins along microtubule filaments (blue lines in Figure 1.2), and their localization to targeted cell destinations is critical for correct patterning of the embryo [13–16] (see bottom panel in Figure 1.2). If localization is incomplete 1-2 days into this early stage in egg formation, the cell develops abnormally or development is compromised. mRNA localization in *Xenopus laevis* oocytes is therefore an important model for understanding how maternal molecules are localized to influence pattern and polarity [17].

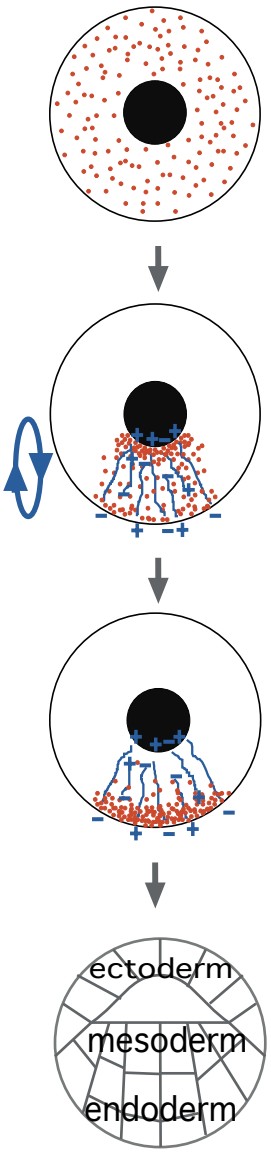

Figure 1.2: Cartoon of the mRNA localization process, and influence on healthy embryo development.

Understanding the mechanisms of mRNA localization would prove useful in determining factors that prevent complete localization. In particular, uncovering the contribution of localization mechanisms such as diffusion, active transport by motor proteins, and anchoring at the cell periphery is required in order to understand how the distribution of mRNA in the vegetal cytoplasm is regulated in space and time during localization. Challenges in understanding these mechanisms include the experimental obstacles for analysis of transport directionality in vivo, the lack of information on the microtubular structure, and the limited insight into the anchoring mechanism at the cell bottom. In collaboration

Table 1.1: Time and length scales in mRNA localization of *Xenopus laevis* oocytes.

Parameter	Value and references
VLE RNA localization time	24-48 hours [3]
Oocyte diameter (stage III)	300 μm [17]
Vg1 localization element length	340 nt $\approx$ 115 nm [12]
Diameter of FRAP bleach spot	5 μm [3]
Average microtubule length	9 μm [18]
Microtubule lifetime before catastrophe	100-400 s [18, 19]
Average kinesin run length	1-2 μm [20, 21]
Average dynein run length	1-2 μm [20]

with the Mowry lab at Brown University, we sought to understand how the long-term dynamics of mRNA molecules varies across the oocyte and to uncover the mechanisms of mRNA localization. Key time and length scales for mRNA localization in *Xenopus laevis* oocytes are provided in Table 1.1.

Our approach in quantifying the contributions of different mechanisms to localization in *Xenopus* oocytes is informed by parameter estimation using fluorescence recovery after photobleaching (FRAP) data. These experiments are widely used to understand molecular transport in living cells [22, 23], primarily through modeling diffusion and binding kinetics [1]. As a result, previous work models FRAP data using linear reaction-diffusion partial differential equations and estimates diffusion coefficients and binding rates by fitting the fluorescence recovery data to analytical expressions [24–33] or by using numerical optimization for more complex geometries and models [34–36]. In addition to movement by diffusion, macromolecules are actively transported on cytoskeletal networks by molecular motors such as myosin, kinesin, or dynein in many cell types. Therefore, our goal was to extract movement information from FRAP data in cells where transport is a key mechanism for the dynamics. To capture active transport, we use advection-reaction-diffusion partial differential equations that account for binding, diffusion, and active transport of particles. We carry out parameter estimation through optimization of numerical

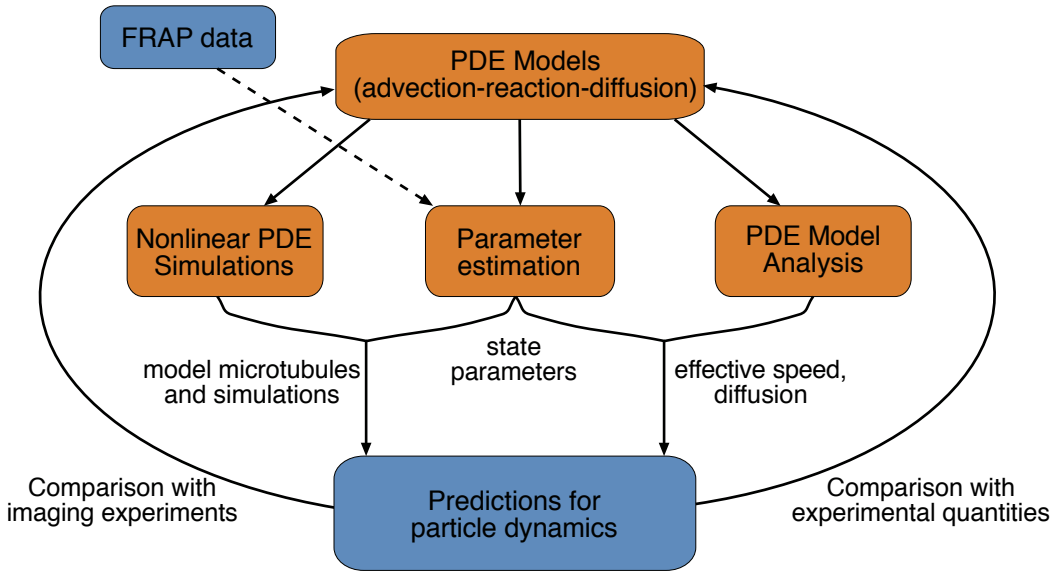

Figure 1.3: Cartoon of the proposed approach to drawing predictions for particle dynamics from FRAP data (adapted from [1]). Input (FRAP data) and output (Predictions for mRNA dynamics) are marked in blue, while modeling and analysis frameworks are marked in orange.

solutions of the PDE models and demonstrate that our approach allows efficient extraction of consistent estimates for movement, diffusion, and transition rate parameters from FRAP data using models of two or four particle states. Advection has been included in previous active transport models to describe spatial localization of RNA in *Drosophila* oocytes and embryos [6, 37], neurofilament transport along axons in neurons [38, 39] and motor-driven transport along filaments [40, 41]. However, such models were not applied to FRAP experimental data [1].

Given estimates of the parameters in the transport models considered, we show how these parameters can be used to predict effective velocities and diffusion rates for long-term dynamics [1] (see Figure 1.3). Consider for instance a system where particles switch between movement and diffusion. In one state, particles move with speed c, and in the other, they diffuse with rate d. Since the particles can switch between states with transition rates β_1 and β_2, the effective velocity and diffusion of the particles in the long run are different from the individual state parameters. Dynamical systems analysis of general advection-reaction-diffusion models allows us

to provide general formulas for these large time quantities, that go beyond specific examples [38, 39] and extend the treatment of reaction-hyperbolic systems [40, 42–44] and of diffusion in one population [41]. Mathematical derivation of these large-time solutions, as well as calculation of expected run lengths of motor-cargo complexes on microtubules (see Table 1.1), allow us to compare parameter estimation predictions with experimental observations and to refine or validate the models in the context of the mRNA dynamics in *Xenopus laevis* [1] (see Figure 1.3 for our approach).

In oocytes of the frog *Xenopus laevis*, transport of Vg1 mRNA to the vegetal cortex of the *Xenopus* oocyte requires a sequence element termed the VLE (vegetal localization element) RNA [1]. Since microinjected VLE RNA localizes to the vegetal cortex [45], we refer to it as "localizing RNA" (see Figure 1.4, RNA in red). Noting that most RNAs do not localize, we refer to RNA that is not actively transported but instead only diffuses as "nonlocalizing RNA". Our approach combines the parameter estimates obtained using modeling of FRAP data with the analytical methods to provide insights into the mechanisms of transport for both localizing and nonlocalizing RNA in *Xenopus* oocytes (see Figure 1.3). For non-localizing RNA, our method shows that RNAs in oocytes treated with nocodazole (which disrupts the cytoskeletal network) may be highly immobile, possibly due to their retention in large granules of RNAs. On the other hand, nonlocalizing β-globin RNAs switch between reaction and diffusion in the cytoplasm. For localizing VLE RNA, we confirm distinct directionality in different regions in the cell as suggested by photoactivation experiments in [3]. The transport in the vegetal direction is predicted to be faster and primarily unidirectional close to the nucleus, while the effective diffusion is expected to be higher in the lower vegetal cytoplasm. Importantly, our results suggest that some movement in both the animal and vegetal pole directions occurs throughout the vegetal cytoplasm in *Xenopus* oocytes [1]. Moreover, moving RNAs are predicted to

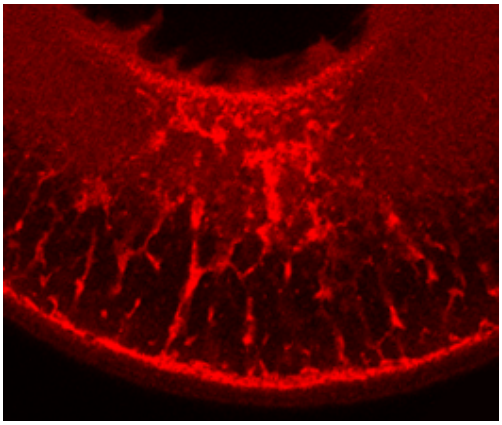

Figure 1.4: mRNA localization during stage III of oogenesis [2] (fluorescently labeled mRNA in red; from the Mowry lab).

remain in paused states for extended times, as observed in experiments [46].

The parameter estimates are also useful in a modeling framework that accounts for the fact that active transport is restricted to the microtubule cytoskeleton (see Figure 1.3). To model the transport on microtubules in the cytoplasm, we consider nonlinear advection-reaction-diffusion PDE models, where the transition rate from diffusion to active transport is dependent on the density of microtubules at that location. In particular, we extend the dynamical systems analysis to a two-state nonlinear model of transport for mRNA particles, and derive the large time effective velocity and diffusion of the molecules under the assumption of parallel microtubules. In *Xenopus* oocytes, imaging experiments show that the filaments have a mixture of random and radially outward orientations [3, Figure S3], [7, Figures S5, S6]. We therefore develop a computational framework for simulating the nonlinear PDE models with active transport occurring on model microtubule structures similar to [6]. Given that the microtubules experience dynamic instability and that they depolymerize in a shrinkage catastrophe when they age [18] (see also Table 1.1), we use multiple such model microtubule structures to simulate the periods of growth and shrinkage in microtubule dynamics. Our approach shows that bidirectional transport is essential in achieving the observed spatial scales of localization as in Figure 1.4

(see also [3, Figure S2]) and that an anchoring mechanism at the cell periphery is required for reproducing complete transfer of mRNA to the vegetal cortex in the timespan of localization. This method can be used to test for predictions on what dynamical states may get anchored at the periphery and directs future experimental attention to obtaining a better understanding of the anchoring mechanism.

A summary of our main modeling and analysis findings is provided below:

- *FRAP data analysis:* Using active transport models for particle dynamics, we developed efficient algorithms that analyze FRAP data to estimate key parameters such as velocities of motor-RNA complexes, diffusion coefficients of free RNA, and transition rates between different dynamical states.
- *Large time analysis of transport models:* We proved that the dynamics of actively transported molecules at large time is captured by two quantities, namely the effective velocity and effective diffusion, and we related these quantities analytically to parameters extracted from FRAP data using general transport models.
- *Predictions for mRNA dynamics:* Our analysis of FRAP data provided support for the role of paused states and bidirectional transport in *Xenopus* oocyte mRNA dynamics.
- *Simulations of nonlinear PDE models:* Parameter estimates from FRAP data informed our simulations of transport models on computational model microtubule structures, which allowed the comparison of mRNA spatial distributions with experimental data and suggested anchoring at the cell cortex is a key mechanism for the timescale of localization.

We emphasize that in addition to intracellular transport in frog oocytes, our FRAP

data analysis and the calculation of relevant biological quantities is appropriate more generally in understanding the contribution of diffusion, active transport, and binding kinetics in many other organisms.

Chapter Two

Parameter Estimation for Fluorescence Microscopy Data (FRAP)

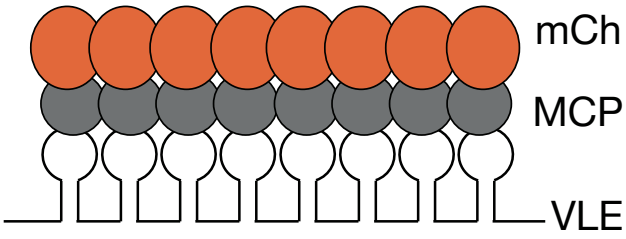

Figure 2.1: Cartoon of VLE RNA (VLE-MS2) tagged with fluorescent protein mCherry (mCh-MCP) (visualization similar to [3, Figure 4A]).

2.1 Overview of FRAP experiments

Fluorescence recovery after photobleaching (FRAP) is a widely used tool for investigating protein mobility and local molecular transport in living cells [1, 22, 23]. Live cell imaging using fluorescent proteins offers a powerful approach to answer both qualitative and quantitative questions about the dynamics and kinetics of RNA localization in the early stages of *Xenopus* oogenesis [17].

In FRAP experiments, the cell membrane sample is initially labeled with a fluorescent tag (see Figure 2.1). Then, a circular spot is irreversibly bleached with a short pulse of light of high intensity [22]. Fluorescence intensity is subsequently measured in the bleach spot and the amount of fluorescence in the region is plotted as time progresses. Note that photobleaching is not perfect, so that the first post-bleach fluorescence intensity is low but not zero (see Figure 2.2, $t = 0$, bottom panel). Fluorescence recovery curves such as the one in Figure 2.2 (bottom panels) are used to determine information on processes such as diffusion, binding reactions, and active transport that are key in determining the mobility in a wide range of living cells [22].

In *Xenopus laevis*, acquiring FRAP data requires the definition of regions of interest (ROI). The ROI of experimental interest is photobleached and additionally there are two control ROIs that are necessary for adjustment and normalization

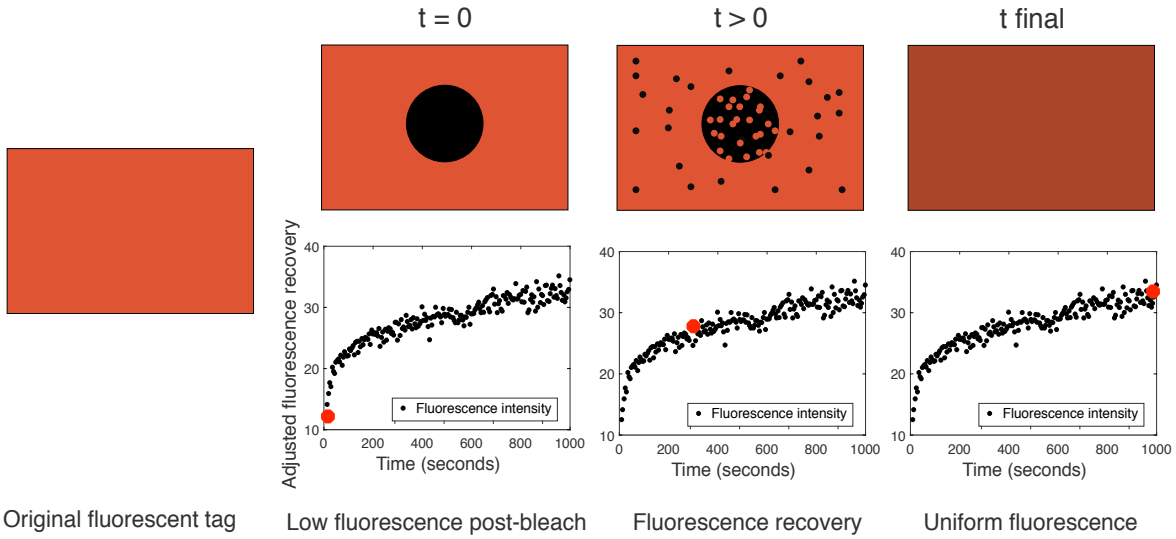

Figure 2.2: Cartoon of FRAP fluorescent labeling and recovery. Top panels show the dynamics of fluorescent molecules into the bleach spot at different time points, and bottom panels show the fluorescence recovery in the bleach spot. Red dots in the bottom panel correspond to the amount of fluorescence in the bleach spot at the time points indicated in the corresponding top panel. Note that even though the $t = 0$ bleach spot appears not to contain any fluorescence (top panel), photobleaching is not perfect and thus there is a low amount of fluorescence in the bleach spot at the initial time (bottom panel).

of the FRAP data. These ROIs must be the same size as the experimental ROIs [17]. First, fluorescence data $\mathrm{ROI}_o(t)$ is collected outside of the sample, and this is used to correct for the black level of the imaging field. Additionally, $\mathrm{ROI}_n(t)$ must be collected from within the sample and does not get bleached. The changes in fluorescence in this area are used to correct for photobleaching throughout the length of the experiment [17].

Since fluorescence is visualized as diffuse staining, FRAP data cannot be used to distinguish or track individual particles [1]; to make predictions about mobility and local transport, further analysis is needed to understand FRAP recovery curves as in the bottom panels of Figure 2.2. A large number of diffusion and reaction-diffusion models have been proposed for the quantitative analysis of FRAP recovery data (reviewed in [22, 23]). Most previous work makes use of linear reaction-diffusion partial differential equations models to predict diffusion and binding in cells. Depending

on the relative timescales of diffusion and binding, these methods involve estimating diffusion coefficients and binding rates by fitting the fluorescence recovery data to analytical solutions of the equations [24–33] or by using optimization and numerical solutions of the partial differential equations for more complex geometries and models [34–36].

2.2 Half-time method for cortical RNA

RNA localization in the *Xenopus* oocyte is responsible for the establishment of polarity during oogenesis as well as the specification of germ layers during embryogenesis. However, the inability to monitor mRNA localization in live vertebrate oocytes has posed a major barrier to understanding the mechanisms driving directional transport. In [17], a method for imaging MS2 tagged RNA in live *Xenopus* oocytes is described and applied to study the dynamics of cortical RNA. Here we provide results published in [17] and focus on methods for analyzing FRAP data by estimating the halftime of recovery. This procedure can be used to measure mobility in different regions of the oocyte, enabling the direct observation of molecular dynamics throughout the oocyte.

FRAP data is collected in different FRAP ROIs in the vegetal cytoplasm. The raw time series fluorescence data for each FRAP ROI is then corrected and normalized. To correct for acquisition photobleaching, the adjusted fluorescence time series $A(t)$ is calculated as

$$A(t) = F(t) \times \frac{F_{\text{pre}}}{F_{\text{n}(t)}} = (\text{ROI}(t) - \text{ROI}_o(t)) \times \frac{(\text{ROI}_n(1) - \text{ROI}_o(1))}{(\text{ROI}_n(t) - \text{ROI}_o(t))}\,, \tag{2.1}$$

where $F(t)$ is the background subtracted fluorescence at time t, F_{pre} is the background subtracted average intensity for all the prebleach frames and $F_n(t)$ is the background subtracted fluorescence intensity value of a neighboring region at time t. Here, $\text{ROI}(t)$ corresponds to the raw fluorescence data from the photobleached cortical region of interest at time t, and $\text{ROI}_o(t)$ and $\text{ROI}_n(t)$ are fluorescence data from the non-photobleached regions outside and inside the oocyte respectively at time t. The second equality in equation (2.1) shows that $F(t)$ corresponds to subtracting the background fluorescence $\text{ROI}_o(t)$ from the fluorescence intensity of interest $\text{ROI}(t)$, F_{pre} corresponds to subtracting the background fluorescence of the prebleach frames $\text{ROI}_o(1)$ from prebleach fluorescence outside the cortical region $\text{ROI}_n(1)$, and $F_n(t)$ is the subtraction of the background fluorescence intensity $\text{ROI}_o(t)$ from the fluorescence at the neighboring region outside the cortical region $\text{ROI}_n(t)$.

The adjusted fluorescence data $A(t)$ can be further normalized against prebleach values, yielding normalized adjusted fluorescence time series $A_{\text{norm}}(t)$:

$$A_{\text{norm}}(t) = \frac{A(t)}{f_{\text{pre}}} = \frac{A(t)}{(\text{ROI}(1) - \text{ROI}_o(1))}\,, \tag{2.2}$$

where f_{pre} is the background subtracted average prebleach intensity in the cortical region ROI. As in the case of adjusted data $A(t)$, we do not calculate the normalized fluorescence in the regions outside the sample (o) and inside the sample (n) since these areas are not photobleached, and thus the mobility of mRNA in these regions cannot be determined.

Either of the resulting time series data $A(t)$ or $A_{\text{norm}}(t)$ can be analyzed to extract binding and diffusion information from FRAP experiments. Other studies have provided a third normalization technique that can be applied to adjusted normalized fluorescence data $A_{\text{norm}}(t)$ to yield time series $N(t)$. This method requires that the

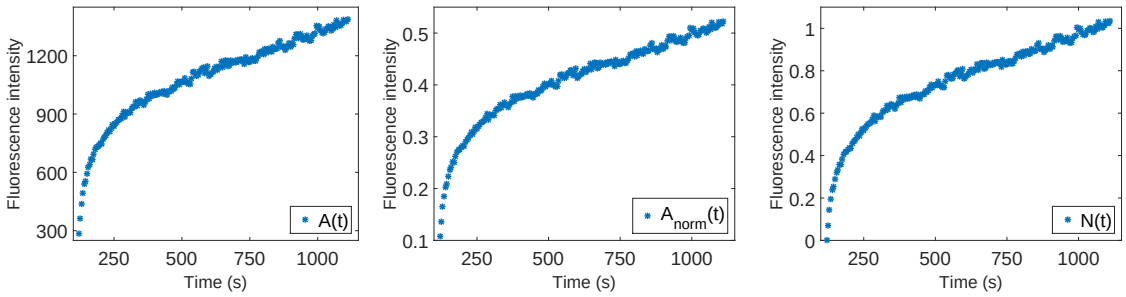

Figure 2.3: Fluorescence intensity normalizations for FRAP recovery curves. Adjusted fluorescence data (left, $A(t)$) is corrected for acquisition photobleaching, normalized adjusted fluorescence data (center, $A_{\text{norm}}(t)$) is further normalized against prebleach values, and normalized data (right, $N(t)$) has its values set to 0 right after the bleach and 1 at the plateau value.

intensity values are further normalized so that the fluorescence intensity right after the bleach is set to 0, and the one at the plateau level is set to 1 [24]. The equation for this normalized time series data $N(t)$ is:

$$N(t) = \frac{A_{\text{norm}}(t) - A_{\text{norm}}(1)}{A_{\text{norm,pl}} - A_{\text{norm}}(1)}\,, \tag{2.3}$$

where $A_{\text{norm}}(1)$ is the photobleach corrected fluorescence intensity at the first time point after bleaching, and $A_{\text{norm,pl}}$ is the average photobleach corrected fluorescence intensity at the plateau. In practice, $A_{\text{norm,pl}}$ can be calculated as the average of the last 10% values of photobleach corrected fluorescence intensity A_{norm} measured in the experiment. We will focus on $A(t)$ and $N(t)$ in our subsequent analysis of the FRAP fluorescence intensity curves (see Figure 2.3).

Assuming that there is no active transport of RNA in the cortical RNA population, the adjusted or normalized FRAP fluorescence intensity curves are fit to the established model in [27]. The equations for the single binding protein reaction involve a population of free molecules that can diffuse, as well as a population of bound complexes. The two populations interact via binding and dissociation rates k_{on} and k_{off}. The main assumptions for the approach in this model are that the RNA population achieves equilibrium before photobleaching, and that the binding

sites are part of a large and immobile complex.

Previous FRAP experiments in this system [3, 47] assume that mRNA diffusion across the bleach spot is fast compared to the timescale of binding of mRNA with sites it interacts with. Under this assumption, the fluorescence intensity recovery is called interaction-limited as described in [48] or, equivalently, it satisfies the reaction dominant simplification of the model in [27]. This assumption means that the adjusted and the normalized fluorescence curves can be fit to the exponential model:

$$\text{frap}(t) = a\left(1 - \frac{k_{\text{on}}}{k_{\text{on}} + k_{\text{off}}} e^{-k_{\text{off}} t}\right), \tag{2.4}$$

or equivalently to the formulation

$$\text{frap}(t) = a - ce^{-k_{\text{off}} t}, \tag{2.5}$$

where $c = a\frac{k_{\text{on}}}{k_{\text{on}}+k_{\text{off}}}$ and $\text{frap}(t) = A(t)$ or $N(t)$.

Half times of recovery $t_{1/2}$ based on the above model can be calculated using the equation:

$$t_{1/2} = \frac{\ln(2)}{k_{\text{off}}}. \tag{2.6}$$

We analyzed the data from the three ROIs for the cortical RNA using the above methodology. Using this method and fitting with the least squares norm, both the adjusted and normalized data sets show very similar halftimes $t_{1/2}$. Estimates of the rates k_{on} and k_{off} can also be calculated by fitting fluorescence intensity curves to equation (2.5). Cortical RNA has been hypothesized to be a highly stable complex as the RNA remains localized from stage II through the end of oogenesis [49]. The FRAP results discussed here confirm that the cortical fraction of RNA is present in

a highly stable complex (see Figure 2.4 and Table 2.1).

Table 2.1: Estimated halftimes for mobility of RNA in the cortical area of two oocytes (three regions). Estimates using both adjusted and normalized data are provided.

	$t_{1/2}$ (s), oocyte 1			$t_{1/2}$ (s), oocyte 2		
Region	1	2	3	1	2	3
Adjusted data	361.72	263.73	350.16	202.65	224.93	201.92
Normalized data	361.70	263.77	350.16	202.66	224.93	201.91

The Matlab routine *fit* was used to fit the model equation (2.5) to experimental data (The MathWorks, Natick, MA). In order to optimize the fluorescence data fits and ensure convergence to reasonable parameter estimates, we can specify initial guesses for the parameters a, k_{off}, and c in equation (2.5) as well as bounds for these parameters. Given the meaning of a and c, we search for positive parameters that are bounded above by a value larger than the maximum fluorescence data (in practice, this can be $2 \times \max(\text{frap}(t))$ and choose the last fluorescence data point in the FRAP recovery as its starting point in the Matlab *fit* command. Parameter k_{off} is most relevant in estimating halftimes and mobility: we search for a positive parameter, relax the upper bound to infinity, and choose 0 as the initial guess in the Matlab *fit* command to optimize the data fitting. The results of the optimization for k_{off} are summarized through estimates of the halftimes (equation (2.6)) in Table 2.1.

The method described here for analysis of FRAP data produces similar rates of k_{off}, and thus halftimes of RNA recovery for different data normalizations and different regions of interest. The halftime and binding/unbinding rate estimates obtained from fitting corrected FRAP data are useful in that they may provide a comparison of mobility of RNA in wild-type oocytes with RNA mobility in oocytes where motor protein function has been disrupted. As noted above, this FRAP analysis approach assumes that there is no active transport of RNA [47]. This is a reasonable assump-

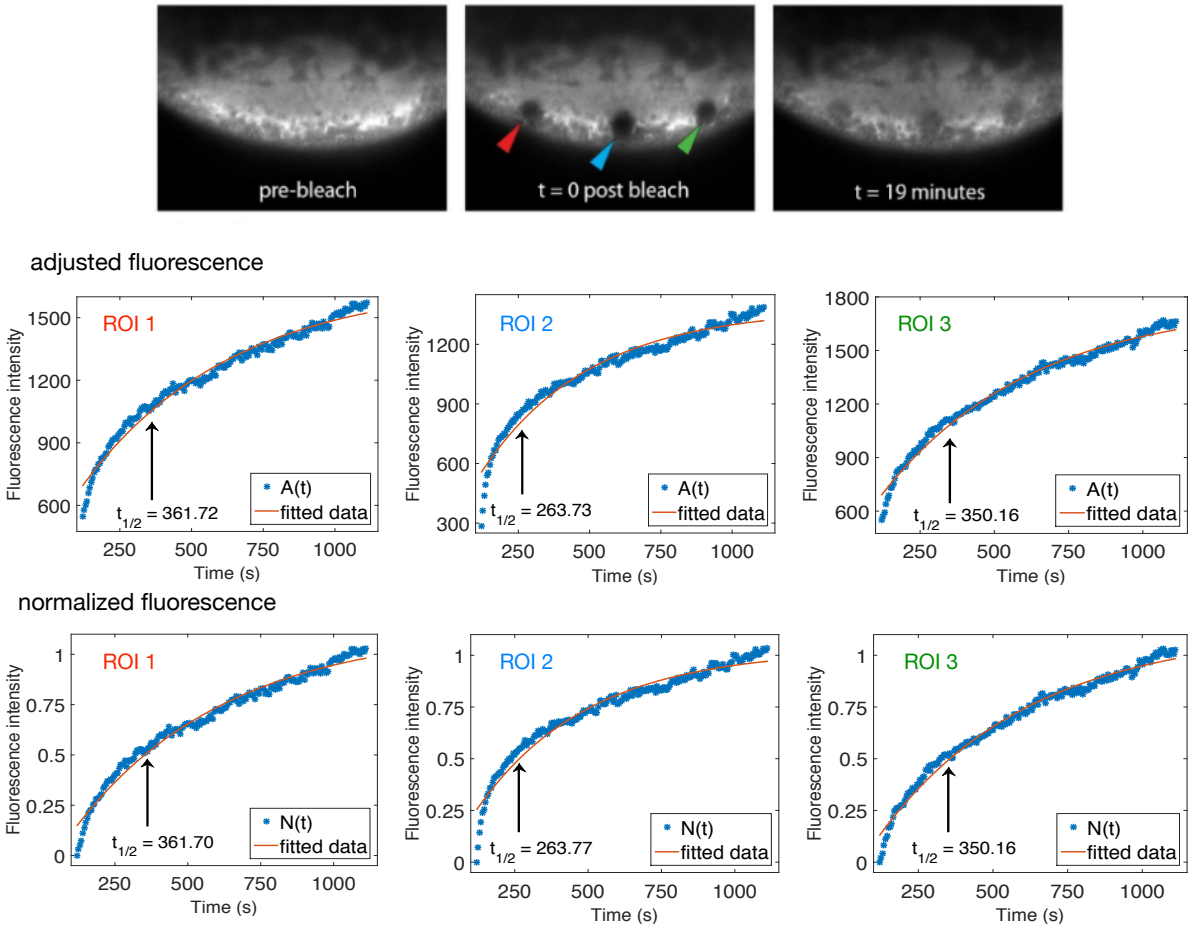

Figure 2.4: Adjusted fluorescence intensity curves (top) and normalized fluorescence intensity curves (bottom) fitted with equation (2.5). ROI 1 (left), ROI 2 (center) and ROI 3 (right) are experimental regions in the cortical area of the same oocyte. The black vertical lines indicate the halftime of recovery for each fluorescence curve.

tion for cortical regions of the *Xenopus* oocyte as shown in the top panels of Figure 2.4, but cannot account for RNA localization in the vegetal cytoplasm area, which may depend on the active transport of RNA by molecular motor proteins kinesin and dynein [3, 7].

2.2.1 Limitations

We recall that the reaction dominant simplification of the reaction-diffusion model derived in [27] is employed for diffusion coefficient estimates for VLE RNA in [3] based on the estimated halftimes of recovery. However, since diffusion is assumed to be very fast compared to the timescale of binding and FRAP experiments, the diffusion coefficient cannot be recovered in this scenario [27], and therefore a method that directly estimates parameters such as the diffusion coefficient and binding rates is required. In the next section, we make no assumptions on the contribution of diffusion to particle mobility and instead use a numerical approach to estimate this coefficient along with other key parameters for FRAP recoveries. Moreover, mRNAs in healthy *Xenopus* egg cells are believed to be actively transported by molecular motor proteins [3]. As a result, in the next section we also consider active transport as a key mechanism in parameter estimation using FRAP experiments for VLE RNA in untreated oocytes.

2.3 Novel FRAP parameter estimation using numerical methods for model PDEs

The parameter estimation methods and algorithms in this section were published in [1].

FRAP data analysis typically involves modeling two particle states (diffusion and binding reactions) and making assumptions about diffusion, the number and type of binding interactions, and their respective timescales in cells [27]. However, in systems with active movement, FRAP data analysis may overestimate diffusion rates

if active transport is not taken into account [22]. The goal is to include transport mechanisms in modeling particle dynamics and develop a numerical approach for extracting movement parameters from FRAP data using these models. To account for binding, diffusion, and active transport of particles, we consider advection-reaction-diffusion partial differential equations. Parameter estimation is carried out through optimization of numerical solutions of the PDE models considered. Using models of two or four particle states (see Figure 2.5), we demonstrate that the framework proposed allows us to efficiently extract consistent parameter estimates for movement, diffusion, and transition rates based on FRAP data. While advection is considered in certain studies for RNA localization in *Drosophila* oocytes and embryos [6, 37], axonal transport [38, 39], and motor-driven transport along filaments [40, 41], these models are not used for quantifying particle dynamics using FRAP experiments. Here we design general techniques for FRAP parameter estimation using transport partial differential equations modeling frameworks and demonstrate their efficacy.

2.3.1 Modeling particle transport and diffusion

We model the particle dynamics using systems of advection-reaction-diffusion partial differential equations (PDEs). In the application to *Xenopus* RNA dynamics, the variables correspond to concentrations of mRNA and mRNA-motor complexes in different dynamical states. The simplest model we consider assumes that particles can be in one of two states: a population u moving with speed c (e.g., carried by molecular motor proteins to the cell cortex) and a population v diffusing in the cytoplasm with diffusion coefficient d. As in [6], particles can transition from moving

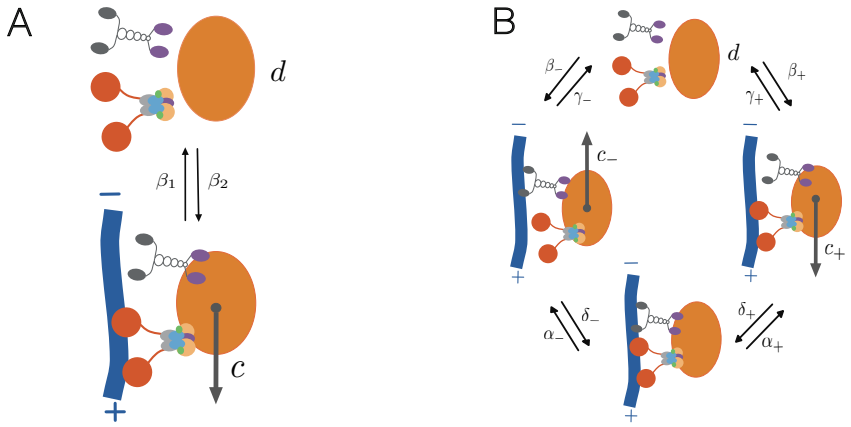

Figure 2.5: A. Cartoon of the 2-state model of active transport, consisting of a population of diffusing particles with diffusion coefficient d, and a population of moving particles with velocity c. B. Cartoon of the 4-state model of active transport, consisting of a population of diffusing particles with diffusion coefficient d, a population of particles moving towards the cortex with velocity c_+ and one moving towards the nucleus with velocity c_-, and a stationary population (from [1]).

to diffusing, and vice versa, so that the dynamics is described by the PDE system:

$$\begin{aligned} u_t &= cu_y - \beta_1 u + \beta_2 v \\ v_t &= d\Delta v + \beta_1 u - \beta_2 v\,, \end{aligned} \tag{2.7}$$

where β_1 is the rate of transition from the moving to the diffusing state, and β_2 is the rate of transition from the diffusing to the moving state (see Figure 2.5A). We note that movement is assumed to be one-dimensional (e.g., along a microtubule), while diffusion can occur in two spatial dimensions. We will refer to this model as the 2-state model from here on. While FRAP experiments do not distinguish between different dynamical states (see Figure 2.6B), our approach allows us to use FRAP recovery curves to estimate transport parameters and rates for each state.

To take into account the possibility of bidirectional transport of particles, we also consider a more complex 4-state model. In this model, we consider a moving population u^+ carried by one type of motor protein (e.g., dynein) to the vegetal cortex, a moving population u^- carried by another type of motor protein (e.g.,

kinesin) to the nucleus, a population v diffusing in the cytoplasm with diffusion coefficient d, and a population w paused on the microtubules. Different mechanisms that may account for the stationary population w are reviewed in [46]. These four states react through binding and unbinding reactions as follows:

$$\begin{aligned} v_t &= d\Delta v - \beta_+ v + \gamma_+ u^+ - \beta_- v + \gamma_- u^- \\ u_t^+ &= c_+ u_y^+ + \beta_+ v - \gamma_+ u^+ + \alpha_+ w - \delta_+ u^+ \\ u_t^- &= -c_- u_y^- + \beta_- v - \gamma_- u^- + \alpha_- w - \delta_- u^- \\ w_t &= \delta_+ u^+ + \delta_- u^- - \alpha_+ w - \alpha_- w \,, \end{aligned} \tag{2.8}$$

with rates α_+, α_-, β_+, β_-, δ_+, δ_-, γ_+, γ_- as in Figure 2.5B. In *Xenopus* oocytes, assuming that dynein moves the mRNA down towards the cortex and kinesin mostly moves the cargo up to the nucleus [3], the rates β_+ and β_- can be interpreted as binding rates for dynein and kinesin, respectively, to microtubules, while the rates γ_+ and γ_- correspond to unbinding rates for dynein and kinesin, respectively, from microtubules (see Table 2.2).

To accurately capture the dynamics of particles that are not actively transported (e.g., nonlocalizing β-globin RNA), we also consider the reaction-diffusion system

$$\begin{aligned} u_t &= -\beta_1 u + \beta_2 v \\ v_t &= d\Delta v + \beta_1 u - \beta_2 v \,, \end{aligned} \tag{2.9}$$

where population u is in the stationary state and population v is diffusing. This model has been previously analyzed in many studies, including [27].

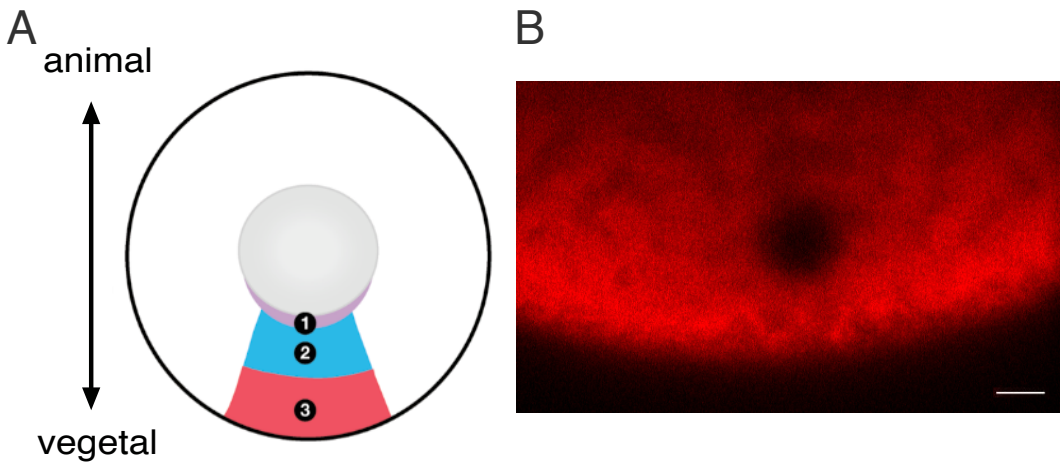

Figure 2.6: A: FRAP bleach regions are numbered according to their location: the perinuclear cup (Region 1), the upper vegetal cytoplasm (Region 2) and the lower vegetal cytoplasm (Region 3). B: Shown is a representative oocyte in which a 5 μm circular ROI of VLE-MS2 RNA bound by MCP-mCh was bleached in the vegetal cytoplasm (Region 3). Scale bar corresponds to 10 μm.

2.3.2 Initial conditions: approximations of postbleach intensity profiles

In order to provide initial conditions for the concentrations of RNA in equations (2.7), (2.8) and (2.9), we analyze the FRAP postbleach intensity profiles from *Xenopus* experiments (see Figures 2.6B and 2.7). The photobleaching process in FRAP is commonly assumed to be an irreversible first-order reaction [22, 24, 26, 31]:

$$\frac{dC(\boldsymbol{r},t)}{dt} = -\alpha I_b(\boldsymbol{r})C(\boldsymbol{r},t)\,,$$

where C is the spatial concentration of fluorophores, $\boldsymbol{r}$ denotes the radial position, and α is a bleaching parameter.

The bleaching distribution $I_b(\boldsymbol{r})$ is assumed to have a Gaussian profile [24]:

$$I_b(\boldsymbol{r}) = \frac{I_0}{\pi r_e^2} e^{-2\frac{\boldsymbol{r}^2}{r_e^2}}\,,$$

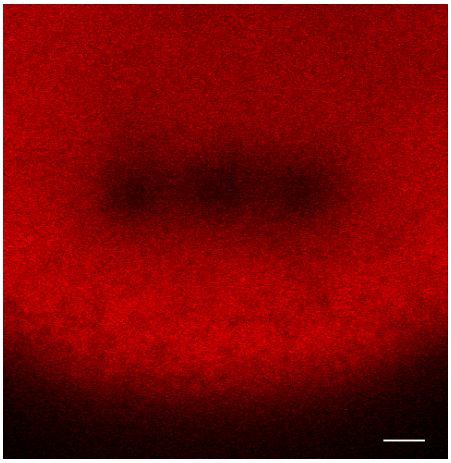

Figure 2.7: Shown is a representative oocyte in which three 5 μm circular ROIs of βG-MS2 RNA (β-globin RNA) bound by MCP-mCh were bleached as detailed in the main text. Scale bar corresponds to 10 μm.

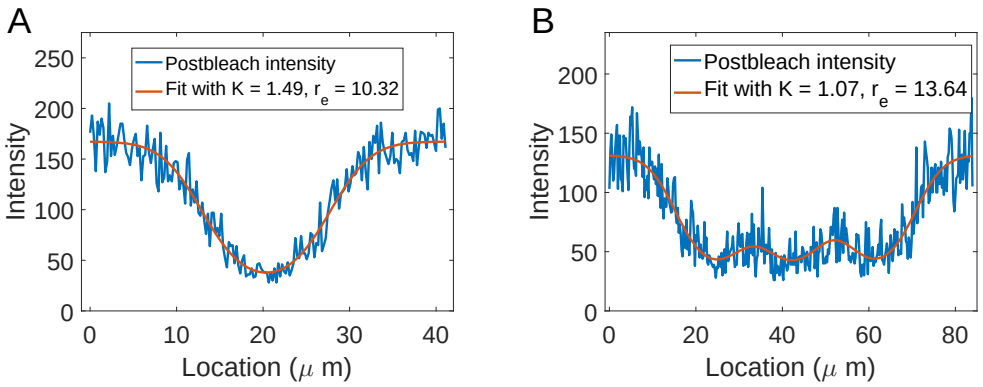

Figure 2.8: Sample fit of FRAP postbleach intensity distribution for VLE RNA (see bleach spot in Figure 2.6B) with equation (2.11) (A) and for β-globin RNA (see bleach images in Figure 2.7) with equation (2.12) (B) (from [1]).

where r_e is the effective radius of the distribution. This profile is due to the limitations of bleaching and scanning in FRAP, which lead to what is referred to as a "corona effect" of the bleached region in [34]. The distribution of fluorophore concentration after photobleaching can therefore be described by

$$C(\boldsymbol{r}) = C_0 e^{-\alpha I_b(\boldsymbol{r})} = C_0 e^{-Ke^{-2\boldsymbol{r}^2/r_e^2}}, \tag{2.10}$$

where K is the bleaching depth parameter [22].

We focus on the focal plane of the fluorescence distribution in the first frame after photobleach, and fit parameters K, r_e and x_0 to the initial profile:

$$C(x) = C_0 e^{-Ke^{-2\frac{(x-x_0)^2}{r_e^2}}}. \tag{2.11}$$

We note that x_0 corresponds to the center of the bleach location. Similar to [32], the fit of the postbleach intensity profile to the exponential of a Gaussian is very good, as can be seen in Figure 2.8A.

Postbleach profiles indicate that the three bleach spots in the cytoplasm are not clearly separated for FRAP experiments in β-globin RNA oocytes (see Figure 2.7). Since parameter estimation is sensitive to the initial condition given by the postbleach profile (see §2.3.5), we decided to treat all three spots together in the initial condition to accurately model the photobleach dynamics (see Figure 2.8B). The advantage of the numerical parameter estimation method described in the next section is that it can be applied to experimental data where more than one bleach spot must be considered. The above fitting procedure is repeated for parameters K, r_e, x_0, x_1 and x_2 using the postbleach distribution model:

$$C(x) = C_0 e^{-K\left(e^{-2\frac{(x-x_0)^2}{r_e^2}}+e^{-2\frac{(x-x_1)^2}{r_e^2}}+e^{-2\frac{(x-x_2)^2}{r_e^2}}\right)}. \tag{2.12}$$

An example of a postbleach intensity profile in the nonlocalizing RNA case and the fit to equations (2.12) is provided in Figure 2.8B. The equations for postbleach distribution are modeled in Matlab (The Math Works, Natick, MA), and the optimization is carried out using the Matlab routine *nlinfit*.

It is worth noting that studies including [24, 26, 27] assume uniform and instantaneous photobleaching of the circular area during FRAP. In this uniform disk

model, the post-bleaching intensity profile is assumed to have a flat circular profile instead of the exponential of a Gaussian in (2.11), as this simplifies the analytical calculations in those approaches.

2.3.3 Numerical FRAP data fitting procedure

We analyzed the FRAP recovery curves by numerically integrating the model equations in (2.7), (2.8), and (2.9) using an efficient exponential time-differencing fourth-order Runge-Kutta scheme [50, 51] for time integration coupled with Fourier spectral methods for space discretization. The boundary conditions for the PDE systems are taken to be periodic in both the x and y dimensions. The spatial domain size is taken to be large relative to the bleach spot size, with length scales of 40 μm in the horizontal direction x, and 60 μm in the vertical direction of movement y. We used 64 Fourier modes in the spectral decomposition in both directions, which is sufficient for the purpose of our simulations. Finally, different time steps were tested, and $\Delta t = 0.1$ was chosen for yielding consistent results while also minimizing computation costs.

We assumed a uniform point spread function [22] for detection of fluorescence by the microscope, which means that the observed fluorescence recovery is obtained by integrating the concentrations of particles in all states over the bleach spot. For the 2-state model, this implies:

$$\mathrm{FRAP}(t) = \int_{\text{bleach spot}} (u+v)(\boldsymbol{r},t)d\boldsymbol{r}\,. \tag{2.13}$$

This is a reasonable assumption given that the microscope resolution is much smaller than the radius of the bleach spots in our experiments [29]. Since photofading during

image acquisition can affect parameter estimation [52], we adjusted the FRAP data by correcting for background fluorescence and dividing the result by the fluorescence intensity of a neighboring image at each time point [17, 52] (see §2.2). Parameter estimation was carried out for individual oocyte data or for an average of data from multiple oocytes, as indicated in the text. It is worth noting that the diffuse fluorescence staining in FRAP experiments does not distinguish between different particle populations. The method proposed here provides insight into the contribution of the different dynamical states by fitting the sum of the particle concentrations over the bleach spot to the ensemble FRAP fluorescence.

The model equations and the fit to experimental data were programmed in Matlab (the Math Works, Natick, MA). The Matlab routine *lsqnonlin* was used to perform L^2-norm fits of the experimental FRAP data to equations similar to (2.13) for the appropriate model. Tests of fitting to FRAP data generated using known parameters for model (2.7) revealed that the initial guess for the unknown parameters is instrumental in convergence to the true parameter fit. Model fits of the experimental data were therefore preceded by ample parameter sweeps, as in [27, 36]. For the 2-state model, we sampled through values of c, d, β_1 and β_2 and chose the parameter combinations that yielded the smallest L^2-norm difference with the experimental data as initial guesses for routine *lsqnonlin*. We used these parameter combinations from all experimental trials for a certain region or type of RNA as initial guesses for multiple starting point search optimization in Matlab using *MultiStart*. This allowed us to reduce the computational cost of the fitting procedure by running the same optimization solver (*lsqnonlin*) using different initial conditions in parallel on a computer cluster.

The parameter estimation procedure is similar when using the reaction-diffusion model (2.9). For the 4-state model (2.8), the estimates for speed c and diffusion

Table 2.2: Fixed parameters for the 4-state model (2.8) from [53, 54] (published in [1]).

Parameter	Kinesin	Dynein
Binding rate (s^{-1})	$\beta_- = 5$	$\beta_+ = 1.6$
Unbinding rate (s^{-1})	$\gamma_- = 1$	$\gamma_+ = 0.27$

coefficient D using model (2.7) are used as initial guesses for c_+ and d in this more complex model. Available parameter values, such as the binding and unbinding rates for kinesin and dynein in [53, 54], are set constant to further reduce the size of the parameter sweep (see Table 2.2). Table 2.2 illustrates a choice of the binding/unbinding rates where dynein moves cargo down to the periphery, and kinesin up to the nucleus, given the evidence for the upper cytoplasm of *Xenopus* oocytes [3]; however, different assumptions on the motors carrying the RNA in either direction do not significantly alter the parameter estimates. Sweeps of the parameter space are then performed for the remaining five model parameters (c_-, α_+, α_-, δ_+, δ_-), and then refined in local parameter regions that yield best data fits. This ample parameter sweep provides initial guesses for the optimization; even though different initial conditions may lead to different estimates of the kinetic rates, the parallel computation setup of our optimization allows us to identify multiple initial conditions that lead to the smallest least squares residuals and consistent estimates of velocities and diffusion.

FRAP data for different types of RNA was fit using three approaches that we outline here. In Approach 1, we considered individual FRAP curves separately for each wild-type oocyte and for each region, and we compare the parameter estimates for VLE RNA mobility in different locations across oocytes. Approach 2 instead considers averages of FRAP data based on sets of 5 oocytes each, so that for each region we obtain only one set of parameter estimates for VLE RNA for each such set. Finally, Approach 3 applies to β-globin RNA, where we consider individual oocyte FRAP data, but we fit all three regions simultaneously, which yields a single set

of parameters for each oocyte. We note that differences in β-globin RNA mobility between regions are not expected [3], and that the initial condition modeling the postbleach dynamics (see Figure 2.8B and §2.3.2) requires that we consider all three bleach spots and therefore that we fit all three FRAP curves simultaneously.

2.3.4 Validation using synthetic FRAP data

To ensure the accuracy of our parameter estimation procedure, we generated synthetic FRAP data using two approaches, denoted as Algorithms 1 and 2.

In Algorithm 1, we fix parameters, and deterministically run the partial differential equations models (2.7) and (2.8) using these fixed parameters and a uniform disk initial condition. The numerical methods for solving these PDEs are described in §2.3.3. Since it is deterministic, Algorithm 1 generates smooth synthetic FRAP recovery curves such as the example in Figure 2.9. We note that the parameters used here are different from the ones we report for real FRAP data since we used the uniform disk initial condition for these validation tests.

Algorithm 2 consists of generating recovery curves from a continuous-time Markov chain (CTMC) in which the states correspond to the populations in model (2.7) or (2.8). Particles are assumed to switch between states using transition rates, and the times spent in each state are assumed to be exponentially distributed random variables with rates given by the transition parameters. This approach requires starting with a large number of particles with locations in a domain of the same size as the one used in Algorithm 1 (see §2.3.3). The bleaching process is modeled by initially assigning particle locations everywhere in the domain except for the designated circular bleaching spot. The FRAP data then corresponds to the number of particles

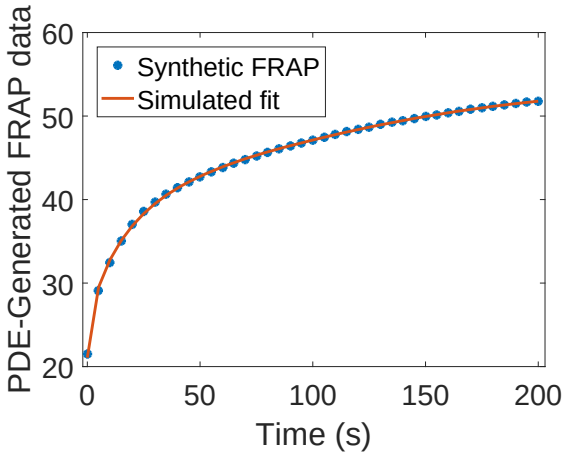

Figure 2.9: Sample synthetic FRAP data generated using Algorithm 1 with equations (2.7) and parameters $c = 0.031\ \mu m/s$, $d = 0.061\ \mu m^2/s$, $\beta_1 = 4\text{e}{-13}\ s^{-1}$ and $\beta_2 = 0.007\ s^{-1}$ is fit using our parameter estimation procedure. The recovered parameters are $c = 0.029\ \mu m/s$, $d = 0.06\ \mu m^2/s$, $\beta_1 = 1\text{e}{-12}\ s^{-1}$ and $\beta_2 = 0.007\ s^{-1}$.

Table 2.3: Table of input and output parameters for Figure 2.10. Input corresponds to parameters used for data generation, Output (10^4) corresponds to parameters estimated using data generated with 10^4 RNAs, and Output (10^6) corresponds to parameters estimated using data generated with 10^6 RNAs (from [1]).

	c_+ $(\mu m/s)$	c_- $(\mu m/s)$	d $(\mu m^2 s^{-1})$	α_+ (s^{-1})	α_- (s^{-1})	δ_+ (s^{-1})	δ_- (s^{-1})
Input	0.157	2e−4	0.11	0.008	4e−6	3e−4	0.12
Output (10^4)	0.114	2e−6	0.05	0.002	0.005	0.03	0.08
Output (10^6)	0.13	3e−5	0.09	3e−5	0.008	0.03	0.05

that move to the circular spot locations at different time intervals. We note that instead of adding Gaussian noise to simulated recovery curves as in [28], Algorithm 2 proposes a stochastic model that generates more realistic noisy FRAP recovery curves (see Figure 2.10). Parameter estimation on these FRAP curves shows that increasing the number of particles modeled yields increasingly better parameter recoveries (see Table 2.3).

We tested our parameter estimation methods on synthetic FRAP data generated using these methods. The generated FRAP curves were used to estimate parameters using the procedure outlined in §2.3.3, assuming unknown parameters and using the parameter sweep approach to generate initial guesses. We compared the results with the original fixed parameters, which were taken from a previous experimental run

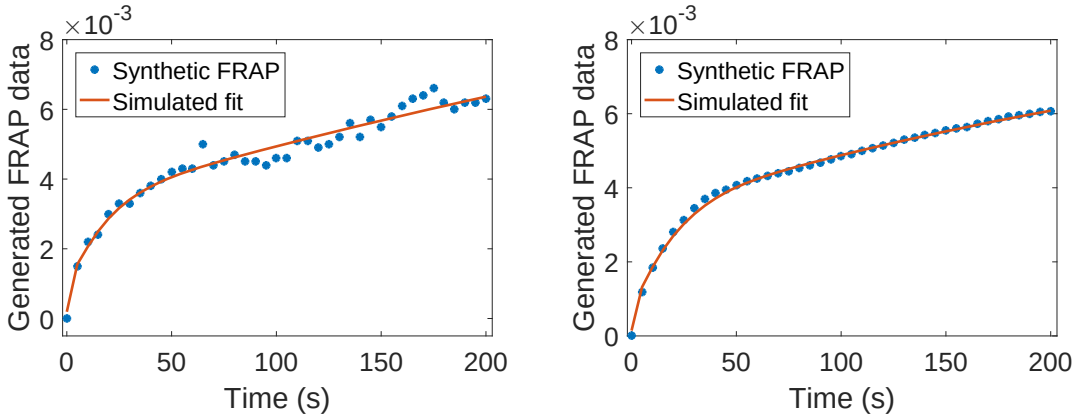

Figure 2.10: Sample synthetic data generated using Algorithm 2 with the Markov Chain approach (10^4, respectively 10^6 RNAs) based on the 4-state model are fit using our parameter estimation procedure. We note that the approximately 10^8 RNAs injected in the oocytes correspond to roughly 10^4-10^6 RNAs given the spatial domain considered in our simulations (from [1]).

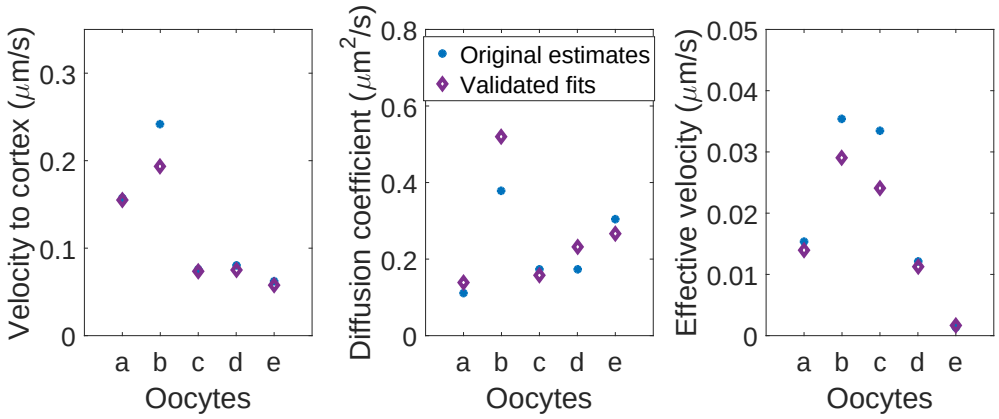

Figure 2.11: Parameter estimates from Region 1 VLE RNA FRAP data using Approach 1 for individual oocyte trials are validated using PDE-generated FRAP recovery curves (5 trials shown, from [1]).

for VLE RNA in Region 1. A fit of the synthetic data generated using the 2-state PDE model (2.7) with Algorithm 1 is illustrated in Figure 2.9. Figure 2.11 shows that estimates of the speed to cortex c_+, diffusion coefficient d and effective velocity v (see derivation in §3.2) using Algorithm 1 are almost identical to the originally-estimated parameters. This confirms that parameter estimates for velocities and diffusion can be recovered using the proposed parameter sweep and multiple start point optimization methods. In addition, Figure 2.10 shows the fit of the synthetic FRAP data generated using the CTMC approach in Algorithm 2.

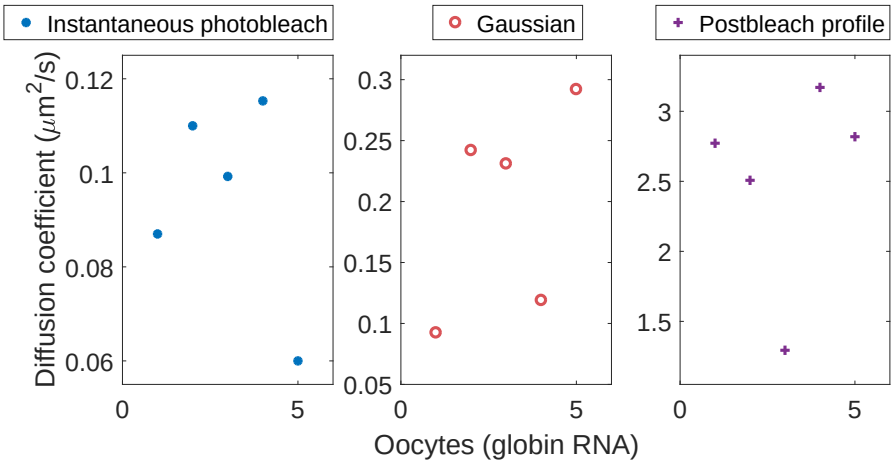

Figure 2.12: Estimated diffusion coefficients for β-globin RNA using the same set of 5 oocytes with instantaneous photobleach (flat circular disk initial conditions), Gaussian initial conditions, and the experimental photobleach profile (exponential of Gaussian initial condition), left to right. Note the different scales of the vertical axis (from [1]).

2.3.5 Sensitivity to initial postbleach profile

In analyzing FRAP recovery curves, the initial condition for equations (2.7), (2.8), and (2.9) can play a significant role on parameter estimates. The uniform disk model studied in [24] and [26] assumes that the photobleach and dynamics during the photobleach process are instantaneous. The post-bleaching intensity is thus assumed to have a flat circular disk profile, which eases the derivation of analytical solutions for FRAP recovery. We compared the results of our parameter estimation methods for β-globin non-localizing RNA using the uniform disk model initial conditions with results obtained using a Gaussian initial condition of the form $C_0 e^{-\frac{(x-x_0)^2}{r^2}}$ and with results using the spatial fluorescence distribution after photobleach for initial conditions (fit with a biophysically-justified exponential of Gaussian, see Figure 2.8).

Figure 2.12 shows that the diffusion coefficient estimates for the same set of 5 oocytes are sensitive to the initial condition. The uniform disk profile assumption yields diffusion coefficients 35 times smaller on average than the ones that use

spatial information from postbleach intensity profiles (see additional estimates in Figure 2.13). Similarly, uniform disk model estimates of the diffusion coefficient in particles engaged in active transport (localizing VLE RNA) were on average 33 times smaller than the postbleach profile ones for individual FRAP data (see Figure 2.14). These observations indicate that the assumption of instantaneous dynamics during the photobleach process leads to significant under-estimation of the diffusion coefficient, as previously noted and explored in multiple studies [28, 30–32, 34]. This suggests that a more practical approach for FRAP data analysis consists of using the spatial distribution of fluorescence from the first postbleach image as the initial condition for the FRAP dynamical equations [22, 29, 30, 32, 35].

Binding rates estimates are also believed to be affected by the assumption of instantaneous diffusion during photobleaching [31]. Our tests using advection-reaction-diffusion models such as (2.7) and (2.8) reveal similar estimates for velocities towards the vegetal cortex (c and c_+) for the two initial conditions. However, estimates of velocities in the animal pole direction, as well as some transport rates, are affected by the uniform disk model assumption. Predictions of the speed of transport c_- in the animal pole direction are underestimated about twenty-fold when the bleaching dynamics is assumed to be instantaneous (see Figure 2.15). This suggests that the photobleach process should be taken into account in order to obtain more accurate orders of magnitude estimates for model parameters.

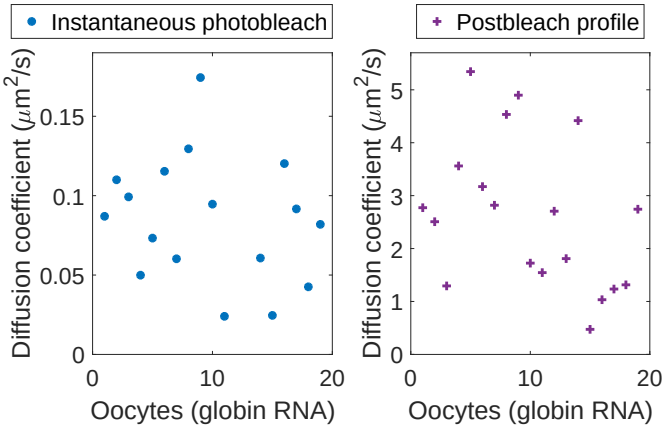

Figure 2.13: Estimated diffusion coefficients for β-globin RNA using the instantaneous photobleach and experimental photobleach profile initial conditions (from [1]).

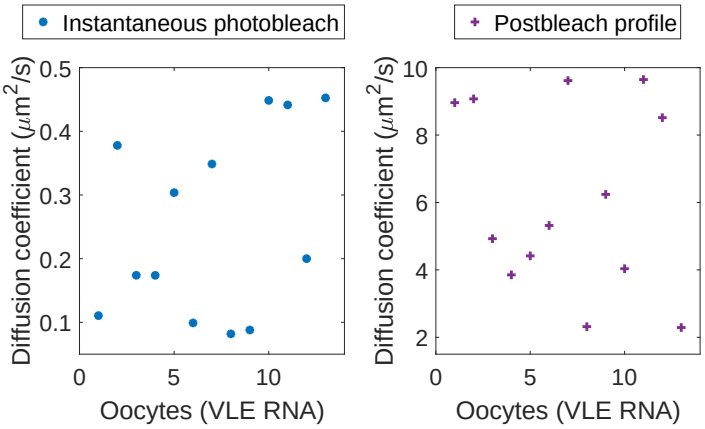

Figure 2.14: Estimated diffusion coefficients for VLE RNA using the instantaneous photobleach and experimental photobleach profile initial conditions (using Approach 1, from [1]).

2.3.6 Parameter estimation for non-localizing and localizing RNA

Non-localizing RNA: We first apply the parameter fitting procedure to FRAP data for particles that are not actively transported in the cell. The attempts to fit FRAP data for β-globin RNA using the 2-state active transport model (2.7) either fail or provide inconsistent results across trials. The parameter sweeps result in initial guesses for the unknown parameters that overestimate velocity c and underestimate

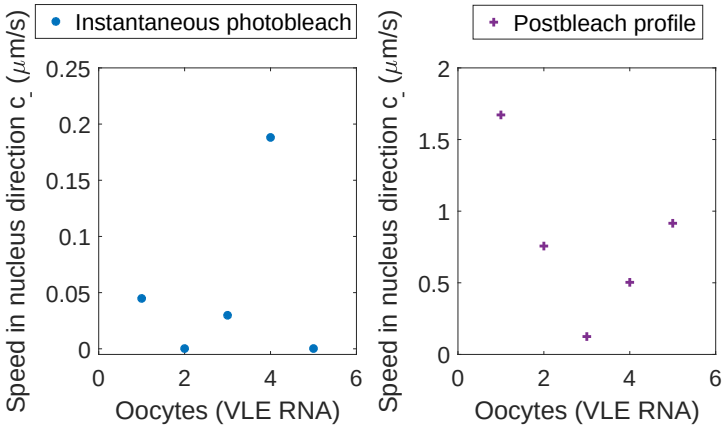

Figure 2.15: Estimated speeds in the animal pole direction in region 3 for VLE RNA using the instantaneous photobleach and experimental photobleach profile initial conditions (using Approach 1, from [1]).

diffusion coefficient d, leading to fits with no useful biological conclusions. This suggests that stationary states are essential in the dynamics of β-globin RNA in the cytoplasm, and that active transport is unlikely to occur for this nonlocalizing RNA.

We therefore perform parameter estimation separately for FRAP data from each individual oocyte using reaction-diffusion equations (2.9) and initial conditions (2.12) corresponding to three bleach spots (see Figure 2.7). Note that given the experimental postbleach profile (Figure 2.7) and initial conditions (2.12), we use Approach 3 described in §2.3.3, so that for each individual oocyte we consider all three regions simultaneously. This approach yields consistent results for diffusion coefficient d, which averages 2 $\mu m^2/s$ (with standard deviation 1.3 $\mu m^2/s$) across 9 oocyte trials in an experimental set, and 2.8 $\mu m^2/s$ (with standard deviation 2 $\mu m^2/s$) for a second set of oocytes. Sample fits of the data are included in Figure 2.16.

The reaction-diffusion model (2.9) is also applied to individual nocodazole-treated oocyte FRAP recoveries (using Approach 1, see §2.3.3). This predicts an average RNA diffusion estimate of 2.3 $\mu m^2/s$ (with standard deviation 1.38 $\mu m^2/s$), similar to the estimate for non-localizing β-globin RNA.

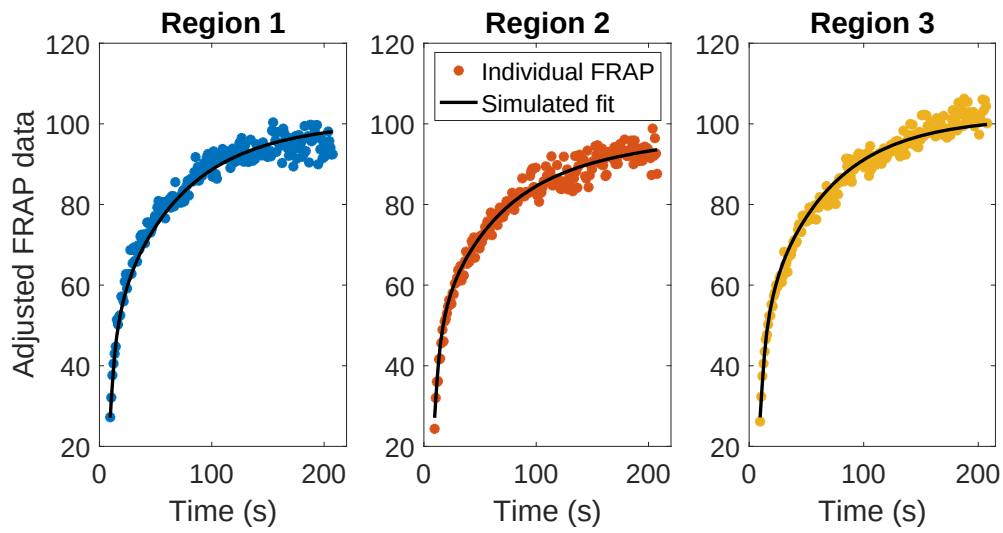

Figure 2.16: Sample oocyte β-globin RNA FRAP data for the 3 bleached regions is fit using Approach 3 and our parameter estimation procedure. Estimated parameters are $D = 2.77\ \mu m^2/s$, $\beta_1 = 0.03\ s^{-1}$ and $\beta_2 = 0.05\ s^{-1}$. Note that all three bleach spots are fit simultaneously and yield a single set of estimated parameters for one oocyte (from [1]).

Localizing VLE RNA: VLE RNA localizes at the vegetal cortex of *Xenopus* oocytes, and its dynamics are driven by both diffusion and transport by molecular motor proteins. Therefore, we first fit FRAP data from three regions in the vegetal cytoplasm (see Figure 2.6A) to the 2-state model (2.7). It is worth noting that we perform parameter estimation fits for VLE RNA by considering individual data (Approach 1) or averaged data (Approach 2), as specified. We recall from §2.3.3 that Approach 1 considers FRAP data from a specific cytoplasm region in an individual oocyte, and Approach 2 considers FRAP data from a specific region, averaged over a set of oocytes.

Using Approach 2, we estimate speed c, diffusion coefficient d, unbinding rate β_1 and binding rate β_2 (see Table 2.4 and fit in Figure 2.17A). Figure 2.17B shows the fit using Approach 1 to determine the same key parameters for an individual oocyte. We note that estimates of velocity c and diffusion coefficient d are fairly consistent across regions in the cytoplasm, and we provide estimates for averages of additional average experimental data in Table 2.5. The RNA is predicted to be moving towards

Table 2.4: Estimated parameters for FRAP WT average data (set of 5 oocytes) using Approach 2 and model (2.7) for VLE RNA (from [1]).

Region	c $(\mu m/s)$	d $(\mu m^2 s^{-1})$	β_1 (s^{-1})	β_2 (s^{-1})
1	0.05	0.26	2.3e−14	0.006
2	0.09	1.42	0.003	0.0007
3	0.07	0.83	4e−5	1.4e−6

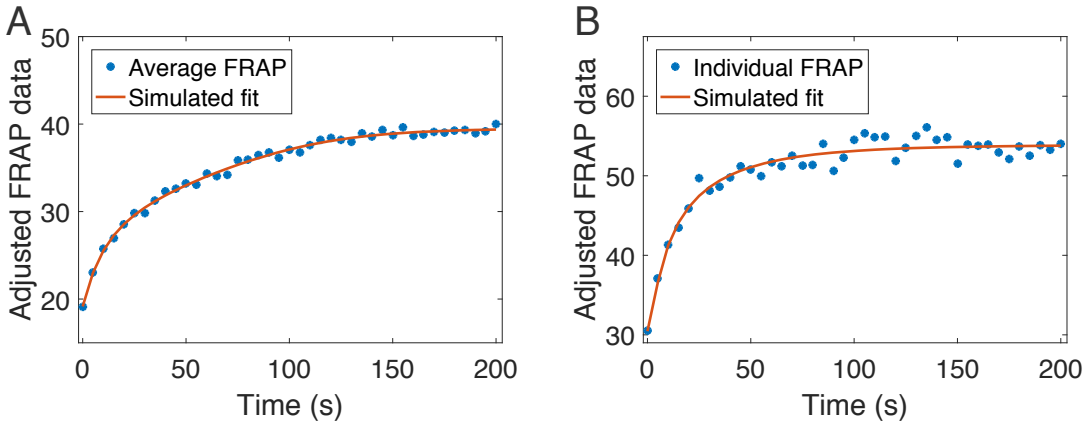

Figure 2.17: A: Average FRAP data for VLE RNA in Region 2 in 5 WT oocytes is fit with model (2.7) using Approach 2; estimated parameters are $c = 0.09\ \mu m/s$, $d = 1.42\ \mu m^2/s$, $\beta_1 = 0.003\ s^{-1}$ and $\beta_2 = 8\text{e}{-4}\ s^{-1}$. B: Sample FRAP data for VLE RNA in Region 3 in an individual WT oocyte is fit with model (2.8) using Approach 1; estimated parameters are $c_+ = 0.51\ \mu m/s$, $c_- = 0.91\ \mu m/s$, $d = 2.29\ \mu m^2/s$, $\alpha_+ = 1\text{e}{-5}\ s^{-1}$, $\alpha_- = 2\text{e}{-5}\ s^{-1}$, $\delta_+ = 0.1\ s^{-1}$ and $\delta_- = 0.02\ s^{-1}$.

the cortex in Region 1 (upper vegetal cytoplasm), while in Regions 2 and 3 (mid and lower vegetal cytoplasm), a higher proportion of particles are diffusing. This is consistent with the hypothesis in [3], which states that transport by motor proteins in the direction of the vegetal cortex is more effective closer to the nucleus (Regions 1 and 2).

As outlined in 2.3.3, we then use these estimates to inform initial guesses for velocity c_+ in the vegetal cortex direction and diffusion d in the 4-state model (2.8). We perform ample parameter sweeps for rates α_+, α_-, δ_+, δ_- and velocity c_- in the nucleus direction. Parameter estimation for individual FRAP curves using Approach 1 yields consistent estimates, and the results for averaged FRAP trial data using Approach 2 are summarized in Table 2.6. We note that the diffusion coefficients are consistent across regions for all sets of oocytes, and that the velocity towards the

Table 2.5: Estimated parameters for FRAP WT average data based on 3 sets of 5 oocytes each using model (2.7) and Approach 2 for VLE RNA, for regions 1-3 in the cytoplasm (see Figure 2.6A). The superscripts I, II, and III correspond to the index of the average FRAP data set under consideration. Quantities v (effective speed) and σ^2 (effective diffusion) are introduced in §3.2.1 and the predicted values are discussed in §4.3.

	Region								
Parameter	1^{I}	1^{II}	1^{III}	2^{I}	2^{II}	2^{III}	3^{I}	3^{II}	3^{III}
c $(\mu m/s)$	0.05	0.12	0.08	0.09	0.1	0.12	0.07	3e−4	0.04
d $(\mu m^2 s^{-1})$	0.26	1.63	0.37	1.42	1.02	0.99	0.83	2.64	1.85
v $(\mu m/s)$	0.05	0.12	0.05	0.02	0.1	0.1	0.002	0.0001	4e−6
σ^2 $(\mu m^2 s^{-1})$	5e−12	1e−7	0.36	3.03	0.005	0.37	8.53	2.62	3.71

Table 2.6: Estimated parameters for FRAP WT average data based on 3 sets of 5 oocytes each using model (2.8) and Approach 2 for VLE RNA, for all three regions in the cytoplasm (see Figure 2.6). The superscripts I, II, and III correspond to the index of the average FRAP data set under consideration. The sets considered are the same as used for parameter estimation using model (2.7) in Table 2.4. Quantities v (effective speed) and σ^2 (effective diffusion) are introduced in §3.2.1 and the predicted values are discussed in §4.3.

	Region								
Parameter	1^{I}	1^{II}	1^{III}	2^{I}	2^{II}	2^{III}	3^{I}	3^{II}	3^{III}
c_+ $(\mu m/s)$	0.37	1.06	0.08	1.16	0.9	0.18	2.39	0.62	0.27
c_- $(\mu m/s)$	0.04	4e−9	4e−6	0.15	0.43	0.009	2.93	1.96	0.89
d $(\mu m^2 s^{-1})$	1.82	9.85	6.98	7.45	7.446	3.41	1.51	9.9	0.95
v $(\mu m/s)$	0.07	0.12	0.05	0.14	0.08	0.077	-0.39	-0.22	-0.18
σ^2 $(\mu m^2 s^{-1})$	0.39	1.22	0.38	1.36	1.68	0.67	4.18	6.5	2.42

animal pole is consistently higher in region 3 compared to regions 1 and 2. A sample FRAP data fit for an individual oocyte is provided in Figure 2.17B.

We note that the ample parameter sweeps and local refinements allow us to determine best choices for initial guesses in parameter estimation. Optimizations carried out in parallel for these guesses yield best parameter fits for velocities and the diffusion coefficient that are consistent across trials. While the estimated kinetic rates have a larger variance across individual oocyte fits, they yield consistent biological predictions when combined in relevant quantities such as fractions of particles in each state and asymptotic velocity and diffusion (see §4.3).

2.3.7 Sensitivity analysis

In this section, we illustrate results of a global sensitivity analysis performed on the input parameters for fitting average VLE RNA data using the 2-state model (2.7) and the 4-state model (2.8) with Approach 2. Given that parameters such as speed c, diffusion coefficient d, and reaction rates β_1 and β_2 are either not estimated in the literature for this system or their values vary on orders of magnitude, we use a global method which investigates sensitivity across a whole range of input parameter variation [55].

In particular, we use the Sobol indices method, which requires the ANOVA representation of the model function of interest, and consists of the decomposition of this function into sums of elementary functions [55, 56]. In this setting, we assume that a model has n parameters, and that the model results are described by the function $y = f(x)$, where $x = (x_1, x_2, \ldots, x_n)$ is a point in an n-dimensional hypercube with each x_i a parameter of the system [56]. In our case, x consists of the parameters that we are estimating for each model, and y is the residual of the actual FRAP data and the FRAP curve generated by the model with these parameters. The ANOVA representation below assumes that each parameter $x_i \in I = [0, 1]$, however in practice this can easily be extended to the interval appropriate for each parameter.

The Sobol method considers the representation for $f(x)$ in I^n [56]:

$$f(x) = f_0 + \sum_i f_i(x_i) + \sum_{i<j} f_{ij}(x_i, x_j) + \ldots + f_{12\cdots n}(x_1, x_2, \ldots, x_n)\,. \tag{2.14}$$

This is called the ANOVA representation if additionally the summand functions satisfy [56]:

$$\int_0^1 f_{i_1\cdots i_s}(x_{i_1}, \ldots, x_{i_s})dx_k = 0 \tag{2.15}$$

for $k = i_1, \ldots, i_s$. This condition leads to a unique representation where $f_{i_1 \cdots i_s}(x_{i_1}, \ldots, x_{i_s})$ can be expressed as [56]:

$$f_0 = \int f(x) dx \,, \tag{2.16}$$

$$f_i(x_i) = \int f(x) \prod_{k \neq i} dx_k - f_0 \,, \tag{2.17}$$

$$f_{ij}(x_i, x_j) = \int f(x) \prod_{k \neq i,j} dx_k - f_0 - f_i(x_i) - f_j(x_j) \,, \tag{2.18}$$

etc.

Assuming that f is square integrable, so are $f_{i_1,\ldots,i_s}$, so that squaring and integration of equation (2.14) over I^n gives [56]:

$$\int f^2(x) dx - f_0^2 = \sum_{s=1}^{n} \sum_{i_1 < \ldots < i_s}^{n} \int f_{i_1 \cdots i_s}^2 dx_{i_1} \ldots dx_{i_s} = \sum_{s=1}^{n} \sum_{i_1 < \ldots < i_s}^{n} D_{i_1 \cdots i_s} \,. \tag{2.19}$$

$D_{i_1 \cdots i_s}$ and $D = \int f^2(x) dx - f_0^2$ are called variances [56], and are used to determine the Sobol sensitivity indices [55, 56], which allow us to examine both the ranking of the individual parameter sensitivities and the parameter coupling. The sensitivity indices are given by:

$$S_{i_1 \cdots i_s} = \frac{D_{i_1 \cdots i_s}}{D} \,, \tag{2.20}$$

so that $\sum_{s=1}^{n} \sum_{i_1 < \ldots < i_s}^{n} S_{i_1 \cdots i_s} = 1$.

The meaning of these indices is that $S_{i_1 \cdots i_s}$ measures the global sensitivity to the set of parameters $x_{i_1}, \ldots, x_{i_s}$. In particular, the indices of interest for our setting are S_i, which measure the global sensitivity of the residual function to each parameter i, and $S_{i,j}$, which measure the effect of varying x_i and x_j simultaneously, in addition to the effect of the individual variations of these parameters. Following [4], the parameter space is explored in this implementation of the Sobol method by using

stochastic collocation [5, 56–58], which means that this method uses sample points in the hypercube consisting of quadrature points from a tensor product of the 1D quadrature rule computed with the algorithm in [58].

Figure 2.18 shows the first and second order Sobol sensitivity indices as well as a visualization tool [4, 5] for these indices for parameter estimates obtained using the 2-state model on wild-type average FRAP data from Region 2 (mid-cytoplasm) in the oocytes. Recall that the objective function for this sensitivity analysis is the residual between the data generated with the tested parameters and the actual FRAP data. We note that in the figure on the right, the sensitivity of a single parameter is denoted by a circle, whose diameter is equal to the sensitivity of that parameter [4, 5]. The fill color of the circles indicates whether the residual increases with an increase in the parameter (white) or decreases with an increase in the parameter (black), respectively. The lines connecting two circles show the interaction of two parameters, where the thickness of the segment is equal to the sensitivity of the interaction pair. These lines measure how much the results will be changed when two parameters are changed simultaneously [4, 5].

Figure 2.18 thus predicts that in region 2, the speed and diffusion coefficient of the mRNA particles are the most sensitive parameters. It is worth noting that in our simulations, the estimates for speed and diffusion are also the most consistent across individual oocytes. In addition, in §3.2.2 we introduce quantities such as the effective speed and diffusion of mRNA particles, which combine these individual parameters into more complex expressions; in §4.3, we observe that these quantities are consistent within regions as well. We note that the pairwise sensitivity of speed c with rates β_1 and β_2 is consistent with the formula for the effective velocity $v = c\frac{\beta_2}{\beta_1+\beta_2}$. Moreover, the sensitivity of the diffusion coefficient d, the reaction rates β_1 and β_2, and their interaction is also consistent with the influence of these parameters on the expression

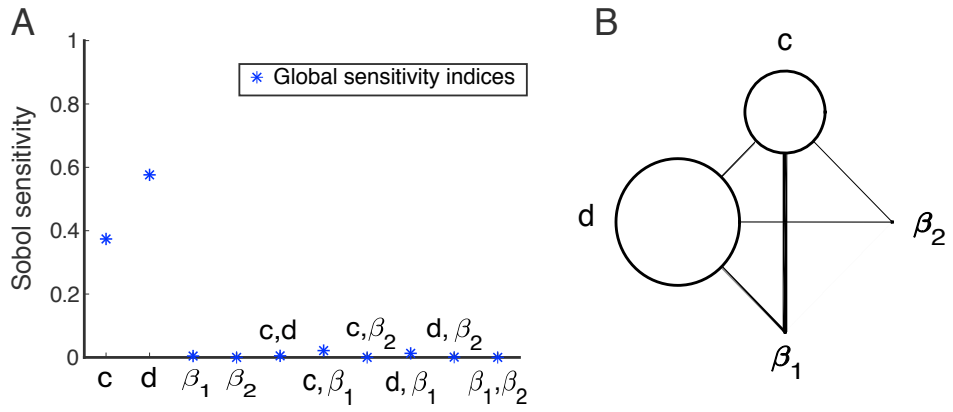

Figure 2.18: A: Sobol sensitivity indices for parameter estimates in Region 2 of average VLE RNA data using the 2-state model (2.7). B: Visualization of the sensitivity indices where circles show single parameter sensitivity, color shows whether the residual increases with an increase in the parameter (white) or decreases with an increase in the parameter (black), and line thickness shows sensitivity of the two-parameter interaction pair [4, 5].

for the effective diffusion of the particles: $\sigma^2 = 2d\frac{\beta_1}{\beta_1+\beta_2} + 2c^2\frac{\beta_1\beta_2}{(\beta_1+\beta_2)^3}$.

Given that effective speed v and effective diffusion σ^2 are consistent quantities that determine the dynamics of mRNA particles for large time, an alternative way to carry out the sensitivity analysis is by considering the sensitivity with respect to these quantities. Knowledge of v, σ^2, β_1/β_2 and β_2 determines the values of speed c and diffusion coefficient d using:

$$c = v\left(1 + \frac{\beta_1}{\beta_2}\right), \tag{2.21}$$

$$d = \frac{\sigma^2}{2}\left(1 + \left(\frac{\beta_1}{\beta_2}\right)^{-1}\right) - \frac{v^2}{\beta_2}, \tag{2.22}$$

which are necessary in solving the PDE models numerically. Figure 2.19 shows the results of this sensitivity analysis for the 2-state model using the same average FRAP data used for the analysis in Figure 2.18. We note that in this approach, the interactions of the effective speed and effective diffusion with the variations in transition rates (in particular the binding rate β_2) are more evident.

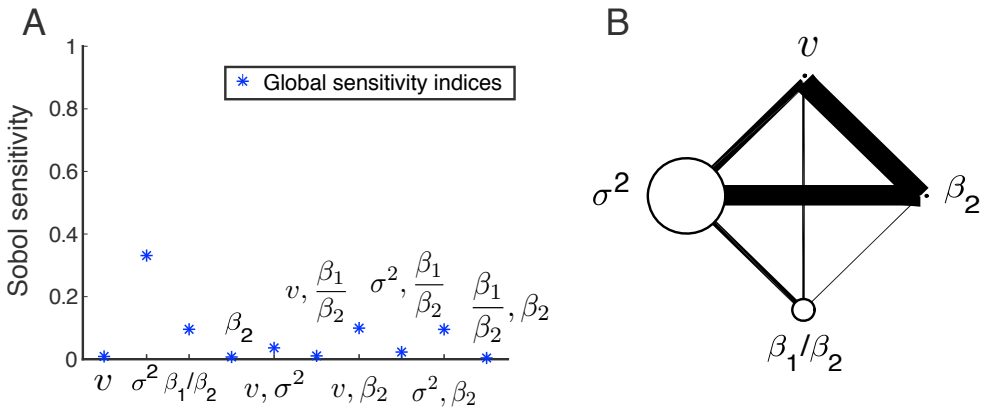

Figure 2.19: A: Sobol sensitivity indices for parameter estimates in Region 2 of average VLE RNA data using the 2-state model (2.7). B: Visualization of the sensitivity indices is described in the legend of Figure 2.18 and in [4, 5].

For the 4-state model, we carried out two sensitivity analyses on two sets of parameters following the approach in [4, 5]; this allowed us to reduce the computation time while still assessing the sensitivity to all the reaction rates that we estimate. We first focused on a subset of 5 parameters (c_+, c_-, d, α_+ and α_-), and Figure 2.20 illustrates the sensitivity indices of these parameters and their interactions for average VLE RNA data in Region 1 (under the nucleus), Region 2 (mid-cytoplasm), and Region 3 (lower cytoplasm), see Figure 2.6. It is worth noting that the velocity in the vegetal direction c_+ is most sensitive in Region 1, where experiments suggest that the mRNA transport may be primarily unidirectional [3], while the interaction of velocities in the animal and vegetal directions (c_- and c_+) becomes considerably more sensitive in Region 3, where bidirectional transport is hypothesized [3]. Furthermore, we recall that parameters α_+ and α_- correspond to the rates of switching from the paused state to one of the moving states (see Figure 2.5B); the sensitivity of the system to these parameters for which there are virtually no values in the literature highlights the importance of using a parameter estimation approach as described in §2.3.3 to determine these values. These parameters determine the expected time spent in a paused state ($\frac{1}{\alpha_+ + \alpha_-}$), which is difficult to obtain experimentally and is derived in §4.1. Given that the expressions for effective velocity and diffusion are

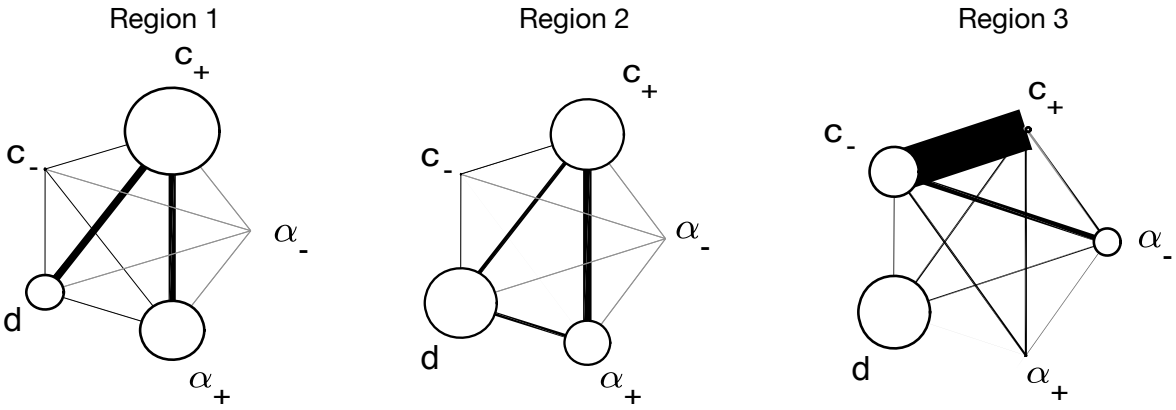

Figure 2.20: Visualization of the sensitivity indices for parameter estimates for c_+, c_-, d, α_+ and α_- in Region 1 (left), Region 2 (center) and Region 3 (right) of average WT data using the 4-state model (2.8). A description of this visualization method from [4, 5] is included in Figure 2.18 and in the text.

more complex in the case of the 4-state system (see §3.2.2), it is difficult to assess the insights the Sobol indices provide on these asymptotic quantities for this model.

We then focused on calculating the sensitivity indices for the subset of parameters c_+, c_-, d, δ_+ and δ_-. Figure 2.21 confirms that the interaction between the velocities in the animal and vegetal directions (c_- and c_+) is most sensitive in Region 3, closer to the cortex, as suggested by the results in Figure 2.20 as well. We note that the pairwise interaction between speed c_+ in the moving down state and rate δ_+ leaving this state to a paused state is significant in Regions 2 and 3; these parameters appear together in the expression for the average run length of the RNA cargo moving down on microtubules: $\frac{c_+}{\gamma_+ + \delta_+}$, see 4.1; note that γ_+ is taken as a fixed parameter from the literature (see Table 2.2). Given that our predictions of average run lengths agree well with experimentally-reported average distances covered by kinesin and dynein motors on microtubule filaments (see §4.3), it is important to note that the model is sensitive to the interaction between these estimated parameters. Similarly, the interaction between speed c_- in the moving up state and rate δ_- leaving this state to a paused state is considerable in Region 3 of Figure 2.21. Since most movement in the animal direction is hypothesized to occur in this region, it is not surprising

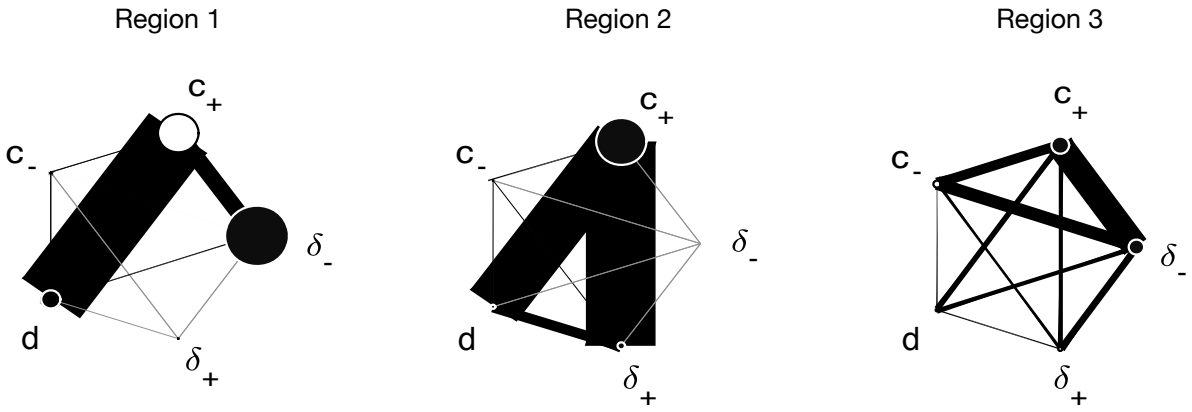

Figure 2.21: Visualization of the sensitivity indices for parameter estimates for c_+, c_-, d, δ_+ and δ_- in Region 1 (left), Region 2 (center) and Region 3 (right) of average VLE RNA data using the 4-state model (2.8). A description of this visualization method from [4, 5] is included in Figure 2.18 and in the text.

that this interaction is important in the estimation of the average run lengths in this direction.

The stochastic collocation method requires choosing a number of collocation points. Given that sensitivity indices are identical when using 5 and 6 collocation points, we used 5 collocation points for all sensitivity analyses in this section. For instance, in the case of the 2-state model with 4 key parameters, this means that the algorithm for calculating the Sobol indices requires 5^4 function evaluations. The parameters were varied between $\hat{x}/10$ and $10\hat{x}$ (where $\hat{x}$ is the parameter estimate for $x = c$, d, β_1 and β_2), to test the sensitivity of the residual fit with respect to the magnitude of the estimated parameter values. To obtain Figure 2.19, the effective speed and diffusion were varied in the smaller range $[\hat{x}/2, 2\hat{x}]$ to ensure that the diffusion coefficient in equation (2.22) remains positive.

Chapter Three

Large-time Analysis of Deterministic and Stochastic Models of mRNA Dynamics

3.1 Overview of previous results

In this chapter we consider the behavior of mRNA particles undergoing advection, diffusion, and reaction for large time. The methods in §3.2.1 will allow us to conclude that the PDE model systems considered have approximate traveling wave solutions in the limit of large time, so that the position of an mRNA particle is completely determined by its mean and variance.

These approximations for PDE models of active transport have been previously explored in [42] using singular perturbation theory and rigorously proven in [43] for linear reaction-hyperbolic equations. [44] extended the analytical convergence results to a linear reaction-hyperbolic system coupled with a diffusion equation. [40, 41, 59] introduce a quasi-steady-state (QSS) method for reducing linear reaction-diffusion-advection equations to a scalar Fokker Plank equation under certain assumptions on the magnitude of the state transition rates relative to the velocities of the moving states. Probabilistic methods are also used to study the approximate traveling wave solutions for axonal transport in [60, 61]. Specific systems of equations are primarily studied in the context of one-dimensional intracellular transport of organelles and vesicles along the axon of nerve cells in [38, 39, 62, 63]. Our work in §3.2 goes beyond these approaches by considering a general two-dimensional advection-reaction-diffusion system with no assumptions on the parameters. The main result provides analytical expressions for the effective velocity and diffusion of the particles at large times given arbitrary numbers of advection, reaction, and diffusion states.

3.2 Large-time Fourier analysis

The results in this section are published in [1].

3.2.1 Calculation of effective speed and velocity for general models

In this section, our goal is to use large-time Fourier analysis to determine the asymptotic behavior of solutions to PDE models of active transport for large time. We have the following result:

Theorem 3.1. *Consider the advection-reaction-diffusion equation system:*

$$\frac{\partial \boldsymbol{u}(y,t)}{\partial t} = A\boldsymbol{u} + C\partial_y \boldsymbol{u} + D\Delta \boldsymbol{u}\,, \tag{3.1}$$

where $\boldsymbol{u}$ is an n-by-1 column vector of all populations of particles with different dynamic behavior, and $A, C, D \in \mathbb{R}^{n\times n}$. Assume that C and D are diagonal with real entries corresponding to velocities, and with positive real entries for diffusion coefficients, respectively, of the n populations. Further assume that A is the singular matrix of transition rates between the n states, admitting one zero eigenvalue with algebraic and geometric multiplicity 1.

Given a δ-function or Gaussian initial condition (say, $u_l(y,0) = \delta(y)$ for each population l), the solution for large time for each population has the form

$$u_l(y,t) = \frac{1}{\sqrt{2\pi a_2 t}} e^{-\frac{(y+a_1 t)^2}{2a_2 t}}\,,$$

where a_1 is the effective velocity and $2a_2$ is the effective diffusion of the particles in

the moving Gaussian solution. In the limit of $t \to \infty$, the effective velocity (denoted as v) and effective diffusion (denoted as σ^2) are given by:

$$v = a_1 = \frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} \tag{3.2}$$

$$\sigma^2 = 2a_2 = 2\frac{\langle \boldsymbol{\psi_0}, (D - C\tilde{A}^{-1}\tilde{C})\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle}\,, \tag{3.3}$$

where $\boldsymbol{u_0}$ is the eigenvector of the zero eigenvalue of A, $\boldsymbol{\psi_0}$ is the eigenvector corresponding to the zero eigenvalue of the adjoint matrix A^*, $\tilde{A}$ corresponds to the projection of matrix A on its range, and $\tilde{C}\boldsymbol{u_0} = \left(C - \frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} I\right)\boldsymbol{u_0}$.

Proof. Considering the ansatz

$$(u_1, u_2, \ldots, u_n)^T(y,t) = e^{\lambda t}e^{\nu y}\tilde{\boldsymbol{u}}_{\boldsymbol{0}}\,, \tag{3.4}$$

with $\nu = ik$, equation (3.1) becomes:

$$(A + \nu C + \nu^2 D - \lambda I)\tilde{\boldsymbol{u}}_{\boldsymbol{0}} = 0\,. \tag{3.5}$$

Let $\boldsymbol{u_0}$ be the eigenvector of the zero eigenvalue of A, and $\boldsymbol{v}$ be in the generalized eigenspace $\mathcal{V}$ corresponding to all non-zero eigenvalues of A. Let $\boldsymbol{\psi_0}$ be the eigenvector corresponding to the zero eigenvalue of the adjoint matrix A^*.

Taking $\tilde{\boldsymbol{u}}_{\boldsymbol{0}} = a\boldsymbol{u_0} + \boldsymbol{v}$ allows us to apply a Lyapunov-Schmidt reduction to equation (3.5) by projecting it onto the eigenspace $\mathcal{V}$ and the space spanned by $\boldsymbol{u_0}$. We proceed with these projections as follows:

(i) Projection onto $\mathcal{V}$-space of equation eqn is given by:

$$\text{eqn} - \frac{\langle \boldsymbol{\psi_0}, \text{eqn} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} \boldsymbol{u_0} \,.$$

Here eqn refers to equation (3.5). This gives:

$$\begin{aligned} a(A &+ \nu C + \nu^2 D - \lambda I)\boldsymbol{u_0} + (A + \nu C + \nu^2 D - \lambda I)\boldsymbol{v} \\ &- a\frac{\langle \boldsymbol{\psi_0}, (\nu C + \nu^2 D - \lambda I)\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} \boldsymbol{u_0} - \frac{\langle \boldsymbol{\psi_0}, (\nu C + \nu^2 D - \lambda I)\boldsymbol{v} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} \boldsymbol{u_0} = 0 \,. \end{aligned} \tag{3.6}$$

Note that $A\boldsymbol{u_0} = 0$ and $A^*\boldsymbol{\psi_0} = 0$ by definition, and $\langle \boldsymbol{\psi_0}, \lambda \boldsymbol{v} \rangle = \lambda \langle \boldsymbol{\psi_0}, \boldsymbol{v} \rangle = 0$ $\forall \boldsymbol{v} \in \mathcal{V}$.

It can also be shown for the fourth term in (3.6) that

$$\frac{\langle \boldsymbol{\psi_0}, (\nu C + \nu^2 D)\boldsymbol{v} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} = \nu B_\nu \boldsymbol{v} \sim \mathcal{O}(\nu)$$

using the Cauchy-Schwartz inequality. Here B_ν is the operator sending $\boldsymbol{x} \to \frac{\langle \boldsymbol{\psi_0}, (C+\nu D)\boldsymbol{x} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle}$. Similarly, the third term in (3.6) is

$$a\frac{\langle \boldsymbol{\psi_0}, (\nu C + \nu^2 D - \lambda I)\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} = -a\lambda + a\nu B_\nu \boldsymbol{u_0} \sim -a\lambda + \mathcal{O}(\nu) \,.$$

Combining these observations yields:

$$a(\nu C + \nu^2 D - \lambda I)\boldsymbol{u_0} + (A + \nu C + \nu^2 D - \lambda I - \nu \boldsymbol{u_0} B_\nu)\boldsymbol{v} + a\lambda \boldsymbol{u_0} - a\nu (B_\nu \boldsymbol{u_0})\boldsymbol{u_0} = 0 \,.$$

Simplifying and separating $\boldsymbol{v}$ gives:

$$\begin{aligned}
\boldsymbol{v} &= -a(A+\nu(C-\boldsymbol{u_0}B_\nu)+\nu^2 D-\lambda I)^{-1}(\nu C+\nu^2 D-\nu B_\nu \boldsymbol{u_0} I)\boldsymbol{u_0} \\
&\approx -a(\tilde{A}^{-1}+\mathcal{O}(\nu+\lambda))(\nu C-\nu B_\nu \boldsymbol{u_0} I+\nu^2 D)\boldsymbol{u_0} \\
&\approx -a\nu(\tilde{A}^{-1}+\mathcal{O}(\nu+\lambda))\left(C-\frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle} I+\mathcal{O}(\nu)\right)\boldsymbol{u_0}\,. \qquad (3.7)
\end{aligned}$$

Matrix $\tilde{A}$ corresponds to the projection of matrix A on space $\mathcal{V}$, so that $\tilde{A}$ is invertible. The inversion is allowed because the left-hand side ($\boldsymbol{v}$) is in the range of matrix A, and $\tilde{C}\boldsymbol{u_0} = \left(C-\frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle} I\right)\boldsymbol{u_0}$ is also readily shown to be in the range of A. For this, we evaluate the projections onto $\mathcal{V}$-space and $\boldsymbol{u_0}$-space:

$$\text{eqn}-\frac{\langle \boldsymbol{\psi_0}, \text{eqn}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle}\boldsymbol{u_0}\bigg|_{\text{eqn}=\left(C-\frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle} I\right)\boldsymbol{u_0}} = \left(C-\frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle} I\right)\boldsymbol{u_0}\,,$$

$$\langle \boldsymbol{\psi_0}, \text{eqn}\rangle\bigg|_{\text{eqn}=\left(C-\frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle} I\right)\boldsymbol{u_0}} = 0\,,$$

so that $\tilde{C}\boldsymbol{u_0}$ is indeed in the range of $\tilde{A}$.

(ii) Projection onto $\boldsymbol{u_0}$-space is given by:

$$\langle \boldsymbol{\psi_0}, \text{eqn}\rangle\,,$$

where again eqn denotes equation (3.5)

Here, the projection gives:

$$\langle \boldsymbol{\psi_0}, (A+\nu C+\nu^2 D-\lambda I)(a\boldsymbol{u_0}+\boldsymbol{v})\rangle = 0\,.$$

Since $A^*\boldsymbol{\psi_0} = 0$ and $\langle \boldsymbol{\psi_0}, \boldsymbol{v} \rangle = 0$, and using $\boldsymbol{v}$ from (3.7) results in:

$$\langle \boldsymbol{\psi_0}, (-\lambda I + \nu C + \nu^2 D - \nu^2 (C + \nu D) \left[(\tilde{A}^{-1} + \mathcal{O}(\nu + \lambda)) \left(C - \frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} I \right) \right] \boldsymbol{u_0} \rangle = 0 \,. \tag{3.8}$$

The linearity of the inner product gives:

$$-\lambda \langle \psi_0, \boldsymbol{u_0} \rangle + \nu \langle \boldsymbol{\psi_0}, C\boldsymbol{u_0} \rangle + \nu^2 \langle \boldsymbol{\psi_0}, D\boldsymbol{u_0} \rangle - \nu^2 \langle \boldsymbol{\psi_0}, C\tilde{A}^{-1}\tilde{C}\boldsymbol{u_0} \rangle + \mathcal{O}(\nu^2(\nu + \lambda)) = 0 \,.$$

Using the implicit function theorem around $(\nu, \lambda) = (0, 0)$ allows us to write $\lambda = \mathcal{O}(\nu)$, and the higher order term at the end of the equation is $\mathcal{O}(\nu^3)$. Then λ is given by:

$$\boxed{\lambda = \nu \frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} + \nu^2 \left[\frac{\langle \boldsymbol{\psi_0}, (D - C\tilde{A}^{-1}\tilde{C})\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} \right] + \mathcal{O}(\nu^3) \,.} \tag{3.9}$$

Returning to ansatz (3.4), component l of the vector of particle concentrations $\boldsymbol{u}$ is described by:

$$u_l(y, t) = e^{\left(a_1 \nu + \frac{a_2}{2} \nu^2 + \sum_{j=3}^{\infty} a_j \nu^j\right) t} e^{\nu y} \tilde{u}_{0l}(y) \,, \tag{3.10}$$

where

$$\begin{aligned} a_1 &= \frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} \\ a_2 &= 2 \frac{\langle \boldsymbol{\psi_0}, (D - C\tilde{A}^{-1}\tilde{C})\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} \,. \end{aligned}$$

Assuming a Dirac delta function initial condition $u_{0l} = \delta(y)$ (modeling a single particle located at $y = 0$), its Fourier transform in equation (3.10) is $\tilde{u}_{0l} = 1/(\sqrt{2\pi})$. Similar to the approach in [38], this allows us to calculate the concentration of particle

population l by taking the inverse Fourier transform:

$$u_l = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ik(y+a_1 t) - \frac{a_2}{2}k^2 t} \times e^{\sum_{j=3}^{\infty} a_j \nu^j t} \times \frac{1}{\sqrt{2\pi}} dk \,.$$

As in [38], the change of variables $\tilde{y} = y + a_1 t$ and $\tilde{k} = kt^{1/2}$ gives:

$$u_l = \frac{1}{2\pi\sqrt{t}} \int_{-\infty}^{\infty} e^{i\tilde{k}\frac{\tilde{y}}{t^{1/2}} - \frac{a_2}{2}\tilde{k}^2} \times e^{\sum_{j=3}^{\infty} \frac{a_j (i\tilde{k})^j}{t^{j/2-1}}} d\tilde{k} \,.$$

In the second term in the product above, $j/2-1 > 0$, so that the summation vanishes as $t \to \infty$. It is therefore sufficient to calculate:

$$\begin{aligned} u_l &= \frac{1}{2\pi\sqrt{t}} \int_{-\infty}^{\infty} e^{i\tilde{k}\frac{\tilde{y}}{t^{1/2}} - \frac{a_2}{2}\tilde{k}^2} d\tilde{k} \\ &= \frac{1}{\sqrt{2\pi a_2 t}} e^{-\frac{(y+a_1 t)^2}{2a_2 t}} \,. \end{aligned}$$

Since this holds for each population l, the solution of the advection-reaction-diffusion PDEs for large time thus consists of a spreading Gaussian, and the effective velocity and diffusion of the particle behavior is given by:

$$\text{effective velocity} = a_1 = \frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} \tag{3.11}$$

$$\text{effective diffusion} = a_2 = 2\frac{\langle \boldsymbol{\psi_0}, (D - C\tilde{A}^{-1}\tilde{C})\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} \,. \tag{3.12}$$

■

We note that average transport velocity and spreading for the specific equations modeling neurofilament transport are derived in [38] and [39]. The spreading Gaussian solutions for large time have also been investigated for reaction-hyperbolic systems of PDEs in [40, 42–44]. [41] introduces diffusion in the context of tug-of-war

studies for motor-driven transport, with a focus on diffusion in one particle population. The approach outlined above provides analytical expressions for effective velocity and diffusion for large times for a system with arbitrary numbers of particles undergoing diffusion, active transport and reaction.

3.2.2 Expressions for specific models of mRNA dynamics

Effective velocity and diffusion for the 2-state model

We calculate the expressions for effective velocity and diffusion using the 2-state model of particle dynamics given by equations (2.7).

In this case, $C = \begin{pmatrix} c & \\ & 0 \end{pmatrix}$, $D = \begin{pmatrix} 0 & \\ & D \end{pmatrix}$ and $A = \begin{pmatrix} -\beta_1 & \beta_2 \\ \beta_1 & -\beta_2 \end{pmatrix}$.

The eigenvectors of A and A^* in equation (3.9) are given by $\boldsymbol{u_0} = \begin{pmatrix} \beta_2/(\beta_1+\beta_2) \\ \beta_1/(\beta_1+\beta_2) \end{pmatrix}$ and $\boldsymbol{\psi_0} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

This gives that the $\mathcal{O}(\nu)$ term in (3.9) is:

$$a_1 = \frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} = \frac{c\beta_2/(\beta_1+\beta_2)}{1} = \boxed{c\frac{\beta_2}{\beta_1+\beta_2}}, \tag{3.13}$$

which corresponds to the effective velocity in (3.11).

Similarly, the $\mathcal{O}(\nu^2)$ term in (3.9) is:

$$a_2 = 2\frac{\langle \boldsymbol{\psi_0}, (D - C\tilde{A}^{-1}\tilde{C})\boldsymbol{u_0}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle}\,.$$

Note that the non-zero eigenvalue of A is $\lambda_1 = -(\beta_1 + \beta_2)$, and its corresponding eigenvector is $\boldsymbol{v} = (1, -1)^T$. Then $\tilde{A}\boldsymbol{v} = \lambda_1 \boldsymbol{v}$ and thus $\tilde{A}^{-1} = \lambda_1^{-1} = -\frac{1}{\beta_1+\beta_2}$.

Therefore:

$$\begin{aligned}
a_2 = 2\frac{\langle \boldsymbol{\psi_0}, (D - C\tilde{A}^{-1}\tilde{C})\boldsymbol{u_0}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle} &= 2\frac{\langle \boldsymbol{\psi_0}, (D + (1/(\beta_1+\beta_2))C\tilde{C})\boldsymbol{u_0}\rangle}{1} \\
&= 2\langle \boldsymbol{\psi_0}, \left(D + \frac{1}{\beta_1+\beta_2}C\left(C - \frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle}I\right)\right)\boldsymbol{u_0}\rangle \\
&= 2\langle \boldsymbol{\psi_0}, \left(D + \frac{1}{\beta_1+\beta_2}C\left(C - \frac{c\beta_2}{\beta_1+\beta_2}I\right)\right)\boldsymbol{u_0}\rangle \\
&= 2d\frac{\beta_1}{\beta_1+\beta_2} + 2c^2\frac{\beta_1\beta_2}{(\beta_1+\beta_2)^3}\,.
\end{aligned}$$

Then

$$a_2 = \boxed{2d\frac{\beta_1}{\beta_1+\beta_2} + 2c^2\frac{\beta_1\beta_2}{(\beta_1+\beta_2)^3}}, \tag{3.14}$$

which corresponds to the expression for effective diffusion in (3.12).

[61] derive expressions similar to (3.13) and (3.14) for the effective speed and diffusion of an on/off transport particle using stochastic methods. Our analysis yields the additional first term in equation (3.14) compared to the expression for effective spread in [61], which is due to our assumption of diffusion in the off state.

Effective velocity and diffusion for the 4-state model

We also calculate the expressions for effective velocity and diffusion using the 4-state model of intracellular transport described by equations (2.8).

In this case, we have $C = \begin{pmatrix} c_+ & & & \\ & -c_- & & \\ & & 0 & \\ & & & 0 \end{pmatrix}$, $D = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & D \end{pmatrix}$, and transition rate matrix

$$A = \begin{pmatrix} -(\gamma_+ + \delta_+) & 0 & \alpha_+ & \beta_+ \\ 0 & -(\gamma_- + \delta_-) & \alpha_- & \beta_- \\ \delta_+ & \delta_- & -(\alpha_+ + \alpha_-) & 0 \\ \gamma_+ & \gamma_- & 0 & -(\beta_+ + \beta_-) \end{pmatrix}.$$

The eigenvectors of A and A^* in equation (3.9) can also be easily found: $\boldsymbol{\psi_0} = (1,1,1,1)^T$, and $\boldsymbol{u_0}$ corresponds to the proportion of each population at equilibrium (see § 3.2.1). $\boldsymbol{u_0}$ can be normalized so that $\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle = 1$.

This gives the $\mathcal{O}(\nu)$ term in (3.9):

$$\begin{aligned} \frac{\langle \boldsymbol{\psi_0}, C\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle} &= -(\alpha_-\beta_-c_-\delta_+ + \alpha_-\beta_+c_-\delta_+ - \alpha_+\beta_-c_+\delta_- - \alpha_+\beta_+c_+\delta_- + \alpha_-\beta_-c_-\gamma_+ \\ &- \alpha_-\beta_+c_+\gamma_- + \alpha_+\beta_-c_-\gamma_+ - \alpha_+\beta_+c_+\gamma_+) \\ &/ \quad (\alpha_-\beta_-\delta_+ + \alpha_+\beta_-\delta_- + \alpha_-\beta_+\delta_+ + \alpha_+\beta_+\delta_- + \alpha_-\beta_-\gamma_+ + \alpha_-\beta_+\gamma_- + \beta_-\delta_-\delta_+ \\ &+ \alpha_+\beta_-\gamma_+ + \alpha_+\beta_+\gamma_- + \beta_+\delta_-\delta_+ + \alpha_-\delta_+\gamma_- + \alpha_+\delta_-\gamma_+ + \beta_-\delta_-\gamma_+ + \beta_+\delta_+\gamma_- \\ &+ \alpha_-\gamma_-\gamma_+ + \alpha_+\gamma_-\gamma_+)\,, \end{aligned} \tag{3.15}$$

which is the effective velocity in the 4-state example. Note that the above expression can be calculated using Matlab or Mathematica.

The $\mathcal{O}(\nu^2)$ term in equation (3.9) requires calculation of $a_2 = \frac{\langle \boldsymbol{\psi_0}, (D - C\tilde{A}^{-1}\tilde{C})\boldsymbol{u_0} \rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0} \rangle}$. Noting that $R(A) = (R(\boldsymbol{\psi_0}))^{\perp}$, we seek a matrix representation of $\tilde{A}$ using a basis in the complement of $\boldsymbol{\psi_0} = (1,1,1,1)^T$. A choice for this basis is $\boldsymbol{v_{01}} = (1,0,-1,0)^T$, $\boldsymbol{v_{02}} = (0,1,0,-1)^T$, and $\boldsymbol{v_{03}} = (1,0,0,-1)^T$, yielding:

$$\begin{aligned}
\tilde{A}\boldsymbol{v_{01}} &= \alpha_1 \boldsymbol{v_{01}} + \alpha_2 \boldsymbol{v_{02}} + \alpha_3 \boldsymbol{v_{03}} \,, \\
\tilde{A}\boldsymbol{v_{02}} &= \beta_1 \boldsymbol{v_{01}} + \beta_2 \boldsymbol{v_{02}} + \beta_3 \boldsymbol{v_{03}} \,, \\
\tilde{A}\boldsymbol{v_{03}} &= \gamma_1 \boldsymbol{v_{01}} + \gamma_2 \boldsymbol{v_{02}} + \gamma_3 \boldsymbol{v_{03}} \,.
\end{aligned}$$

Note that α_i, β_i, γ_i have simple expressions that Matlab's or Mathematica's symbolic environments can readily find. This is done by solving equations of the form $V_0(\alpha_1, \alpha_2, \alpha_3)^T = \tilde{A}\boldsymbol{v_{01}}$, with $V_0 = (\boldsymbol{v_{01}}, \boldsymbol{v_{02}}, \boldsymbol{v_{03}})$.

Since we are interested in $\tilde{A}^{-1}\tilde{C}\boldsymbol{u_0}$, we seek $\bar{\boldsymbol{x}} = \bar{x}_1 \boldsymbol{v_{01}} + \bar{x}_2 \boldsymbol{v_{02}} + \bar{x}_3 \boldsymbol{v_{03}}$ such that $\tilde{A}\bar{\boldsymbol{x}} = \tilde{C}\boldsymbol{u_0}$. Writing $\tilde{C}\boldsymbol{u_0} = \boldsymbol{x} = x_1 \boldsymbol{v_{01}} + x_2 \boldsymbol{v_{02}} + x_3 \boldsymbol{v_{03}}$ gives:

$$\begin{aligned}
\alpha_1 \bar{x}_1 + \beta_1 \bar{x}_2 + \gamma_1 \bar{x}_3 &= x_1 \,, \\
\alpha_2 \bar{x}_1 + \beta_2 \bar{x}_2 + \gamma_2 \bar{x}_3 &= x_2 \,, \\
\alpha_3 \bar{x}_1 + \beta_3 \bar{x}_2 + \gamma_3 \bar{x}_3 &= x_3 \,.
\end{aligned}$$

Note that x_i can also be readily found for this example by solving $V_0(x_1, x_2, x_3)^T =$

$\tilde{C}\boldsymbol{u_0}$ in Matlab. The equation for $\bar{x}_i$ is therefore:

$$\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} . \tag{3.16}$$

Given that α_i, β_i, γ_i and x_i have expressions that can be determined as described above, this linear system can be solved in Matlab or Mathematica. This recovers $\bar{\boldsymbol{x}} = \tilde{A}^{-1}\tilde{C}\boldsymbol{u_0} = \bar{x}_1\boldsymbol{v_{01}} + \bar{x}_2\boldsymbol{v_{02}} + \bar{x}_3\boldsymbol{v_{03}}$.

The $\mathcal{O}(\nu^2)$ term in the expression for λ is:

$$\begin{aligned} a_2 = 2\frac{\langle \boldsymbol{\psi_0}, (D - C\tilde{A}^{-1}\tilde{C})\boldsymbol{u_0}\rangle}{\langle \boldsymbol{\psi_0}, \boldsymbol{u_0}\rangle} &= 2\frac{\langle \boldsymbol{\psi_0}, (D - C\tilde{A}^{-1}\tilde{C})\boldsymbol{u_0}\rangle}{1} \\ &= 2\langle \boldsymbol{\psi_0}, D\boldsymbol{u_0} - C\bar{\boldsymbol{x}}\rangle \\ &= \boxed{2{\boldsymbol{\psi_0}}^T(D\boldsymbol{u_0} - C\bar{\boldsymbol{x}})\,.} \end{aligned} \tag{3.17}$$

An analytical expression for this term can be found using the symbolic environments in Matlab or Mathematica.

3.3 Large-time analysis using stochastic processes

3.3.1 Renewal reward theory approach for 2-state models

An alternative approach to obtain approximations of the mean and variance of displacement of an mRNA particle after large time t is to consider the stochastic processes theory of renewal rewards.

In this framework, we let T_i be times spent by a particle in movement, and $\tilde{T}_{i+1}$ times spent by the particle in diffusion. Then we model the distributions of the times in each of these states as exponential random variables as follows: $T_i \sim \text{Exp}(\beta_1)$, and $\tilde{T}_{i+1} \sim \text{Exp}(\beta_2)$.

We consider the advection-diffusion cycle $S_i = T_i + \tilde{T}_{i+1}$, so that S_1 are independent and identically-distributed positive random variables and $\{S_i; i = 1, 2, 3...\}$ forms a renewal process. We note that $E[S_1] = \frac{\beta_1+\beta_2}{\beta_1\beta_2}$. We further define:

$$J_n = S_1 + S_2 + ... + S_n \,, \tag{3.18}$$

for $n \geq 1$ and $J_0 = 0$. J_n thus represents the amount of time after n advection-diffusion state jumps. The assumption of equal number of movement and diffusion steps is valid in the limit of large target time T considered in our analysis.

We consider

$$X_t = \sup\{n : J_n \leq t\} \tag{3.19}$$

which represents the number of jumps (and state cycles) by time t, and forms a renewal counting process.

In this setting, the elementary renewal theorem [64] states that

$$\lim_{t\to\infty} \frac{E[X_t]}{t} = \frac{1}{E[S_1]} \,, \tag{3.20}$$

which means that for large target time T, the expected number of jumps is

$$E[X_T] = T\frac{\beta_1\beta_2}{\beta_1+\beta_2} = \frac{T}{\frac{1}{\beta_1}+\frac{1}{\beta_2}} \,.$$

This makes intuitive sense for the number of steps in each state, given the rates of jumping between states.

In the context of renewal processes, we can also define a renewal-reward process as follows:

$$Y_t = \sum_{i=1}^{X_t} W_i \,, \tag{3.21}$$

where W_i are the rewards associated with each cycle, which can depend on the renewal process S_i. In order to model displacement of mRNA particles through movement and diffusion states, we define W_i to depend on the components of S_i (T_i and $\tilde{T}_{i+1}$) as follows:

$$W_i = cT_i + G(0, 2d\tilde{T}_i) \,, \tag{3.22}$$

so that Y_T represents the position of the particle at time T.

Then the elementary renewal theorem for renewal reward processes gives:

$$\lim_{t\to\infty} \frac{E[Y_t]}{t} = \frac{E[W_1]}{E[S_1]} \,, \tag{3.23}$$

so that the mean displacement of the particle after large time T is given by:

$$E[Y_t] = c\frac{\beta_2}{\beta_1 + \beta_2} T \,. \tag{3.24}$$

We note that this agrees with the result using the large time Fourier analysis in §3.2.2 (see (3.13)).

The variance of the particle displacement can also be approximated for large time using a consequence of the central limit theorem for renewal processes. The result

[65] states that

$$\lim_{t\to\infty} Y_t \approx N\left(\frac{E[W_1]}{E[S_1]}t, \frac{\gamma^2 t}{E[S_1]^3}\right). \tag{3.25}$$

provided that

$$\gamma^2 = Var[E[S_1]W_1 - E[W_1]S_1] > 0\,. \tag{3.26}$$

For large time T, this implies that

$$Var[Y_T] = \frac{\gamma^2 T}{E[S_1]^3}\,.$$

We proceed by verifying that condition (3.26) is satisfied in our setting for particles going through movement-diffusion cycles. Since S_1 and W_1 are not linearly independent, we cannot immediately conclude that $\gamma^2 > 0$. We therefore explicitly calculate γ^2:

$$\begin{aligned}\gamma^2 &= Var\left[\frac{\beta_1+\beta_2}{\beta_1\beta_2}(cT_1 + G(0, 2d\tilde{T}_1)) - \frac{c}{\beta_1}(T_1+\tilde{T}_1)\right]\\ &= Var\left[\frac{c}{\beta_2}T_1 + \frac{\beta_1+\beta_2}{\beta_1\beta_2}G(0,2d\tilde{T}_1) - \frac{c}{\beta_1}\tilde{T}_1\right].\end{aligned}$$

The independence of T_1 and $\tilde{T}_1$ gives:

$$\gamma^2 = \frac{c^2}{\beta_2^2}Var[T_1] + Var[Y] = \frac{c^2}{\beta_1^2\beta_2^2} + Var[Y]\,,$$

where $Y = \frac{\beta_1+\beta_2}{\beta_1\beta_2}G(0,2d\tilde{T}_1) - \frac{c}{\beta_1}\tilde{T}_1$.

To calculate $Var[Y] = E[Y^2] - E[Y]^2$, we observe that $E[Y] = -\frac{c}{\beta_1\beta_2}$, and

$$E[Y^2] = E\left[\frac{(\beta_1+\beta_2)^2}{\beta_1^2\beta_2^2}G^2(0,2d\tilde{T}_1) + \frac{c^2}{\beta_1^2}\tilde{T}_1^2 - \frac{2c(\beta_1+\beta_2)}{\beta_1^2\beta_2}\tilde{T}_1 G(0,2d\tilde{T}_1)\right].$$

Using that $E[G^2] = (E[G])^2 + Var[G] = Var[G]$ and $E[\tilde{T}_1^2] = E[\tilde{T}_1]^2 + Var[\tilde{T}_1] = \frac{2}{\beta_2^2}$

yields

$$E[Y^2] = \frac{(\beta_1+\beta_2)^2}{\beta_1^2\beta_2^2}\frac{2d}{\beta_2} + \frac{2c^2}{\beta_1^2\beta_2^2} - \frac{2c(\beta_1+\beta_2)}{\beta_1^2\beta_2}E[\tilde{T}_1 G(0, 2d\tilde{T}_1)]\,.$$

For the last term $E[\tilde{T}_1 G(0, 2d\tilde{T}_1)]$, the joint distribution of the exponential and Gaussian random variables is given by:

$$f_{\tilde{T}_1, G_1(0,2d\tilde{T}_1)}(x,y) = \begin{cases} \frac{\beta_2}{2\sqrt{d\pi x}} e^{-bx-\frac{y^2}{4dx}}, & \text{if } x \geq 0 \\ 0, & \text{else}\,. \end{cases}$$

Then $E[\tilde{T}_1 G(0, 2d\tilde{T}_1)] = \int_0^\infty \int_{-\infty}^\infty xy f_{\tilde{T}_1, G_1(0,2d\tilde{T}_1)}(x,y)dydx = 0$, so that this term brings no contribution to $E[Y^2]$. Therefore:

$$E[Y^2] = \frac{(\beta_1+\beta_2)^2}{\beta_1^2\beta_2^2}\frac{2d}{\beta_2} + \frac{2c^2}{\beta_1^2\beta_2^2}\,,$$

and

$$Var[Y] = \frac{(\beta_1+\beta_2)^2}{\beta_1^2\beta_2^2}\frac{2d}{\beta_2} + \frac{2c^2}{\beta_1^2\beta_2^2} - \frac{c^2}{\beta_1^2\beta_2^2} = \frac{(\beta_1+\beta_2)^2}{\beta_1^2\beta_2^2}\frac{2d}{\beta_2} + \frac{c^2}{\beta_1^2\beta_2^2}\,.$$

This gives

$$\gamma^2 = \frac{(\beta_1+\beta_2)^2}{\beta_1^2\beta_2^2}\frac{2d}{\beta_2} + \frac{2c^2}{\beta_1^2\beta_2^2}\,, \tag{3.27}$$

so that $\gamma^2 > 0$ unless $d = c = 0$, which is not of interest since it corresponds to diffusion and active transport playing no role in the particle dynamics.

A consequence of Theorem (3.25) gives the variance formula

$$Var[Y_T] = \gamma^2 \frac{\beta_1^3\beta_2^3}{(\beta_1+\beta_2)^3}T\,,$$

and using (3.27) yields:

$$Var[Y_T] = 2d\frac{\beta_1}{\beta_1+\beta_2}T + 2c^2\frac{\beta_1\beta_2}{(\beta_1+\beta_2)^3}T\,. \tag{3.28}$$

This variance for large time also agrees with the result using the Fourier large time analysis in §3.2.2 (see (3.14)).

3.3.2 Extension and challenges for general models

While the renewal reward theory in §3.3.1 can be applied to confirm the expressions for effective velocity and diffusion for large time for the 2-state model of dynamics, it is more challenging to use this approach for models with more than 2 states such as the 4-state model in Figure 2.5B. In particular, the time in the i^{th} cycle was easily expressed in the case of the 2-state system as $S_i = T_i + \tilde{T}_{i+1}$, however when considering the 4-state model this first passage time cannot simply be written as a sum of times in individual states.

To model the 4-state system using stochastic processes, we express the dynamics as a process with regenerative increments. We consider a continuous-time stochastic process $\{X(t) : t \geq 0\}$ defined on state space $S = \{0, 1, 2, 3\}$, where each index corresponds to one of the dynamic states modeled: diffusion, movement down, pausing, and movement up. We assume that the process starts in state 0 ($X(0) = 0$) and define renewal times T_n denoting the times when the process returns to state 0. The inter-renewal times $\xi_n = T_n - T_{n-1}$ are i.i.d., so that the process $X(t)$ is regenerative over times T_n [66]. We also denote the number of renewals in $(0, t]$ by $N(t)$, which forms a renewal process with inter-occurrence times ξ_n [66].

We also consider the real-valued process $\{Z(t) : t \geq 0\}$, which corresponds to the reward (displacement) at time t. This process also has regenerative increments over T_n. The strong law of large numbers then gives the following result:
Assume $E[M_n]$ is finite, where $M_n = \sup_{T_{n-1} \leq t \leq T_n} |Z(t) - Z(T_{n-1})|$, with $n \geq 1$. Further assume that $\mu = E[T_1]$ and $a = E[Z(T_1)]$ both exist and are not both infinite. Then

$$t^{-1} Z(t) \to a/\mu \,, \tag{3.29}$$

a.s. as $t \to \infty$ [66]. Since in our case the stochastic process $X(t)$ is a continuous-time Markov chain with reward structure Z_j (where j corresponds to the state index), the calculation of the long-run average reward is simplified to:

$$\lim_{t \to \infty} \frac{Z(t)}{t} = \sum_{j \in S} Z_j p_j \,. \tag{3.30}$$

Here, p_j corresponds to the stationary distribution of the embedded Markov chain of the states [67]. Applying this approach to the 4-state model yields the same expression for the effective velocity as obtained in the previous section with dynamical systems methods (see equation (3.15)).

The central limit theorem for regenerative processes provides insight into the variance of the reward:
Assume that $E[M_1]$ and $\sigma^2 = Var[Z(T_1) - aT_1]$ are finite, and $\sigma > 0$. Then:

$$\frac{Z(t) - at}{\sqrt{t}} \xrightarrow{d} N(0, \sigma^2/\mu) \,, \tag{3.31}$$

as $t \to \infty$ [66]. Therefore, the long run variance of the reward can be obtained by calculating σ^2/μ. Noting that $E[Z(t) - at] = 0$ from (3.29), this is equivalent to

calculating the second moment:

$$E[(Z(T_1) - aT_1)^2] = E[Z(T_1)^2] - 2aE[T_1 Z(T_1)] + a^2 E[T_1^2]\,. \tag{3.32}$$

However, as mentioned at the beginning of this section, the second moments of the first passage time T_1, cycle reward Z_1, and the expected value of the product $T_1 Z_1$ are challenging to calculate, and there are no existing methods for explicitly determining these quantities. As suggested in [66], the variance σ^2 can nonetheless be determined by observing the regenerative process $Z(t)$ up to a large time t, and calculating a confidence interval around the mean a.

We therefore approach this calculation numerically, by setting up a Markov chain of the states in the 4-state model and keeping track of the renewal time T_n and the reward at each renewal time $Z(T_n)$. We then estimate the second moment of $Z(t) - at$ in the following way:

$$\sigma^2 = E[(Z(T_1) - aT_1)^2] = \lim_{t\to\infty} \frac{\sum_{k=1}^{N(t)-1} Z(T_k + 1) - Z(T_k) - a(T_{k+1} - T_k)}{N(t) - 1}\,. \tag{3.33}$$

To calculate the variance of the reward $\frac{\sigma^2}{\mu}$, we also need the mean cycle length $\mu = E[T_1]$. We can use the following well-known proposition:
If $X(t)$ is a positive recurrent CTMC, then the limiting probability distribution p exists, is unique, and is calculated as follows:

$$p_j = \frac{E[H_j]}{E[T_{jj}]}\,, \tag{3.34}$$

where H_j is the amount of time spent in state j during a cycle, and T_{jj} is the cycle length when starting and returning to state j. If we assume that $X(0) = 0$ (the

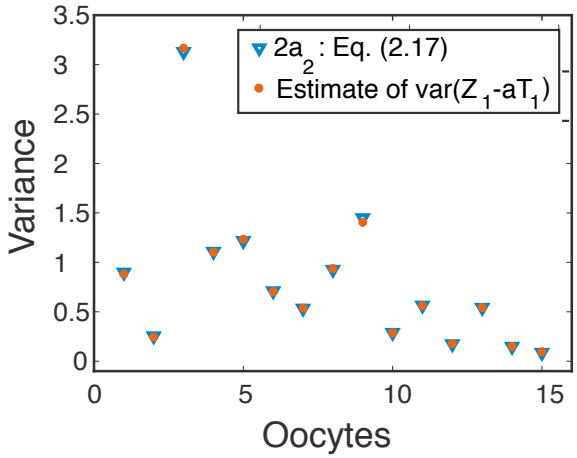

Figure 3.1: Comparison of the reward (displacement) variance predicted with the expressions from the large time Fourier analysis in §3.2.2 (blue triangles) and the long run reward variance estimated using observing the Markov process and renewal reward theory (red circles). The parameters used for the specific values plotted here correspond to FRAP data for VLE RNA in wild-type oocytes from all 3 regions of the oocytes.

particle starts in diffusion), then we have that

$$E[T_{00}] = \mu = \frac{1}{p_0(\beta_+ + \beta_-)} .$$

We therefore calculate the reward variance σ^2/μ using FRAP parameter estimates from 15 trials and plot the values in Figure 3.1. In our computation, we use time $t = 10^6$ s as the large time when we stop the process and compare our predictions for the variance of the displacement with the values obtained using the analytical expression derived with dynamical systems methods in §3.2.2. The good agreement of these estimated effective diffusions further suggests that the analytical approach proposed in §3.2.1 is a good alternative to potentially costly computations of the stochastic process up to a large time.

Notes on an alternative approach

We note that an analytical approach for calculating the effective velocity and diffusion of variable-length stepping of kinesins on microtubules has been developed in [68]. This approach makes use of renewal reward theory as well, however our attempts to adapt it to our model have revealed that in calculating second moments of the cycle time or reward, the approach in [68] does not consider certain cross-terms in the calculations. This observation points to some implicit assumptions in [68] that do not apply in our setting. In the following, we show that this approach does not agree with the central-limit theorem result for renewal rewards in the 2-state system.

Following [68], the 2-state system can also be modeled by considering the sequence (T_i, Z_i) of times and rewards (displacements) in each state, which are independent of each other when conditioned on the particle states X_i. Using the formulation in [68], the mean and covariance matrices are set up as follows:

$$\mu^T_{T|X} = \left(\frac{1}{\beta_1}, \frac{1}{\beta_2}\right) , \tag{3.35}$$

$$\mu^T_{Z|X} = \left(\frac{c}{\beta_1}, 0\right) , \tag{3.36}$$

$$\eta^T_{T|X} = \left(\frac{2}{\beta_1^2}, \frac{2}{\beta_2^2}\right) , \tag{3.37}$$

$$\eta^T_{Z|X} = \left(\frac{2c^2}{\beta_1^2}, \frac{2D}{\beta_2}\right) , \tag{3.38}$$

$$\mu^T_{T,Z|X} = \left(\frac{2c}{\beta_1^2}, 0\right) . \tag{3.39}$$

The transition probability matrix for the embedded Markov chain of states is in this case simply $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and therefore the stationary distribution for the embedded chain is $\boldsymbol{\pi_X} = (1/2, 1/2)$ (principal left eigenvector). The algorithm in [68] uses the

following means and variances:

$$\mu_T = \mu^T_{T|X}\pi_X = \frac{1}{2}\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right), \tag{3.40}$$

$$\mu_Z = \mu^T_{Z|X}\pi_X = \frac{c}{2\beta_1}, \tag{3.41}$$

$$\begin{aligned}\sigma^2_T &= \eta^T_{T|X}\pi_X - \mu^2_T = \frac{1}{\beta_1^2}+\frac{1}{\beta_2^2}-\frac{1}{4}\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right)^2 = \\ &= \frac{3}{4\beta_1^2}+\frac{3}{4\beta_2^2}-\frac{1}{2\beta_1\beta_2},\end{aligned} \tag{3.42}$$

$$\begin{aligned}\sigma^2_Z &= \eta^T_{Z|X}\pi_X - \mu^2_Z = \frac{c^2}{\beta_1^2}+\frac{D}{\beta_2}-\frac{c^2}{4\beta_1^2} = \\ &= \frac{3c^2}{4\beta_1^2}+\frac{D}{\beta_2},\end{aligned} \tag{3.43}$$

$$\begin{aligned}\sigma_{T,Z} &= \mu^T_{T,Z|X}\pi_X - \mu_T\mu_Z = \frac{c}{\beta_1^2}-\frac{1}{4}\frac{c}{\beta_1}\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right) = \\ &= \frac{3c}{4\beta_1^2}-\frac{c}{4\beta_1\beta_2}.\end{aligned} \tag{3.44}$$

The effective speed calculation using the renewal rewards approach is then given by the following expression [68, 69]:

$$\begin{aligned}V_\infty &= \frac{\mu_Z}{\mu_T} = \frac{\mu^T_{Z|X}\pi_X}{\mu^T_{T|X}\pi_X} \\ &= \frac{c/\beta_1}{1/\beta_1+1/\beta_2} = c\frac{\beta_2}{\beta_1+\beta_2},\end{aligned} \tag{3.45}$$

which agrees with equation (3.13) derived in § 3.2.2 and equation (3.24) derived in § 3.3.1. The effective diffusion calculation is [68]:

$$\begin{aligned}D_\infty &= \frac{1}{2\mu_T}(V^2_\infty\sigma^2_T+\sigma^2_Z-2V_\infty\sigma_{T,Z}) \\ &= D\frac{\beta_1}{\beta_1+\beta_2}+2c^2\frac{\beta_1\beta_2}{(\beta_1+\beta_2)^3}.\end{aligned} \tag{3.46}$$

Note that this expression has a factor of 2 in the second term, making it different from the direct renewal rewards calculation in equation (3.28) of §3.3.1 and the Fourier

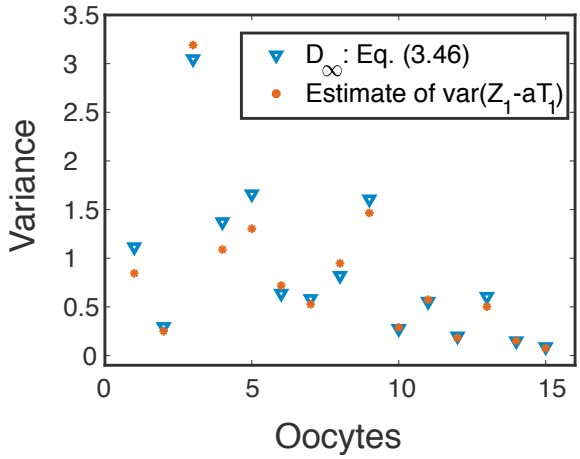

Figure 3.2: Comparison of the reward (displacement) variance predicted with the expression in (3.46) (blue triangles) and the long run reward variance estimated using observing the Markov process and renewal reward theory (red stars). The parameters used for the specific values plotted here correspond to FRAP data for VLE RNA in wild-type oocytes from all 3 regions of the oocytes (the same estimates are used in generating Figure 3.1).

analysis result in equation (3.14) of § 3.2.2.

The same framework applied to the 4-state model also illustrates that the expression for the effective diffusion does not agree with either the Markov process numerical observation or the large time Fourier analysis in §3.2.2. Using the same FRAP parameter estimates from 15 trials used in Figure 3.1, we plot the values given by the analytical expression in (3.46) derived with this alternative renewal rewards approach and compare it with the numerical estimates from observing the Markov process described in § 3.3.2. Figure 3.2 shows that while some trials lead to agreement of the estimated effective diffusions, others result in values that differ from the numerical observation of the Markov process up to $t = 10^6$ s.

Chapter Four

Application of FRAP Parameter Estimation and Model Analysis to mRNA Dynamics in *Xenopus* Oocytes

In this chapter, we combine the parameter estimation results in §2 and the analytical quantities derived in §3 in order to obtain insights into the dynamics of localizing and non-localizing RNA in different regions of the *Xenopus* oocyte. In addition, in §4.1 we derive the equilibrium fractions of particles in each state based on the models of active transport considered in §2 and determine some measures of dissociation (such as run times and lengths of mRNA particles on microtubules) that are useful in performing model validation in §4.3.

4.1 Equilibrium fractions of particles in each state and measures of dissociation

The results in this section are published in [1].

The equilibrium distribution of particles in different states given the general model (3.1) is readily obtained by solving

$$A\boldsymbol{u} = 0\,. \tag{4.1}$$

Then the additional assumption: $\sum_{i=1}^{n} u_i = 1$ yields the percentages of particles in each dynamic state at equilibrium.

An alternative approach to modeling particle mobility is by using a continuous-time Markov chain (CTMC) of the times and states of a particle undergoing intracellular transport. In this framework, we introduce matrix A with A_{ij} the rate of the transition from state i to state j, which corresponds to the transition matrix of the CTMC. Solving the linear system (4.1) becomes equivalent to solving the

equilibrium or balance equations of the Markov process [64]. We recall that the advection-diffusion 2-state model introduced in §2.3.1 is given by:

$$u_t = cu_y - \beta_1 u + \beta_2 v \tag{4.2a}$$

$$v_t = d\Delta v + \beta_1 u - \beta_2 v\,. \tag{4.2b}$$

In this case, the fractions of particles in each state are simply:

$$\text{fraction moving} = \frac{\beta_2}{\beta_1 + \beta_2} \tag{4.3a}$$

$$\text{fraction diffusing} = \frac{\beta_1}{\beta_1 + \beta_2}\,. \tag{4.3b}$$

The 4-state model expressions for fractions in each state are computed in a similar way and depend on all model transition rates.

The CTMC modeling approach is also useful in determining the dissociation-based quantities that appear in experimental literature, such as distances and times spent on microtubules before a motor-cargo complex unbinds [68]. Sojourn times of a homogeneous Markov chain in each state are defined as the mean time spent in the state before switching to another state. It is well established that sojourn times of a homogeneous Markov chain i are exponentially distributed with parameter q_i, where q_i is the transition rate of leaving state i for any other state [64]. This means that the mean sojourn times for the 2-state model are:

$$\text{expected run time moving} = \frac{1}{\beta_1}\,,$$

$$\text{expected run time diffusing} = \frac{1}{\beta_2}\,.$$

Similarly, the mean times in the states of the 4-state model are given by:

$$\begin{aligned}
\text{expected run time up} &= \frac{1}{\gamma_- + \delta_-}, \qquad \text{expected time diffusing} &= \frac{1}{\beta_- + \beta_+}, \\
\text{expected run time down} &= \frac{1}{\gamma_+ + \delta_+}, \qquad \text{expected time pausing} &= \frac{1}{\alpha_- + \alpha_+}.
\end{aligned}$$

The expected run length of motor-cargo complexes on microtubule filaments is then simply the speed in the desired direction times the mean sojourn time in the corresponding moving state. For the 4-state model, this yields:

$$\begin{aligned}
\text{expected run length up} &= \frac{c_-}{\gamma_- + \delta_-}, \\
\text{expected run length down} &= \frac{c_+}{\gamma_+ + \delta_+}.
\end{aligned}$$

4.2 Predictions for non-localizing RNA

Most results in this section are published in [1].

The parameter estimation results in §2.3.6 for non-localizing RNA (β-globin RNA and VLE RNA in nocodazole-treated oocytes) reveal similar estimates for the diffusion coefficient of the molecules, as illustrated in Figure 4.1. However, to accurately compare the mobility of the two types of RNA, we must take into account the estimates for the transition rates β_1 and β_2. We therefore combine the parameter estimates in §2.3.6 and the expressions for fractions of particles in different states in §4.1 to analyze the mobility of RNA, assuming that the concentration of RNA particles has reached steady state throughout the time course of a FRAP experiment. The estimates for rates β_1 and β_2, together with equations (4.3), suggest that non-localizing β-globin RNA does not solely diffuse, and may instead spend on average

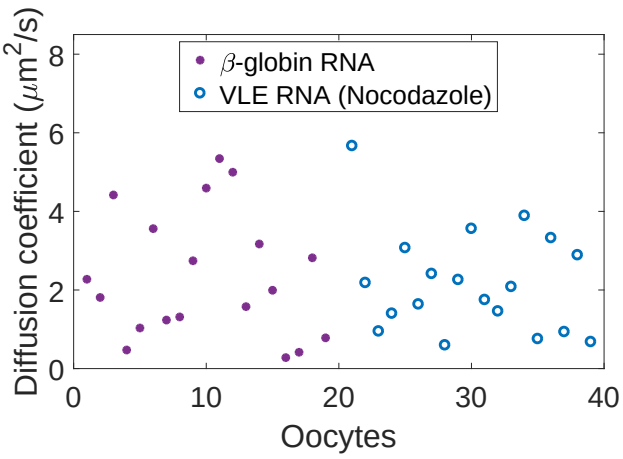

Figure 4.1: Diffusion coefficient estimates for nonlocalizing β-globin RNA and VLE RNA treated with nocodazole in 19 oocytes each.

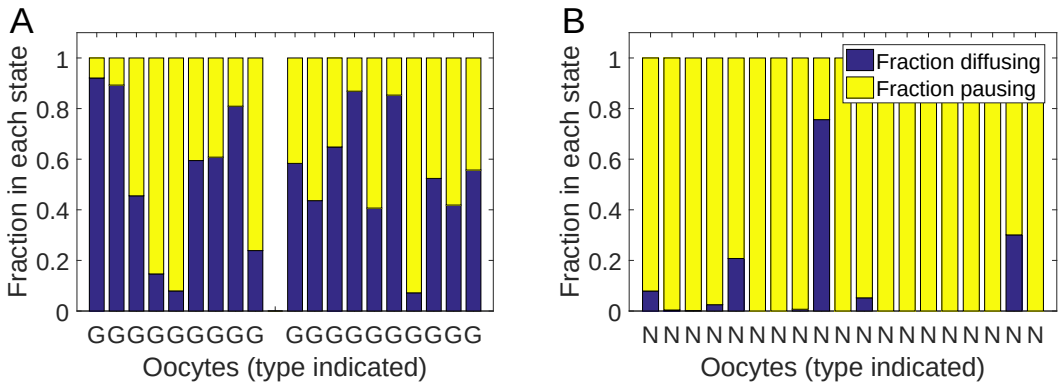

Figure 4.2: Predicted fractions of (A) nonlocalizing β-globin RNA (G) in sets of 9 and 10 untreated oocytes and (B) VLE RNA in 19 nocodazole-treated oocytes (N) in diffusing and stationary states for individual oocyte trials. Parameter estimation is performed with a three bleach spot initial condition for β-globin RNA using Approach 3, and with a one bleach spot initial condition for nocodazole-treated VLE RNA using Approach 1 (see §2.3.3, figure from [1]).

about 47% of time (with standard deviation 32%) in a paused state for the first set (see Figure 4.2A), and 46% of time (with standard deviation 23%) for the second set (see Figure 4.3). The predicted fractions of β-globin RNA in each state for two sets of individual oocytes are displayed in the bar graph in Figure 4.2A, and similar results are provided in Figure 4.3 for two additional sets of oocytes.

At equilibrium, an average of 92.45% (with standard deviation 18.4%) of RNA molecules in nocodazole-treated oocytes are stationary, compared to about 46% in the β-globin RNA case (see Figure 4.2B). We note that the distinct diffusing behavior

predicted in a few of the oocytes in Figure 4.2B is a result of the diluted fluorescence signal under nocodazole treatment, where there is a lower effective concentration of mRNA distributed throughout the cytoplasm rather than an accumulation at the cell periphery. These results suggest that, in nocodazole-treated oocytes, active transport of VLE RNA is blocked due to microtubule impairment, and diffusion may be restricted due to the large size of VLE RNA granules.

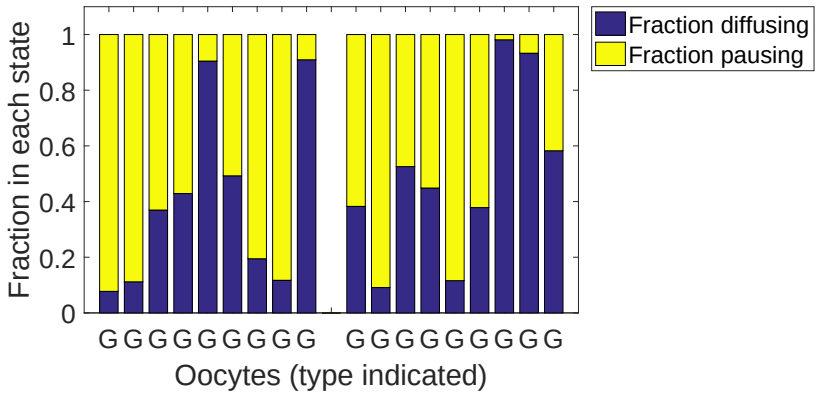

Figure 4.3: Predicted fractions of nonlocalizing β-globin RNA (G) in diffusing and stationary states for two additional sets of 9 healthy oocytes each (using Approach 3). The sets provided here and in Figure 4.2A are each from FRAP experiments carried out on different days. The average diffusion coefficient d for the first set is 1.8 $\mu m^2/s$ (with standard deviation 1.1 $\mu m^2/s$), and for the second set 3.1 $\mu m^2/s$ (with standard deviation 1.5 $\mu m^2/s$). mRNA particles are predicted to spend on average 60% of time in a paused state (with standard deviation 32%) for the first set, and 51% of time (with standard deviation 31%) for the second set (from [1]).

4.3 Predictions for localizing RNA and model validation

Similarly, the parameter estimation results in §2.3.6 for localizing VLE RNA reveal certain spatial differences in the dynamics in the oocyte. Given that localizing RNA models account for active transport of particles, we begin by surveying the results for speeds of mRNA as it is transported along microtubule filaments. Figure 4.4 shows that the 2-state model (4.2) predicts smaller speeds in the vegetal direction in

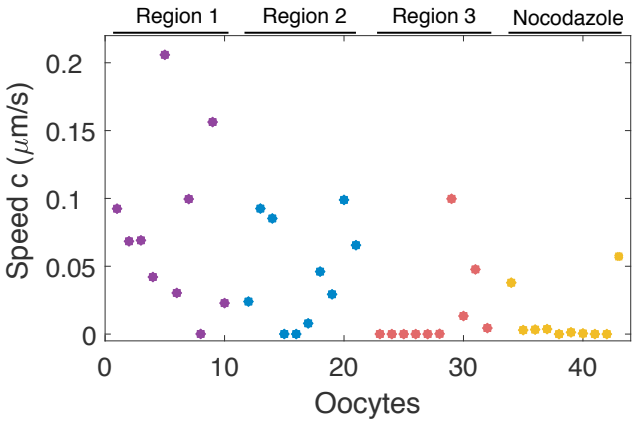

Figure 4.4: Speed c estimated using the 2-state model and Approach 1 for individual FRAP data in healthy and nocodazole-treated (N) oocytes. Regions are numbered and colored as in Figure 2.6A or Figure 4.7A.

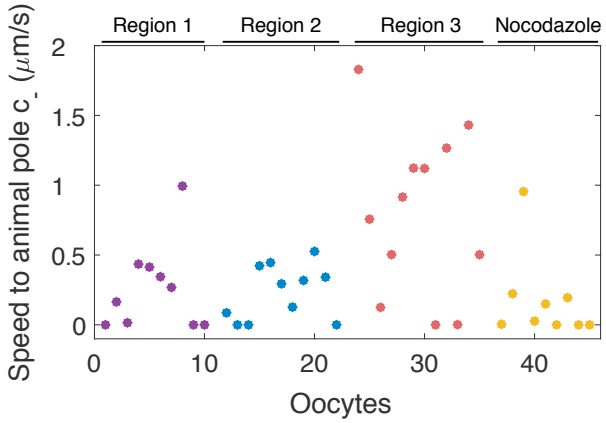

Figure 4.5: Speed in the animal pole direction c_- estimated using the 4-state model and Approach 1 for individual FRAP data in healthy and nocodazole-treated (N) oocytes. Regions are numbered and colored as in Figure 2.6A or Figure 4.7A.

Region 3 than in Regions 1 and 2. In addition, the disruption of microtubules with nocodazole mostly leads to negligible speeds for VLE RNA in these oocytes. The 4-state model (2.8) considers two moving states to test the hypothesis of bidirectional transport; we therefore plot the estimated speeds in the animal pole direction in Figure 4.5, and note that a considerable number of the trials predict higher speed in the animal pole direction in Region 3. These results are consistent with the hypothesis in [3] that the dynamics in the lower cytoplasm may be slower than in the upper cytoplasm of the vegetal wedge.

While the results in Figures 4.4 and 4.5 show some differences in the dynamics in different regions of the cytoplasm, the insights from these graphs is limited since they do not reflect the overall dynamics of the mRNA molecules. Instead, these plots show only the speeds of the particles when they are in a specific moving state. In the following, we outline predictions for mRNA dynamics and model selection that take into account all parameter estimates (including the transition rates between states pictured in Figure 2.5A-B) and combine them with the theoretical quantities derived in §4.1 and §3.2.1.

Mobility of localizing VLE RNA can be investigated using the predicted fractions of particles in each state derived in §4.1. Using the parameter estimates in §2.3.6, our results confirm the hypothesis that bidirectional transport plays an important role in the dynamics of VLE RNA [3], and further suggest that particles might spend on average about 72.5% of time in a paused state (with standard deviation 21%, see Figure 4.6). The RNA in nocodazole-treated oocytes are predicted to spend most of the time in a paused state, consistent with the results using model (2.9) in Figure 4.2B.

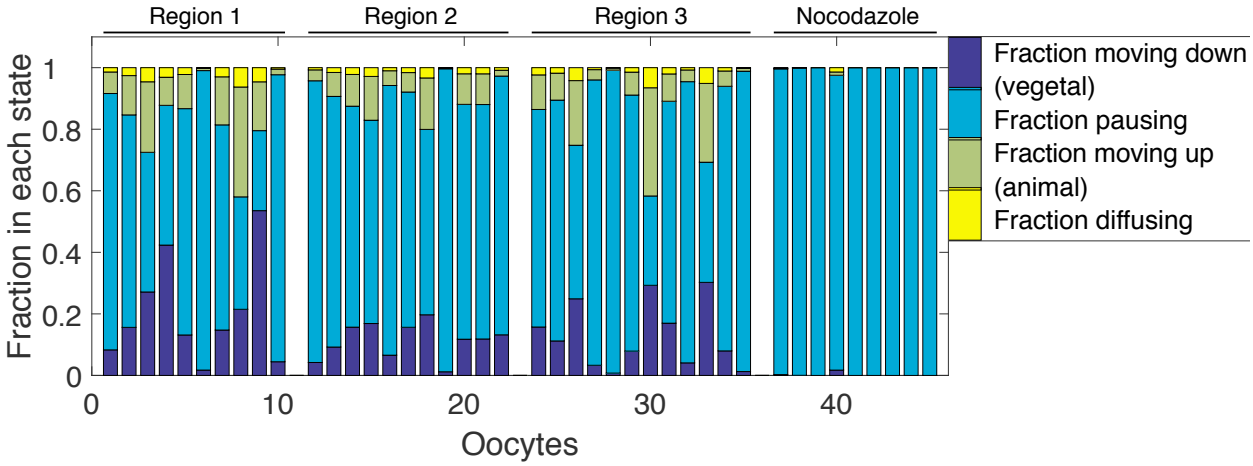

Figure 4.6: Predicted fractions of localizing VLE RNA in different states for individual oocyte trials (using Approach 1) in healthy and Nocodazole-treated (N) oocytes. The fits are carried out using the 4-state model (2.8) (from [1]).

Given that particle dynamics can be captured using multiple PDE models of active transport, we investigate the question of model selection for localizing RNA in *Xenopus* oocytes. The analysis of these PDE systems in §3.2.1 allows us to compute and compare quantities of interest predicted by the 2-state model (4.2) and the 4-state model (2.8). These quantities include the effective velocity and diffusion of a particle for large time, the percentage of particles in each state at equilibrium, and the expected run length and time of motor-cargo complexes on microtubules (that is, how far and how long a motor travels on average before dissociating from a microtubule). These asymptotic and dissociation-based quantities are mentioned in the experimental literature [68] and are thus very useful in evaluating models of active transport.

The effective velocity v and effective spread σ^2 for the 2-state model are the actual long-term speeds and diffusion of particles given the transition rates between particle states. These quantities are given by (see §3.2.2):

$$v = c\frac{\beta_2}{\beta_1+\beta_2}\,, \tag{4.4}$$

$$\sigma^2 = 2d\frac{\beta_1}{\beta_1+\beta_2} + 2c^2\frac{\beta_1\beta_2}{(\beta_1+\beta_2)^3}\,. \tag{4.5}$$

Equivalent quantities can be calculated for the 4-state model using Mathematica [70] to yield more complex expressions that depend on all the model parameters. These asymptotic quantities are then evaluated using estimated parameters for averaged FRAP data: the results are summarized in Tables 2.5 and 2.6. For both the 2-state and the 4-state models, we note that the effective velocity is either consistently smaller (2-state model) or negative indicating net movement in the animal pole direction (4-state model) for Region 3 in the lower cytoplasm. The effective diffusion also shows a consistent increase in Region 3. Both models therefore support the

hypothesis of faster movement in the upper vegetal cytoplasm (Regions 1-2) and higher spread of particles (suggesting bidirectional transport) in the lower vegetal cytoplasm (Region 3). Note that a negative effective velocity corresponds to net movement in the nucleus direction. The observations in Tables 2.5 and 2.6 show results for average FRAP data from different sets of oocytes, but the differences between regions hold for parameter estimates in individual FRAP data trials as well (see Figure 4.7B). Given that the FRAP bleaching experiments are performed at the same time for all three regions in the cytoplasm, we do not expect these differences to be a result of experimental variability.

The derivation of the effective velocity and diffusion for large time also provides a tool for comparison of mobility of localizing VLE RNA with mobility of RNA in cells treated with nocodazole. Figure 4.7B shows predictions for these particle displacement quantities when fitting both types of data to the 4-state model (2.8). We note that the nocodazole-treated trials (Figure 4.7B, yellow) predict almost no net movement, as is expected when the microtubule structure is disrupted. In untreated oocytes, the transport to the vegetal cortex is more significant close to the nucleus (Region 1, purple) than close to the vegetal cortex (Region 3, red). While we predict that the majority of untreated oocytes spend similar amounts of time in transport in the animal and vegetal directions (see Figure S3.6), Figure 4.7B also incorporates the velocity predictions to highlight directional bias through effective velocity and diffusion calculations.

The low order of magnitude estimates for binding/unbinding rates using the 2-state model (see Table 2.4) would lead to the unlikely suggestion that either all particles are diffusing or they are all being transported at a given location (see equations (4.3)). By contrast, the 4-state model predicts a more uniform distribution of particles in different states (see Figure 4.6), with a slight bias to a higher percentage

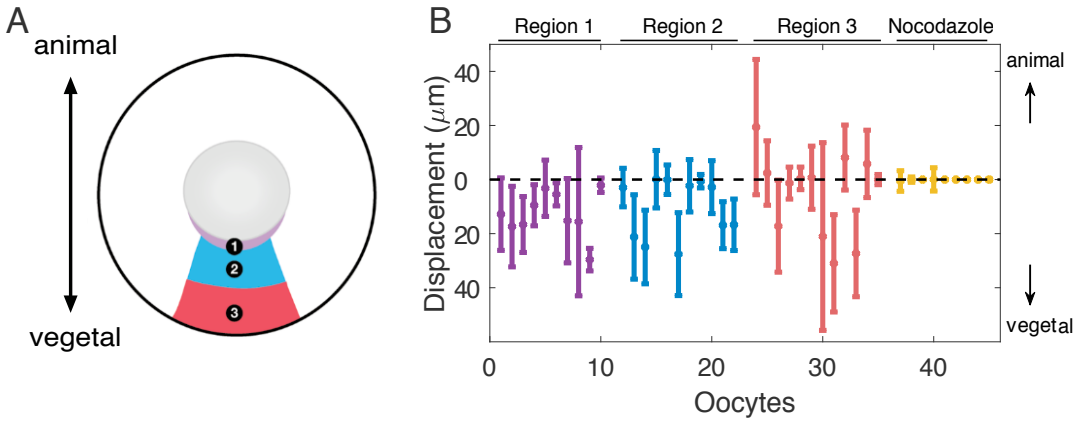

Figure 4.7: A: FRAP bleach regions are numbered according to their location: the perinuclear cup (Region 1), the upper vegetal cytoplasm (Region 2) and the lower vegetal cytoplasm (Region 3). B: Predicted effective displacement and spread at $T = 200$s using the 4-state model and Approach 1 for individual FRAP data in healthy and nocodazole-treated (N) oocytes. The distance between the dots and the zero axis corresponds to the predicted average displacement of a particle towards the nucleus or the vegetal cortex, and error bars correspond to the predicted spread of displacement due to diffusion for each trial (from [1]).

of particles moving in Regions 1 and 2. This suggests that the additional complexity of considering another moving population and a stationary state in the 4-state model (2.8) is necessary in order to model particle mobility in these experiments.

In addition, we compare predictions of the expected run lengths and times of an RNA particle on microtubules to experimental results. The average moving run time and distance for the 2-state model ($1/\beta_1$, respectively c/β_1) are very large compared to experimental observations of the processivity of molecular motor proteins. Kinesin and dynein motors have been shown to have average run lengths of roughly 1-2 μm [20, 21], and velocities of about 0.5-1 $\mu m/s$ [71], so that they are expected to spend a few seconds on an individual microtubule filament. These run lengths may be larger when multiple motors are attached to and transport cargo [53, 54]. The time and distance spent by a particle on a microtubule for the 4-state model are given by (as in §4.1):

$$\begin{aligned} \text{expected run time up} &= \frac{1}{\gamma_- + \delta_-}, \qquad \text{expected run length up} &= \frac{c_-}{\gamma_- + \delta_-}, \\ \text{expected run time down} &= \frac{1}{\gamma_+ + \delta_+}, \qquad \text{expected run length up} &= \frac{c_+}{\gamma_+ + \delta_+}. \end{aligned}$$

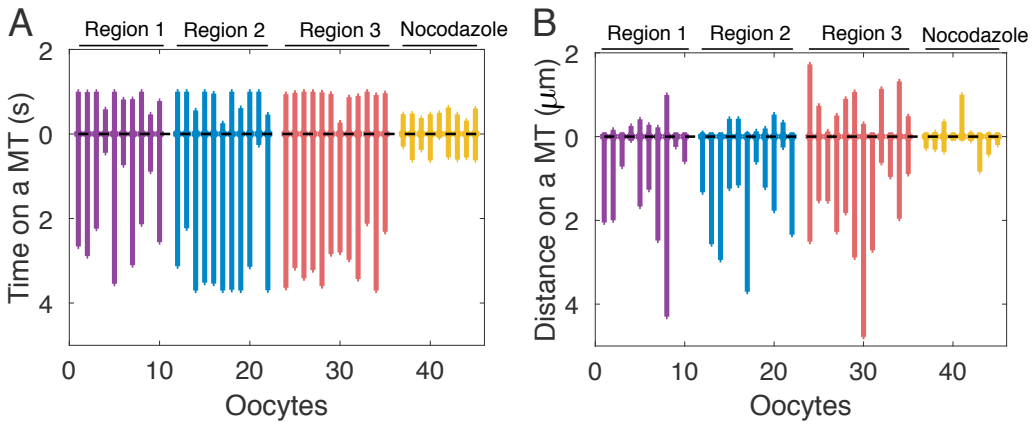

Figure 4.8: Predicted expected run time (A) and length (B) spent before dissociating from a microtubule for VLE RNA. The length of the segment extending up corresponds to the time or distance for movement up a MT towards the nucleus, and the length of the segment extending down corresponds to the time or distance for movement down a MT towards the vegetal cortex. Purple, blue and red correspond to regions 1, 2 and 3, respectively, as in Figure 4.7A; yellow corresponds to nocodazole-treated oocytes (from [1]).

The predictions for these moving states quantities given estimated parameters for experimental FRAP data from individual oocytes are displayed in Figure 4.8. We note that run times are on the order of seconds, and run lengths are on the order of μm, as expected from previous experimental measurements [54]. It is also worth noting that the results for untreated oocytes (Figure 4.8, purple, blue and red) further support the hypothesis of bidirectional transport of RNA in the cytoplasm [3], with a bias to movement in the vegetal cortex direction. In [3], this hypothesis was revealed through a more complicated photoactivation experiment, since the standard FRAP data analysis did not account for active transport processes. In the case of nocodazole-treated oocytes (Figure 4.8, yellow), both run times and lengths are predicted to be considerably shorter in both transport directions.

Chapter Five

Nonlinear Dynamics: Accounting for the Microtubule Structure

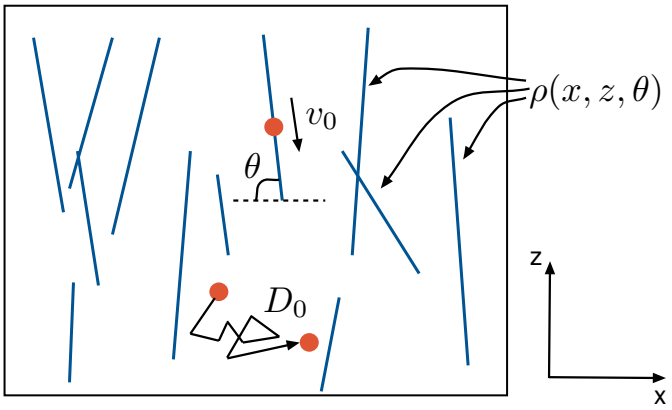

Figure 5.1: Illustration of a microtubular structure with density ρ (blue lines). Particles (red dots) can either be transported with velocity v_0 along microtubule filaments, or freely diffuse in the cytoplasm with diffusion coefficient D_0.

5.1 Nonlinear PDE models

5.1.1 Review of previous results

In [72], Bressloff *et al.* study a stochastic model of active vesicular transport and its influence on cell polarization. The motor-cargo complex is treated as a particle that randomly switches between a free diffusion state and a ballistic motion state with velocity $\boldsymbol{V}(\theta)$, where the direction θ is determined by the orientation of the cytoskeletal filament to which the complex is bound at location (x, z) (see Figure 5.1). Assuming constant speed of movement v_0, the velocity is given by [72]:

$$\boldsymbol{V}(\theta) = -v_0 \cos\theta \boldsymbol{e}_x - v_0 \sin\theta \boldsymbol{e}_z \,.$$

The main assumption here is that there is a density $\rho(x, z, \theta)$ of filaments with the given orientation θ (see Figure 5.1). In [72], ρ may also be dependent on the concentration of signaling molecules at the membrane u, so that it can also be time-dependent.

The equations for particle movement are given by:

$$\begin{aligned} \frac{\partial p(\boldsymbol{r},\theta,t)}{\partial t} &= -\boldsymbol{V}(\theta)\cdot\nabla p(\boldsymbol{r},\theta,t) - \frac{\beta}{\varepsilon}p(\boldsymbol{r},\theta,t) + \frac{\alpha\rho(\boldsymbol{r},\theta)}{\varepsilon}p_0(\boldsymbol{r},t)\,, \\ \frac{\partial p_0(\boldsymbol{r},t)}{\partial t} &= \varepsilon D\nabla^2 p_0(\boldsymbol{r},t) + \frac{\beta}{\varepsilon}\int_0^\pi p(\boldsymbol{r},\theta,t)d\theta - \frac{\alpha\bar{\rho}(\boldsymbol{r})}{\varepsilon}p_0(\boldsymbol{r},t)\,, \end{aligned} \tag{5.1}$$

where $p_0(\boldsymbol{r},t)$ denotes the probability density that the particle is at position $\boldsymbol{r} = (x,z)$ at time t and is diffusing, and $p(\boldsymbol{r},\theta,t)$ is the probability that the particle is bound to a microtubule at location $\boldsymbol{r}$ and moving with velocity $\boldsymbol{V}(\theta)$ [72]. Here $\bar{\rho}(\boldsymbol{r},t) = \int_0^\pi \rho(\boldsymbol{r},0)d\theta$. This system is an extension of the model initially considered by Hawkins *et al.* [73]. In [73], there is an additional assumption that particles switch very fast between diffusion and active transport by motor proteins, so that a deterministic advection-diffusion equation is considered instead of equations (5.1).

The small ε in this approach incorporates the assumption that the switching rates are very fast and diffusion is slow compared to typical motor velocities. In the limit $\varepsilon \to 0$, the total probability density is conserved. In addition, this assumption is necessary in order to proceed with the quasi-steady-state (QSS) approximation in [72], which yields the effective velocity as well as the nonuniform effective diffusion results summarized in Table 5.1. Given our parameter estimation results in §2.3.6, these assumptions on diffusion and binding constant parameters do not hold in the case of mobility of mRNA in *Xenopus* oocytes. Our goal is thus to use dynamical systems tools as in §3.2.1 to obtain estimates of the effective velocity and diffusion of an mRNA particle based on the dynamics of the nonlinear system (5.1) and under no assumptions of small diffusion coefficient and large reaction rates. In addition, our analysis does not require an assumption of small microtubule density as in [72]: $a(\boldsymbol{r}) = \frac{\alpha\rho(\boldsymbol{r})}{\beta} \ll 1$.

5.1.2 Large-time Fourier analysis for parallel filaments

We consider the case where filaments are oriented parallel to each other, which is observed for instance in microtubules of neural growth cones. In this case, [72] considers the density of filaments as a function of the concentration of signaling molecules on the membrane: $\rho(\boldsymbol{r},\theta,t) = Ku(x,t)\delta(\theta-\pi/2)$ (see Figure 5.2). The equations for the concentration of particles (or equivalently, the probability of a particle to be in) the two states (active transport and diffusion) become:

$$\begin{aligned}\frac{\partial p(\boldsymbol{r},\theta,t)}{\partial t} &= -\boldsymbol{V}(\theta)\cdot\nabla p(\boldsymbol{r},\theta,t) - \beta p(\boldsymbol{r},\theta,t) + \alpha K u(x,t)\delta(\theta-\pi/2)p_0(\boldsymbol{r},t)\,,\\ \frac{\partial p_0(\boldsymbol{r},t)}{\partial t} &= D\nabla^2 p_0(\boldsymbol{r},t) + \beta\int_0^\pi p(\boldsymbol{r},\theta,t)d\theta - \alpha K u(x,t)p_0(\boldsymbol{r},t)\,. \end{aligned} \tag{5.2}$$

Since $\theta = \pi/2$, movement can only occur along the vertical dimension: $\boldsymbol{V}(\theta) = -v_0\boldsymbol{e_z}$. Letting $\bar{p}(\boldsymbol{r},t) = \int_0^\pi p(\boldsymbol{r},\theta,t)d\theta$ and $\bar{\rho} = \int_0^\pi p(\boldsymbol{r},\theta,t)d\theta = Ku(x,t)$ yields:

$$\begin{aligned}\frac{\partial \bar{p}(\boldsymbol{r},t)}{\partial t} &= v_0\boldsymbol{e_z}\cdot\nabla\bar{p}(\boldsymbol{r},t) - \beta\bar{p}(\boldsymbol{r},t) + \alpha K u(x,t)p_0(\boldsymbol{r},t)\,,\\ \frac{\partial p_0(\boldsymbol{r},t)}{\partial t} &= D\nabla^2 p_0(\boldsymbol{r},t) + \beta\bar{p}(\boldsymbol{r},t) - \alpha K u(x,t)p_0(\boldsymbol{r},t)\,. \end{aligned} \tag{5.3}$$

We note that these equations are an extension of the linear 2-state model in equations (2.7). We use variable and parameter notations that are different from §2 and §3 so as to be consistent and provide a comparison with the approach and results in [72]. We prove the following theorem for the behavior of solutions to equations (5.3):

Theorem 5.1. *Consider the advection-reaction-diffusion equation system in* (5.3) *modeling the dynamics of particles with transport restricted to parallel microtubules in two spatial dimensions. Assuming the spatial domain* $x \in [0,1]$ *and an infinite*

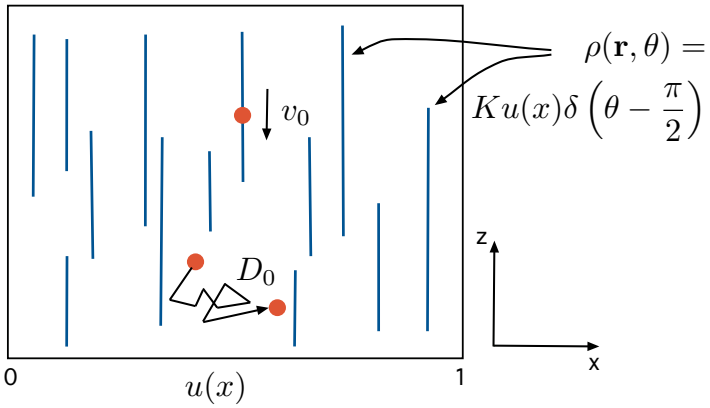

Figure 5.2: Illustration of a parallel microtubular structure with density $\rho(\boldsymbol{r},\theta) = Ku(x)\delta(\theta - \pi/2)$ (blue lines). Particles (red dots) can either be transported with velocity v_0 down along microtubule filaments, or freely diffuse in the cytoplasm with diffusion coefficient D_0.

domain in the z direction, we consider homogenous Neumann boundary conditions:

$$\begin{aligned} \frac{\partial \bar{p}}{\partial x}(x=0,z,t) &= \frac{\partial \bar{p}}{\partial x}(x=1,z,t) = 0\,, \\ \frac{\partial p_0}{\partial x}(x=0,z,t) &= \frac{\partial p_0}{\partial x}(x=1,z,t) = 0\,. \end{aligned} \tag{5.4}$$

Assume that the concentration of signaling molecules is a space-dependent (in the x dimension) continuously differentiable function $u(x)$. Further assume that the reaction rates between particle populations satisfy $\beta \neq 0$ and $\beta + \alpha K u \neq 0$. See Figure 5.2 for an illustration of the particle dynamics.

Given a δ-function or Gaussian initial condition for $\bar{p}$ and p_0, the effective velocity and effective diffusion of the particles in the direction of transport (z) in the limit $t \to \infty$ are given by:

$$\begin{aligned} v &= \frac{\alpha K \langle u(x), 1\rangle}{\beta + \alpha K \langle u(x), 1\rangle} v_0\,, \\ \sigma^2 &= \frac{\beta}{\beta + \alpha K \langle u(x), 1\rangle} D + \frac{\beta \alpha K \langle u(x), 1\rangle}{\left(\beta + \alpha K \langle u(x), 1\rangle\right)^3} v_0^2 + \frac{\beta \alpha K v_0}{\left(\beta + \alpha K \langle u(x), 1\rangle\right)^2} \langle u(x), w_1(x)\rangle\,, \end{aligned}$$

where $\langle f(x), g(x)\rangle = \int_0^1 f(x)g(x)dx$ *and* w_1 *satisfies*

$$D(w_1)_{xx} = \frac{\alpha K}{\beta + \alpha K\langle u(x), 1\rangle} v_0 \left(\langle u(x), 1\rangle - u(x)\right) . \tag{5.5}$$

Proof.

Insights from $\mathbf{u}(\mathbf{x}) \equiv \mathbf{u}$

We start with the assumption that the concentration of signaling molecules at the cortex is a constant, and consider the following ansatz:

$$\bar{p}(\boldsymbol{r}, t) = v_1(x)e^{\lambda t + ikz}\tilde{u_0} \tag{5.6}$$

$$p_0(\boldsymbol{r}, t) = v_2(x)e^{\lambda t + ikz}\tilde{v_0}\,, \tag{5.7}$$

Plugging these into equation (5.3) and simplifying yields:

$$\lambda\tilde{u_0} = v_0 ik\tilde{u_0} - \beta\tilde{u_0} + \alpha K u\frac{v_2(x)}{v_1(x)}\tilde{v_0}\,, \tag{5.8}$$

$$\lambda\tilde{v_0} = D\left(\frac{(v_2)_{xx}}{v_2} - k^2\right)\tilde{v_0} + \beta\frac{v_1(x)}{v_2(x)}\tilde{u_0} - \alpha K u\tilde{v_0}\,. \tag{5.9}$$

The goal is to solve for v_2 and v_1, so we start by isolating $\frac{v_2}{v_1}$ in both equations. (5.8) gives:

$$\frac{v_2}{v_1} = \frac{\lambda - v_0 ik + \beta}{\alpha K u}\frac{\tilde{u_0}}{\tilde{v_0}}\,.$$

Plugging this into (5.9) yields the equation for v_2:

$$\lambda = D\left(\frac{(v_2)_{xx}}{v_2} - k^2\right) + \alpha K u\left(\frac{\beta}{\lambda - v_0 ik + \beta} - 1\right)\,,$$

thus

$$(v_2)_{xx} = \frac{1}{D}\left(\lambda + k^2 D + \alpha K u\left(\frac{\lambda - v_0 ik}{\lambda - v_0 ik + \beta}\right)\right) v_2\,. \tag{5.10}$$

We therefore seek $v_2(x) = c_1 e^{r_1 x} + c_2 e^{r_2 x}$, with $r_j = \pm il$, where

$$r_j = \sqrt{\frac{1}{D}\left(\lambda + k^2 D + \alpha K u\left(\frac{\lambda - v_0 ik}{\lambda - v_0 ik + \beta}\right)\right)}\,.$$

Re-arranging gives:

$$\begin{aligned}\lambda &= D(r^2 - k^2) - \alpha K u \frac{\lambda - v_0 ik}{\lambda - v_0 ik + \beta}\,,\\ &= -D(l^2 + k^2) - \alpha K u \frac{\lambda - v_0 ik}{\lambda - v_0 ik + \beta}\,.\end{aligned} \tag{5.11}$$

Multiplying through by $\lambda - v_0 ik + \beta$ and denoting $\nu_1 = ik$, $\nu_2 = il$ yields:

$$\lambda^2 + \lambda\left(-D(\nu_1^2 + \nu_2^2) - v_0\nu_1 + \beta + \alpha K u\right) - D(\beta - v_0\nu_1)(\nu_1^2 + \nu_2^2) - \alpha K u \nu_1 v_0 = 0\,. \tag{5.12}$$

We note that this equation has the form $f(\nu_1, \nu_2, \lambda) = 0$, where $f(0,0,0) = 0$ and f is a continuously differentiable function. Since $\frac{\partial f}{\partial \lambda}|_{(0,0,0)} = \beta + \alpha K u \neq 0$ by assumption, the implicit function theorem gives that $\lambda = g(\nu_1, \nu_2)$ near $(0,0,0)$ with g a unique continuously differentiable function, and further that we can consider the expansion $\lambda = a_0 + a_1\nu_1 + a_2\nu_2 + a_3\nu_1\nu_2 + a_4\nu_1^2 + a_5\nu_2^2 + \ldots$. As in §3.2.1, $a_0 = 0$ corresponds to the dominant eigenvalue. This expansion allows us to obtain expressions for coefficients a_1, a_4 and a_5 using an approach similar to the one in §3.2.1.

Spatial dependence $\mathbf{u}(\mathbf{x})$

We are interested in the case where the concentration of signaling molecules at the cortex has spatial dependence: $\mathbf{u}(\mathbf{x}, \mathbf{t}) \equiv \mathbf{u}(\mathbf{x})$. Equation (5.10) still holds, since plugging in the ansatz (5.6) and (5.7) up to that point did not use the constant concentration assumption. We start by considering the case of wavenumber $k = 0$.

Wavenumber $k = 0$

The case $u = u(x)$ with $k = 0$ leads to

$$(Dv_2)_{xx} = \left(\lambda + \alpha K u(x) \frac{\lambda}{\lambda + \beta}\right) v_2 \tag{5.13}$$

$$= \alpha K u(x) v_2 + \lambda v_2 - \alpha K \frac{\beta}{\lambda + \beta} u(x) v_2 \,. \tag{5.14}$$

We assume that λ is small given the analogy with the constant u case. This gives that $\frac{1}{\lambda+\beta} = \frac{1}{\beta} - \frac{1}{\beta^2}\lambda + \mathcal{O}(\lambda^2)$, so that equation (5.14) becomes:

$$(Dv_2)_{xx} = \alpha K u(x) v_2 + \left(\lambda - \alpha K u(x) + \frac{\alpha K u(x)}{\beta}\lambda + \mathcal{O}(\lambda^2)\right) v_2 \tag{5.15}$$

$$= \lambda \left(1 + \frac{\alpha K u(x)}{\beta}\right) v_2 \,. \tag{5.16}$$

We then consider the general Sturm-Liouville equation $(p\varphi')' + q\varphi = -\tilde{\lambda}\sigma\varphi$,

where we let:

$$\tilde{\lambda} = -\lambda$$
$$p = D > 0$$
$$q(x) = 0$$
$$\sigma(x) = 1 + \frac{\alpha K u(x)}{\beta} > 0\,.$$

Applying results from Sturm-Liouville theory, we conclude that there is a smallest eigenvalue $\tilde{\lambda}_0$, and therefore a largest eigenvalue λ_0 for our system. Moreover, since $q(x) \leq 0$ and since the homogeneous Neumann boundary conditions in equations (5.4) extend to homogeneous Neumann boundary conditions $v_2(x)$, we have that these conditions satisfy:

$$\alpha_1 v_2(0) + \alpha_2 v_2'(0) = 0\,,$$
$$\beta_1 v_2(1) + \beta_2 v_2'(1) = 0$$

for $\alpha_1 = \beta_1 = 0$ and $\alpha_2 = \beta_2 = 1$, so that $\alpha_1\alpha_2 \leq 0$ and $\beta_1\beta_2 \geq 0$. Therefore, we additionally know that $\tilde{\lambda}_k \geq 0\,\forall k$, and:

$$\lambda_n < \lambda_{n-1} < ... < \lambda_0 \leq 0\,.$$

We also know that $\lambda_0 = 0$ is an eigenvalue with the corresponding eigenvector $\varphi_0 = c$ (a constant), which satisfies equation (5.14) and the homogeneous Neumann boundary conditions; we therefore conclude that the largest eigenvalue λ_0 is given by $\lambda_0 = 0$. The assumption of small filament density in [72] ($\frac{\alpha K u(x)}{\beta} \ll 1$) is not necessary in this approach.

Small wavenumber k

For small wavenumber $k \neq 0$, we consider the equation $Lv = 0$, where the operator L is defined on the space $\mathcal{C}^2(\mathbb{R})$ and

$$L(\lambda, \nu)! = D\partial_x^2 + D\nu^2 + \alpha K u(x) \left(\frac{\beta}{\lambda - v_0 \nu + \beta} - 1 \right) - \lambda \tag{5.17}$$

with $\nu = ik$ and homogeneous Neumann boundary conditions. Note that the operator is self-adjoint and

$$N(L)|_{\lambda=0,\nu=0} = N(L^*)|_{\lambda=0,\nu=0} = 1\,. \tag{5.18}$$

Since the null space is finite dimensional at $(\lambda, \nu) = (0, 0)$, we can apply the Lyapunov Schmidt reduction theory, and project the equation onto both $\mathrm{ran}(L)$ and $\mathrm{ran}(L)^{\perp}$. We write:

$$v(x) = a + w(x)\,, \tag{5.19}$$

where a is any constant and $w \in N(L)^{\perp}$, thus $< 1, w(x) >= 0$.

(i) Projection onto the kernel gives:

$$\langle 1, Lv \rangle = 0 \implies \langle 1, La + Lw \rangle = 0\,, \implies \tag{5.20}$$

$$\langle 1, aD\nu^2 + a\alpha K u(x) \left(\frac{\beta}{\lambda - v_0 \nu + \beta} - 1 \right) - a\lambda \rangle +$$

$$\langle 1, Dw_{xx} + \left(D\nu^2 + \alpha K u(x) \left(\frac{\beta}{\lambda - v_0 \nu + \beta} - 1 \right) - \lambda \right) w \rangle = 0\,. \tag{5.21}$$

Using $\langle 1, Lw\rangle = \langle L^*1, w\rangle$ and noting that $\langle c, w\rangle = 0\ \forall c$ yields:

$$-a\lambda + aD\nu^2 + a\alpha K\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right)\langle u(x), 1\rangle + \alpha K\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right)\langle u(x), w\rangle = 0\,. \tag{5.22}$$

(ii) Projection onto the range gives

$$\left(Lv - \frac{\langle 1, Lv\,\rangle}{\langle 1, 1\rangle}1\right) = 0 \tag{5.23}$$

$$\implies La + Lw - \langle 1, La\rangle - \langle 1, Lw\rangle = 0\,. \tag{5.24}$$

This further yields:

$$\begin{aligned} &a\left(D\nu^2 + \alpha K u\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right)\right) - a\lambda + Dw_{xx} \\ &\quad + \left(D\nu^2 + \alpha K u\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right)\right) w \\ &-\lambda w - aD\nu^2 + a\lambda - a\alpha K\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right)\langle 1, u\rangle \\ &\quad -\langle D\nu^2 + \alpha K u\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right) - \lambda, w\rangle = 0\,. \end{aligned} \tag{5.25}$$

Simplifying gives:

$$\begin{aligned} Dw_{xx} + \left(D\nu^2 + \alpha K u(x)\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right) - \lambda\right) w + a\alpha K\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right) u(x) + \\ -a\alpha K\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right)\langle u(x), 1\rangle - \alpha K\left(\frac{\beta}{\lambda - v_0\nu + \beta} - 1\right)\langle u(x), w\rangle = 0\,. \end{aligned} \tag{5.26}$$

Since we assume small wavenumber k, we expand around $\nu = ik$ as follows:

$$\frac{\beta}{\lambda - v_0\nu + \beta} = \frac{\beta}{\lambda + \beta}\frac{1}{1 - \frac{v_0\nu}{\lambda + \beta}} = \frac{\beta}{\lambda + \beta}\sum_{n=0}^{\infty}\left(\frac{v_0\nu}{\lambda + \beta}\right)^n\,. \tag{5.27}$$

Therefore

$$\frac{\beta}{\lambda - v_0\nu + \beta} - 1 = -\frac{\lambda}{\lambda+\beta} + \frac{v_0\beta}{(\lambda+\beta)^2}\nu + \frac{v_0^2\beta}{(\lambda+\beta)^3}\nu^2 + \mathcal{O}(\nu^3)\,. \tag{5.28}$$

Plugging (5.28) into equation (5.22) and denoting $w = a\tilde{w}$ simplifies to:

$$\lambda = \alpha K\left(-\frac{\lambda}{\lambda+\beta} + \frac{v_0\beta}{(\lambda+\beta)^2}\nu + \frac{v_0^2\beta}{(\lambda+\beta)^3}\nu^2 + \mathcal{O}(\nu^3)\right)\langle u(x), 1+\tilde{w}(x)\rangle + D\nu^2\,, \tag{5.29}$$

thus

$$\begin{aligned}\lambda(\lambda+\beta)^3 + \alpha K\langle u(x), v(x)/a\rangle\lambda(\lambda+\beta)^2 - \alpha K v_0\beta\langle u(x), v(x)/a\rangle(\lambda+\beta)\nu \\ - \alpha K v_0^2\beta\langle u(x), v(x)/a\rangle\nu^2 + D\nu^2(\lambda+\beta)^3 + \mathcal{O}(\nu^3) = 0\end{aligned} \tag{5.30}$$

This has the form $f(\nu,\lambda) = 0$, where $f(0,0) = 0$ and f is a continuously differentiable function. Noting that $\frac{\partial f}{\partial\lambda}|_{(0,0)} = \beta^3 \neq 0$ by assumption, the implicit function theorem gives that $\lambda = g(\nu)$ near $(0,0)$ with g a unique continuously differentiable function, and further that we can consider the expansion $\lambda = a_1\nu + a_2\nu^2 + \mathcal{O}(\nu^3)\,.$

We similarly plug (5.28) into equation (5.26) to obtain:

$$D\tilde{w}_{xx} + \left(D\nu^2 - \lambda + \alpha K u(x)\left(-\frac{\lambda}{\lambda+\beta} + \frac{v_0\beta}{(\lambda+\beta)^2}\nu + \frac{v_0^2\beta}{(\lambda+\beta)^3}\nu^2 + \mathcal{O}(\nu^3)\right)\right)\tilde{w} \tag{5.31}$$

$$+ \alpha K\left(-\frac{\lambda}{\lambda+\beta} + \frac{v_0\beta}{(\lambda+\beta)^2}\nu + \frac{v_0^2\beta}{(\lambda+\beta)^3}\nu^2 + \mathcal{O}(\nu^3)\right)(u(x) - \langle u(x), v(x)/a\rangle) = 0\,. \tag{5.32}$$

This is again of the form $h(\nu, \tilde{w}(x)) = 0$, where $h(0,1) = 0$ and $\lambda(0) = 0$. Noting that $\frac{\partial h}{\partial\tilde{w}}|_{(0,1)} = D\partial_{xx}$ and recalling the homogeneous Neumann boundary conditions, this

operator is invertible. Therefore, the implicit function theorem gives that $\tilde{w} = g'(\nu)$ and thus that $w = g'(\nu)$ near $(1, 0)$ with g' a unique continuously differentiable function, and further that we can consider the expansion $v_2(x) = 1 + w(x) = 1 + w_1(x)\nu + w_2(x)\nu^2 + \mathcal{O}(\nu^3)$.

Before proceeding with finding the coefficients in the expansions for λ and $v_i(x)$ $(i = 1, 2)$, we examine the recovery of the concentration of particles in each state for large time given these expansions. Using the inverse Fourier transform gives:

$$\begin{aligned} p_0(\boldsymbol{r}, t) &= \frac{1}{2\pi} \int_{-\delta}^{\delta} (1 + w(x)) e^{\lambda t + ikz} dk \\ &= \frac{1}{2\pi} \int_{-\delta}^{\delta} (1 + w(x)) e^{ik(z+a_1 t) - a_2 k^2 t} e^{\sum_{j=3}^{\infty} a_j (ik)^j t} dk \,, \end{aligned} \tag{5.33}$$

where δ is small. Making the change of variable $\tilde{k} = kt^{1/2}$ and $\tilde{z} = z + a_1 t$ yields:

$$\begin{aligned} p_0(\boldsymbol{r}, t) &= \frac{1}{2\pi\sqrt{t}} \int_{-\delta\sqrt{t}}^{\delta\sqrt{t}} (1 + w(x)) e^{i\tilde{k}\frac{\tilde{z}}{t^{1/2}} - \frac{a_2}{2}\tilde{k}^2} e^{\sum_{j=3}^{\infty} a_j \frac{(i\tilde{k})^j}{t^{j/2-1}}} d\tilde{k} \\ &= \frac{1}{2\pi\sqrt{t}} \int_{-\delta\sqrt{t}}^{\delta\sqrt{t}} \left(1 + i\frac{\tilde{k}}{t^{1/2}} w_1(x) - \frac{\tilde{k}^2}{t} w_2(x) + \sum_{j=1}^{\infty} \frac{(i\tilde{k})^j}{t^{j/2}} w_j(x) \right) e^{i\tilde{k}\frac{\tilde{z}}{t^{1/2}} - \frac{a_2}{2}\tilde{k}^2} e^{\sum_{j=3}^{\infty} a_j \frac{(i\tilde{k})^j}{t^{j/2-1}}} d\tilde{k} \,. \end{aligned} \tag{5.34}$$

Since we are interested in long-term asymptotic behavior at $t \to \infty$, we note that the dominant term in the expansion of $w(x)$ is simply 1 and that the last exponential term in (5.34) converges to 1, so that:

$$p_0(\boldsymbol{r}, t) \approx \frac{1}{2\pi\sqrt{t}} \int_{-\infty}^{\infty} e^{i\tilde{k}\frac{\tilde{z}}{t^{1/2}} - \frac{a_2}{2}\tilde{k}^2} d\tilde{k} = \frac{1}{\sqrt{2\pi a_2 t}} e^{\frac{(z+a_1 t)^2}{2a_2 t}} \,. \tag{5.35}$$

Therefore, the Gaussian form of the solution for large time provides the effective velocity and effective diffusion of the particles, given by a_1 and a_2, respectively. The coefficients in the expansion of $w(x)$ are therefore necessary in determining the

solution behavior only if they appear in the expressions for a_1 and a_2.

We now consider

$$\lambda = a_1\nu + a_2\nu^2 + \mathcal{O}(\nu^3)\,, \tag{5.36}$$

$$v_2(x) = 1 + w(x) = 1 + w_1(x)\nu + w_2(x)\nu^2 + \mathcal{O}(\nu^3)\,. \tag{5.37}$$

Expanding for small ν now gives:

$$\begin{aligned}\frac{\beta}{\lambda - v_0\nu + \beta} - 1 &= \frac{\beta}{\beta - ((v_0 - a_1)\nu - a_2\nu^2 + \mathcal{O}(\nu^3))} \\ &= \frac{v_0 - a_1}{\beta}\nu - \frac{a_2}{\beta}\nu^2 + \left(\frac{v_0 - a_1}{\beta}\nu - \frac{a_2}{\beta}\nu^2 + \mathcal{O}(\nu^3)\right)^2 + \mathcal{O}(\nu^3) \\ &= \frac{v_0 - a_1}{\beta}\nu + \left(\left(\frac{v_0 - a_1}{\beta}\right)^2 - \frac{a_2}{\beta}\right)\nu^2 + \mathcal{O}(\nu^3)\,, \end{aligned} \tag{5.38}$$

where we performed a Taylor expansion around $\nu = 0$ in the second equality.

Plugging this into the kernel projection (5.22) yields:

$$-a(a_1\nu + a_2\nu^2 + \mathcal{O}(\nu^3)) + aD\nu^2 + a\alpha K\langle u(x), 1\rangle\left(\frac{v_0 - a_1}{\beta}\nu + \left(\left(\frac{v_0 - a_1}{\beta}\right)^2 - \frac{a_2}{\beta}\right)\nu^2 + \mathcal{O}(\nu^3)\right) \tag{5.39}$$

$$+ \alpha K\left(\frac{v_0 - a_1}{\beta}\right)\langle u(x), w_1\rangle\nu^2 + \mathcal{O}(\nu^3) = 0\,, \tag{5.40}$$

and we set the $\mathcal{O}(\nu)$ term equal to 0:

$$-aa_1 + a\alpha K\langle u(x), 1\rangle\left(\frac{v_0 - a_1}{\beta}\right) = 0\,, \tag{5.41}$$

thus

$$a_1 = \frac{\alpha K \langle u(x), 1\rangle}{\beta + \alpha K \langle u(x), 1\rangle} v_0 \,. \tag{5.42}$$

Similarly, the $\mathcal{O}(\nu^2)$ term gives:

$$-aa_2 + aD + a\alpha K \langle u(x), 1\rangle \left(\left(\frac{v_0 - a_1}{\beta} \right)^2 - \frac{a_2}{\beta} \right) \nu^2 + \alpha K \left(\frac{v_0 - a_1}{\beta} \right) \langle u(x), w_1 \rangle = 0 \,. \tag{5.43}$$

Noting that $\frac{v_0 - a_1}{\beta} = \frac{v_0}{\beta + \alpha K \langle u(x), 1\rangle}$ further gives:

$$a_2 = \frac{\beta}{\beta + \alpha K \langle u(x), 1\rangle} D + \frac{\beta \alpha K \langle u(x), 1\rangle}{(\beta + \alpha K \langle u(x), 1\rangle)^3} v_0^2 + \frac{\beta \alpha K v_0}{(\beta + \alpha K \langle u(x), 1\rangle)^2} \langle u(x), w_1(x)/a \rangle \,. \tag{5.44}$$

To fully determine the expression for the effective diffusion a_2, we need to determine the function $w_1(x)$ in the expansion (5.37). Returning to the projection onto the range in equation (5.26) and using approximation (5.38), the $\mathcal{O}(\nu)$ term becomes:

$$a\alpha K u(x) \left(\frac{v_0 - a_1}{\beta} \right) + D(w_1)_{xx} - a\alpha K \langle u(x), 1\rangle \left(\frac{v_0 - a_1}{\beta} \right) = 0 \,. \tag{5.45}$$

Therefore, w_1 satisfies:

$$D(w_1)_{xx} = \frac{a\alpha K}{\beta + \alpha K \langle u(x), 1\rangle} v_0 \left(\langle u(x), 1\rangle - u(x) \right) \,. \tag{5.46}$$

Since the operator in (5.17) describes the equation for $v = v_2(x)$, we turn our

attention to $v_1(x)$ in the ansatz for the active transport population (5.6). Recalling that $v_1(x) = \frac{\alpha K u(x)}{\lambda - v_0\nu + \beta}\frac{\tilde{v}_0}{\tilde{u}_0}v_2(x)$, and the expansion

$$\frac{1}{\lambda - v_0\nu + \beta} = \frac{1}{\beta} + \frac{1}{\beta(\beta + \alpha K\bar{u})}v_0\nu + \mathcal{O}(\nu^2)$$

for small ν, we obtain:

$$v_1(x) = \frac{\alpha K\bar{u}}{\beta}\frac{\tilde{v}_0}{\tilde{u}_0} + \mathcal{O}(\nu)\,. \tag{5.47}$$

Given the results of the inverse Fourier transform calculation in (5.34) and (5.35), determining the $\mathcal{O}(1)$ term is sufficient and suggests that the active transport population admits the same solution as the diffusing population for large time, scaled by the scalar $\frac{\alpha K\bar{u}}{\beta}\frac{\tilde{v}_0}{\tilde{u}_0}$. Note that here $\bar{u} = \langle u(x), 1\rangle$.

To summarize, the effective velocity and diffusion of the particles for large time for both populations is given by:

$$\begin{aligned} a_1 &= \frac{\alpha K\langle u(x),1\rangle}{\beta + \alpha K\langle u(x),1\rangle}v_0\,, \\ a_2 &= \frac{\beta}{\beta + \alpha K\langle u(x),1\rangle}D + \frac{\beta\alpha K\langle u(x),1\rangle}{(\beta + \alpha K\langle u(x),1\rangle)^3}v_0^2 + \frac{\beta\alpha K v_0}{(\beta + \alpha K\langle u(x),1\rangle)^2}\langle u(x), \tilde{w_1}(x)\rangle\,, \end{aligned} \tag{5.48}$$

with w_1 satisfying $D(w_1)_{xx} = \frac{\alpha K}{\beta + \alpha K\langle u(x),1\rangle}v_0\left(\langle u(x),1\rangle - u(x)\right)$. Given a concentration of signaling molecules at the cell periphery $u(x)$, this fully characterizes the behavior of the solutions for large time. ■

5.1.3 Comparison with previous results

In [72], the equation describing $c = \bar{p} + p_0$ for parallel filaments becomes:

$$\frac{\partial c}{\partial t} = \eta v_0 u \frac{\partial c}{\partial z} + D(x,t)\nabla^2 c(\boldsymbol{r},t) + Q_{zz}(x,t)\frac{\partial^2 c}{\partial z^2}\,,$$

where $\eta = \frac{\alpha K}{\beta}$. The assumption $\alpha K u \ll \beta$ simplifies the analysis and indicates that the density of microtubule filaments is assumed to be sufficiently small. The authors mention that a detailed analysis of the interaction between motor-cargo complexes and the cytoskeletal network is needed in order to provide a biophysical justification of this assumption [72].

Table 5.1 summarizes the comparison of our results for the effective velocity and diffusion of particles with active transport on parallel microtubules with the results in [72]. Recall that $\bar{u} = \langle u(x), 1\rangle$ and w_1 satisfies $D(w_1)_{xx} = \frac{\alpha K}{\beta + \alpha K \langle u(x),1\rangle} v_0 \left(\langle u(x), 1\rangle - u(x)\right)$. We also note that our analysis was carried out for $u = u(x)$ and no assumptions on the scale of the diffusion coefficient and transition rate parameters, while [72] analyzes equations (5.1) for $u = u(x)$ (and extends to $u = u(x,t)$), large switching rates and slow diffusion, and under the assumption of small microtubule density $\alpha K u \ll \beta$.

Table 5.1: Comparison of expressions for effective velocity and diffusion of particles undergoing diffusion and advection on a parallel array of microtubules using the Fourier analysis approach in §5.1.2 and the quasi-steady-state approximation approach in [72].

	Fourier analysis §5.1.2	QSS approximation [72]
Velocity (z)	$a_1 = \frac{\alpha K \bar{u}}{\beta + \alpha K \bar{u}} v_0$	$\boldsymbol{v}(\boldsymbol{r}) = \frac{\alpha K u}{\beta} v_0$
Diffusion (z)	$a_2 =$ $D\left(1 - \frac{\alpha K \bar{u}}{\beta + \alpha K \bar{u}}\right) +$ $\frac{\alpha K \bar{u}}{(\beta + \alpha K \bar{u})^3}\beta v_0^2 +$ $\frac{\beta v_0}{(\beta + \alpha K \bar{u})^2}\langle \alpha K u(x), w_1(x)\rangle$	$D(x,t) + Q_{zz}(x,t) =$ $D\left(1 - \frac{\alpha K u}{\beta}\right) +$ $\frac{\alpha K u}{\beta^3}\beta v_0^2$

It is worth noting that the results are very similar; however, the expressions for

the effective velocity and diffusion in [72] maintain the spatial dependence through the concentration of signaling molecules $u(x)$, whereas our results depend on the space-averaged quantity $\bar{u} = \langle u(x), 1 \rangle$. Excluding this consideration, the assumption $\alpha K u \ll \beta$ yields agreement in the effective velocities in the two approaches. Moreover, this assumption eliminates the third term in our expression for the effective diffusion and yields an expression equivalent to [72] (see table 5.1). The advantage of our method is that the complete expressions for effective velocity and diffusion we derive hold under no assumptions for the density of the microtubules, or for the magnitude of the reaction rates and diffusion coefficient.

In [72], the space-dependent expression for the effective diffusion is interpreted as a reflection of the stochastic nature of motor transport, yielding anisotropic diffusion. This is different from the motivating model for cytoskeleton transport in [73], where diffusion is assumed to be isotropic in the deterministic advection-diffusion equation for the concentration of molecules in the cytoplasm. A potential explanation for the space-dependent diffusion predicted by [72] may lie in the assumptions of the QSS approximation. In particular, the diffusion coefficient is assumed to be small ($\mathcal{O}(\varepsilon)$) so that moving by active transport allows particles to cover more of the space than by diffusion [74]; in addition, the transition rates are assumed to be very fast ($\mathcal{O}(1/\varepsilon)$) so that particles switch often between diffusion and movement as they explore the spatial domain [74]. For $0 < \varepsilon \ll 1$, the QSS approximation relies on the assumption that solutions remain close to the steady-state solution [72, 74]. Therefore, it is intuitive that the effective diffusion under these assumptions will have little influence from the diffusion state itself and more contributions from frequently switching to the moving state (which depends on the microtubule density $\rho(\boldsymbol{r}, \theta)$, or in the parallel filaments case on the concentration $u(x)$). This may explain why our approach, which does not make these assumptions on the parameters and calculates

the asymptotic quantities as $t \to \infty$, does not reflect spatial differences in the effective diffusion.

5.2 Numerical simulations of mRNA localization

The previous section provides an analytical approach to determining the large time mobility behavior of mRNA particles under the assumption of a parallel array of microtubular filaments. However, *Xenopus laevis* oocytes display microtubules that are randomly oriented in the cytoplasm, with a bias towards a radially outward orientation (see [3, Figure 1 and S3]). A choice of the microtubular density $\rho(\boldsymbol{r}, \theta)$ that reflects this outward orientation can be constructed with an approach similar to the one proposed in [72]. Since the large time analysis is more challenging under the assumption of radially outward microtubules, here we use a numerical approach similar to [6] for investigating the mRNA spatial distribution at different times during localization. We begin this section by outlining the assumptions we consider in creating model microtubule structures for this numerical approach. We then briefly describe the numerical methods used for simulating the PDE models with transport restricted to these model microtubules and outline the insights our approach provides on mechanisms of localization. The numerical framework developed allows us to compare model predictions with results from imaging experiments where fluorescently labeled mRNA is microinjected in oocytes in stage III of development [3] (see Figure 5.3).

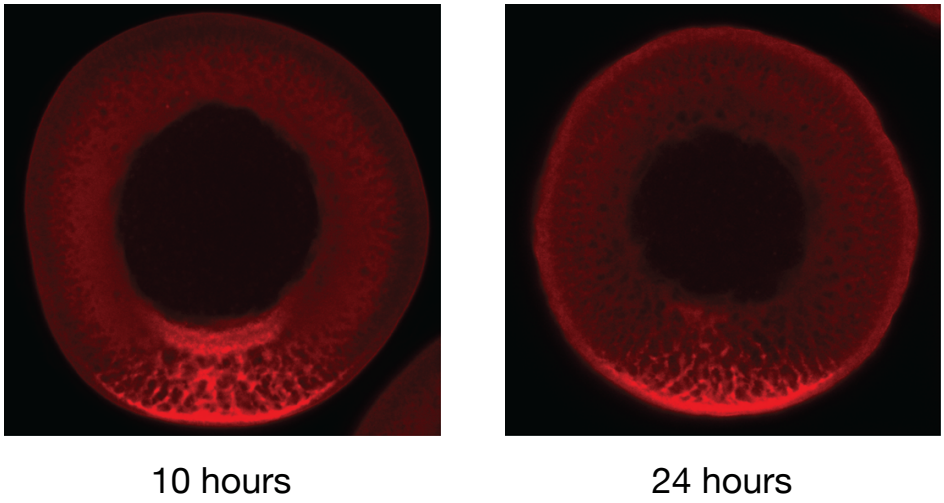

Figure 5.3: Images of fluorescently labeled VLE RNA microinjected into oocytes [3] and observed 10 hours, respectively 24 hours after injection in the localization process.

5.2.1 Models of microtubule structure

In order to generate model microtubule structures, we adapt the algorithm in [6] to reflect a 2-dimensional geometry and the assumptions needed for *Xenopus laevis* oocytes. Since the mRNA first accumulates under the oocyte nucleus prior to localization, we restrict our attention to a half oocyte (see Figure 5.4) for the purpose of these simulations.

While in *Drosophila* oocytes microtubules nucleate at the oocyte cortex [6], in *Xenopus laevis* the seeding points appear to be randomly distributed across the oocyte cytoplasm. Therefore we model the seeding points of the microtubules by implementing a random distribution of their location in a circular half-annulus (since the nucleus is excluded from the surface where microtubules are seeded). In particular, if U_1 and U_2 are uniform distributions on $[0, 1]$ and r_N is the radius of the

nucleus, we choose:

$$\rho = \sqrt{U_1(1^2 - r_N^2) + r_N^2}\,, \tag{5.49}$$

$$\varphi = \pi U_2\,, \tag{5.50}$$

so that $x = \rho\cos\varphi$ and $z = \rho\sin\varphi$ determine the location of the microtubule seeding points. Note that the radius of the oocyte is considered to have non-dimensional length 1.

Research on the distribution of microtubule lengths suggests that the Gamma distribution may accurately describe the catastrophe lengths of microtubules observed experimentally [19]. We note however that both our results and the findings in [6] are not changed by considering a simpler exponential distribution of the microtubule lengths. In the following, we therefore consider an exponential distribution of the filament lengths and further assume that the mean length of the microtubules is $9\mu m$, as informed by [18, 19, 75, 76].

As in [6], we model each microtubule as a sequence of straight segments (each of some constant length dL), with some variability in the orientations of consecutive segments. The orientation of the first segment is drawn randomly and accepted only if its direction is pointing inside the allowed geometry and in a radial direction [6] (half-oocyte annulus in our case). For the remaining orientations, we use the algorithm in [6] to draw the directions from the von Mises distribution on a circle. Note that this distribution requires knowledge of a parameter κ corresponding to the concentration (likelihood) around the previous segment orientation [6]. This parameter measures how close the orientation of subsequent segments is to the previous segments; given that microtubules are neither completely straight nor completely curled up in both *Xenopus* and *Drosophila* oocytes [3, 6, 76, 77], we follow the ap-

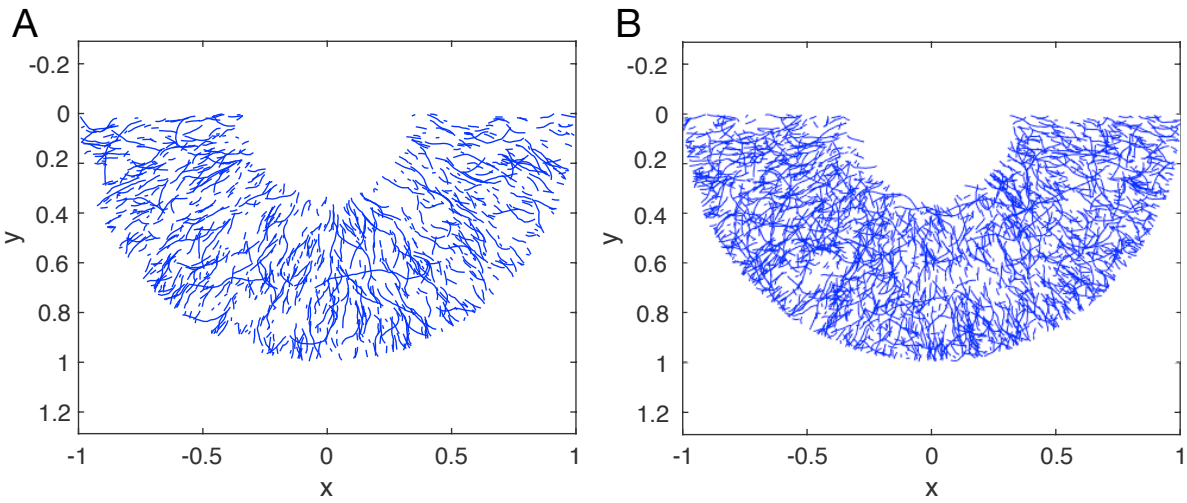

Figure 5.4: Sample microtubular structures with (A) 1000 microtubule filaments and (B) 2000 microtubule filaments (algorithm adapted from [6]).

proach in [6] and choose $\kappa = 18$ which reflects the randomness in orientations (see Figure 5.4). Moreover, the bias for a radially outward orientation of microtubules in *Xenopus* oocytes leads to our modeling assumption that half of the microtubules have random orientations and half have radially outward orientations for the segments following the first.

Figure 5.4 shows two sample microtubule structures with 1000, respectively 2000 microtubules and reflecting the assumptions in this section. While it is challenging to quantify the differences between these model microtubule structures and images of the microtubule cytoskeleton as in [3, Figure S3], these structures compare well visually with the observed experimental structures. It is worth noting that the microtubule filaments are dynamic and undergo catastrophes on the time scale of localization, with mean lifetimes of roughly 5 minutes on average [19]. As a result, we generate multiple such structures and draw from them at random to set a distinct background microtubule structure at different times in the simulation of the partial differential equations models considered.

5.2.2 Numerical methods and parameter choices

As mentioned in the previous section, we consider a two-dimensional domain consisting of a half-circle with nondimensional radius $R = 1$ (thus $x \in [-1, 1]$ and $y \in [0, 1]$). This radius corresponds to roughly 150 μm, which is the average radius for oocytes in stage III of development [3]. In order to carry out simulations of the dynamics of the mRNA, we further exclude the nucleus from the domain by assuming its radius is roughly $r_N = 0.33$ (non-dimensionalized), or 50 μm.

Given this domain choice and the model microtubule structures for *Xenopus* oocytes, we follow the approach in [6] and calculate the velocity field where active transport by molecular motor proteins can occur. To this end, [6] sets up a grid of nondimensional length $dG = 0.04$, calculates the midpoint of each microtubule segment, and sums vectorially the orientations of all such segments whose midpoints are located in the same grid area. After normalizing, this yields a local motor-velocity vector field V_m [6] representing the direction along which mRNA molecules can be actively transported by kinesin motors (towards the positive ends of microtubules). In our simulations, we take the finer grid length $dG = 0.02$ to capture the dynamics more accurately and consider model microtubule structures with 5000 filaments each; Figure 5.4 illustrates cytoskeletons with fewer microtubules for ease in visualization. Since in the 4-state model we assume bidirectional transport along microtubules, our work extends the approach in [6] by also considering a motor velocity field V_r with the opposite orientation from V_m. This additional field can be interpreted as active transport of mRNA by dynein motors (towards the negative ends of microtubules).

In [6], the model microtubule cytoskeleton also provides insight into the cytoplasmic flows driven by kinesin motor transport. In *Drosophila* oocytes, the mRNAs are synthesized in the attached nurse cells and transported into the oocyte, so that

the cytoplasmic flows may be important in the mobility of RNA. By contrast in *Xenopus*, all of the mRNAs are synthesized in the oocyte nucleus, thus we do not consider flows in our simulations of mRNA localization. It is also worth noting that [6] concludes that the cytoplasmic flows do not play a key role in the localization process.

The dynamics of mRNA is then simulated using the following equations for the 2-state model:

$$\begin{aligned} \frac{\partial p(\boldsymbol{r},\theta,t)}{\partial t} &= -\boldsymbol{V}(\theta)\cdot\nabla p(\boldsymbol{r},\theta,t) - \beta p(\boldsymbol{r},\theta,t) + \alpha\rho(\boldsymbol{r},\theta)p_0(\boldsymbol{r},t)\,, \\ \frac{\partial p_0(\boldsymbol{r},t)}{\partial t} &= D\nabla^2 p_0(\boldsymbol{r},t) + \beta\int_0^\pi p(\boldsymbol{r},\theta,t)d\theta - \alpha\bar{\rho}(\boldsymbol{r})p_0(\boldsymbol{r},t)\,, \end{aligned} \tag{5.51}$$

which are identical to the system (5.2) with no assumptions on the parameters and with density $\rho(\boldsymbol{r},\theta)$ given by the model microtubule structure. The partial differential equations are solved using a finite-volume discretization on staggered grids with no flux boundary conditions [6, 78]. The 4-state system of equations is easily extended from equations (5.51) to include two active transport states as well as an additional paused state, and the implementation of the numerical methods extends the algorithm in [6] to account for these states. In addition to accounting for bidirectional transport in the 4-state model, our simulations use realistic spatially-dependent parameters obtained using parameter estimation in the 3 regions where FRAP is performed in the oocyte cytoplasm (see Figure 4.7A). We assume that Region 1 extends 20 μm from the nucleus into the cytoplasm, and Region 3 extends 40 μm from the vegetal cortex into the cytoplasm. We further make the assumption that the region boundaries are set radially.

The initial conditions for the nonlinear PDEs in (5.51) are set to model the initial early accumulation of mRNA in the perinuclear cup region (under the nucleus) [3].

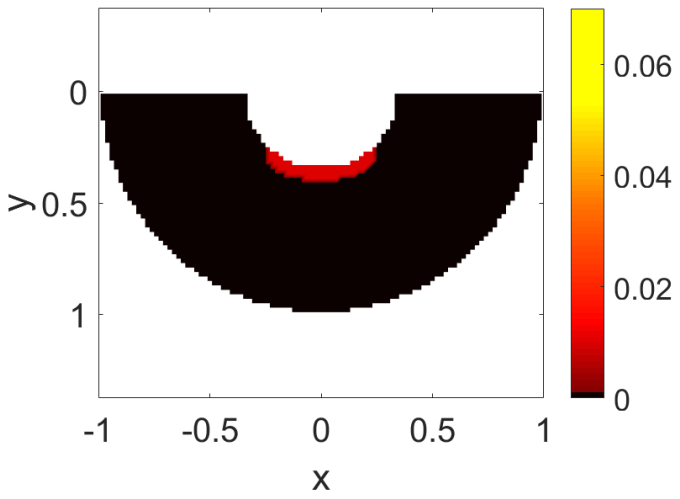

Figure 5.5: Initial condition for the mRNA distribution assuming that the mRNA initially localizes next to the nucleus or is injected experimentally in the perinuclear cup. Color bar for the constructed color map is included.

We therefore set a uniform positive value initial condition in the region $R_1 \leq r \leq R_2$ and $x_1 \leq x \leq x_2$, where $R_1 = r_N = 0.33$, $R_2 = 0.4$, $x_1 = -0.25$ and $x_2 = 0.25$, and 0 everywhere else in the domain. Figure 5.5 illustrates this initial condition, and also includes the Matlab color map created to visualize the results of our simulations and to compare with [3, Figures 1C, 3D and S2, A-C] and Figure 5.3. We note that in the following simulations, we pick the color bar consistently the same within simulations, but the maximum value on the color bar axis may be chosen differently in simulations where the concentration of mRNA is more spread out in the cytoplasm, for ease in visualization.

Finally, in §5.2.3 we will be interested in the amount of mRNA localized at the target location (vegetal cortex). To define this quantity, we introduce δ as a parameter that determines the width of the target region. In particular, our choice of $\delta = 0.9$ means that we calculate the fraction of mRNA localized in the annular region $0.9 \leq R \leq 1$ right next to the cortex. In addition, we use another parameter $\delta_{\text{anchor}} = 0.95$ to separate a potential region where mRNA may get anchored at

the cell cortex. While in stage 9 *Drosophila* oocytes anchoring is not required for localization [6], our findings in the next section show that this mechanism may be key in achieving the predicted timescales of localization in *Xenopus* oocytes.

5.2.3 Predictions for mRNA dynamics

We first test the algorithm adapted from [6] using the 2-state model and parameters estimated using an average FRAP dataset based on 5 wild-type oocytes. Given an initial condition as in Figure 5.5 for both the moving and the diffusing populations of VLE RNA, we determine numerically the spatial distribution of VLE RNA over 24 hours of localization using 24 different microtubule structures as in Figure 5.4 (with 5000 microtubules each); we note that this corresponds to randomly changing the microtubule cytoskeleton every hour during the simulation. Figure 5.6A-C shows the distribution of mRNA at 4, 10, and respectively 24 hours during localization. Panel D illustrates the evolution of the fraction of mRNA localized within $1 - \delta = 10\%$ of the cortex, with more than 95% of the molecules localized by 15 hours after the accumulation at the perinuclear cup. Therefore, the 2-state model predicts that the localization of mRNA at the vegetal cortex is faster than observed experimentally, where mRNA localizes in 24-48 hours after injection [3].

To determine the localization dynamics predicted by the 4-state model, we use the extended algorithm and parameters estimated using an individual FRAP oocyte. Since this model accounts for bidirectional transport along microtubules, it is crucial to also consider the orientation of the microtubules driving animal and vegetal transport in the cytoplasm of *Xenopus* oocytes. The direction of growth of microtubules is denoted as positive, while the opposite direction is denoted as negative. Immunostaining experiments in [7] find that there exists a population of growing plus

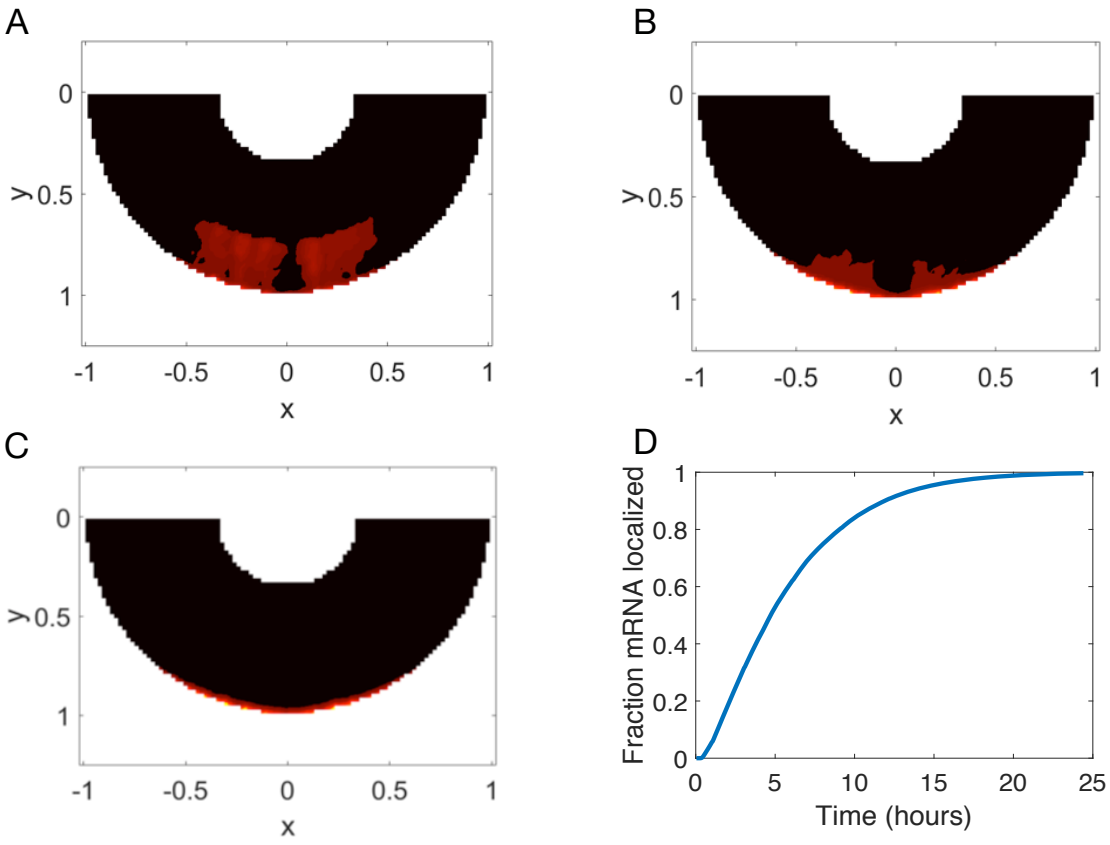

Figure 5.6: Spatial distribution of mRNA predicted by the 2-state model at (A) 4, (B) 10, and (C) 24 hours after injection at the perinuclear cup (under the nucleus, see initial condition in Figure 5.5). (D) illustrates the time evolution of the fraction of mRNA localized 10% from the vegetal cortex.

ends (for kinesin-mediated transport) at the vegetal cortex of *Xenopus* oocytes, and that minus ends (for dynein-mediated transport) are present throughout the vegetal cytoplasm but appear considerably more dense at the side wedges (see Figure 5.7B). Therefore, in the simulations for the 4-state model we consider a microtubule structure for downward transport with uniform nucleation as in Figures 5.4 and 5.7A, as well as a microtubule structure for upward transport that accounts for the distribution of minus ends discovered in [7, Figure 6G] as in Figure 5.7B. In the structure corresponding to upward transport, two thirds of the microtubules are uniformly nucleated throughout the half-cytoplasm, and a third is concentrated in the side wedges. In addition, we assume that all microtubules in this structure have an overall radial orientation. Figure 5.7 shows the density of the microtubules in the

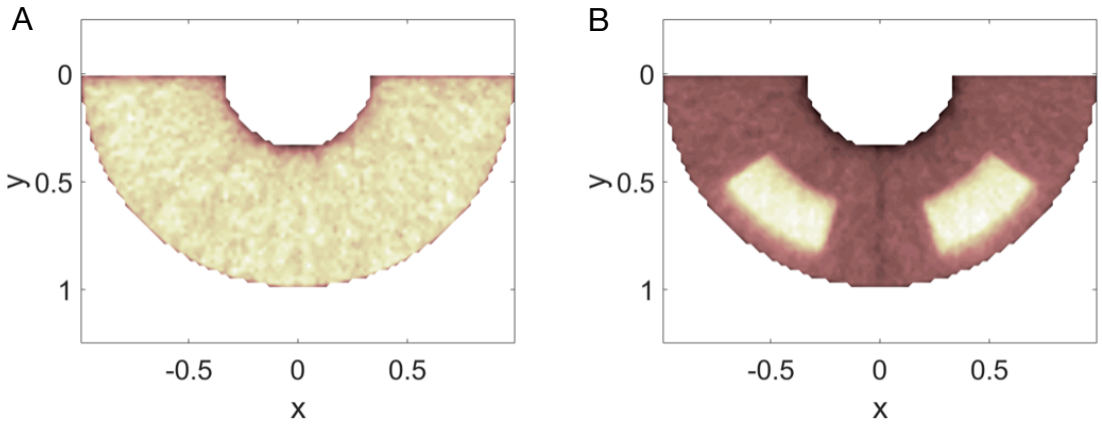

Figure 5.7: Spatial density of microtubules modeling (A) transport to the vegetal cortex and (B) transport to the animal pole in agreement with observations in [7] (visualization uses *pink* color map in Matlab).

cytoplasm for both of the microtubule cytoskeletons considered.

Figure 5.8 shows that the spatial distributions of mRNA 4 and 10 hours into localization more closely resemble the imaging experiments in [3, Figures S2] than the results of the 2-state model (Figure 5.6A-C), but the mRNA is not fully localized 24 or even 48 hours after mRNA accumulation next to the nucleus. We overlay the microtubule density for animal pole transport from Figure 5.7B to these distributions to show that the location where we model more intense bidirectional transport limits the spread of localization, especially in the upper cytoplasm. Figure 5.8D further confirms that mRNA does not localize in the 24 hours simulated, and shows that the fraction of localization converges to roughly 38% in a day. As a result, the 4-state model parameters estimated using FRAP data predict that achieving complete localization of mRNA in *Xenopus* oocytes may require anchoring at the cell cortex, as hypothesized in [3]. We also note that simulations considering the same microtubule structures for both vegetal and animal pole transport (as in Figure 5.7A) lead to similar timescales of RNA localization, however the spatial distribution of RNA extends throughout the vegetal cortex and therefore does not agree with the imaging experiments in [3] (see also Figure 5.3).

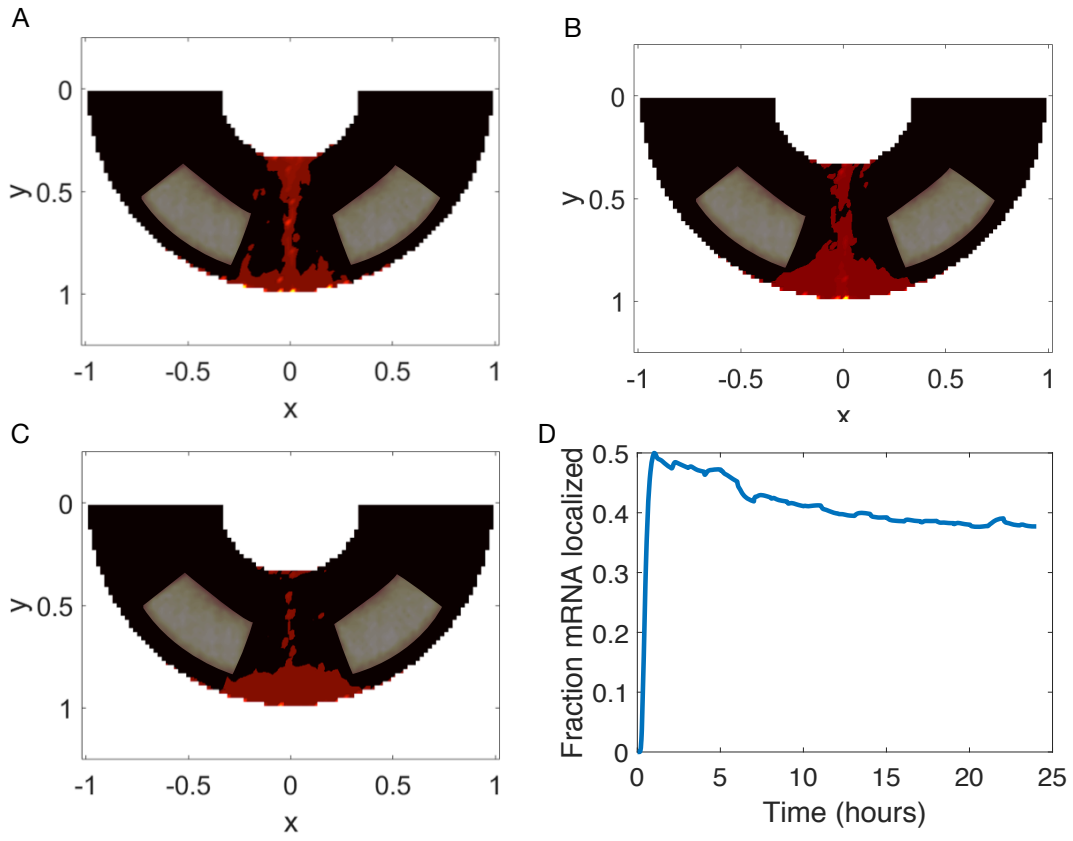

Figure 5.8: Spatial distribution of mRNA predicted by the 4-state model at (A) 4, (B) 10, and (C) 24 hours after injection at the perinuclear cup (see initial condition in Figure 5.5). (D) illustrates the time evolution of the fraction of mRNA localized 10% from the vegetal cortex.

We further test the 4-state model under different assumptions for the initial condition and microtubule orientations. Figure 5.9B-D shows that the distributions of mRNA throughout localization are not considerably changed when considering a wider initial condition as in panel A with $x \in [-0.35, 0.35]$ (compared to $x \in [-0.25, 0.25]$ in Figure 5.5). Though not shown, the fraction of mRNA localized closely resembles Figure 5.8D. To better understand how the wider initial condition leads to similar narrow mRNA localization spread, we provide additional plots of the dynamics 0-4 hours after RNA injection in Figure 5.10. We note that the mRNA starts spread out under the nucleus (see also initial condition in Figure 5.9A) but progresses towards a more narrow distribution, due to the assumed locations for upward radial microtubule transport.

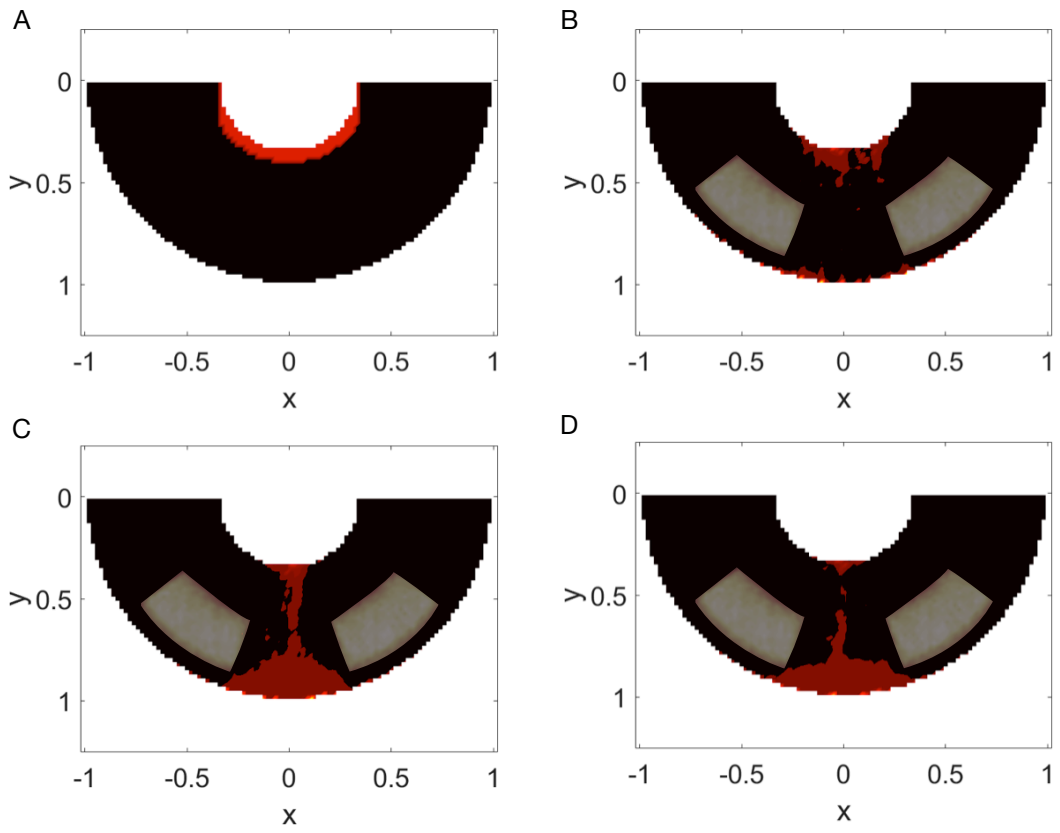

Figure 5.9: Spatial distribution of mRNA predicted by the 4-state model at (B) 4, (C) 10, and (D) 24 hours after injection at the perinuclear cup, with initial condition given by a wider distribution of mRNA in the perinuclear cup as shown in (A).

On the other hand, taking the microtubule structure for upward transport to consist of half randomly-oriented and half radially-oriented filaments changes the spatial distribution by predicting that the mRNA spreads out more in the lower cytoplasm (see Figure 5.11A-C). It is worth noting that this is not consistent with experimental observations in [3], where the width of the mRNA spread at the vegetal cortex is limited and does not extend to the entire half-oocyte. These results also suggest that mRNA localization may be dependent on the microtubule orientations in the cytoplasm, whose organization likely depends on previous cell signaling processes.

Finally, we also consider a setting for the microtubule structure that matches the model proposed in [3, Figure 7]. Based on experiments that knock down the

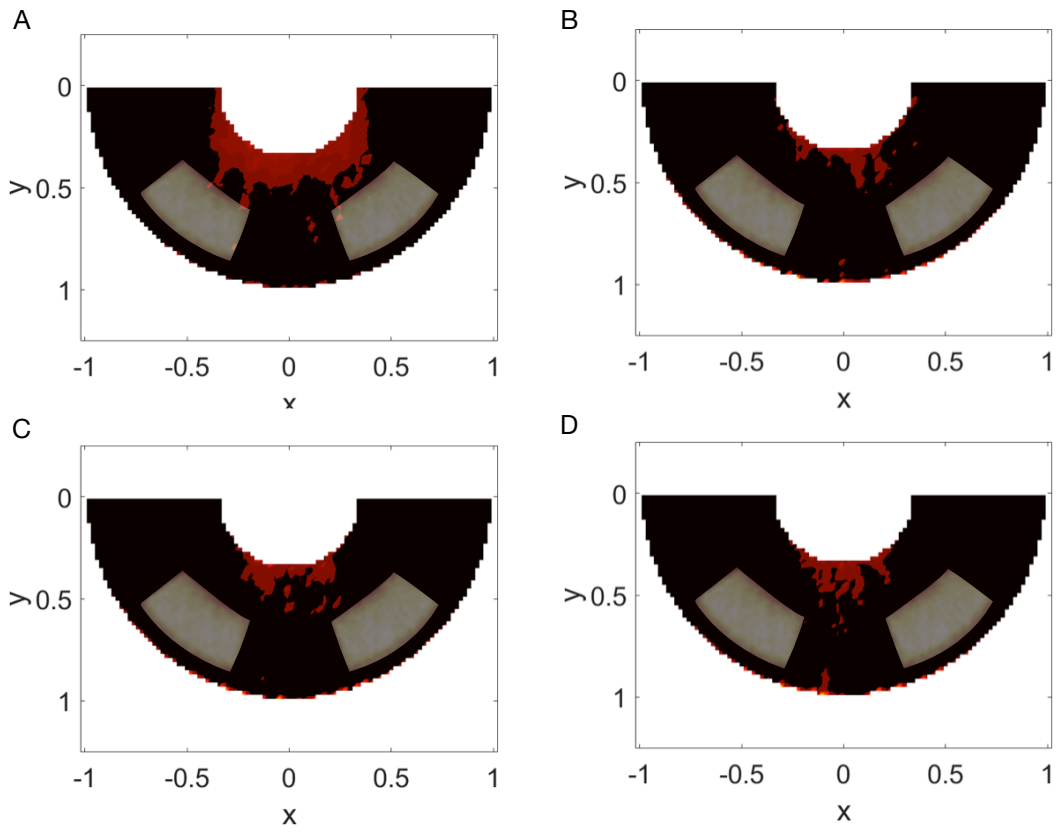

Figure 5.10: Spatial distribution of mRNA predicted by the 4-state model at (A) 0.5, (B) 1, (C) 2, and (D) 3 hours after injection at the perinuclear cup, with initial condition given by a wider distribution of mRNA in the perinuclear cup as shown in Figure 5.9A.

kinesin and dynein motor proteins separately, [3] suggests a hypothesis of unidirectional vegetal transport in the upper cytoplasm (driven by dynein), and bidirectional animal and vegetal transport in the lower cytoplasm (driven by kinesin) prior to a potential anchoring step at the the cortex. We model this proposed set-up using model microtubule structures for upward transport that are more dense in a wedge centrally located in the lower vegetal cytoplasm (see Figure 5.12D). The results in Figure 5.12A-C are similar to the mRNA distributions we observe when using side wedge densities (see Figures 5.7 and 5.8), and the amount of RNA localized similarly converges to roughly 38% in 24 hours. However, having upward pointing microtubules in a central wedge means that mRNA may be slower to accumulate in the lower vegetal cytoplasm (above the vegetal cortex), with multiple simulations

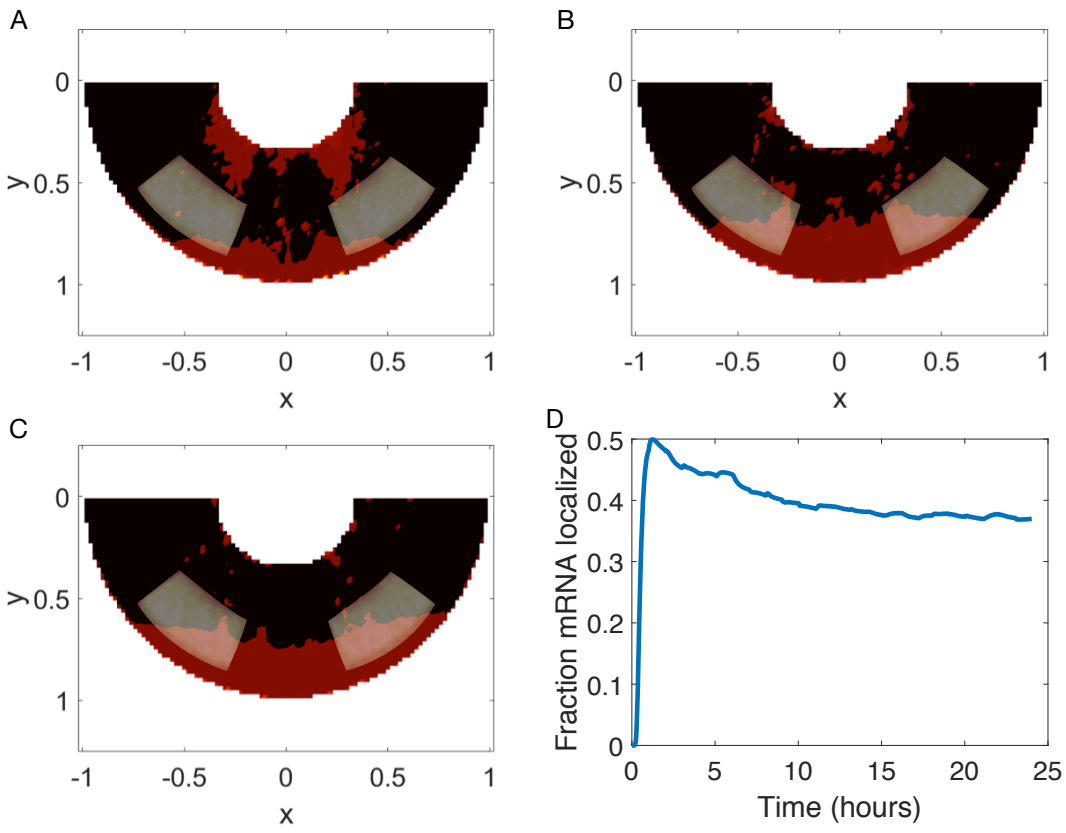

Figure 5.11: Spatial distribution of mRNA predicted by the 4-state model at (A) 4, (B) 10, and (C) 24 hours after injection at the perinuclear cup (see initial condition in Figure 5.5) with upward microtubule structures assumed to be oriented half randomly and half radially outward. (D) illustrates the time evolution of the fraction of mRNA localized 10% from the vegetal cortex.

showing no significant mRNA concentration in this location 10-12 hours into localization (see Figure 5.12A,B). Additional imaging experiments from the Mowry lab show mRNA distributed throughout the vegetal wedge 8 hours into localization and seem to support the set-up in Figure 5.7. Future experiments may provide more insight into the microtubule cytoskeleton that yields the observed localization patterns.

5.2.4 Insights on anchoring mechanisms

Despite the limited experimental insight into the anchoring mechanism at the cell bottom, the numerical approach developed in this section allows us to investigate an-

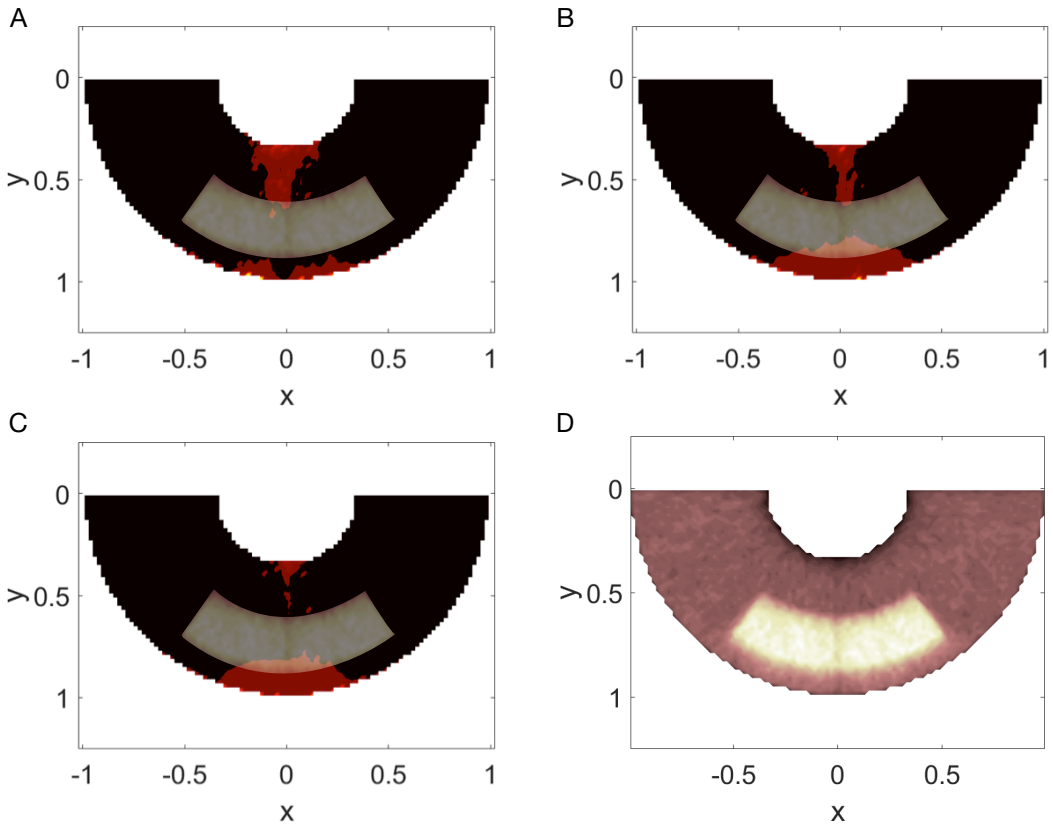

Figure 5.12: Spatial distribution of mRNA predicted by the 4-state model at (A) 4, (B) 10, and (C) 24 hours after injection at the perinuclear cup (see initial condition in Figure 5.5) with upward microtubule structures assumed to be more dense in a central vegetal wedge as in (D).

choring mechanisms that may be responsible for the spatial and timescale of healthy localization and development in *Xenopus* oocytes. In particular, modifying the rates between different dynamical states in an anchoring location (assumed to be 5% from the cortex in our simulations) can reveal the states of mRNA or mRNA-motor protein complexes involved in anchoring. Recalling the 4-state model cartoon in §2.3.1 (Figure 2.5B), we test the anchoring of the paused state by setting rates $\alpha_+ = \alpha_- = 0$. The results are almost unchanged from the baseline simulation with no anchoring (Figure 5.8), indicating that anchoring of the stationary state alone is not sufficient to achieve localization in 24-48 hours after RNA injection.

While [6] notes that anchoring is not predicted to play a role in oskar RNA localization simulations for *Drosophila* oocytes, this study tests anchoring for a model

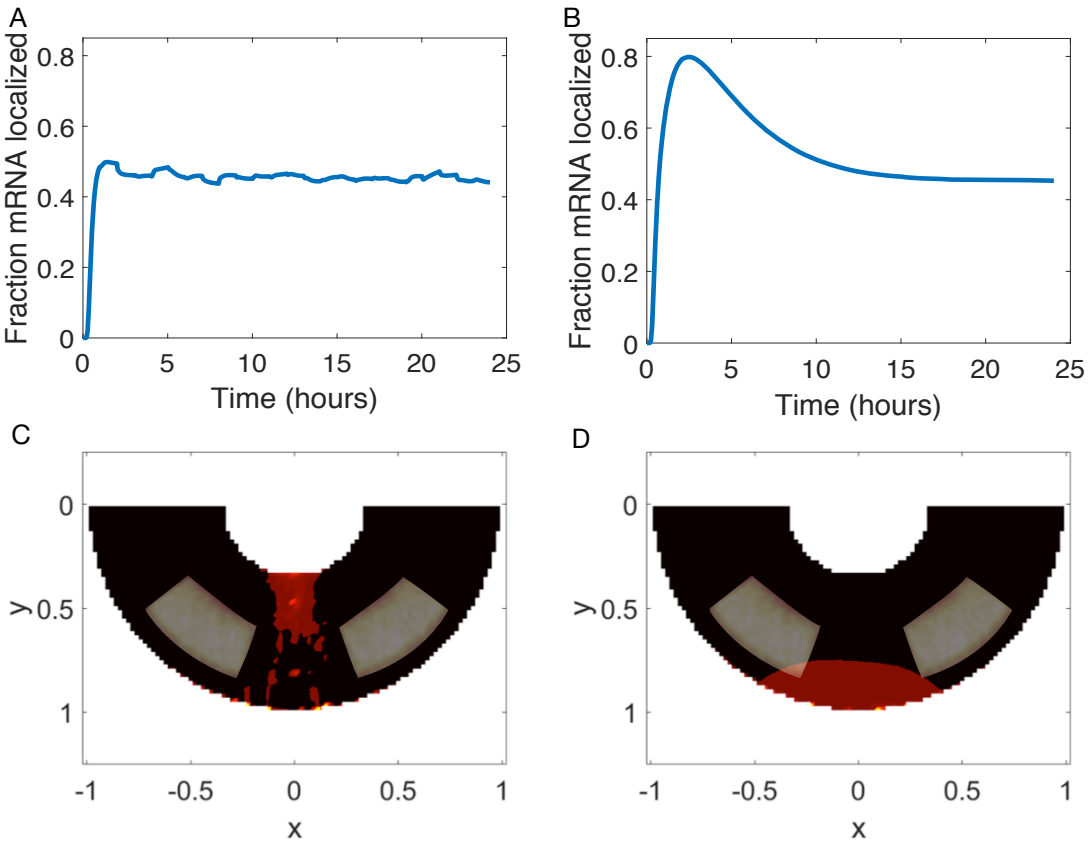

Figure 5.13: Top: Time evolution of the fraction of mRNA localized 10% from the vegetal cortex when setting $\delta_+ = 0$ (A) and $\gamma_+ = 0$ (B) in the 4-state model. Bottom: Spatial distribution of mRNA predicted by the 4-state model with $\delta_+ = 0$ (C) and $\gamma_+ = 0$ (D) 10 hours after injection at the perinuclear cup.

similar to our 2-state model by setting the unbinding rate for the moving state to 0. In addition, [3] suggests that mRNA that reaches the vegetal cortex may be captured and stably anchored at this location through the activity of the dynein motor protein, which has been shown to be able to transition from its transport function to a stable anchor function for localised RNAs in *Drosophila* blastoderm embryos and oocytes [79, 80]. We therefore test whether reducing the transition rates out of the moving down state (δ_+ and γ_+) in an anchoring region has an influence on the localization timescale and patterns predicted.

When setting the rate from the moving down state to the stationary state $\delta_+ = 0$ close to the cortex, our simulations show that the fraction of mRNA localized quickly

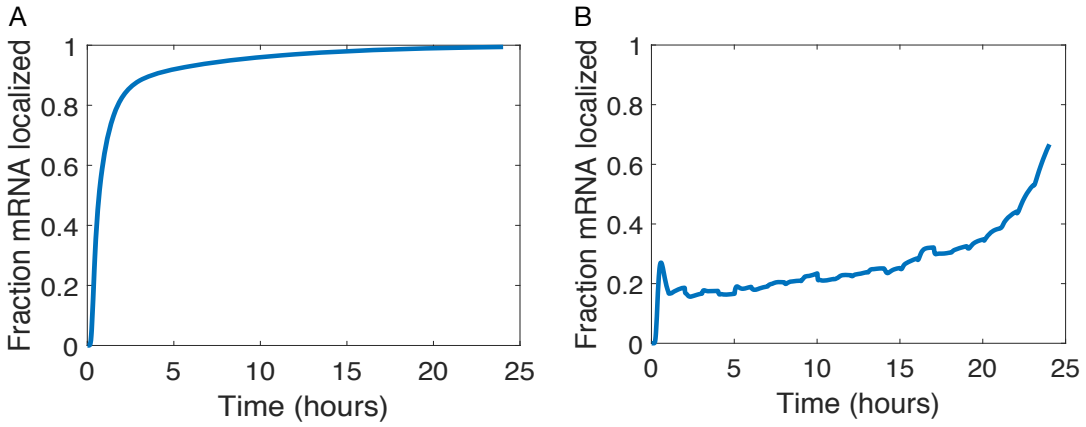

Figure 5.14: Time evolution of the fraction of mRNA localized 10% from the vegetal cortex using the 4-state model with $\delta_+ = 0$ and $\gamma_+ = 0$ (A) and γ_+ decreasing linearly from $10\gamma_+^*$ to 0 in 24 hours (B) (where γ_+^* is the value of the transition rate in Region 3 of the cytoplasm).

converges to roughly 44%, a small increase from the baseline (no anchoring) case (see Figure 5.13A). If on the other hand we set the rate from the moving down state to the diffusion state $\gamma_+ = 0$, Figure 5.13B shows that the percentage of RNA localized 10% from the cortex increases quickly to 78% and then converges to a stable value of roughly 45%. Figure 5.13C-D also shows the distribution of the mRNA in the cytoplasm 10 hours after mRNA injection for both anchoring scenarios.

Setting both transition rates out of the moving down state to 0 ($\delta_+ = \gamma_+ = 0$) in the anchoring region leads to complete localization achieved in 24 hours (see Figure 5.14A). These simulations suggest that anchoring at the cell cortex may depend on stabilization of mRNA-motor complexes while moving down towards the vegetal pole. This hypothesis is supported by experiments in [3, Figure S8], where dynein motor proteins appear to be co-localized with VLE RNA at the cell cortex 24 hours into localization. However, the dynamics when setting both transition rates to 0 is undoubtedly very fast, with 90% of localization occurring in less than 4 hours into the localization process. Therefore, besides pointing to a necessary anchoring mechanism in the vegetal cortex, another potential insight from our simulations is that the transition rates in this anchoring region may be regulated to vary with space

or time. Our approach allows to test such hypotheses; for example, Figure 5.14B shows the predicted localization when $\delta_+ = 0$ and γ_+ is set to decrease linearly with time from $10\gamma_+^*$ to 0 over 24 hours of simulation time. It is worth noting that the evolution of localization is considerably changed in this setting. While images of localization fraction as in Figure 5.14 are not currently available from experiments in *Xenopus* oocytes, this work directs future experimental attention to the mechanisms of anchoring at the cell cortex and to a potential role of motor proteins in stabilizing RNA.

Chapter Six

Conclusion

In this work we have presented several modeling, analytical and computational frameworks that provide insight into the dynamics of messenger RNA in *Xenopus laevis* oocytes. Since complete localization of mRNA is key in healthy development, we investigated several mechanisms that may be responsible for the accumulation of mRNA at the cell cortex. The methods developed here can extend more generally to understand diffusion, active transport, and binding kinetics of molecules in many living cells.

To summarize from the Introduction, our contributions include:

- *FRAP data analysis:* Using active transport models for particle dynamics, we developed efficient algorithms that analyze FRAP data to estimate key parameters such as velocities of motor-RNA complexes, diffusion coefficients of free RNA, and transition rates between different dynamical states.

- *Large time analysis of transport models:* We proved that the dynamics of actively transported molecules at large time is captured by two quantities, namely the effective velocity and effective diffusion, and we related these quantities analytically to parameters extracted from FRAP data using general transport models.

- *Predictions for mRNA dynamics:* Our analysis of FRAP data provided support for the role of paused states and bidirectional transport in *Xenopus* oocyte mRNA dynamics.

- *Simulations of nonlinear PDE models:* Parameter estimates from FRAP data informed our simulations of transport models on computational model microtubule structures, which allowed the comparison of mRNA spatial distributions with experimental data and suggested anchoring at the cell cortex is a

key mechanism for the timescale of localization.

We elaborate on these findings in the following sections.

FRAP data analysis

A new numerical approach to parameter estimation was introduced for FRAP (fluorescence recovery after photobleaching) data for models of active transport. These methods apply to intracellular dynamics in any organism where directed movement (e.g., by molecular motor proteins) is believed to play a key role in particle localization. The advantage of the proposed method is that it can be applied to any FRAP bleach spot geometry, it does not require normalization of the FRAP data, and it can take into account the post-bleach intensity profile distribution. This approach allows us to determine transport parameters for different particle states (active transport, diffusion, etc.) even though FRAP data does not distinguish between these populations. As in [35, 36], the model equations we consider cannot be solved analytically (equations (2.7) and (2.8)) or have a complex initial condition geometry (equations (2.9) with initial condition (2.12)). This motivates the use of efficient numerical integration for systems of advection-reaction-diffusion PDEs [50, 51], and of large parameter sweeps for the model parameters. The computational cost of both these parameter sweeps and optimization with multiple initial conditions is significantly reduced through the use of parallel computation. The method is validated using FRAP recovery curves simulated using the PDE models (Algorithm 1) as well as using Continuous-time Markov chain models (Algorithm 2) of the dynamical processes [1].

Initial conditions for these dynamical systems have been shown to be affected by the temporal limitations of bleaching and scanning in confocal FRAP experiments [34]. We used the experimental spatial distribution of fluorescence right after pho-

tobleaching in order to take into account the dynamics that may occur during the bleaching process. In our experimental setup, these initial postbleach profiles were well described by the exponential of a Gaussian (Figure 2.8). Our results suggest that failing to account for the postbleach dynamics in our experiments under-estimates diffusion coefficients roughly thirty-fold, and can have an impact on the order of magnitude of parameters such as velocities and rate constants. This observation is similar to conclusions in studies of diffusion and binding-diffusion models, where the assumption of negligible diffusion during bleaching leads to significant under-estimation of the magnitude of the diffusion coefficient [1, 32, 34].

Large time analysis of transport models

While analytical solutions for PDE models of active transport are not available, we provide rigorous mathematical derivations for solutions of the general system:

$$\frac{\partial \boldsymbol{u}(y,t)}{\partial t} = A\boldsymbol{u} + C\partial_y \boldsymbol{u} + D\Delta \boldsymbol{u}\,, \tag{6.1}$$

for large time. Here $\boldsymbol{u}$ is a column vector for the distributions of n populations of particles with different dynamics, and $A, C, D \in \mathbb{R}^{n\times n}$. C is a diagonal matrix of the velocities of the n populations, while D is a diagonal matrix of their diffusion coefficients. Considering a Fourier mode ansatz and carrying out the Lyapunov Schmidt reduction of the resulting equation allow us to obtain expressions for the effective velocity and diffusion of a particle for large time, given bidirectional movement, diffusion and binding dynamics in an arbitrary number of states. These quantities correspond to the bulk movement and diffusion of the particles given their transition rates between states, and provide a useful comparison to observed timescales of intracellular movement. Model selection is facilitated by the derivations of fractions of particles in each state, as well as of expected run times and lengths of cargo

on microtubules. Knowledge of biologically-relevant timescales and length scales for microtubule travel provides comparison and validation for the models of active transport appropriate in different experimental settings (Figure 4.8). This is particularly important in applications where the regime of the parameter space is not known, so that simplifications of the PDE models as considered in [27] are difficult. For these situations, we propose using general PDE models that build up to the complete range of mechanisms that are believed to influence particle dynamics. These models can then be further validated and compared using derivations of biological quantities of relevance as described above [1].

Predictions for mRNA dynamics

In measuring mobility of mRNA using FRAP data from *Xenopus* oocytes, our results for effective particle speed and diffusion (see Figure 1.3) confirm the differences in dynamics between localizing VLE RNA and non-localizing RNA (Figure 4.7). Moreover, we confirm that distinct kinetics and transport directionality can be expected for RNA transport in different cytoplasmic regions of a single cell, as hypothesized in [3]. We provide further evidence that bidirectional transport occurs in all regions of the cytoplasm (Figure 4.8), but the different velocities of the molecular motor proteins in different areas of the cell suggest that vegetal transport has a more clear directional bias in the upper vegetal cytoplasm (Regions 1 and 2) than in the lower vegetal cytoplasm (Region 3) [1]. This is consistent with the finding that RNA is transported by kinesin motors on a bidirectional array of microtubules close to the vegetal cortex as suggested in [7].

Simulations of nonlinear PDE models

The parameter estimates for distinct cytoplasm regions also allow us to observe

mRNA localization through accurate two-dimensional simulations [6] of the transport models with movement restricted to the microtubule cytoskeleton discovered in [7]. This numerical framework provides a direct comparison with experimentally-observed patterns of RNA localization as in [3], and further suggests that bidirectional transport and an anchoring mechanism at the cell cortex are required in order to reproduce time and spatial scales of localization. Our work points to the need to direct experimental attention to anchoring of mRNA at the cell cortex, and to the possible role that motor proteins may play in anchoring the molecules at the periphery. By extending the dynamical systems analysis to the nonlinear PDE advection-diffusion model:

$$\begin{aligned} \frac{\partial p(\boldsymbol{r},\theta,t)}{\partial t} &= -\boldsymbol{V}(\theta)\cdot\nabla p(\boldsymbol{r},\theta,t) - \beta p(\boldsymbol{r},\theta,t) + \alpha\rho(\boldsymbol{r},\theta)p_0(\boldsymbol{r},t)\,, \\ \frac{\partial p_0(\boldsymbol{r},t)}{\partial t} &= D_0\nabla^2 p_0(\boldsymbol{r},t) + \beta\int_0^\pi p(\boldsymbol{r},\theta,t)d\theta - \alpha\bar{\rho}(\boldsymbol{r})p_0(\boldsymbol{r},t)\,, \end{aligned} \tag{6.2}$$

we complemented the numerical approach with the analysis of how the mRNA distribution behaves at large time given that microtubules are oriented in an array parallel to the animal-vegetal axis. A natural next step for this work would be to provide insight into the influence of the microtubular density $\rho(\boldsymbol{r},\theta)$ on localization patterns by studying equations (6.2) analytically for different microtubule geometries. In particular, considering a density of microtubules pointing outward would be relevant to the *Xenopus* oocyte cytoskeleton and would provide a comparison to the numerical and experimental results.

6.1 Open problems

Our methods for extracting velocity, diffusion and binding rate information from FRAP recovery data are broadly applicable to other systems where active transport is involved in intracellular dynamics, such as mRNA localization in *Drosophila* oocytes [81] or neurofilament transport along axons in neurons [38, 39]. The numerical approach to estimate parameters based on FRAP data can be applied to any biomolecules that are believed to undergo diffusion, active transport, and binding kinetics. In addition, the modeling and parameter estimation of FRAP experiments can also be extended to complementary experiments such as fluorescence correlation spectroscopy (FCS) and photoactivation (PA), which would provide additional validation of parameters for transport mechanisms [1].

To improve predictions for the mechanisms of mRNA localization in *Xenopus* oocytes, our macroscopic models of mRNA localization may be combined with microscopic modeling of interactions of the competing motor proteins kinesin and dynein. Various studies reviewed in [46] point to hypotheses such as a tug of war between motors directed in opposite directions as well as a potential co-dependence of motors in staying attached to microtubules to achieve transport. As a better understanding of the antagonistic behavior of kinesin and dynein motors becomes available [46], this information may allow for better predictions of how the localization spread changes and could provide valuable insights into the spatio-temporal regulation of the molecular motors in the egg cell. In particular, this approach would be useful in studying abnormal localization when transport by motor proteins is inhibited, and could be combined with parameter estimation using fluorescence and imaging data from dynein- or kinesin-impaired oocytes [3].

In addition, prior to its transport toward the vegetal cortex, mRNA moves from its uniform distribution throughout the cytoplasm to localize under the nucleus in a region denoted the perinuclear cup (see Figure 1.2). Our work may provide a starting point for understanding the mechanisms through which mRNA accumulates next to the nucleus before localizing in *Xenopus laevis* oogenesis.

Bibliography

[1] M.-V. Ciocanel, J. Kreiling, J. Gagnon, K. Mowry, B. Sandstede, Analysis of active transport by fluorescence recovery after photobleaching, Biophysical Journal (accepted).

[2] K. Mowry, Mowry lab research images, available at `http://www.brown.edu/Departments/Molecular_Biology/mowry/photos.html`.

[3] J. A. Gagnon, J. A. Kreiling, E. A. Powrie, T. R. Wood, K. L. Mowry, Directional transport is mediated by a dynein-dependent step in an RNA localization pathway, PLOS Biol 11 (4) (2013) e1001551.

[4] A. Witthoft, J. A. Filosa, G. E. Karniadakis, Potassium buffering in the neurovascular unit: models and sensitivity analysis, Biophysical Journal 105 (9) (2013) 2046–2054.

[5] A. Witthoft, Models of neurovascular coupling in the brain, Ph.D. thesis, Brown University (2015).

[6] P. K. Trong, H. Doerflinger, J. Dunkel, D. St Johnston, R. E. Goldstein, Cortical microtubule nucleation can organise the cytoskeleton of *Drosophila* oocytes to define the anteroposterior axis, eLife 4 (2015) e06088.

[7] T. J. Messitt, J. A. Gagnon, J. A. Kreiling, C. A. Pratt, Y. J. Yoon, K. L. Mowry, Multiple kinesin motors coordinate cytoplasmic RNA transport on a subpopulation of microtubules in *Xenopus* oocytes, Developmental Cell 15 (3) (2008) 426–436.

[8] S. Gilbert, Developmental Biology 6th edition, Sunderland (MA): Sinauer Associates, 2000.

[9] A. Ajduk, M. Zernicka-Goetz, Polarity and cell division orientation in the cleavage embryo: from worm to human, Molecular Human Reproduction 22 (10) (2016) 691–703.

[10] E. R. Gavis, Expeditions to the pole: RNA localization in *Xenopus* and *Drosophila*, Trends in Cell Biology 7 (12) (1997) 485–492.

[11] K. L. Mowry, C. A. Cote, RNA sorting in *Xenopus* oocytes and embryos, The FASEB Journal 13 (3) (1999) 435–445.

[12] M. L. King, T. J. Messitt, K. L. Mowry, Putting RNAs in the right place at the right time: RNA localization in the frog oocyte, Biology of the Cell 97 (1) (2005) 19–33.

[13] L. Dale, G. Matthews, A. Colman, Secretion and mesoderm-inducing activity of the TGF-beta-related domain of *Xenopus* Vg1., The EMBO Journal 12 (12) (1993) 4471.

[14] G. H. Thomsen, D. A. Melton, Processed Vg1 protein is an axial mesoderm inducer in *Xenopus*, Cell 74 (3) (1993) 433–441.

[15] B. Birsoy, M. Kofron, K. Schaible, C. Wylie, J. Heasman, Vg1 is an essential signaling molecule in *Xenopus* development, Development 133 (1) (2006) 15–20.

[16] C. Medioni, K. Mowry, F. Besse, Principles and roles of mRNA localization in animal development, Development 139 (18) (2012) 3263–3276.

[17] E. A. Powrie, M.-V. Ciocanel, J. A. Kreiling, J. A. Gagnon, B. Sandstede, K. L. Mowry, Using in vivo imaging to measure RNA mobility in *Xenopus* laevis oocytes, Methods 98 (1) (2015) 60–65.

[18] R. Milo, R. Phillips, Cell Biology by the Numbers, available at `http://book.bionumbers.org/`.

[19] M. K. Gardner, M. Zanic, C. Gell, V. Bormuth, J. Howard, Depolymerizing kinesins Kip3 and MCAK shape cellular microtubule architecture by differential control of catastrophe, Cell 147 (5) (2011) 1092–1103.

[20] M. Sirajuddin, L. M. Rice, R. D. Vale, Regulation of microtubule motors by tubulin isotypes and post-translational modifications, Nature Cell Biology 16 (4) (2014) 335–344.

[21] K. S. Thorn, J. A. Ubersax, R. D. Vale, Engineering the processive run length of the kinesin motor, The Journal of Cell Biology 151 (5) (2000) 1093–1100.

[22] N. Lorén, J. Hagman, J. K. Jonasson, H. Deschout, D. Bernin, F. Cella-Zanacchi, A. Diaspro, J. G. McNally, M. Ameloot, N. Smisdom, et al., Fluorescence recovery after photobleaching in material and life sciences: putting theory into practice, Quarterly Reviews of Biophysics 48 (03) (2015) 323–387.

[23] J. G. McNally, Quantitative FRAP in analysis of molecular binding dynamics in vivo, Methods in Cell Biology 85 (2008) 329–351.

[24] D. Axelrod, D. Koppel, J. Schlessinger, E. Elson, W. Webb, Mobility measurement by analysis of fluorescence photobleaching recovery kinetics., Biophysical Journal 16 (9) (1976) 1055.

[25] J. Blonk, A. Don, H. V. Aalst, J. Birmingham, Fluorescence photobleaching recovery in the confocal scanning light microscope, Journal of Microscopy 169 (3) (1993) 363–374.

[26] K. Braeckmans, L. Peeters, N. N. Sanders, S. C. De Smedt, J. Demeester, Three-dimensional fluorescence recovery after photobleaching with the confocal scanning laser microscope, Biophysical Journal 85 (4) (2003) 2240–2252.

[27] B. Sprague, R. Pego, D. Stavrera, J. McNally, Analysis of binding reactions by fluorescence recovery after photobleaching, Biophysical Journal 86 (2004) 3473–3495.

[28] J. Braga, J. M. Desterro, M. Carmo-Fonseca, Intracellular macromolecular mobility measured by fluorescence recovery after photobleaching with confocal laser scanning microscopes, Molecular Biology of the Cell 15 (10) (2004) 4749–4760.

[29] J. K. Jonasson, N. Lorén, P. Olofsson, M. Nydén, M. Rudemo, A pixel-based likelihood framework for analysis of fluorescence recovery after photobleaching data, Journal of Microscopy 232 (2) (2008) 260–269.

[30] F. Mueller, P. Wach, J. G. McNally, Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching, Biophysical Journal 94 (8) (2008) 3323–3339.

[31] M. Kang, C. A. Day, K. Drake, A. K. Kenworthy, E. DiBenedetto, A generalization of theory for two-dimensional fluorescence recovery after photobleaching applicable to confocal laser scanning microscopes, Biophysical Journal 97 (5) (2009) 1501–1511.

[32] M. Kang, C. A. Day, E. DiBenedetto, A. K. Kenworthy, A quantitative approach to analyze binding diffusion kinetics by confocal FRAP, Biophysical Journal 99 (9) (2010) 2737–2747.

[33] K. S. Zadeh, H. J. Montas, A class of exact solutions for biomacromolecule diffusion–reaction in live cells, Journal of Theoretical Biology 264 (3) (2010) 914–933.

[34] M. Weiss, Challenges and artifacts in quantitative photobleaching experiments, Traffic 5 (9) (2004) 662–671.

[35] J. Beaudouin, F. Mora-Bermúdez, T. Klee, N. Daigle, J. Ellenberg, Dissecting the contribution of diffusion and interactions to the mobility of nuclear proteins, Biophysical Journal 90 (6) (2006) 1878–1894.

[36] J. Braga, J. G. McNally, M. Carmo-Fonseca, A reaction-diffusion model to study RNA motion by quantitative fluorescence recovery after photobleaching, Biophysical Journal 92 (8) (2007) 2694–2703.

[37] J. Chen, J. Lippincott-Schwartz, J. Liu, Intracellular spatial localization regulated by the microtubule network, PLOS One 7 (4) (2012) e34919.

[38] P. Jung, A. Brown, Modeling the slowing of neurofilament transport along the mouse sciatic nerve, Physical Biology 6 (4) (2009) 046002.

[39] Y. Li, A. Brown, P. Jung, Deciphering the axonal transport kinetics of neurofilaments using the fluorescence photo-activation pulse-escape method, BMC Neuroscience 15 (Suppl 1) (2014) P132.

[40] J. M. Newby, P. C. Bressloff, Quasi-steady state reduction of molecular motor-based models of directed intermittent search, Bulletin of Mathematical Biology 72 (7) (2010) 1840–1866.

[41] J. Newby, P. C. Bressloff, Random intermittent search and the tug-of-war model of motor-driven transport, Journal of Statistical Mechanics: Theory and Experiment 2010 (04) (2010) P04014.

[42] M. C. Reed, S. Venakides, J. J. Blum, Approximate traveling waves in linear reaction-hyperbolic equations, SIAM Journal on Applied Mathematics 50 (1) (1990) 167–180.

[43] A. Friedman, G. Craciun, Approximate traveling waves in linear reaction-hyperbolic equations, SIAM Journal on Mathematical Analysis 38 (3) (2006) 741–758.

[44] A. Friedman, H. BEI, Uniform convergence for approximate traveling waves in linear reaction-hyperbolic systems, Indiana University Mathematics Journal 56 (5) (2007) 2133–2158.

[45] K. L. Mowry, D. A. Melton, Vegetal messenger RNA localization directed by a 340-nt RNA sequence element in *Xenopus* oocytes, Science 255 (5047) (1992) 991–994.

[46] W. O. Hancock, Bidirectional cargo transport: moving beyond tug of war, Nature Reviews Molecular Cell Biology 15 (9) (2014) 615–628.

[47] P. Chang, J. Torres, R. A. Lewis, K. L. Mowry, E. Houliston, M. L. King, Localization of RNAs to the mitochondrial cloud in *Xenopus* oocytes through entrapment and association with endoplasmic reticulum, Molecular Biology of the Cell 15 (10) (2004) 4669–4681.

[48] J. Rino, R. M. Martin, T. Carvalho, M. Carmo-Fonseca, Imaging dynamic interactions between spliceosomal proteins and pre-mRNA in living cells, Methods 65 (3) (2014) 359–366.

[49] J. A. Gagnon, K. L. Mowry, Molecular motors: directing traffic during RNA localization, Critical reviews in biochemistry and molecular biology 46 (3) (2011) 229–239.

[50] S. Cox, P. Matthews, Exponential time differencing for stiff systems, Journal of Computational Physics 176 (2) (2002) 430–455.

[51] A.-K. Kassam, L. Trefethen, Fourth-order time-stepping for stiff PDE's, SIAM Journal on Scientific Computing 26 (4) (2005) 1214–1233.

[52] M. Kang, M. Andreani, A. K. Kenworthy, Validation of normalizations, scaling, and photofading corrections for FRAP data analysis, PLOS ONE 10 (5) (2015) e0127966.

[53] S. Klumpp, R. Lipowsky, Cooperative cargo transport by several molecular motors, Proceedings of the National Academy of Sciences of the United States of America 102 (48) (2005) 17284–17289.

[54] M. J. Müller, S. Klumpp, R. Lipowsky, Tug-of-war as a cooperative mechanism for bidirectional cargo transport by molecular motors, Proceedings of the National Academy of Sciences 105 (12) (2008) 4609–4614.

[55] B. Iooss, P. Lemaître, A review on global sensitivity analysis methods, in: Uncertainty Management in Simulation-Optimization of Complex Systems, Springer, 2015, pp. 101–122.

[56] I. M. Sobol, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and Computers in Simulation 55 (1) (2001) 271–280.

[57] S. A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, in: Dokl. Akad. Nauk SSSR, Vol. 4, 1963, p. 123.

[58] F. Heiss, V. Winschel, Estimation with numerical integration on sparse grids, Department of Economics Discussion paper 2006-15, University of Munich, http://econpapers.repec.org/paper/lmumuenec/916.htm.

[59] P. C. Bressloff, J. M. Newby, Stochastic models of intracellular transport, Reviews of Modern Physics 85 (1) (2013) 135.

[60] E. A. Brooks, et al., Probabilistic methods for a linear reaction-hyperbolic system with constant coefficients, The Annals of Applied Probability 9 (3) (1999) 719–731.

[61] L. Popovic, S. A. McKinley, M. C. Reed, A stochastic compartmental model for fast axonal transport, SIAM Journal on Applied Mathematics 71 (4) (2011) 1531–1556.

[62] D. Smith, R. Simmons, Models of motor-assisted transport of intracellular particles, Biophysical Journal 80 (1) (2001) 45–68.

[63] G. Craciun, A. Brown, A. Friedman, A dynamical system model of neurofilament transport in axons, Journal of theoretical biology 237 (3) (2005) 316–322.

[64] F. Beichelt, Stochastic processes in Science, Engineering and Finance.

[65] A. Gut, Cumulative shock models, Advances in Applied Probability 22 (02) (1990) 504–507.

[66] R. Serfozo, Basics of applied stochastic processes, Springer Science & Business Media, 2009.

[67] H. C. Tijms, A first course in stochastic models, John Wiley and sons, 2003.

[68] J. Hughes, W. O. Hancock, J. Fricks, Kinesins with extended neck linkers: a chemomechanical model for variable-length stepping, Bulletin of Mathematical Biology 74 (5) (2012) 1066–1097.

[69] J. Hughes, W. O. Hancock, J. Fricks, A matrix computational approach to kinesin neck linker extension, Journal of theoretical biology 269 (1) (2011) 181–194.

[70] I. Wolfram Research, Mathematica, Champaign, Illinois, version 10.0 Edition (2014).

[71] L. Conway, J. L. Ross, Kinesin motor transport is altered by macromolecular crowding and transiently associated microtubule-associated proteins, arXiv preprint arXiv:1409.3455.

[72] P. C. Bressloff, B. Xu, Stochastic active-transport model of cell polarization, SIAM Journal on Applied Mathematics 75 (2) (2015) 652–678.

[73] R. J. Hawkins, O. Benichou, M. Piel, R. Voituriez, Rebuilding cytoskeleton roads: Active-transport-induced polarization of cells, Physical Review E 80 (4) (2009) 040903.

[74] P. C. Bressloff, J. M. Newby, Quasi-steady-state analysis of two-dimensional random intermittent search processes, Physical Review E 83 (6) (2011) 061139.

[75] E. Schulze, M. Kirschner, Microtubule dynamics in interphase cells., The Journal of Cell Biology 102 (3) (1986) 1020–1031.

[76] D. L. Gard, B. J. Cha, E. King, The organization and animal–vegetal asymmetry of cytokeratin filaments in stage VI *Xenopus* oocytes is dependent upon F-actin and microtubules, Developmental Biology 184 (1) (1997) 95–114.

[77] D. L. Gard, Organization, nucleation, and acetylation of microtubules in *Xenopus laevis* oocytes: a study by confocal immunofluorescence microscopy, Developmental Biology 143 (2) (1991) 346–362.

[78] W. Malalasekera, K. Versteeg, Computational fluid dynamics, the finite volume method (2007).

[79] R. Delanoue, I. Davis, Dynein anchors its mRNA cargo after apical transport in the *Drosophila* blastoderm embryo, Cell 122 (1) (2005) 97–106.

[80] R. Delanoue, B. Herpers, J. Soetaert, I. Davis, C. Rabouille, *Drosophila* squid/hnRNP helps dynein switch from a *gurken* mRNA transport motor to an ultrastructural static anchor in sponge bodies, Developmental Cell 13 (4) (2007) 523–538.

[81] V. L. Zimyanin, K. Belaya, J. Pecreaux, M. J. Gilchrist, A. Clark, I. Davis, D. St Johnston, In vivo imaging of oskar mRNA transport reveals the mechanism of posterior localization, Cell 134 (5) (2008) 843–853.