INFERENCE AND ANALYSIS FOR STOCHASTIC DENSITY-DEPENDENT POPULATION DYNAMICS, WITH APPLICATION TO DRUG RESISTANCE

by

LINH HUYNH

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Dissertation Advisor: Prof. Peter J. Thomas

Department of Mathematics, Applied Mathematics and Statistics

CASE WESTERN RESERVE UNIVERSITY

2022

CASE WESTERN RESERVE UNIVERSITY

SCHOOL OF GRADUATE STUDIES

We hereby approve the thesis of

Linh Huynh

Candidate for the Doctor of Philosophy degree*.

Prof. Peter J. Thomas

Committee Chair, Professor
Department of Mathematics, Applied Mathematics, and Statistics

Prof. Jenny Brynjarsdottir

Associate Professor
Department of Mathematics, Applied Mathematics, and Statistics

Prof. David Gurarie

Professor
Department of Mathematics, Applied Mathematics, and Statistics

Prof. Dr. Jacob G. Scott

External Faculty, Associate Professor and and Staff Physician-Scientist
Department of Translational Hematology and Oncology Research

Date: May 18, 2022

*We also certify that written approval has been obtained for any proprietary material contained therein.

Table of Contents

List of Tables

List of Figures

Acknowledgement

The most important acknowledgement goes to Prof. Peter Thomas for his incredible guidance and support throughout my PhD journey. Scientifically, he has taught me stochastic processes and pushed me through sophisticated, long mathematical computations. My rebellious self always tried to find a shorter way to solve the problem, but I have learned so much from doing those rigorous computations. Professionally, Prof. Thomas has been a strong advocate for my academic success in many ways, which include, but are not limited to, providing networking opportunities, endorsing advanced course teaching assignment (which was unsuccessful but beyond his control), and giving me feedback on my practice job talk. Personally, he provided research support so that I could live with my husband and finish my degree remotely in my final year, which really improved my productivity in different ways. It is hard to imagine a better mentor and teacher, and I feel honored and grateful to call him my PhD advisor.

I am also very thankful for the mentorship of Prof. Dr. Jacob Scott. In addition to evolutionary dynamics and mathematical oncology, I have learned so much about grant writing, scientific communication, and lab management from him. Moreover, it has been a joy to be part of his resourceful, fun, inclusive lab, and I am thankful for friends such as Nara, Raoul, Kyle, Steph, Vish, etc.

I would like to thank Prof. David Gurarie, Prof. Jenny Brynjarsdottir, and Prof. Alethea

Barbaro for their time being on my area exam and/or dissertation committee, and for evaluating my annual academic progress. I greatly appreciate their mathematical and statistical feedback on my work.

It is a privilege to be part of the Department of Mathematics, Applied Mathematics, and Statistics. I would like to thank Prof. Nick Gurski for giving me an opportunity to lecture advanced mathematical topics and to serve on the Search Committee. Moreover, I greatly appreciate that he advocated for a stipend raise for all graduate students. I would like to thank the following professors for helping me build a strong foundation in numerical analysis: Prof. Daniela Calvetti, Prof. Julia Dobrosotskaya, Prof. Weihong Guo, Prof. Wanda Strychalski, and Prof. Longhua Zhao. I feel truly inspired to be a successful female mathematician like them. I also would like to thank Prof. Erkki Somersalo for his helpful inverse problem course, and Prof. Wojbor Woyczynski for his informative statistics courses. I will always remember Prof. Woyczynski passing by my office and always asking "What are you doing here on Friday evening?" The hallway is not the same any more without him. Many thanks to Sakeenah Bari-Harold, Donyear Thomas, Eric Jackson, and Joe Peterson for helping me with administrative paperwork. I thank my academic sibling, Shusen Pu, for the template of this thesis. I had a wonderful, balanced life during my time at Case Western Reserve University thanks to all the amazing on-campus activities such as Zumbathon, Farm Harvest Festival, Mid-Autumn Festival, Lunar New Year firework, free massage, free yoga, rock painting, bubble tea giveaway, and many more. Among all the universities I have studied at, the student-life organization at CWRU is probably the best–including the most effective COVID mask design! Finally, I thank my calculus students for their curiosity and good questions. Being the instructor of record for their classes was an absolute joy.

My journey since college has been as adventurous as it can be, which really tested the patience of many people in addition to my own. I am indebted to the following formal or informal, long-time or short-time mentors: Prof. Turner Hogan, Prof. Lybeth Hodges, Prof. Lindsay Renfro, Prof. Alexey Kuznetsov, Prof. Yaroslav Molkov, Prof. Bünyamin Sarı, Prof. Caner Kazanci, Prof. Beatrice Rivière, Prof. Richard Tapia, and Prof. Janet Best. They were very generous with their time and advice, and have shaped my academic growth in many impactful ways.

My family and friends thought I was crazy for turning down good PhD offers to be in two master programs before doing a PhD, but I have learned so much from being at different institutions. At Rice, I learned people skills and scientific communication from Arturo, Emily, and Prof. Paul Hand, in addition to industry collaboration skills from the department. Coming from the small department at Rice, I truly enjoyed having many graduate friends at OSU and being in a large TA office, discussing and doing math until late night. I gave credit to Roman Nitze for my decision to join OSU, as he gave an impression of a nurturing department during recruitment. I thank Dan Boros, Deborah Stout, John Lewis, Daniel Glasscock, and Amit Vutha for organizing probably the most remarkable teaching training program in the country, which really shaped my teaching philosophy and style. I think Matt Thompson at the MBI did the best organizational work among all the workshops and conferences I have been to. While at Rice, I gained phenomenal scientific computing knowledge; at OSU, I gained knowledge in dynamical systems; at CWRU, I gained knowledge in stochastic processes. Finally, my life in Ohio couldn't be more fun. Both Columbus and Cleveland (and places in between) are so culturally rich with arts, music, theater, trails, lakes, flowers, cuisines, festivals, movies (locations for Shawshank

Redemption and Christmas Story), etc. I treasure my time at Steiner House, where I learned democratic group decision making, and enjoyed daily authentic international foods and conversations with friends from various departments such as law, social work, accounting.

Throughout my journey, I was so lucky to be surrounded by great friends, and I am extremely grateful for their support. Lan Thanh, chi Giao, anh Dat, anh Tri, co Huyen, Likeleli, Abdul, Sebastian, Richard, Jizhou, Martha, chi Thu, Evelyn, I have missed seeing you around! I also thank Duc Huy, Hoang Duy, Viet Duc for offering a helping hand at some critical moments. I thank my high school math teacher, thay Cuong, for being very strict on when to write $\Rightarrow$ vs. $\Leftrightarrow$, and the like, which really helped foster my mathematical thinking.

Last but not least, I dedicate this thesis to my family. Words cannot describe how thankful I am to have them in my life. Not only have my parents and my brother taught me scientific reasoning and creativity in daily life activities, but they have also set a model of kindness, thoughtfulness, sincerity, and generosity. They have instilled in me a love for learning without putting any pressure. Most importantly, I thank my parents, brother, grandmothers, and aunts for their unwavering support. I thank my baby niece and nephew for their cute imagination and creativity, which really helped me get through some dark moments of graduate school. Finally, a very special thank to my husband for taking care of me, especially when I had some health issues, and for having been deterministic to all of my stochastic fluctuations!

Inference and Analysis for Stochastic Density-Dependent Population Dynamics, with Application to Drug Resistance

Abstract

by

LINH HUYNH

This thesis addresses two aspects of the stochastic dynamics of pathogenic populations: time scale and density dependence. Ecological and evolutionary processes are fundamentally distinguished by their time scales. One of the mathematical tools to quantify time scales of stochastic processes is first-passage time analysis. However, we show that cell populations with the same *net* growth rates, but different birth and death rates, exhibit different mean extinction times. This finding calls for the need to parse out birth and death rates from *net* growth rates, which is also important in the second aspect: density-dependent dynamics. Observations of only the *net* growth rates cannot inform whether density dependence manifest in the birth process, death process, or both. Therefore, this thesis aims to separately infer birth and death rates from cell number time series that follow logistic birth-death processes. We develop two methods, both of which harness stochasticity in cell population dynamics to make inferences. The first method utilizes the mean and variance of stochastic fluctuations in cell number increments, which provides a novel perspective on the problem of stochastic parameter estimation. We validate this method by deriving analytical expressions for the theoretical means and variances of estimation

errors as functions of the discretization bin size, and discover that intermediate bin sizes are optimal and that the estimation is less effective for fast-producing cell types. This approach requires sufficiently large sample sizes. To overcome this requirement, we consider the maximum likelihood approach, in which we solve a one-dimensional constrained nonlinear optimization problem to identify the most likely density dependence parameter, assuming the other parameters are known. The computational complexity of adding more unknown parameters leads us to answer the question of "more likely" instead. Towards this goal, we derive the theoretical time-dependent mean and variance of the state in our birth-death process, and the theoretical mean of the log-likelihoood. Together, all the results in this thesis contribute to disambiguating mechanisms (especially drug resistance) underlying the same *net* growth rate for biological systems at different scales.

Chapter 1

Introduction

1.1 Background

Drug resistance is a concerning global crisis that hinders the efficacy of treatments of pathogens [86, 95]. In this thesis, we mathematically study drug resistance from the population-level dynamics perspective (as opposed to other scales such as the cellular level). In doing so, we need to determine the growth law of a target population, for instance exponential, logistic, or other. However, the choice of growth law is not too straightforward, and may depend on the problems we want to analyze [51]. This thesis focuses on logistic growth, one of the simplest (but still realistic [33]) models for density dependence. Density dependence has been shown to have effects on the ecology and evolution of microbial and cancer cells [20, 46, 64, 24, 27]. We analyze the ecological and evolutionary density-dependent dynamics of pathogenic populations (e.g. harmful bacteria or cancer cells), especially under drug treatments, and infer mechanisms underlying the dynamics. This biological application motivates analysis and development of inference

methods for stochastic processes. Two important components of this thesis are (1) mathematical models and analysis, and (2) inference methods.

Mathematical models and analysis. As surveyed in the book by Bacaer [3], mathematical models of population dynamics were developed at least as early as 1202 with the Fibonacci sequence. In 1838, Verhulst published a logistic differential equation model for single-species/homogeneous populations [92], which has been one of the classical mathematical models for density-dependent population dynamics. Between 1920-1926, Lotka and Volterra developed a two-species/heterogeneous logistic competitive model [60, 93]. In these models, density dependence is assumed to have a negative effect on the population–in particular, reducing its net growth rate and subsequently population size. Density dependence, however, can also have a positive effect on population growth. The positive density dependence, commonly referred to as the Allee effect [1] and typically modeled with cubic growth [44], can affect population evolution [75].

Despite the tractability of deterministic models, ecological and evolutionary dynamics are fundamentally stochastic as discussed in [50, 13, 69, 56, 91]. Steinar Engen and Bernt-Erik Sæther (1998) [26] discusses stochastic population dynamics models in terms of demographic and environmental stochasticity, diffusion approximation, density dependence, and age-structured models with no density dependence, which are some factors one may take into consideration when modeling stochastic population dynamics [10]. When building stochastic models for population dynamics, one must pay attention to the types of noise incorporated into the system, because different noise classes can result in different dynamics–for example, in extinction [83] or invasion probabilities [11]. Figure 1.1 illustrates that there are three main kinds of noise: demographic, environmental, and measurement in ecological

models [79]. Roughly speaking, demographic stochasticity is noise from demographic processes intrinsic to the cell populations such as birth and death, while environmental stochasticity is noise from the environment [79]. For example, in the context of the problem we are interested in, drug resistance, environmental noise can be the stochasticity in drug absorption. Engen et al. (1998) [25] provides mathematical definitions of demographic and environmental stochasticity in terms of the variance of a general stochastic analog of a family of density-dependent deterministic models. In this thesis, we utilize binomially distributed demographic noise (i.e. from the birth and death events of individuals) to derive the parameter identification method in Section 3.3.3, and Gaussian noise for the simulation in Section 3.3.2, the statistical inference in Section 4.2, and the derivation of theoretical mean and variance of the state $X(t)$ in our one-species logistic birth-death model in Section 4.3.

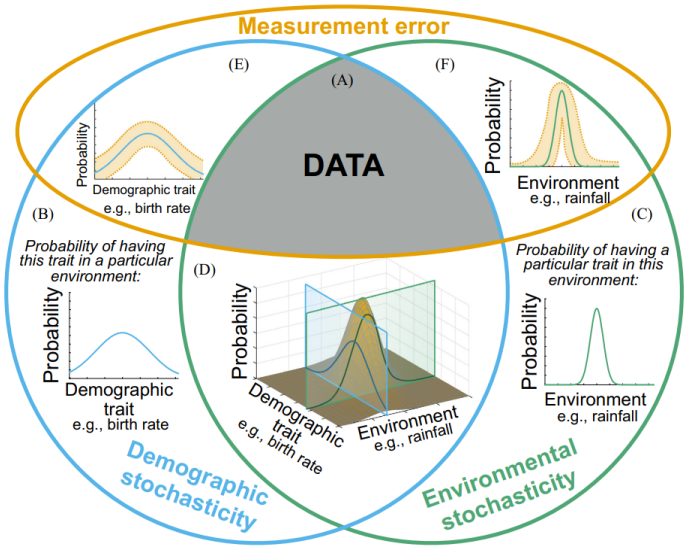

Figure 1.1: **There are three main kinds of noise in ecological models: demographic, environmental, and measurement.** This figure is taken from [79] as an illustration. The corresponding author, Prof. Lauren Sullivan (sullivanll@missouri.edu), has granted permission to include this figure in this thesis.

Inference methods. Stochasticity allows us to make inference of the population dynamics. Existing statistical inference methods [90] are mainly classified into Bayesian inference and maximum likelihood. Parameter estimation under these two approaches commonly require a numerical solution to a high-dimensional optimization problem [14]. While there have been many advances in developing efficient optimization algorithms [72], in this thesis, we want to focuses on fundamental principles of probability theory and stochastic processes.

Broad research questions. We are broadly interested in the role of stochastic fluctuations in the answers to the following biological/medical questions:

Q1. What are the mechanisms of drug resistance?

Q2. How do ecological and evolutionary processes in population dynamics play a role in development of drug resistance?

Q3. Can we develop optimal treatments to overcome drug resistance?

Baguley et al. (2010) [4] gives a good review of different drug resistance mechanisms for cancer. To uncover drug resistance mechanisms from data, this thesis takes the first step of disambiguating whether drug resistance happens through the birth process, death process, or some combination of both, which motivates the work of Chapter 3 and Chapter 4. In these chapters, we also answer the ecological part of Question Q2. We would like to use the tools built from these chapters to answer Questions Q2 and Q3 with a focus on evolutionary process. Chapter 2 provides some preliminary results towards this goal.

We provide additional detailed literature review and description of mathematical models and methods in each of Chapter 2, Chapter 3, and Chapter 4.

1.2 Thesis Contributions

The main goal of this thesis is to separately infer birth and death rates from cell number time series using two methods that both harness stochasticity in cell population dynamics. We have submitted the results from Chapter 3 and part of Chapter 4 for publication [41]. Key results of this thesis are:

C1. First-passage time analysis of extinction in a birth-death-mutation process model and demonstration of the need for separate birth and death rates; see Section 2.4.

C2. Development of a direct estimation method to separately identify birth and death rates from cell number time series data, giving a novel perspective on stochastic parameter identifiability; see Section 3.3.3.

C3. Derivation of analytical expressions for the theoretical means and variances of errors in estimating birth and death rates as functions of the discretization bin size; see Section 3.3.4 and Appendix A.2.

C4. Application of the direct estimation method to disambiguate underlying autoregulation, drug efficacy, and drug resistance mechanisms, and to infer whether these processes happen through birth, death, or some combination of the two. The direct estimation method can be applied to other biological systems at different scales; see Section 3.4.

C5. Development of a likelihood-based inference method to infer the most likely density dependence parameter corresponding to a given single cell number time series; see Section 4.2.

C6. Derivation of a theoretical probability distribution for the system state in our logistic

birth-death process model; see Section 4.3.

C7. Derivation of the theoretical mean of the log-likelihood function for our logistic birth-death process model; see Section 4.4.

The rest of the thesis is organized as follows. In Chapter 2, we discuss first-passage time analysis of pathogen eradication and demonstrate the significance of separate birth and death rates. In Chapter 3, we discuss our novel stochastic parameter identification method, error analysis of the method (details in Appendix A.2), and application of the method to describe drug resistance mechanisms. In Chapter 4, we discuss the likelihood-based approach, in which we answer the questions of "most likely" and "more likely" density-dependent dynamics, and analytically derive the probability distribution of the state in our logistic birth-death process, and the theoretical mean of the log-likelihood. Finally, in Chapter 5, we provide an overview of future research directions.

Chapter 2

First-Passage Time: Significance of Separate Birth and Death Rates

2.1 Motivation

Ecology and evolution are two major aspects of population dynamics. Ecological and evolutionary processes are fundamentally distinguished by their time scales, with ecological processes generally occurring on shorter time scales, and evolutionary processes unfolding on longer time scales [22]. However, "shorter" time scales and "longer" time scales may be difficult to quantify. Moreover, sometimes rapid evolution can happen on ecological time scales [15, 89]. Recently, Gerlee [34] has introduced a method to separate ecology and evolution using perturbation analysis. Since ecology and evolution are fundamentally stochastic as discussed in [62, 66, 65], another mathematical tool that one can use to study the time scale of ecology and evolution is first-passage time analysis [5, 74, 38]. Under birth-death process models, the computation of first-passage times requires knowing the

birth and death rates of the processes. Towards the broad goal of understanding ecological and evolutionary time scales in biological systems, in this chapter, we analyze how cell populations with the same *net* growth rate, but different birth and death rates, can exhibit different mean extinction times. Then, in Chapters 3, we show how to separately identify birth and death rates from data. Here, we consider the expected values of extinction times, motivated by the problem of pathogen eradication. For future work, we can consider other statistical measures of first-passage times such as extreme first-passage times (i.e. the time it takes for the fastest searcher to reach a target) [18, 52, 53]–especially in the context of optimal treatments.

2.2 Mathematical Model

We consider a cell population consisting of two cell types: drug-sensitive cells, which we refer to as S-cells, and drug-resistant cells, which we refer to as R-cells. One of the classical models describing density-dependent dynamics of two-species populations is the competitive Lotka–Volterra model [60, 93]. In our work, since we are interested in the evolution of cell populations under drug treatments, we add a mutation term to the model as follows:

$$\frac{d\phi_S}{dt} = b_{0S}\phi_S - \underbrace{b_{0S}\mu\phi_S}_{\text{mutation}} - d_{0S}\phi_s - \frac{r_S}{K_S}\phi_S^2 - \alpha_{SR}\frac{r_S}{K_S}\phi_S\phi_R, \tag{2.1}$$

$$\frac{d\phi_R}{dt} = b_{0R}\phi_R + \underbrace{b_{0S}\mu\phi_S}_{\text{mutation}} - d_{0R}\phi_R - \frac{r_R}{K_R}\phi_R^2 - \alpha_{RS}\frac{r_R}{K_R}\phi_R\phi_S, \tag{2.2}$$

where ϕ_S and ϕ_R are numbers of S-cells and R-cells respectively; b_{0S} and b_{0R} are the intrinsic *per capita* birth rates of S-cells and R-cells respectively; d_{0S} and d_{0R} are the intrinsic *per capita* death rates of S-cells and R-cells respectively; $r_S = b_{0S} - d_{0S}$ and

$r_R = b_{0R} - d_{0R}$ are intrinsic *per capita* net growth rates of S-cells and R-cells respectively; K_S and K_R are carrying capacities of S-cells and R-cells respectively, in the absence of the competing species. The presence of species R decreases the effective carrying capacity for species S at a rate α_{SR}. That is, the effective carrying capacity for species S is $K_S - \alpha_{SR}\phi_R$. Similarly, the carrying capacity of species R decreases at a rate α_{RS} per unit increase in the species S. We consider a mutation scenario where mutation happens through the birth (or division) of S-cells, and interpret the mutation rate μ as the probability that an S-cell divides into one S-cell and one R-cell. We can combine the two terms $b_{0S}\phi_S$ and $(-b_{0S}\mu\phi_S)$ in Equation (2.1) into $b_{0S}(1-\mu)\phi_S$ and interpret $(1-\mu)$ as the probability that mutation does *not* happen when an S-cell divides. In Equations (2.1) and (2.2), we parameterize the intra-species density-dependent terms with the density dependence parameters $0 \leq \gamma_S \leq 1$ and $0 \leq \gamma_R \leq 1$ as follows:

$$\frac{d\phi_S}{dt} = \underbrace{b_{0S}(1-\mu)\phi_S - \gamma_S \frac{r_S}{K_S}\phi_S^2}_{\text{birth rate}} - \underbrace{\left(d_{0,S}\phi_s + (1-\gamma_S)\frac{r_S}{K_S}\phi_S^2 + \alpha_{SR}\frac{r_S}{K_S}\phi_S\phi_R\right)}_{\text{death rate}}, \quad (2.3)$$

$$\frac{d\phi_R}{dt} = \underbrace{b_{0,R}\phi_R + b_{0S}\mu\phi_s - \gamma_R \frac{r_R}{K_R}\phi_R^2}_{\text{birth rate}} - \underbrace{\left(d_{0R}\phi_R + (1-\gamma_R)\frac{r_R}{K_R}\phi_R^2 + \alpha_{RS}\frac{r_R}{K_R}\phi_R\phi_S\right)}_{\text{death rate}}. \quad (2.4)$$

Here, we assume that the inter-species density dependence is only in the death rate for each cell type. Note that we only consider nonnegative birth and death rates; in case an expression such as $b_{0S}(1-\mu)\phi_s - \gamma_S \frac{r_S}{K_S}\phi_S^2$ should become negative, we treat it as zero. For notational convenience we omit writing $\max\left\{b_{0S}(1-\mu)\phi_s - \gamma_S \frac{r_S}{K_S}\phi_S^2, 0\right\}$ for such expressions in this section.

Table 2.1 provides a stochastic analog of the deterministic model (2.1) and (2.2): a two-

species birth-death process with quadratic competition. Let N_S and N_R be integer-valued random variables representing numbers of S-cells and R-cells respectively.

r	Reaction/Transition	Rate $\lambda_r([i,j]^T)$	Stoich. ζ_r	Interpretation
1	$N_S \to 2N_S$	$b_{0_S}(1-\mu)N_S - \gamma_S \frac{r_S}{K_S} N_S^2$	$[1,0]^T$	Birth (S)
2	$N_S \to \emptyset$	$d_{0_S} N_S$	$[-1,0]^T$	Death (S)
3	$N_S + N_S \to N_S$	$(1-\gamma_S)(r_S/K_S)N_S^2$	$[-1,0]^T$	Death (S)
4	$N_S \to N_S + N_R$	$b_{0_S}\mu N_S$	$[0,1]^T$	Birth with Mutation
5	$N_R \to 2N_R$	$b_{0_R} N_R - \gamma_R \frac{r_R}{K_R} N_R^2$	$[0,1]^T$	Birth (R)
6	$N_R \to \emptyset$	$d_{0_R} N_R$	$[0,-1]^T$	Death (R)
7	$N_R + N_R \to N_R$	$(1-\gamma_R)(r_R/K_R)N_R^2$	$[0,-1]^T$	Death (R)
8	$N_S + N_R \to N_R$	$\alpha_{SR} \frac{r_S}{K_S} N_S N_R$	$[-1,0]^T$	Competition
9	$N_R + N_S \to N_S$	$\alpha_{RS} \frac{r_R}{K_R} N_R N_S$	$[0,-1]^T$	Competition

Table 2.1: Two-Species Birth-Death-Mutation Process Model. The state in this model is the pair (N_S, N_R), where N_S represents the number of S-cells and N_R represents the number of R-cells. There are nine reactions/transitions: reaction $r = 1$ represents natural birth of S-cells; reaction $r = 2$ represents natural death of S-cells; reaction $r = 3$ represents death of S-cells due to interactions among the S-cells; reaction $r = 4$ represents birth of R-cells due to mutation; reaction $r = 5$ represents natural birth of R-cells; reaction $r = 6$ represents natural death of R-cells; reaction $r = 7$ represents natural death of R-cells due to interactions among the R-cells; reaction $r = 8$ represents natural death of S-cells due to interactions with the R-cells; reaction $r = 9$ represents natural death of R-cells due to interactions with the S-cells. The stoichiometry vector ζ_r shows the changes in S-cells and R-cells following reaction/transition r. The notation T represents the transpose of the vector.

Note that the terms that are quadratic in N_S or N_R in Table 2.1 can be interpreted as either direct or indirect competition. In the context of ecology, direct competitive interactions are referred to as interference competition and indirect competitive interactions are referred to as exploitation competition [80]. All reactions except reaction 1 and reaction 5 may be interpreted using standard mass-action kinetics. For example, the quadratic term in the death rate (reactions 3 and 7) may be interpreted as homogeneous bimolecular collision reactions; reactions 8 and 9 may be interpreted as heterogeneous bimolecular collision

reactions. Reactions 2 and 6 may be interpreted as unimolecular decay reactions; reaction 4 may be interpreted as a unimolecular source reaction. The rates for reactions 1 and 5 should be understood as nonnegative. The third column of Table 2.1 gives the stoichiometry vector for each reaction.

The *per capita* birth rates and death rates for S-cells are thus given by

$$b_{N_S} = b_{0S}(1-\mu) - \gamma_S \frac{r_S}{K_S} N_S \tag{2.5}$$

$$d_{N_S} = d_{0S} + (1-\gamma_S)\frac{r_S}{K_S} N_S + \alpha_{SR}\frac{r_S}{K_R} N_R, \tag{2.6}$$

which means

$$b_{N_S} + d_{N_S} = b_{0S}(1-\mu) + d_{0S} + (1-2\gamma_S)\frac{r_S}{K_S} N_S + \alpha_{SR}\frac{r_S}{K_R} N_R. \tag{2.7}$$

When $\gamma_S = 0.5$, $b_{N_S} + d_{N_S}$ does not depend on N_S (but still depend on N_R). For fixed values of N_R, N_S, and the other parameters in Equation (2.7), decreasing γ_S from the value 0.5 makes $b_{N_S} + d_{N_S}$ bigger than the case $\gamma_S = 0.5$, and $\gamma_S > 0.5$ makes $b_{N_S} + d_{N_S}$ smaller. Similarly, we have the same properties for R-cells.

2.3 Mean Extinction Time Computation Method

The extinction time random variable T for this system is the first-passage time to the absorbing state $[N_S, N_R]^T = [0, 0]^T$ starting from some initial state $[N_S, N_R]^T = [n_S, n_R]^T$.

Mathematically, T is defined as follows:

$$T([n_S, n_R]^T) = \inf_{t>0} \left\{ t \,\middle|\, [N_S(t), N_R(t)]^T = [0,0]^T, \text{ given } [N_S(0), N_R(0)]^T = [n_S, n_R]^T \right\}. \tag{2.8}$$

To compute the mean of the extinction time, we use one-step analysis. Intuitively, the mean extinction time from the current state is weighted by the mean extinction times from all the next states that the current state can transition into after one transition, plus the expected time to stay in the current state. Assuming that the mean of $T([n_S, n_R]^T)$ is finite, let $\tau\left([n_S, n_R]^T\right) := \mathbb{E}\Big[T([n_S, n_R]^T)\Big]$ denote the mean first-passage time from state $[n_S, n_R]^T$ to the absorbing state $[0,0]^T$. Then, following e.g. [32, 68]

$$\tau\left([n_S, n_R]^T\right) = \frac{1}{\Lambda}\left[1 + \sum_{r=1}^{n_r} \tau\left([n_S, n_S]^T + \boldsymbol{\zeta_r}\right) \lambda_r\left([n_S, n_R]^T\right)\right] \tag{2.9}$$

$$\Rightarrow \left(\sum_{r=1}^{n_r} \tau\left([n_S, n_S]^T + \boldsymbol{\zeta_r}\right) \lambda_r\left([n_S, n_R]^T\right) \right) - (\Lambda) \cdot \tau\left([n_S, n_R]^T\right) = -1 \tag{2.10}$$

where n_r denotes the total number of outgoing reactions from the current state; $\Lambda = \sum_{r=1}^{n_r} \lambda_r\left([n_S, n_R]^T\right)$ denotes the total rate of all reaction rates $\lambda_r\left([n_S, n_R]^T\right)$; and $\boldsymbol{\zeta_r}$ denotes the stoichiometry vector. In Equation (2.10), $\frac{1}{\Lambda}$ is the mean dwell time in state $[n_S, n_R]^T$, and $\frac{\lambda_r\left([n_S, n_R]^T\right)}{\Lambda}$ is the probability of going from state $[n_S, n_R]^T$ to state $\left([n_S, n_R]^T + \boldsymbol{\zeta_r}\right)$. Denote $\boldsymbol{\tau}$ as a vector of all mean extinction times. Combining Equation (2.10) for all the states leads to a nonhomogeneous linear system of equations

$$\boldsymbol{Q\tau} = -\mathbf{1} \quad \text{with boundary condition} \quad \tau([0,0]^T) = 0. \tag{2.11}$$

If $\hat{N}_S$ and $\hat{N}_R$, respectively, represent the upper limits for N_S and N_R imposed to define the computational boundary, then there are $\mathcal{M} = [(\hat{N}_S + 1)(\hat{N}_R + 1) - 1]$ total states (excluding the state $(0, 0)$). Then $\boldsymbol{Q}$ is a sparse $\mathcal{M} \times \mathcal{M}$ matrix. Solving the System of Equations (2.11), we obtain the vector of mean times to extinction starting from the all states in the system.

2.4 Effect of Density Dependence

In this section, we apply the method described in Section 2.3 to analyze the density dependence effect and demonstrate the significance of parsing out birth and death rates from *net* growth rate. We consider pairs of birth and death rates that have the same net growth rate, but different density dependence distribution (i.e. different values of γ_S and γ_R). The main purpose of this section is to show that mean extinction times are different for different values of density dependence parameter, even though these parameters result in the same net growth rate. When γ_X $(X = S, R)$ is equal to 0, density dependence is fully in the death rate. When γ_X $(X = S, R)$ is equal to 1, density dependence is fully in the birth rate. When γ_X $(X = S, R)$ is equal to 0.5, density dependence is split equally between the two rates. Figures 2.1, 2.2, 2.4, and 2.3 all show that the mean extinction times are different for different combinations of γ_S and γ_R, even though the *net* growth rates for S-cells and R-cells are the same for all the scenarios. In these figures, for purposes of illustration, we artificially cap the population size at a maximum of $\hat{N}_{\mathrm{R}} = \hat{N}_{\mathrm{S}} = 20$, thereby limiting the size of our computational domain. Truncating the domain in this way is equivalent to setting the per capita birth rates $b_{N,X} = 0$ for $N \geq 20$. In Figures 2.1 and 2.2, we use the parameter values: $\alpha_{SR} = \alpha_{RS} = 0.75$, $b_{0S} = b_{0R} = 2/10$, $d_{0S} = d_{0R} = 1/10$, $K_S = K_R = 10^2$. We set the mutation probability per cell division to $\mu = 0$ for Figure 2.1

and $\mu = 25 \times 10^{-4}$ for Figure 2.2.

Under parameters given above, the intrinsic growth rates are initially positive: $r_S > 0$ and $r_R > 0$. Thus the unconstrained S-cell and R-cell populations would grow to their carrying capacities, as illustrated in Figure 2.5 **(A, B)** for the case $\mu = 0$. In contrast to deterministic models where the populations will never go extinct for positive net growth rates, the stochastic fluctuations in cell number increments eventually drive the populations to extinction [17, 88]. When the variance of cell number increments is larger, the quasistationary distribution will be broader and extinction will happen sooner. Quasistationary distribution is stationary distribution reached after a long period time. (Note, however, that in Figures 2.1–2.4, by capping the total population at a value much lower than the unconstrained carrying capacity, we greatly reduce the expected extinction time.) In a later chapter of this thesis (Section 3.3.3) we will explain in detail that the variance of cell number increments (i.e. change in cell number after a sufficiently small time period) depends on the sum of birth and death rates. We formulate our mathematical model and parameterization with the assumption that the density dependence (i.e. quadratic) term either reduces the birth rate or increases the death rate, or both. Hence, among the three scenarios, the sum of birth and death rates, is the largest for γ_X $(X = S, R)$ = 0 and smallest for γ_X $(X = S, R)$ = 1, as explained above at the end of Section 2.2. That is why for the scenario $\gamma_S = \gamma_R = 0$, as in plot **(A)** in all Figures 2.1, 2.2, 2.4, and 2.3, the mean extinction time is the smallest compared to the other cases. Similarly, for the scenario $\gamma_S = \gamma_R = 1$, as in plot **(I)** in all Figures 2.1, 2.2, 2.4, and 2.3, the mean extinction time is the largest compared to the other cases. Figure 2.2 shows that even with mutation, the pattern is similar to the case without mutation, cf. Figure 2.1. All the nine cases in Figure 2.1 and Figure 2.2 have the same net growth rates for the S-cell population and

R-cell population, but we can see that the mean extinction times are different for any initial state (n_S, n_R). Even within one row or one column, where the parameters for one species (e.g. S-cells or R-cells) are fixed, we can see that the mean extinction times are different for different birth and death rates of the other species.

In Figure 2.3, we use the same parameters as those in Figure 2.1, except $b_S = 1/10$ and $d_S = 2/10$, and hence the intrinsic net growth rate $r_S < 0$. That is why we see the decay in the S-cell population in Figure 2.5 **(C, D)**, while the R-cell population is still growing. We can interpret this scenario as when the whole two-species population is treated with a drug that kills the drug-sensitive cells fast but does not have any effect on the drug-resistant cells. In Figure 2.1 and Figure 2.2, we have to wait for both S-cell population and R-cell population to die out, while for this case (Figure 2.3), we only need to wait for the R-cell population to die out. That is why the values of mean extinction time in Figure 2.3 are smaller than those in Figure 2.1 and Figure 2.2. Moreover, in Figure 2.3, the mean extinction time only changes when γ_R changes. The same idea applies to Figure 2.3, where it is the R-cell population that has negative growth rate r_R, not S-cells. Figure 2.4 shows that when we switch the parameters (i.e. $r_R < 0$, instead of r_S), the plots of the mean extinction times flip.

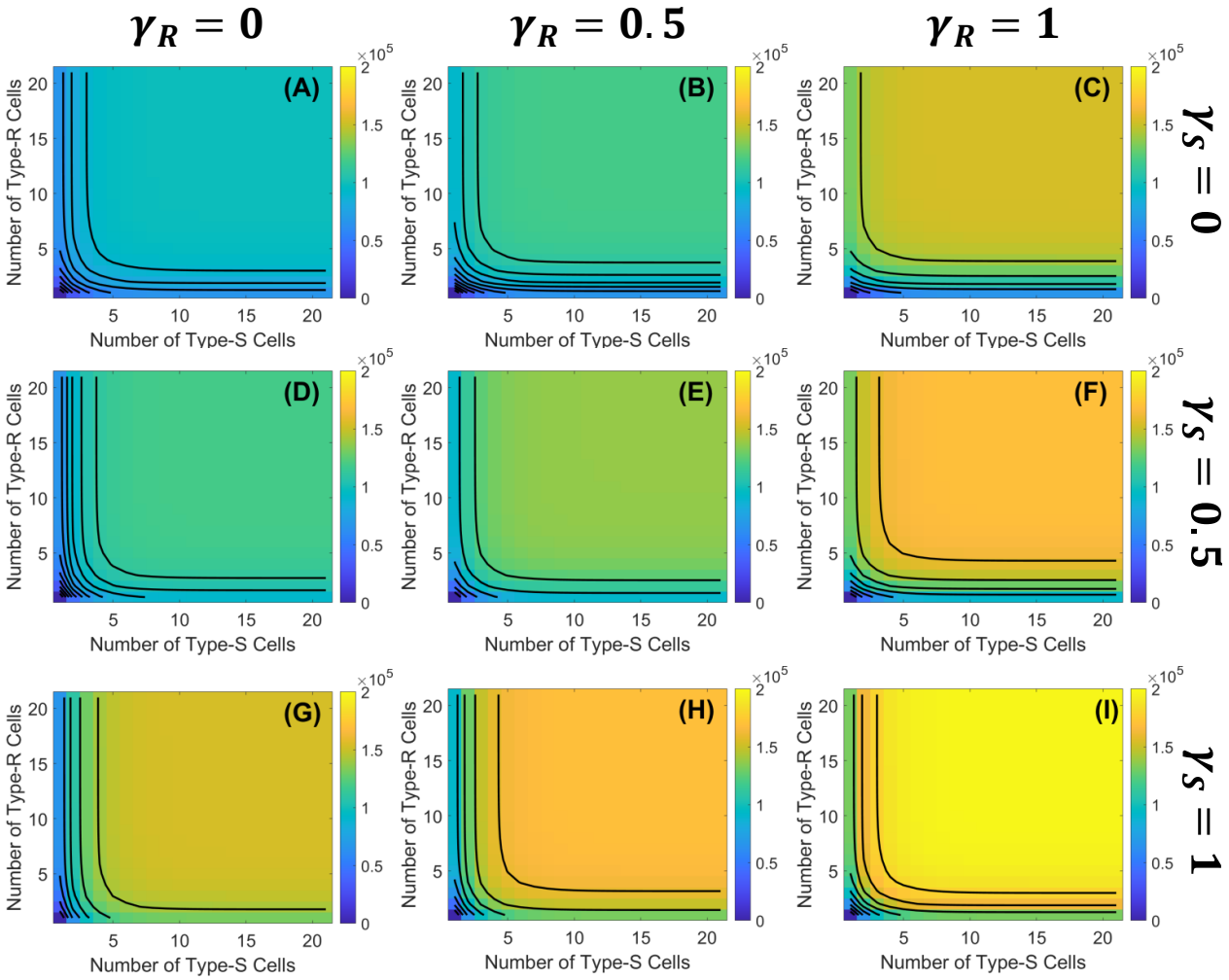

Figure 2.1: **Longer extinction times resulting from systems with smaller birth and death rates, despite the same net growth rate.** The gradient of colors correspond to different values of mean extinction time starting from the location $[n_S, n_R]^T$ on the (N_S, N_R) coordinate. The black lines are part of the corresponding contour plot. We compute the mean extinction times using the parameter values: $\mu = 0$, $\alpha_{SR} = \alpha_{RS} = 0.75$, $b_{0S} = b_{0R} = 2/10$, $d_{0S} = d_{0R} = 1/10$, $K_S = K_R = 10^2$. The differences between the plots in this figure is the values of the intra-species density dependence parameters γ_S and γ_R, which determine how much density dependence is in the birth and death rates. **(A)**: $\gamma_S = 0, \gamma_R = 0$. **(B)**: $\gamma_S = 0, \gamma_R = 0.5$. **(C)**: $\gamma_S = 0, \gamma_R = 1$. **(D)**: $\gamma_S = 0.5, \gamma_R = 0$. **(E)**: $\gamma_S = 0.5, \gamma_R = 0.5$. **(F)**: $\gamma_S = 0.5, \gamma_R = 1$. **(G)**: $\gamma_S = 1, \gamma_R = 0$. **(H)**: $\gamma_S = 1, \gamma_R = 0.5$. **(I)**: $\gamma_S = 1, \gamma_R = 1$. We observe that when $\gamma_S = \gamma_R$, the mean extinction times are symmetric with respect to the diagonal, i.e. the line where $N_S = N_R$. Moreover, in each row from left to right, the mean extinction time gets larger as the birth and death rates of the R-cell population gets smaller. We observe the same pattern for each column from top to bottom.

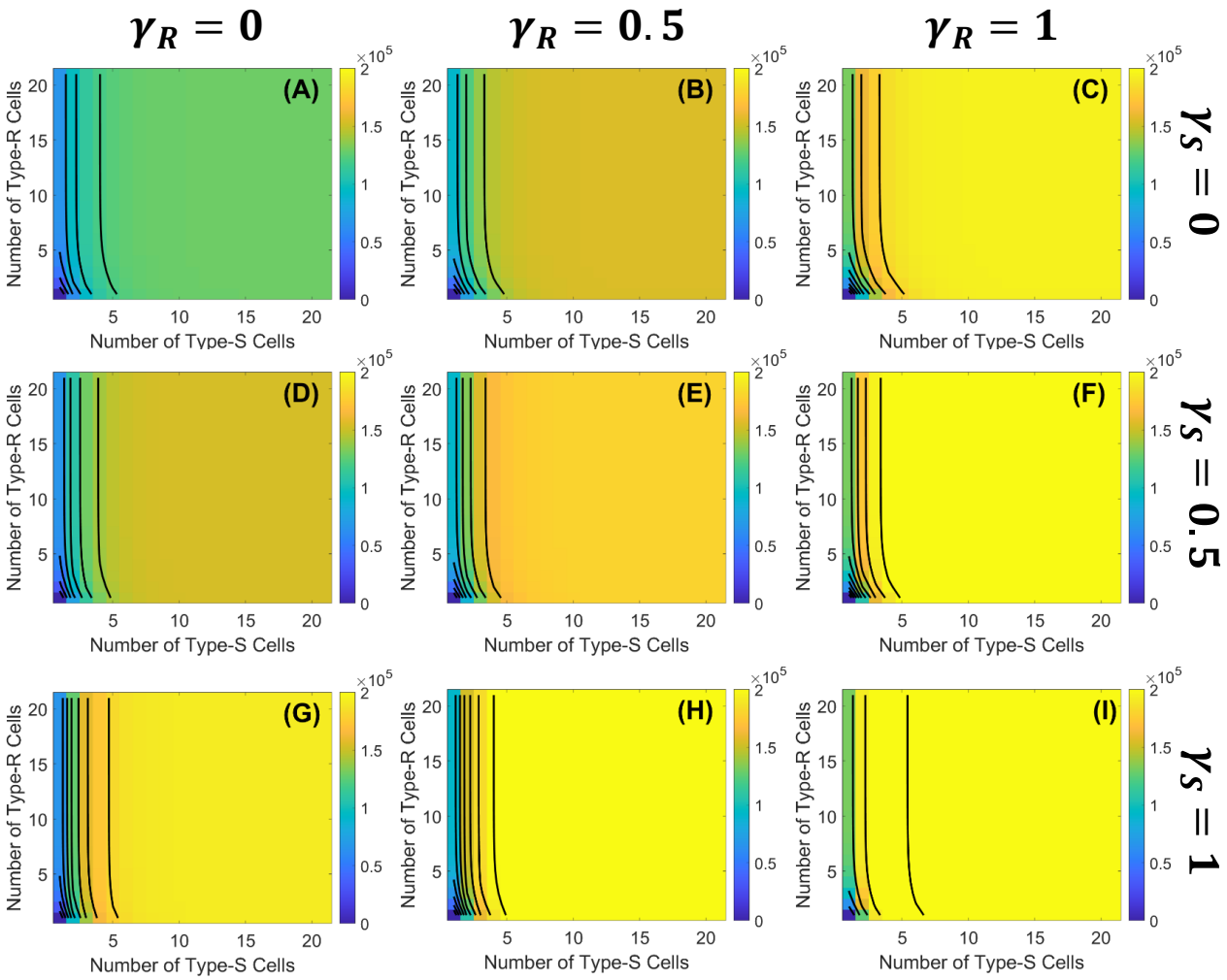

Figure 2.2: **Nonzero mutation rates change the birth rates and subsequently mean extinction times.** The gradient of colors correspond to different values of mean extinction time starting from the location $[n_S, n_R]^T$ on the (N_S, N_R) coordinate. The black lines are part of the corresponding contour plot. We compute the mean extinction times using the parameter values: $\mu = 25 \times 10^{-4}$, $\alpha_{SR} = \alpha_{RS} = 0.75$, $b_{0S} = b_{0R} = 2/10$, $d_{0S} = d_{0R} = 1/10$, $K_S = K_R = 10^2$. The differences between the plots in this figure is the values of the intra-species density dependence parameters γ_S and γ_R, which determine how much density dependence is in the birth and death rates. **(A)**: $\gamma_S = 0, \gamma_R = 0$. **(B)**: $\gamma_S = 0, \gamma_R = 0.5$. **(C)**: $\gamma_S = 0, \gamma_R = 1$. **(D)**: $\gamma_S = 0.5, \gamma_R = 0$. **(E)**: $\gamma_S = 0.5, \gamma_R = 0.5$. **(F)**: $\gamma_S = 0.5, \gamma_R = 1$. **(G)**: $\gamma_S = 1, \gamma_R = 0$. **(H)**: $\gamma_S = 1, \gamma_R = 0.5$. **(I)**: $\gamma_S = 1, \gamma_R = 1$. We observe the same pattern as in Figure 2.1: the smaller the birth and death rates, the larger the mean extinction times. In addition, this figure shows that the nonzero mutation rate $\mu = 25 \times 10^4$, which decreases the birth rate of S-cells and increases the birth rate of R-cells, alter the mean extinction times (compared with Figure 2.1).

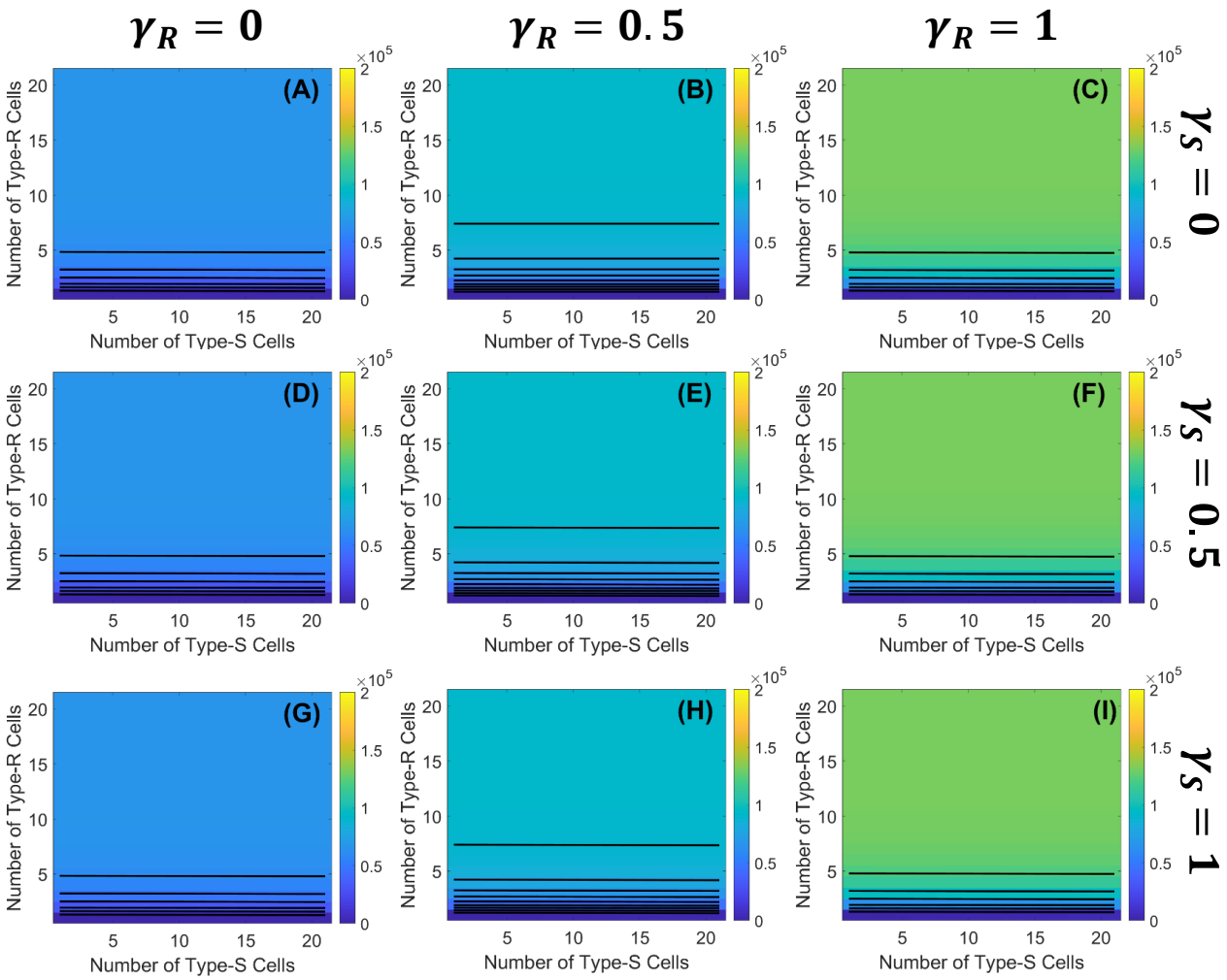

Figure 2.3: **Negative intrinsic *net* growth rate of S-cells reduces mean time to extinction.** The gradient of colors correspond to different values of mean extinction time starting from the location $[n_S, n_R]^T$ on the (N_S, N_R) coordinate. The black lines are part of the corresponding contour plot. We compute the mean extinction times using the parameter values: $\mu = 0$, $\alpha_{SR} = \alpha_{RS} = 0.75$, $b_{0S} = 1/10$, $b_{0R} = 2/10$, $d_{0S} = 2/10$, $d_{0R} = 1/10$, $K_S = K_R = 10^2$. The differences between the plots in this figure is the values of the intra-species density dependence parameters γ_S and γ_R, which determine how much density dependence is in the birth and death rates. **(A)**: $\gamma_S = 0, \gamma_R = 0$. **(B)**: $\gamma_S = 0, \gamma_R = 0.5$. **(C)**: $\gamma_S = 0, \gamma_R = 1$. **(D)**: $\gamma_S = 0.5, \gamma_R = 0$. **(E)**: $\gamma_S = 0.5, \gamma_R = 0.5$. **(F)**: $\gamma_S = 0.5, \gamma_R = 1$. **(G)**: $\gamma_S = 1, \gamma_R = 0$. **(H)**: $\gamma_S = 1, \gamma_R = 0.5$. **(I)**: $\gamma_S = 1, \gamma_R = 1$. We observe that the mean extinction time only changes when the rates of R-cells change, as evident by the same mean extinction time in each column, and different mean extinction times in each row.

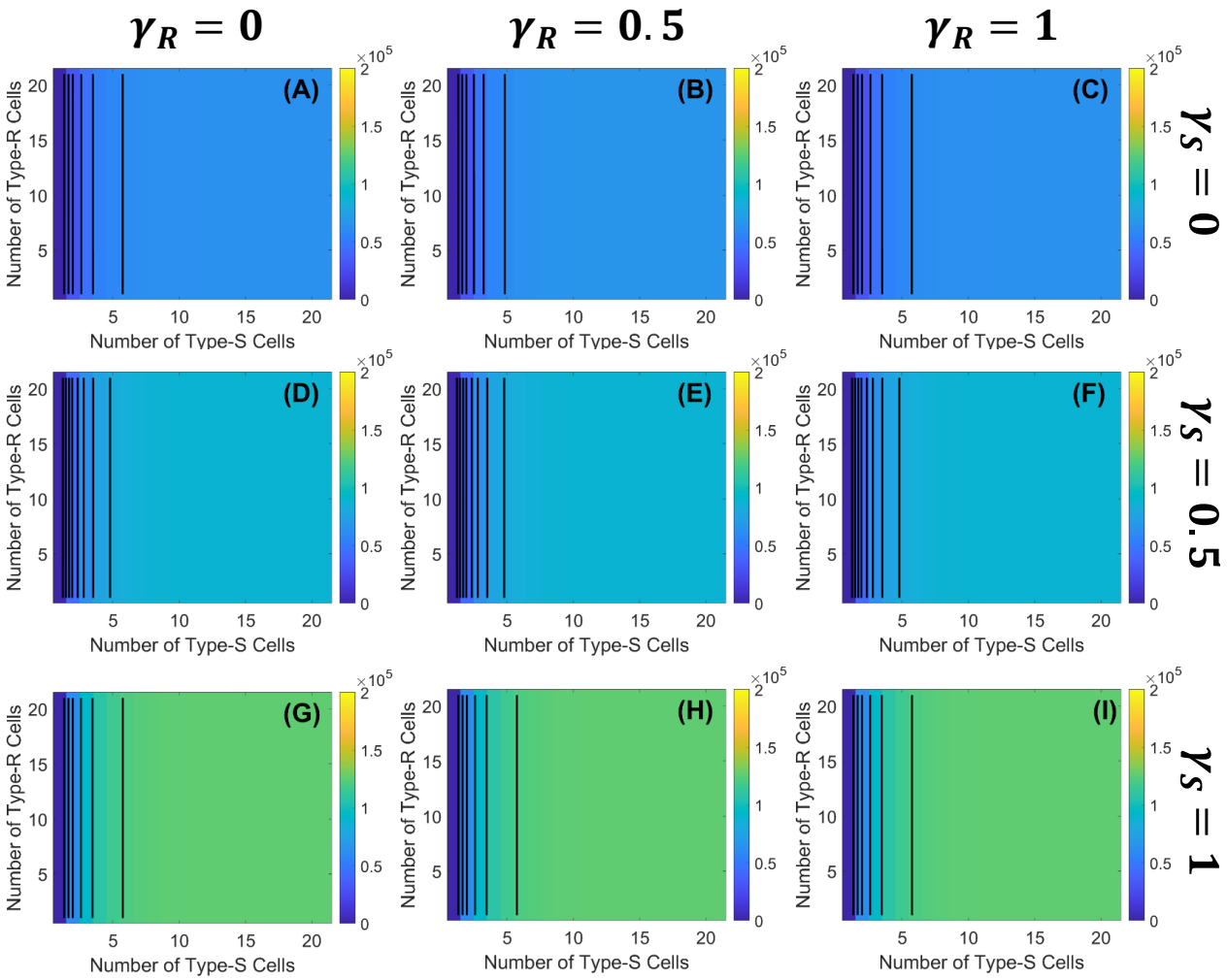

Figure 2.4: **Negative intrinsic *net* growth rate r_R of R-cells reduces mean time to extinction.** The gradient of colors correspond to different values of mean extinction time starting from the location $[n_S, n_R]^T$ on the (N_S, N_R) coordinate. The black lines are part of the corresponding contour plot. We compute the mean extinction times using the parameter values: $\mu = 0$, $\alpha_{SR} = \alpha_{RS} = 0.75$, $b_{0S} = 2/10$, $b_{0R} = 1/10$, $d_{0R} = 1/10$, $d_{0R} = 2/10$, $K_S = K_R = 10^2$. The differences between the plots in this figure is the values of the intra-species density dependence parameters γ_S and γ_R, which determine how much density dependence is in the birth and death rates. **(A)**: $\gamma_S = 0, \gamma_R = 0$. **(B)**: $\gamma_S = 0, \gamma_R = 0.5$. **(C)**: $\gamma_S = 0, \gamma_R = 1$. **(D)**: $\gamma_S = 0.5, \gamma_R = 0$. **(E)**: $\gamma_S = 0.5, \gamma_R = 0.5$. **(F)**: $\gamma_S = 0.5, \gamma_R = 1$. **(G)**: $\gamma_S = 1, \gamma_R = 0$. **(H)**: $\gamma_S = 1, \gamma_R = 0.5$. **(I)**: $\gamma_S = 1, \gamma_R = 1$. We observe that the mean extinction time only changes when the rates of S-cells change, as evident by the same mean extinction time in each row, and different mean extinction times in each column.

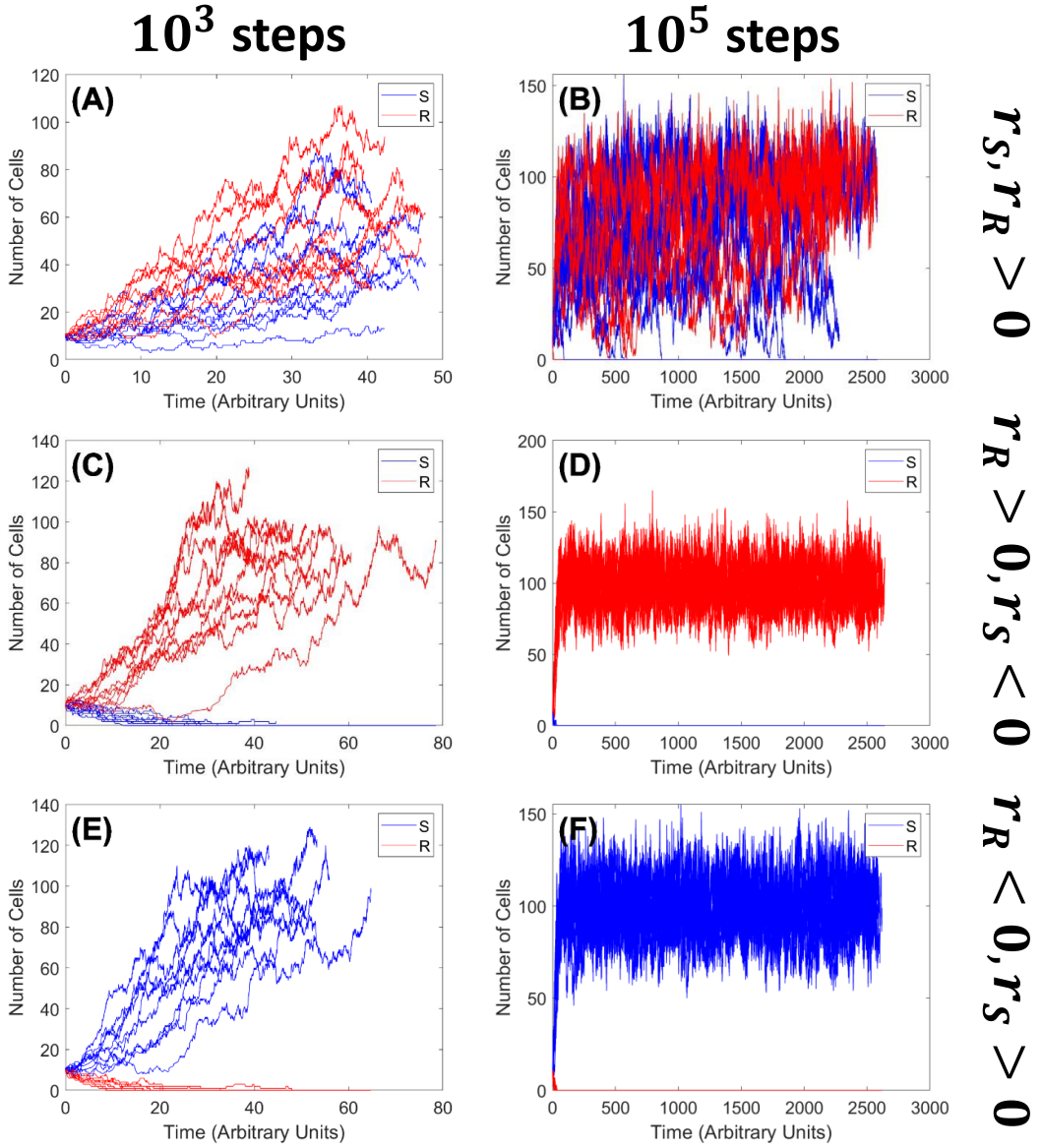

Figure 2.5: **Simulated cell number trajectories for unconstrained cell populations.** This figure compares the dynamics for three scenarios: in scenario **(A, B)** $b_{0S} = b_{0R} = 2/10$ and $d_{0S} = d_{0R} = 1/10$; in scenario **(C, D)** $b_{0S} = 1/10$, $d_{0S} = 2/10$, $b_{0R} = 2/10$, $d_{0R} = 1/10$; in scenario **(E, F)** $b_{0R} = 1/10$, $d_{0R} = 2/10$, $b_{0S} = 2/10$, and $d_{0S} = 1/10$. The trajectories in the left column **(A, C, E)** are simulated for 10^3 timesteps. The trajectories in the right column **(B, D, F)** are simulated for 10^5 timesteps.

2.5 Conclusion and Discussion

In summary, we demonstrate the effect of separate birth and death rates on mean extinction times. This chapter illustrates the need for parsing out *net* growth rate into separate birth and death rates, which leads to our work in Chapter 3 and Chapter 4. In the next chapter, we are going to show that with the additional information of the mean of cell number increments, we can go one step further from this chapter to separately identify birth and death rates from data.

Our broad goal is to analyze the ecological and evolutionary processes involved in drug resistance. The relation between ecology and evolution is fundamentally characterized by their disparate time scales. Towards the broad research question, we would like to make a connection between the work of this chapter and the optimal treatments developed by Nara Yoon et al. (2018) [97] and Jeff Maltas et al. (2019) [61]. We would like to analyze how pathogenic populations evolve under the two treatments described in these two papers, both of which are characterized by certain time scales. Both Nara Yoon et al. (2018) and Jeff Maltas et al. (2019) consider cell populations consisting of different collaterally sensitive cell types following exponential growth. Yoon's model is deterministic, while Maltas' model is stochastic. Collateral sensitivity occurs when the development of an adaptation conferring resistance to one drug induces sensitivity to another drug [67, 78]. In particular, in Yoon's model, there are two cell types: type-A is sensitive to drug A and resistant to drug B, and type-B is resistant to drug A and sensitive to drug B. Yoon's proposed treatment minimizes population size, while Maltas' proposed treatment minimizes drug resistance. In Maltas' paper, drug resistance level is defined as $\log_2\left(\frac{IC_{50,\text{mutant}}}{IC_{50,\text{wild type}}}\right)$, where IC_{50} is the drug concentration that reduces cell population growth to 50%. Yoon's optimal treatment

is characterized by a time point called $T_{\min}$: first we apply one drug until the time point $T_{\min}$ at which time the net growth rates of both cell types from an ODE model are equal to each other, then we switch between the two drugs repeatedly. In contrast, Maltas' optimal treatment is based on the Markov Decision Process dynamic programming method.

A Markov Decision Process is a Markov Chain with a set of actions and a reward or cost function associated with the actions. It consists of four components: states $s \in \mathcal{S}$, actions $a \in \mathcal{A}$, a set of action dependent transition probabilities $p(s, s', a)$, and an immediate objective function $VI(s, a)$. A policy π is a map from the state set $\mathcal{S}$ to the action set $\mathcal{A}$. We want to find an optimal policy π^* that specifies which action is optimal for each state. By "optimal", we mean the cumulative reward or cost $VC(s)$ is maximized or minimized. In doing so, we use the Value Iteration Algorithm, which iterates backward from a final target state. The importance of the objective value from the future is weighted by a discount factor, σ with $0 \leq \sigma \leq 1$. The iteration of the algorithm is based on the Bellman equations described in [7] and in equation (2.12) below. The algorithm is terminated at iteration K if $\|VC_K - VC_{K-1}\|_\infty < \varepsilon$.

$$VC_N(s) = \max_a(\text{ or } \min_a)\left(VI(s,a) + \sigma \sum_{j=1}^{|\mathcal{S}|} p(s, \mathcal{S}(j), a) VC_{N+1}(s)\right), \tag{2.12}$$

where $\mathcal{S}(j)$ denote an element of the state set $\mathcal{S}$. The actions in Maltas' system are six drugs to be used. The states s of the process are vectors of the form $[r_1, \ldots, r_6]^T$, where $r_a(s) \in \{-1, 0, 1, \ldots, 9\}$ is the resistance level to drug a with $a \in \mathcal{A} = \{1, \ldots, 6\}$, when the cell is in state s. Denote the 6-dimensional hypercube comprising the set of all states by $\mathcal{S}$. Maltas et al. defines the "immediate cost" to be the resistance level to the currently applied drug, r_i. Explicitly, if the current state is $s_c = [r_1, \ldots, r_6]^T$ then the

immediate cost R_I for applying drug $a \in \mathcal{A}$ is r_a. The paper aims to minimizes the cost, which is equivalent to maximize the negative of the cost. The time scale in Maltas' optimal treatment is characterized by the discount factor σ in the Markov Decision Process, because σ indicates how much we weight the value of the objective function in the long term. In his paper, Maltas concludes that long-term treatments are better, i.e. the higher the discount factor, the lower the final drug resistance level. However, when we apply the Markov Decision Process dynamic programming method to minimize population size in Yoon's cell population system of two collaterally sensitive cell types, our preliminary results show that the higher the discount factor, the higher the population size. In particular, the actions in our work are the two drugs A and B. The states of the system are the pairs of the number of sensitive cells N_S and resistant cells N_R: (N_S, N_R). The transition probabilities are the birth, death, and mutation rates defined similarly as in Yoon's paper. The cost is the total population size $N_{\text{total}} := (N_S + N_R)$ that we want to minimize. The result regarding discount factor is consistent with our intuition: the longer we let the cell populations grow, the larger their population size becomes (under the exponential growth). For future work, we would like to develop an optimal treatment that balances between the two approaches put forward by Yoon et al. and Maltas et al.'s optimal treatments. In addition, we would like to analyze the time-related parameters in their papers, namely the time point $T_{\min}$ and discount factor σ, and their roles in the evolution of pathogens under the treatments–by taking separate birth and death rates into consideration.

Chapter 3

Stochastic Parameter Identifiability: Birth vs. Death Process Disambiguation

3.1 Motivation

Density dependence, a phenomenon in which a population's *per capita* growth rate changes with population density [40], plays an important role in the ecology and evolution of microbial and cancerous populations, especially under drug treatments. For example, Karslake et al. 2016 [45] shows experimentally that changes in *E.coli.* cell density can either increase or decrease the efficacy of antibiotics. Existing work such as [46], [24], [70], [84], and [27] shows that interactions between drug sensitive and resistant cancerous cells can shape the population's evolution of drug resistance. To analyze the role of density dependence, especially in drug resistance, we consider one of the first, classical mathematical models of density-dependent population dynamics, Verhulst's logistic growth model [92], which

describes the dynamics of a homogeneous population in terms of its net growth rate:

$$\frac{d\phi}{dt} = r\Big(1 - \frac{\phi}{K}\Big)\phi = r\phi - \frac{r}{K}\phi^2. \tag{3.1}$$

In Equation (3.1), ϕ denotes population size, r denotes intrinsic *per capita* net growth rate, and K denotes carrying capacity. The density dependence term $\frac{r}{K}\phi^2$ describes the direct or indirect interactions between individuals in the population. The minus $(-)$ sign indicates the interactions have a negative *net* effect on the population–in particular, reducing the population size. We may consider this negative *net* effect as the difference between a positive effect and a negative effect by introducing a parameter $c \geq 0$:

$$\frac{d\phi}{dt} = r\phi - \frac{r}{K}\phi^2 = r\phi + \underbrace{c\frac{r}{K}\phi^2}_{\text{cooperation}} \underbrace{- (1+c)\frac{r}{K}\phi^2}_{\text{competition}}. \tag{3.2}$$

In the context of ecology, we may interpret the term $c\frac{r}{K}\phi^2$ as cooperation, the term $(1 + c)\frac{r}{K}\phi^2$ as competition, and the parameter c as a measure of cooperation. In this chapter, we consider only competition (i.e. $c = 0$). For future work on the cases where $c > 0$, please Section 3.5. Competitive interactions between individuals can hinder the growth of population size through either the birth process, death process, or some combination of the two. However, the formulation in Equation (3.1) leaves the underlying nature of the density dependence unclear. Density dependence can be manifest in the birth process, death process, or some combination of the two processes. To disambiguate birth-related vs. death-related mechanisms, we rewrite the density dependence term with the parameter γ as follows:

$$\frac{r}{K}\phi^2 = \gamma\frac{r}{K}\phi^2 + (1-\gamma)\frac{r}{K}\phi^2, \quad 0 \leq \gamma \leq 1. \tag{3.3}$$

We interpret the term $\gamma\frac{r}{K}\phi^2$ as the reduction in the population's growth rate due to competition-regulated mechanisms affecting the birth process, and $(1-\gamma)\frac{r}{K}\phi^2$ as the population's competition-regulated mechanisms affecting the death process.

For completion, we also disentangle the intrinsic net growth rate $r\phi$ into birth and death as follows:

$$r\phi = b_0\phi - (b_0 - r)\phi = b_0\phi - d_0\phi, \qquad \text{with } b_0 \geq r > 0 \text{ and } d_0 := b_0 - r \geq 0, \tag{3.4}$$

and interpret b_0 as the population's intrinsic *per capita* birth rate and d_0 as the population's intrinsic *per capita* death rate. We can also interpret b_0 as the population's low-density *per capita* birth rate and d_0 as the population's low-density *per capita* death rate, because for low density, or $N \approx 0$, then the birth and death rates are b_0 and d_0 respectively. Hence, we parameterize Equation (3.1) with γ, b_0, and d_0 as follows:

$$\frac{d\phi}{dt} = \underbrace{\left(b_0 - \gamma\frac{r}{K}\phi\right)\phi}_{\text{birth}} - \underbrace{\left(d_0 + (1-\gamma)\frac{r}{K}\phi\right)\phi}_{\text{death}}, \quad 0 \leq \gamma \leq 1, \quad r = b_0 - d_0. \tag{3.5}$$

For fixed K, b_0, and d_0 (or r), while different values of γ in Equation (3.5) result in equations algebraically equivalent to Equation (3.1), they describe different density-dependent biological processes. For example, in ecology, one distinguishes exploitative competition, where limited resources hinder the growth of the populations, from interference competition, where individuals fight against one another [43]. The former is manifest in density-dependent birth rates, while the latter leads to density-dependent death rates. The term $\gamma\frac{r}{K}\phi^2$ in Equation (3.5) can be interpreted as the case where individuals have to compete for resources and experience reduced birth due to unfavorable living conditions. In contrast,

the term $(1-\gamma)\frac{r}{K}\phi^2$ in Equation (3.5) can be interpreted as the case where interactions between individuals lead to increased death. Nevertheless, both cases may result in the same net growth rates. This example motivates us to ask the following question:

[Q]: In the context of density-dependent population dynamics, how much of a population's change in net growth is through mechanisms affecting birth and how much is through mechanisms affecting death?

The significance of the answer to this question can also be seen in other contexts. The Allee effect [1] of density-dependent dynamics (a positive correlation between population density and *per capita* net growth rate) provides another example. Although the Allee effect is typically modeled with cubic growth [44] instead of logistic growth, answering question [Q] would contribute to understanding the mechanisms that give rise to the effect. Increasing *per capita* net growth rates with increased population density could result from increased cooperation or mating among individuals (increased birth rates) or from a reduction in fighting due to habitat amelioration (decreased death rates) [23]. This distinction is important because populations that experience the Allee effect can become extinct if the population sizes fall below the Allee threshold [83]. Extinction problems are of interest because, for example, we hope to eventually eradicate tumors and harmful bacteria within individual hosts. Clinically, bactericidal drugs such as penicillin promote cell death, while bacteriostatic drugs such as chloramphenicol, clindamycin, and linezolid inhibit cell division [71]. [59] shows that bactericidal and bacteriostatic drugs affect cellular metabolism differently, and the bacterial metabolic state in turn influences drug efficacy. Identifying "-cidal" versus "- static" drugs may help contribute to developing more efficacious drug treatments. From an evolutionary perspective, [31] shows that assuming a zero death rate

leads to overestimating bacterial mutation rates under stress, which in turn can lead to incorrect conclusions about the evolution of bacteria under drug treatments. The authors point out that it is important to separately identify birth and death rates. In another context of evolutionary dynamics, one may compute probability of extinction/escape and mean first-passage time to extinction/escape for cell populations under certain drug treatments such as [42, 49, 29]. Under birth-death process models, computing the probability and mean first-passage time involves separate birth and death rates [5, 68, 32], and cell populations with the same net growth rates–but different birth and death rates–can have different extinction/escape probabilities and mean first-passage times, as demonstrated in Chapter 2. Therefore, the significance of disambiguating birth and death rates underlying a given net growth rate is clear across multiple biological contexts at different scales.

In this chapter, we aim to answer question [Q] by extracting birth and death rates from observations of density-dependent population dynamics. One type of population dynamics information that we can easily observe is population size. However, deterministic dynamical models of populations of size $\hat{\phi}$ do not allow us to disentangle birth rate $b_{\hat{\phi}}$ and death rate $d_{\hat{\phi}}$ from net growth rate $(b_{\hat{\phi}} - d_{\hat{\phi}})$, as the transformations $b_{\hat{\phi}} \to b_{\hat{\phi}} + a_{\hat{\phi}}$ and $d_{\hat{\phi}} \to d_{\hat{\phi}} + a_{\hat{\phi}}$ leave $(b_{\hat{\phi}} - d_{\hat{\phi}})$ unchanged. At a fundamental level, population growth is driven by the birth/division[1] and death of individual cells. At this level, cell birth and death are discrete rather than continuous processes, and may involve stochastic elements such as molecular fluctuations in the reactions within individual cells [55]. Therefore, although the tractability of deterministic population equations has made them attractive as a framework for modeling the growth of pathogenic populations and their responses to therapeutic agents

[1] Although cells do not give birth to offspring in the biological sense, for the rest of the manuscript, we refer to cell division as birth to be consistent with the birth-death process model we use.

[97, 96, 77], a stochastic modeling framework is more appropriate for the research question we consider. Specifically, we consider a birth-death process describing a homogeneous cell population. (By "homogeneous", we mean all cells in the population have the same birth and death rates.) We describe density dependence with logistic growth because it is one of the simplest form of density-dependent dynamics and still captures some realistic cell population dynamics such as the dynamics of cancer cells [33]. Therefore, we will consider a logistic birth-death process model in this chapter.

The remainder of this chapter is structured as follows. In Section 3.2, we describe the mathematical model. Then, we describe our direct estimation method in Section 3.3, where we also validate our method and analyze estimation errors. The method is to estimate birth and death rates Next, in Section 3.4, we use our direct estimation method to answer question [Q] with a focus on disambiguating autoregulation, drug efficacy, and drug resistance mechanisms. Finally, in Section 3.5 we compare our approach to related existing methods [19, 58, 28], and discuss future directions.

3.2 Mathematical Model

We consider systems of homogeneous cells described by a birth-death process, that is, a discrete-state continuous-time Markov chain tracking the number of individual cells $N(t)$ in the system over time t, with state transitions comprising either "birth" ($N \to N+1$) or "death" ($N \to N-1$), as shown in Figure 3.1. In linear birth-death processes, *per capita* birth and death rates are constants that do not depend on N. In contrast, here we consider birth-death processes whose *per capita* birth and death rates depend on N, in order to incorporate density-dependent population dynamics. Specifically, motivated by Equation

(3.5), we define the *per capita* birth rate b_N and death rate d_N in our model as follows:

$$b_N = \max\left\{b_0 - \gamma\frac{r}{K}N, 0\right\}, \tag{3.6}$$

$$d_N = d_0 + (1-\gamma)\frac{r}{K}N, \tag{3.7}$$

where $b_0 > 0$ and $d_0 \geq 0$ are intrinsic (low-density) *per capita* birth and death rates respectively, $r = b_0 - d_0 \geq 0$ is the intrinsic (low-density) *per capita* net growth rate, $K > 0$ is the population's carrying capacity, and $\gamma \in [0, 1]$ determines the extent to which the nonlinear or density-dependent dynamics arises from the *per capita* birth versus death rates. When $\gamma = 0$, the birth process is density-independent; all density dependence lies in the death process. Conversely, when $\gamma = 1$, the density-dependent dynamics is fully contained in the birth process. When $0 < \gamma < 1$, the density-dependent dynamics is split between birth and death. We use the $\max$ function in Definition (3.6) to ensure b_N is nonnegative. The total birth and death rates of the population are $b_N N$ and $d_N N$.

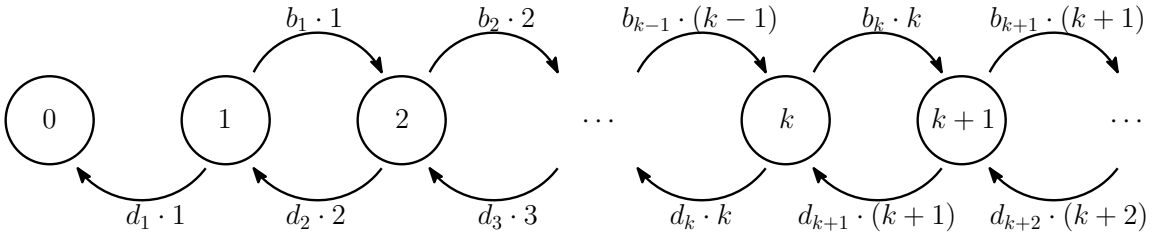

Figure 3.1: **Schematic representation of our birth-death process model.** The *per capita* birth rate b_k, and *per capita* death rate d_k depend on cell number $N = k$ with $k = 0, 1, \ldots$. State $N = k$ transitions to state $N = k + 1$ at rate $b_k \cdot k$ and transitions to state $N = k - 1$ at rate $d_k \cdot k$. At state $N = 0$, the system cannot transition to state $N = 1$, because there is no individual to give birth.

.

For a single-species birth-death process of this form, with $d_1 > 0$ (d_1 is the death rate when $N = 1$) and no immigration, it is well known that the unique stationary probability distribution gives $N(t) \to 0$ as $t \to \infty$ with probability one [2]. Rather than concern

ourselves with the long-term behavior, here we are interested in answering question [Q] by estimating b_N and d_N. Therefore we will focus on the analysis of transient population behavior rather than long-time, asymptotic behavior.

3.3 Direct Estimation of Birth and Death Rates

In this section, we describe our method of estimating the birth and death rates of cell populations that follow the logistic birth-death process model described in Section 3.2. We would like to disambiguate different pairs of birth and death rates for the same observed mean change in population size.

3.3.1 Mathematical Derivation

Let $N(t) \geq 0$ be an integer-valued random variable representing the number of cells at time t. We consider a small time increment Δt, within which each cell can either divide (i.e. one cell is replaced by two cells), die (i.e. one cell disappears and is not replaced), or stay the same (i.e. there is still one cell). Focusing on a single timestep, let $(\Delta N_+|N, \Delta t)$ and $(\Delta N_-|N, \Delta t)$ be two random variables representing the numbers of cells gained and lost, respectively, from an initial population of N cells, after a period of time Δt. The number of cells that neither die nor divide is thus equal to $N - \Delta N_+ - \Delta N_-$. Although the two random variables $(\Delta N_+|N, \Delta t)$ and $(\Delta N_-|N, \Delta t)$ are not strictly independent (as one cell cannot both die and reproduce at the same time), we work in a regime in which the correlation between them is small enough to be neglected. Among N cells, ΔN_+ cells are "chosen" to divide and ΔN_- cells are "chosen" to die. On a time interval of length Δt, the probabilities that a cell divides and dies are $b_N \Delta t + o(\Delta t)$ and $d_N \Delta t +$

$o(\Delta t)$ respectively.[2] For convenience, we will omit the $o(\Delta t)$ correction where possible without introducing inaccuracies. The random variables $(\Delta N_+|N, \Delta t)$ and $(\Delta N_-|N, \Delta t)$ are binomially distributed. In particular,

$$(\Delta N_+|N, \Delta t) \sim \text{ Binomial}(N, b_N \Delta t) \text{ with mean } N b_N \Delta t \text{ and variance } N b_N \Delta t(1 - b_N \Delta t), \tag{3.8}$$

$$(\Delta N_-|N, \Delta t) \sim \text{ Binomial}(N, d_N \Delta t) \text{ with mean } N d_N \Delta t \text{ and variance } N d_N \Delta t(1 - d_N \Delta t). \tag{3.9}$$

Define a random variable $(\Delta N|N, \Delta t)$ to be the net change in population size from N cells after a period of time Δt, i.e. $(\Delta N|N, \Delta t) = (\Delta N_+|N, \Delta t) - (\Delta N_-|N, \Delta t)$. Typically, experimental or clinical measurements reflect only the net change $(\Delta N|N, \Delta t)$ rather than the increase $(\Delta N_+|N, \Delta t)$ or decrease $(\Delta N_-|N, \Delta t)$ separately. Because $(\Delta N_+|N, \Delta t)$ and $(\Delta N_-|N, \Delta t)$ are approximately independent, for sufficiently small Δt, we have

$$\mathbb{E}[\Delta N|N, \Delta t] = \mathbb{E}[\Delta N_+|N, \Delta t] - \mathbb{E}[\Delta N_-|N, \Delta t] = (b_N - d_N) N \Delta t, \tag{3.10}$$

$$\mathbb{V}[\Delta N|N, \Delta t] = \mathbb{V}[\Delta N_+|N, \Delta t] + \mathbb{V}[\Delta N_-|N, \Delta t] - 2\underbrace{\text{Cov}\Big(\Delta N_+|N, \Delta N_-|N\Big)}_{\text{negligible}} \tag{3.11}$$

$$= N b_N \Delta t(1 - b_N \Delta t) + N d_N \Delta t(1 - d_N \Delta t) \tag{3.12}$$

$$= N b_N \Delta t + N d_N \Delta t + O(\Delta t^2) \tag{3.13}$$

$$\approx (b_N + d_N) N \Delta t. \tag{3.14}$$

[2] We adopt the standard convention $\dfrac{o(\Delta t)}{\Delta t} \to 0$ as $\Delta t \to 0$.

Therefore, to estimate birth and death rates $b_N N$ and $d_N N$, we solve the linear system:

$$(b_N - d_N)N = \frac{\mathbb{E}[\Delta N|N, \Delta t]}{\Delta t} \text{ and } (b_N + d_N)N = \frac{\mathbb{V}[\Delta N|N, \Delta t]}{\Delta t}. \tag{3.15}$$

In Section 3.3.3, we discuss how we obtain approximations to $\mathbb{E}[\Delta N|N, \Delta t]$ and $\mathbb{V}[\Delta N|N, \Delta t]$ from discretely sampled finite time series.

3.3.2 Data Simulation

To validate our method, we use simulated *in silico* data. While our underlying model is time-continuous, in experimental and clinical settings, one can only observe cell numbers at discrete time points. In order to efficiently generate an ensemble of trajectories of the birth-death process, we construct a τ-leaping approximation [35] as follows.

Given $N(t)$ individuals at time t, we approximate the number of individuals after a short time interval Δt as

$$N(t + \Delta t) \approx N(t) + \Delta N_+(t) - \Delta N_-(t), \tag{3.16}$$

where $\Delta N_+ \sim \text{Binomial}\Big(N(t), b_{N(t)}\Delta t\Big)$ and $\Delta N_- \sim \text{Binomial}\Big(N(t), d_{N(t)}\Delta t\Big)$ representing the number of cells added to and lost from the system after a period of time Δt. We approximate ΔN_+ and ΔN_- as if they were independent random variables; see discussion in Section 3.3.1. When $N(t)$ is sufficiently large, we approximate the binomial distributions with Gaussian distributions that have the same means and variances as the binomial distributions, for computational efficiency and data format consistency with the experimental data in our lab. Because our discrete-state process in Section 3.2 is now

approximated with a continuous-state process, we replace $N(t)$ with a different notation, $X(t)$, to make this approximation clear. We have

$$\Delta X(t)_+ \sim \text{Normal}\left(X(t)b_{X(t)}\Delta t, X(t)b_{X(t)}\Delta t\Big(1 - b_{X(t)}\Delta t\Big) \right), \tag{3.17}$$

$$\Delta X(t)_- \sim \text{Normal}\left(X(t)d_{X(t)}\Delta t, X(t)d_{X(t)}\Delta t\Big(1 - d_{X(t)}\Delta t\Big) \right). \tag{3.18}$$

Thus, the net change in number of cell after a timestep Δt is

$$\begin{aligned} X(t+\Delta t) - X(t) \approx & X(t)b_{X(t)}\Delta t + \Delta W_+(t)\sqrt{X(t)b_{X(t)}\Big(1 - b_{X(t)}\Delta t\Big)} \\ & - X(t)d_{X(t)}\Delta t - \Delta W_-(t)\sqrt{X(t)d_{X(t)}\Big(1 - d_{X(t)}\Delta t\Big)} \end{aligned} \tag{3.19}$$

$$= \Big(b_{X(t)} - d_{X(t)}\Big)X(t)\Delta t + \sqrt{\Big(b_{X(t)} + d_{X(t)}\Big)X(t)}\Delta W(t) + o\Big(\Delta t\Big), \tag{3.20}$$

where $\Delta W_\pm$ are independent Wiener process increments, and ΔW is a Wiener process increment derived from a linear combination of the $\Delta W_\pm$. Equation (3.20) is the τ-leaping approximation used in our data simulation, which is analogous to the forward Euler algorithm in the deterministic setting. Taking the limit $\Delta t \to dt$, we obtain a version of our population model as a continuous-time Langevin stochastic differential equation

$$dX(t) = \Big(b_{X(t)} - d_{X(t)}\Big)X(t)dt + \sqrt{\Big(b_{X(t)} + d_{X(t)}\Big)X(t)}\, dW(t). \tag{3.21}$$

where $dW(t)$ is delta-correlated white noise satisfying $\langle dW(t)dW(t')\rangle = \delta(t-t')$. We use Equation (3.21) under the Ito interpretation.

3.3.3 Direct Estimation

We conduct S experiments to collect an ensemble of S cell number time series and obtain the following dataset

$$\mathcal{D} = \Big\{ \underbrace{[X(t_0^1), \ldots, X(t_{T_1}^1)]^T}_{\text{time series 1}}, \ldots, \underbrace{[X(t_0^s), \ldots, X(t_{T_1}^s)]^T}_{\text{time series s}}, \ldots, \underbrace{[X(t_0^S), \ldots, X(t_{T_S}^S)]^T}_{\text{time series } S} \Big\}, \tag{3.22}$$

each of which has $T_s + 1$ data points, $s = 1, \ldots, S$. Note that we use the notation X to represent data for the continuous random cell number under a Gaussian approximation, as discussed in Section 3.3.2. We use τ-leaping simulation so that for all the time series indices $s \in \{1, \ldots, S\}$ and all the time point indices $j \in \{0, \ldots, T_s - 1\}$, the difference $t_{j+1}^s - t_j^s$ is equal to Δt, which is independent of s and j, which is consistent with the format of the dataset produced from the EVolutionary biorEactor (EVE) experiments in our laboratory [36]. In our simulation, for all time series $s = 1, \ldots, S$, we choose t_0^s to be equal to t_0 and T_s to be equal to T so that each time series has the same number of data points as the others.

In order to obtain the statistics of the cell number increments, conditioned on the population size, we consider the truncated dataset

$$\mathcal{D}_{-1} = \Big\{ \underbrace{[X(t_0^1), \ldots, X(t_{T_1-1}^1)]^T}_{\text{time series 1}}, \ldots, \underbrace{[X(t_0^s), \ldots, X(t_{T_s-1}^s)]^T}_{\text{time series s}}, \ldots, \underbrace{[X(t_0^S), \ldots, X(t_{T_S-1}^S)]^T}_{\text{time series } S} \Big\}, \tag{3.23}$$

in which we omit the last element of each of the S time series in $\mathcal{D}$. The notation $[\cdots]^T$ means "column vector." We put all the data points in $\mathcal{D}_{-1}$ across the whole ensemble of

trajectories into bins along the population axis. Denote the bin size as η. The left end point X_k of the kth bin $[X_k, X_k + \eta)$, $k = 1, 2, \ldots, k_{\max}$, is equal to $X_k := X_{\min} + (k-1)\eta$, where $X_{\min}$ is the smallest value of cell number across the whole dataset $\mathcal{D}_{-1}$. The total number of bins $k_{\max} \in \mathbb{Z}^+$ is equal to $\left\lceil \frac{X_{\max} - X_{\min}}{\eta} \right\rceil$, where $X_{\max}$ is the largest value of cell number across the whole dataset $\mathcal{D}_{-1}$.

Denote $\hat{S}_k$ as the number of elements in the kth bin $[X_k, X_k + \eta)$. Our method requires a sufficiently large bin size so that the bins have at least two entries in $\mathcal{D}_{-1}$ in order to compute the variances of the cell number increments. For each point $X = X_k + \eta_i$ in the kth bin, $0 \leq \eta_i < \eta$, $i = 1, 2, \ldots, \hat{S}_k$, let ΔX_{ki} be the subsequent increment in X, i.e. $\Delta X_{ki} = X(t_* + \Delta t) - X(t_*)$, where t_* is the time corresponding to $X = X_k + \eta_i$. For each kth bin, $k = 1, 2, \ldots, k_{\max}$, we compute the empirical mean and variance of the cell number increments $\{\Delta X_{ki}\}_{i=1}^{\hat{S}_k}$, and use these statistics (e.g. mean and variance) to estimate the birth and death rates corresponding to the population size $X = X_k + \frac{\eta}{2}$.

3.3.4 Validation and Error Analysis

We validate our method by comparing estimated rates with "true" rates that are used to generate the simulated data. Specifically, we simulate $S = 100$ cell number trajectories, using a numerically efficient τ-leaping approximation described in Section 3.3.2, and estimate birth and death rates using Equations (3.15) and the method described in Section 3.3.3. Figure 3.2 shows that the estimated and true rates are well-aligned. Figure 3.2 **(A, C, E)** shows an ensemble of $S = 100$ independent realizations of the logistic birth-death process formulated in Section 3.2 for three scenarios: $\gamma = 0$ (black), $\gamma = 0.5$ (green), and $\gamma = 1$ (magenta), respectively, simulated using the τ-leaping method with the initial condition $N(t_0) = 10$ and the model parameter values in Table A.1, over a time period

of length 3000 (arbitrary units) and timestep $\Delta t = 1/30$. Figure 3.2 **(B, D, F)** shows the corresponding true and estimated birth and death rates, using a bin size of $\eta = 10^3$. The true birth and death rates are solid blue and red lines respectively. Plus signs $(+)$ denote estimated birth rates, and circles $(\circ)$ denote estimated death rates. We observe that the true and estimated rates are well-aligned.

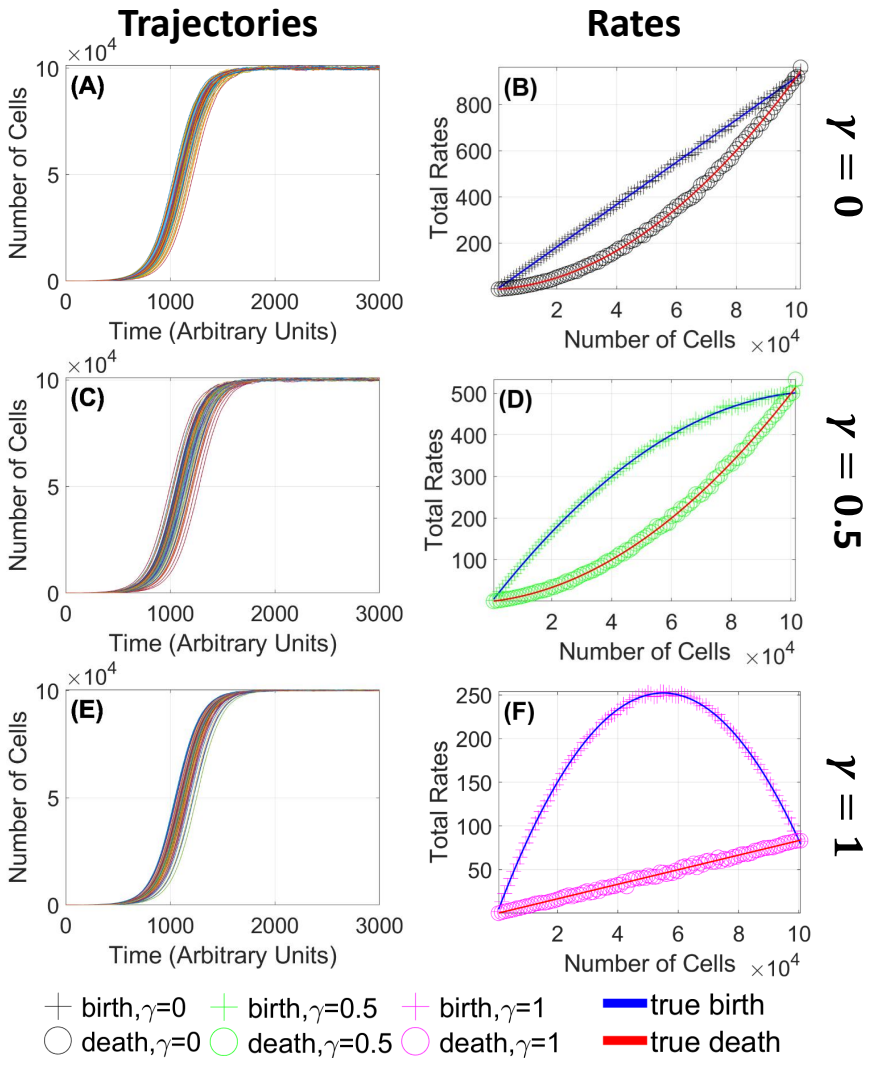

Figure 3.2: **Agreement of estimated and true birth and death rates validates the direct estimation method.** **(A, C, E)**: Time series ensembles simulated using the τ-leaping approximation for the cases $\gamma = 0$ **(A)**, $\gamma = 0.5$ **(C)**, and $\gamma = 1$ **(E)** respectively. Each figure shows $S = 100$ trials. The estimated rates are computed using a bin size of $\eta = 10^3$. Carrying capacity $K = 10^5$ cells; low-density growth rate $r = 1/120$ (arbitrary time units); for other parameters see Table A.1. **(B, D, F)**: Estimated and true birth and death rates, as functions of population size. Blue line: true birth rate. Red line: true death rate. Plus signs $(+)$ denote estimated birth rates; circles $(\circ)$ denote estimated death rates. Throughout the we will use distinct colors to denote values of γ. **(B)** Black: $\gamma = 0$; **(D)** Green: $\gamma = 0.5$; **(E)** Magenta: $\gamma = 1.0$. We observe that the estimated birth and death rates are well-aligned with the true birth and death rates used to simulate the trajectories in **(A)**, **(C)**, and **(E)**.

Using the discretization described in Section 3.3.3, we estimate birth and death rates via the empirical mean $\left\langle \Delta N \middle| N = N_k + \eta_i, 0 \leq \eta_i < \eta, \hat{S}_k \right\rangle$ and empirical variance $\sigma^2\left[\Delta N \middle| N = N_k + \eta_i, 0 \leq \eta_i < \eta, \hat{S}_k\right]$ obtained from an ensemble of $S = 100$ simulated trajectories. Using empirical means and variances to estimate parameters (i.e. birth and death rates) is similar to the Method of Moments [16, 37]. However, our method is slightly different. Instead of using empirical means and variances at the midpoints $\Big($i.e. $\left\langle \Delta N \middle| N = N_k + \frac{\eta}{2}, \hat{S}_k \right\rangle$ and $\sigma^2\left[\Delta N \middle| N = N_k + \frac{\eta}{2}, \hat{S}_k\right]$ respectively $\Big)$, we compute empirical means and variances using all the points in the bins $\Big($i.e. $\left\langle \Delta N \middle| N = N_k + \eta_i, 0 \leq \eta_i \leq \eta, \hat{S}_k \right\rangle$ and $\sigma^2\left[\Delta N \middle| N = N_k + \eta_i, 0 \leq \eta_i \leq \eta, \hat{S}_k\right]$ respectively $\Big)$. Binning helps make sure the sample size is sufficiently large without causing the burden of conducting many experiments. In this section, we analyze the accuracy of our method by showing how the bin size influences the errors in estimating birth and death rates. To quantify the accuracy of our discretization method, we define the error $\mathcal{E}_{k\text{birth}}$ in estimating the birth rate corresponding to population size $N = N_k + \frac{\eta}{2}$, and the error $\mathcal{E}_{k\text{death}}$ in estimating the death rate corresponding to $N =$

$N_k + \frac{\eta}{2}$ as follows:

$$\mathcal{E}_{k\text{birth}} := \frac{\mathbb{E}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big] + \mathbb{V}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big]}{2\Delta t} \tag{3.24}$$

$$- \frac{\Big\langle \Delta N \Big| N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k \Big\rangle + \sigma^2\Big[\Delta N \Big| N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k\Big]}{2\Delta t}, \tag{3.25}$$

$$\mathcal{E}_{k\text{death}} := \frac{\mathbb{V}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big] - \mathbb{E}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big]}{2\Delta t} \tag{3.26}$$

$$- \frac{\sigma^2\Big[\Delta N \Big| N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k\Big] - \Big\langle \Delta N \Big| N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k \Big\rangle}{2\Delta t}. \tag{3.27}$$

$$\tag{3.28}$$

Under the assumption that the samples η_i are iid uniformly distributed on $[0, \eta)$, the theoretical means and variances of the errors $\mathcal{E}_{k\text{birth}}$ and $\mathcal{E}_{k\text{death}}$ are equal to

$$\mathbb{E}\Big[\mathcal{E}_{k\text{birth}}\Big] = \frac{\mathbb{E}\Big[\mathbb{E}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big]\Big] + \mathbb{E}\Big[\mathbb{V}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big]\Big]}{2\Delta t} \tag{3.29}$$

$$- \frac{\mathbb{E}\Big[\Big\langle \Delta N \Big| N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k \Big\rangle\Big] + \mathbb{E}\Big[\sigma^2\Big[\Delta N \Big| N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k\Big]\Big]}{2\Delta t} \tag{3.30}$$

$$= \frac{\mathbb{E}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big] - \mathbb{E}\Big[\Delta N \Big| N = N_k + U, U \sim \text{Unif}[0, \eta)\Big]}{2\Delta t} \tag{3.31}$$

$$+ \frac{\mathbb{V}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big] - \mathbb{V}\Big[\Delta N \Big| N = N_k + U, U \sim \text{Unif}[0, \eta)\Big]}{2\Delta t}, \tag{3.32}$$

$$\mathbb{E}\Big[\mathcal{E}_{k\text{death}}\Big] = \frac{\mathbb{E}\Big[\mathbb{V}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]\Big] - \mathbb{E}\Big[\mathbb{E}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]\Big]}{2\Delta t} \tag{3.33}$$

$$- \frac{\mathbb{E}\Big[\Big\langle \Delta N\Big|N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k\Big\rangle\Big] + \mathbb{E}\Big[\sigma^2\Big[\Delta N\Big|N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k\Big]\Big]}{2\Delta t} \tag{3.34}$$

$$= \frac{\mathbb{E}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big] - \mathbb{E}\Big[\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0,\eta)\Big]}{2\Delta t} \tag{3.35}$$

$$+ \frac{\mathbb{V}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big] - \mathbb{V}\Big[\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0,\eta)\Big]}{2\Delta t}. \tag{3.36}$$

Similarly,

$$\mathbb{V}\Big[\mathcal{E}_{k\text{birth}}\Big] = \mathbb{V}\Big[\mathcal{E}_{k\text{death}}\Big] \tag{3.37}$$

$$= \frac{\mathbb{V}\Big[\Big\langle \Delta N\Big|N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k\Big\rangle\Big] + \mathbb{V}\Big[\sigma^2\Big[\Delta N\Big|N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k\Big]\Big]}{4\Delta t^2} \tag{3.38}$$

$$= \frac{\frac{\mathbb{V}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]}{\hat{S}_k} + \frac{2\Big(\mathbb{V}\Big[\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0,\eta)\Big]\Big)^2}{\hat{S}_k - 1}}{4\Delta t^2}. \tag{3.39}$$

In Figure 3.3, we plot three different kinds curves: the theoretical expected errors, $\Big(\mathbb{E}[\mathcal{E}_{\text{birth}}]$ and $\mathbb{E}[\mathcal{E}_{\text{death}}]\Big)$, represented by blue curves, the theoretical standard deviations of the errors, $\Big(\sqrt{\mathbb{V}[\mathcal{E}_{\text{birth}}]}$ and $\sqrt{\mathbb{V}[\mathcal{E}_{\text{death}}]}\Big)$, represented by red curves, and the empirical errors, $\Big(\mathcal{E}_{\text{empir, birth}}$ and $\mathcal{E}_{\text{empir, death}}\Big)$, represented by black, green, and magenta curves (respectively) as functions of the bin size. For each bin size, the errors are the 2-norms values of the errors

over all bins, particularly,

$$\mathbb{E}[\mathcal{E}_{\text{birth}}] = \sqrt{\sum_k \Big(\mathbb{E}[\mathcal{E}_{\text{kbirth}}]\Big)^2} = \sqrt{\sum_k \Big(\mathbb{E}\Big[\Big(b^{\text{true}}_{N_k+\eta/2} - b^{\text{estimated}}_{N_k+\eta/2}\Big)\Big(N_k+\eta/2\Big)\Big]\Big)^2} \tag{3.40}$$

$$\mathbb{E}[\mathcal{E}_{\text{death}}] = \sqrt{\sum_k \Big(\mathbb{E}[\mathcal{E}_{\text{kdeath}}]\Big)^2} = \sqrt{\sum_k \Big(\mathbb{E}\Big[\Big(d^{\text{true}}_{N_k+\eta/2} - d^{\text{estimated}}_{N_k+\eta/2}\Big)\Big(N_k+\eta/2\Big)\Big]\Big)^2} \tag{3.41}$$

$$\mathbb{V}[\mathcal{E}_{\text{birth}}] = \sqrt{\sum_k \Big(\mathbb{V}[\mathcal{E}_{\text{kbirth}}]\Big)^2} = \sqrt{\sum_k \Big(\mathbb{V}\Big[\Big(b^{\text{true}}_{N_k+\eta/2} - b^{\text{estimated}}_{N_k+\eta/2}\Big)\Big(N_k+\eta/2\Big)\Big]\Big)^2} \tag{3.42}$$

$$\mathbb{V}[\mathcal{E}_{\text{death}}] = \sqrt{\sum_k \Big(\mathbb{V}[\mathcal{E}_{\text{kdeath}}]\Big)^2} = \sqrt{\sum_k \Big(\mathbb{V}\Big[\Big(d^{\text{true}}_{N_k+\eta/2} - d^{\text{estimated}}_{N_k+\eta/2}\Big)\Big(N_k+\eta/2\Big)\Big]\Big)^2} \tag{3.43}$$

$$\mathcal{E}_{\text{empir, birth}} = \sqrt{\sum_k \Big(\mathcal{E}_{\text{kempir, birth}}\Big)^2} = \sqrt{\sum_k \Big(\Big(b^{\text{true}}_{N_k+\eta/2} - b^{\text{estimated}}_{N_k+\eta/2}\Big)\Big(N_k+\eta/2\Big)\Big)^2} \tag{3.44}$$

$$\mathcal{E}_{\text{empir, death}} = \sqrt{\sum_k \Big(\mathcal{E}_{\text{kempir, death}}\Big)^2} = \sqrt{\sum_k \Big(\Big(d^{\text{true}}_{N_k+\eta/2} - d^{\text{estimated}}_{N_k+\eta/2}\Big)\Big(N_k+\eta/2\Big)\Big)^2}, \tag{3.45}$$

where k is the bin index. We derive explicit expressions for $\mathbb{E}\Big[\mathcal{E}_{k\text{birth}}\Big]$, $\mathbb{E}\Big[\mathcal{E}_{k\text{death}}\Big]$, $\mathbb{V}\Big[\mathcal{E}_{k\text{birth}}\Big]$, and $\mathbb{E}\Big[\mathcal{E}_{k\text{death}}\Big]$ as functions of bin size η in Appendix A.2.

For notational convenience, let $\theta_{kb} := b^{\text{true}}_{N_k+\eta/2}\Big(N_k+\eta/2\Big)$, $\hat{\theta}_{kb} := b^{\text{estimated}}_{N_k+\eta/2}\Big(N_k+\eta/2\Big)$,

$\theta_{kd} := d^{\text{true}}_{N_k+\eta/2}\Big(N_k + \eta/2\Big)$, and $\hat{\theta}_{kd} := d^{\text{estimated}}_{N_k+\eta/2}\Big(N_k + \eta/2\Big)$. For each bin k, we have

$$\Big(\mathbb{E}\Big[\mathcal{E}_{k\text{birth}}\Big]\Big)^2 = \Big(\mathbb{E}\Big[\theta_{kb} - \hat{\theta}_{kb}\Big]\Big)^2 = \Big(\mathbb{E}\Big[\hat{\theta}_{kb}\Big] - \theta_{kb}\Big)^2 = \text{bias}^2 \tag{3.46}$$

$$\mathbb{V}\Big[\mathcal{E}_{k\text{birth}}\Big] = \mathbb{V}\Big[\theta_{kb} - \hat{\theta}_{kb}\Big] = \mathbb{V}\Big[\hat{\theta}_{kb}\Big]. \tag{3.47}$$

In statistics, a common measure of estimation errors is mean squared error (MSE). In this context, the mean squared error for the kth bin is $\mathbb{E}\Big[\Big(\mathcal{E}_{\text{kempir, birth}}\Big)^2\Big] = \mathbb{E}\Big[\Big(\hat{\theta}_{kb} - \theta_{kb}\Big)^2\Big]$, which is well-known to be equal to

$$\mathbb{E}\Big[\Big(\mathcal{E}_{\text{kempir, birth}}\Big)^2\Big] = \mathbb{E}\Big[\Big(\hat{\theta}_{kb} - \theta_{kb}\Big)^2\Big] = \underbrace{\Big(\mathbb{E}\Big[\hat{\theta}_{kb}\Big] - \theta_{kb}\Big)^2}_{\text{bias}^2} + \underbrace{\mathbb{V}\Big[\hat{\theta}_{kb}\Big]}_{\text{variance}}. \tag{3.48}$$

The same applies to death rates. The bias-variance tradeoff says that the bias and variance of the estimated parameters are the two conflicting sources of errors [48]. We also observe this conflict in the context of our discretization/binning method. In Figure 3.3, we observe that as the bin size η increases, the expected errors increase, the theoretical variances (or standard deviations) of the errors decreases, and the empirical errors (computed using data from a simulation of $S = 100$ cell number trajectories) balance between the expected values and variances (or standard deviations). The expected values of errors reflect the differences between computing the statistics ΔN at the midpoint $\Big(N = N_k + \frac{\eta}{2}\Big)$ and approximating the statistics using multiple points $\Big(N = N_k + \eta_i, 0 \leq \eta_i < \eta\Big)$. The smaller the bin size, the closer multiple points are to the midpoint, so the error is smaller. However, if the bin is too small, then there are too few samples to accurately estimate theoretical statistics with empirical statistics. The theoretical variances of errors involves sample sizes; the bigger the bin size, the more samples we have. These two competing effects of bin size result in the empirical errors being intermediate values between the two theoretical statistics (expected

values and variances) of the estimation errors. This "Goldilocks principle" is an example of the bias-variance tradeoff common in many estimation problems. We observe in Figure 3.3 that when the bin size is smaller than the optimal bin size, the empirical error curves are close to the sum of the expected error curves and the theoretical error standard deviation curves, which is consistent with Equation (3.48). We expect the same behavior for bin sizes that are larger than the optimal bin size. However, we do not observe what we expect in Figure 3.3. We suspect this inconsistency is due to the fact the uniform distribution approximation in each bin breaks down as the bin size gets larger.

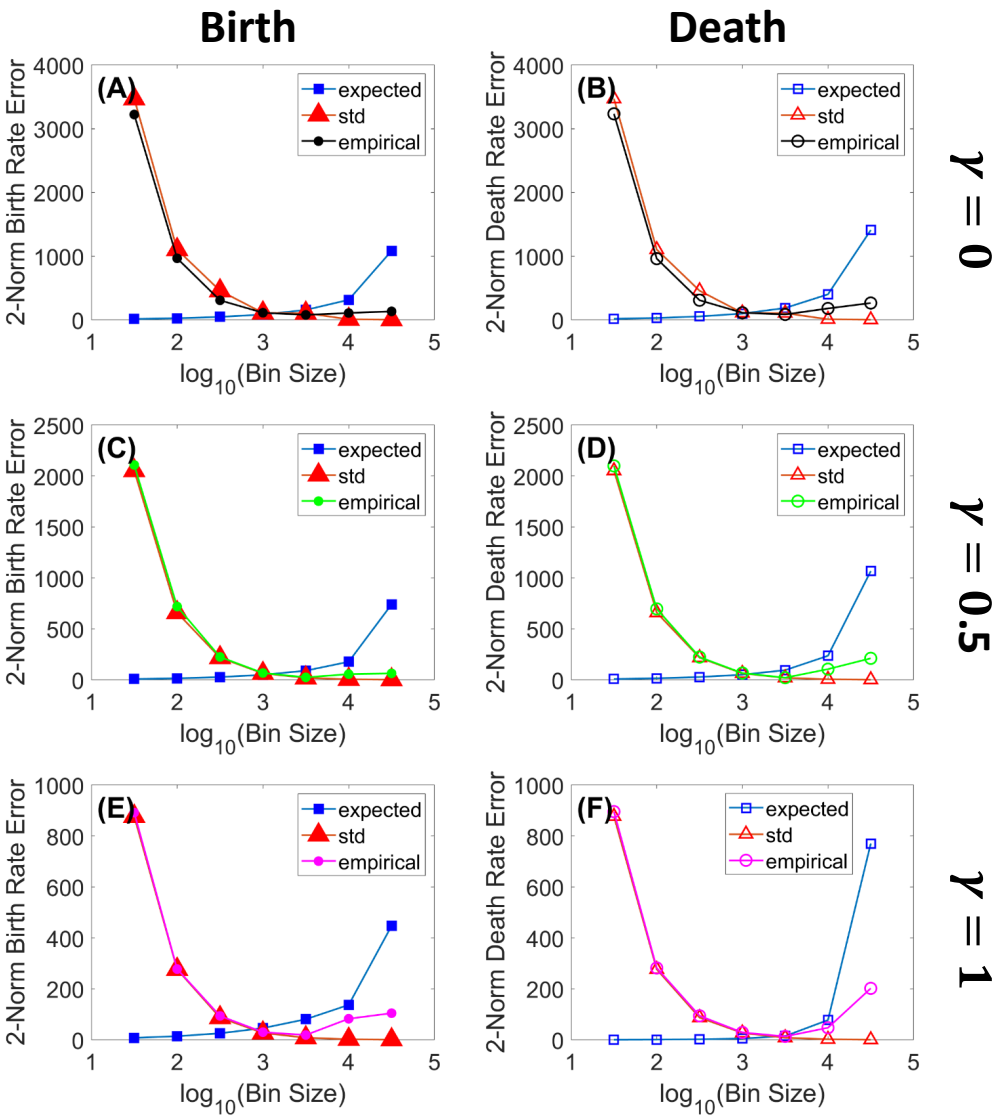

Figure 3.3: **Intermediate bin sizes give optimal estimation performance.** We plot the l_2-norm (over all bins) errors in estimating birth rate (left column) and death rate (right column) as functions of bin size η for carrying capacity $K = 10^5$. Squares ($\square$) denote expected values of errors; triangles ($\triangle$) denote standard deviations of errors; circles ($\circ$) denote empirical errors using data from a simulation of $S = 100$ cell number trajectories. **(A, C, E)**: errors in estimating birth rates. **(B, E, F)**: errors in estimating death rates. **(A, B)**: $\gamma = 0$ (black color); **(C, D)**: $\gamma = 0.5$ (green color); **(E, F)**: $\gamma = 0.5$ (magenta color). We observe that as the bin size η increases, the expected errors increase, the theoretical variances/standard deviations of the errors decreases, and the sample errors balance between the expected values and variances and have convex quadratic shapes.

3.4 Application: Inferring Underlying Mechanisms of Autoregulation, Drug Efficacy, and Drug Resistance

In this section, we apply our direct estimation method (Section 3.3) to shed light on drug resistance mechanisms of pathogenic cell populations (e.g. malignant tumors or harmful bacteria) by disambiguating whether the mechanisms involve the birth process, the death process, or both processes. We consider the scenario where a homogeneous pathogenic cell population grows to its carrying capacity, then is treated with a drug that reduces its carrying capacity, and then overcomes the drug effect to regain its original carrying capacity. Within this scenario, we use "drug resistance" to refer to the pathogenic population's recovery of its original carrying capacity. (For a discussion of different perspectives on drug resistance, please refer to Section 3.5). We divide our analysis into three stages: (1) auto-regulated growth, (2) drug treatment, and (3) drug resistance. The autoregulation stage occurs before the drug treatment stage; during this stage, the cells regulate themselves in such a way that their growth saturates at a given carrying capacity. Such regulation can be due to direct or indirect cell-to-cell interactions, such as exploitation or interference competition. During the drug treatment stage, the cells are regulated by an applied drug, which reduces the population's carrying capacity. The reduced carrying capacity may result either by increasing the density-dependent death rate ("-cidal" effect) or decreasing the density-dependent birth rate ("-static" effect), or both. Finally, in the drug resistance stage, after having been treated with either a "-cidal" or "-static" drug, the cell population fights back and regains to its original carrying capacity by either decreasing its density-dependent death rate or by increasing its density-dependent birth rate. In each of these stages, changes in either birth or death rates could result in the same observed net dynamics. It is important to disambiguate the underlying mechanisms, to appropriately design optimal treatments with

the goal of eventually eradicating the pathogens (i.e. reducing their sizes to zero).

3.4.1 Stage 1: Autoregulation

Cell populations with the same mean net growth rate can grow and reach their carrying capacities through different mechanisms: density-dependent birth dynamics, density-dependent death dynamics, or some combination of the two. The differences between theses scenarios are characterized by different values of the density dependence parameter, γ, in the model described in Section 3.2. We demonstrate this variety with three scenarios:

(I) Density dependence occurs only in the *per capita* death rate, while the *per capita* birth rate is density independent ($\gamma = 0$). In this case, the *per capita* birth rate is $(b_N)_{\text{original}} = b_0$ and the *per capita* death rate is $(d_N)_{\text{original}} = d_0 + \frac{r}{K}N$. The plots corresponding to scenario (I) in all the figures in this chapter are represented by the color black.

(II) Density dependence occurs in both the *per capita* birth and death rates ($\gamma = 0.5$). In this case, the *per capita* birth rate is $(b_N)_{\text{original}} = \max\left\{b_0 - 0.5\frac{r}{K}N, 0\right\}$ and the *per capita* death rate is $(d_N)_{\text{original}} = d_0 + 0.5\frac{r}{K}N$. The plots corresponding to scenario (II) in all the figures in this chapter are represented by the color green.

(III) Density dependence occurs only in the *per capita* birth rate, while the *per capita* death rate is density-independent ($\gamma = 1$). In this case, the *per capita* birth rate is $(b_N)_{\text{original}} = \max\left\{b_0 - \frac{r}{K}N, 0\right\}$ and the *per capita* death rate is $(d_N)_{\text{original}} = d_0$. The plots corresponding to scenario (III) in all the figures in this chapter are represented by the color magenta.

Recall that the random variable $N(t)$ represents the number of cells at time t in the logistic birth-death process described in Section 3.2. Similarly, the parameters b_0, d_0, r, K, and γ

are the same as those described in Section 3.2. Scenarios (I-III) have the same net growth rate, $(b_N)_{\text{original}} - (d_N)_{\text{original}}$, but different magnitudes of birth and death rates. Scenario (I) represents a situation in which the carrying capacity of the population arises through an increase in the *per capita* death rate with population density. Such a scenario could arise, for example, when competition is mediated through cell-to-cell interactions such as predation or other conspecific lethal interactions. Scenario (III), in contrast, represents a situation in which the *per capita* death rate remains constant with increasing population size, but the *per capita* birth (cell division) rate declines. Such a scenario could arise, for example, when competition is mediated by accumulation of waste products or competition for food resources that slow cell division.[3] Scenario (II), intermediate between (I) and (III), represents a combination of such density-dependent mechanisms.

Figure 3.4 shows how our direct estimation method can disambiguate the three autoregulation scenarios (I), (II), and (III). We simulate 100 trajectories of the cell population under each scenario with an initial population $N(t_0) = 10$, and the parameter values in Table A.1, except the carrying capacity value. In addition to using $K = 10^5$ for carrying capacity (Figure 3.4 **(A)**, **(C)**, and **(E)**), we also simulate the population with a carrying capacity $K = 10^2$ (Figure 3.4 **(B)**, **(D)**, and **(F)**). We demonstrate that when the carrying capacity is small (e.g. 10^2), it is easier to see the noise levels than when the carrying capacity is large (e.g. 10^5), as seen in Figure 3.4 **(C)** and **(D)**, because the fluctuations are larger relative to the mean population. After simulating an ensemble of cell number trajectories, we estimate birth and death rates from that ensemble of trajectories using the method given in section 3.3.3, as shown in Figure 3.4 **(A)** and **(B)**. Then, we randomly select one trajectory $N(t)$ from those 100 trajectories, as shown in Figure 3.4 **(C)** and **(D)**, and plot birth and death

[3] Resources depletion can also increase death rates. In this chapter, we neglect such effect.

rates as functions of time, as shown in Figure 3.4 **(E)** and **(F)**. The birth rate $b_{N(t)}$ and death rate $d_{N(t)}$ as functions of time are calculated by treating the rates as composite functions of the cell number $N(t)$, and finding the rates that correspond to the selected cell number time series in Figure 3.4 **(C)** and **(D)**. We couch our model in terms of density-dependent changes in birth and/or death rates (thus, population-number dependent, given a fixed total volume of the cell culture). When the same net growth rate can arise from different density-dependent mechanisms, at the level of birth and death rates, the birth and death rates as functions of *time* can appear markedly different. For example, while in scenarios (I) and (II), the birth and death rates show monotonically increasing, sigmoidal shapes throughout time, in scenario (III), the birth rate has the shape of a concave-down quadratic function as shown by the "$+$" magenta curves in Figure 3.4 **(E)** and **(F)**.

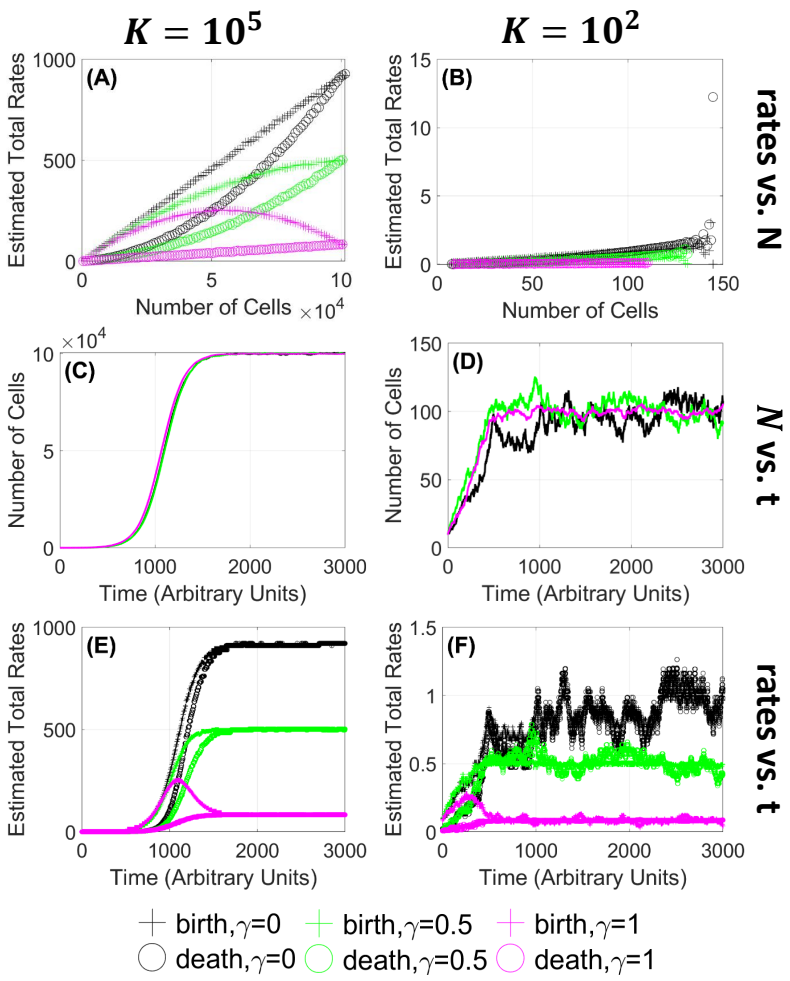

Figure 3.4: **Underlying autoregulation mechanisms are distinguished by separately identified birth and death rates, not necessarily by net changes in total population size.** Plots showing that cell populations with the same net growth rate and carrying capacity grow to the carrying capacity under different density-dependent mechanisms, although the observed dynamics (shown in (C) and (D)) look indistinguishable. Noise levels are more visible for smaller carrying capacities due to smaller scales. **(A, C, E)**: logistic birth-death processes with carrying capacity $K = 10^5$. **(B, D, F)**: logistic birth-death processes with carrying capacity $K = 10^2$. **(A-F)**: black curves correspond to the scenario $\gamma = 0$; green curves correspond to the scenario $\gamma = 0.5$; magenta curves correspond to the scenario $\gamma = 1$. **(A, B)**: estimated birth and death rates for three scenarios using the direct estimation method with an ensemble of 100 trajectories. **(C, D)**: one selected trajectory for each scenario. **(E, F)**: estimated birth and death rates throughout time corresponding to the trajectories in **(C)** and **(D)**. Plus signs $(+)$ denote estimated birth rates; circles $(\circ)$ denote estimated death rates.

3.4.2 Stage 2: Drug Efficacy

In this stage, the cell population is treated with a drug that cuts its carrying capacity in half, either by increasing the *per capita* death rate d_N or by decreasing the *per capita* birth rate b_N. If a drug acts by increasing the *per capita* death rate, we refer to it as a drug with a "-cidal" mechanism. If a drug acts by lowering the *per capita* birth rate, we refer to it as a drug with a "-static" mechanism. If a drug combines both effects, we refer to such a treatment as having a mixed mechanism.

Figure 3.5 shows our disambiguation results for the drug efficacy mechanisms for the three scenarios (I), (II), and (III) described in Section 3.4.1. We simulate 100 trajectories of the cell population under each scenario with an initial population $N(t_0) = 10$, and the parameter values in Table A.1 under three drug efficacy cases: (i) without drug (black curves), (ii) with "-cidal" (death-promoting) drug (red curves), and (iii) with "-static" (birth-inhibiting) drug (blue curves). Both of the drugs reduce the original carrying capacity K to $K/2$. Under the "-cidal" drug, the *per capita* birth and death rates are as follows

- Scenario (I), -cidal: Drug increases the *per capita* death rate, $(d_N)_{\text{cidal}} > (d_N)_{\text{original}}$:

$$(b_N)_{\text{cidal}} = (b_N)_{\text{original}} = b_0, \tag{3.49}$$

$$(d_N)_{\text{cidal}} = d_0 + \frac{r}{(K/2)}N. \tag{3.50}$$

  The density dependence parameter γ remains 0.

- Scenario (II), -cidal: Drug increases the *per capita* death rate, $(d_N)_{\text{cidal}} > (d_N)_{\text{original}}$:

$$(b_N)_{\text{cidal}} = (b_N)_{\text{original}} = \max\left\{b_0 - 0.5\frac{r}{K}N, 0\right\} = \max\left\{b_0 - 0.25\frac{r}{(K/2)}N, 0\right\}, \tag{3.51}$$

$$(d_N)_{\text{cidal}} = d_0 + 0.75\frac{r}{(K/2)}N = d_0 + 1.5\frac{r}{K}N. \tag{3.52}$$

  The density dependence parameter γ changes from 0.5 to 0.25.

- Scenario (III), -cidal: Drug increases the *per capita* death rate, $(d_N)_{\text{cidal}} > (d_N)_{\text{original}}$:

$$(b_N)_{\text{cidal}} = (b_N)_{\text{original}} = \max\left\{b_0 - \frac{r}{K}N, 0\right\} = \max\left\{b_0 - 0.5\frac{r}{(K/2)}N, 0\right\}, \tag{3.53}$$

$$(d_N)_{\text{cidal}} = d_0 + 0.5\frac{r}{(K/2)}N = d_0 + \frac{r}{K}N. \tag{3.54}$$

  The density dependence parameter γ changes from 1 to 0.5.

Under the "-static" drug, the *per capita* birth and death rates are as follows

- Scenario (I), -static: Drug decreases the *per capita* birth rate, $(b_N)_{\text{static}} < (b_N)_{\text{original}}$:

$$(d_N)_{\text{static}} = (d_N)_{\text{original}} = d_0 + \frac{r}{K}N = d_0 + 0.5\frac{r}{(K/2)}N, \tag{3.55}$$

$$(b_N)_{\text{static}} = \max\left\{b_0 - 0.5\frac{r}{K}N, 0\right\}. \tag{3.56}$$

  The density dependence parameter γ changes from 0 to 0.5.

- Scenario (II), -static: Drug decreases the *per capita* birth rate, $(b_N)_{\text{static}} < (b_N)_{\text{original}}$:

$$(d_N)_{\text{static}} = (d_N)_{\text{original}} = d_0 + 0.5\frac{r}{K}N = d_0 + 0.25\frac{r}{(K/2)}N, \tag{3.57}$$

$$(b_N)_{\text{static}} = \max\left\{b_0 - 0.75\frac{r}{(K/2)}N\right\} = \max\left\{b_0 - 1.5\frac{r}{K}N\right\}. \tag{3.58}$$

  The density dependence parameter γ changes from 0.5 to 0.75.

- Scenario (III), -static: Drug decreases the *per capita* birth rate, $(b_N)_{\text{static}} < (b_N)_{\text{original}}$:

$$(d_N)_{\text{static}} = (d_N)_{\text{original}} = d_0, \tag{3.59}$$

$$(b_N)_{\text{static}} = \max\left\{b_0 - \frac{r}{(K/2)}, 0\right\}. \tag{3.60}$$

  The density dependence parameter γ remains 1.

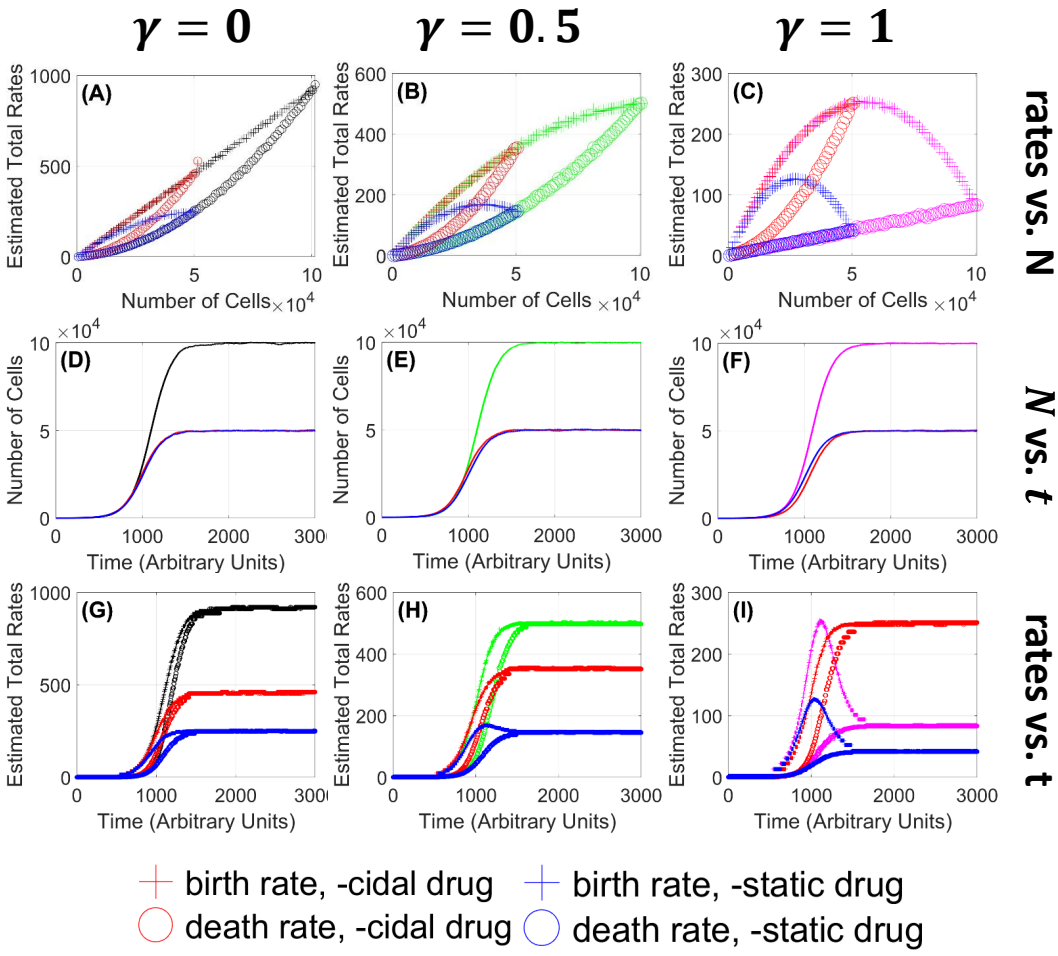

Figure 3.5: **Separating birth and death rates distinguishes the underlying -cidal versus -static action of drugs.** In each of the density-dependent cases (I), (II), (III), two different drugs reduce the cell population's carrying capacity to the same level (shown in red and blues curves in (D), (E), (F)), but the underlying mechanisms are different: increasing death rates (red curves) or decreasing birth rates (blue curves). Black, green, and magenta curves represent scenarios (I), (II), and (III) without drugs. Red curves represent the scenarios under a "-cidal" drug, and blue curves represent the scenarios under a "-static" drug. Plus signs $(+)$ denote estimated birth rates; circles $(\circ)$ denote estimated death rates. **(A, D, G)**: scenario (I) with $\gamma = 0$ (black curves), **(B, E, H)**: scenario (II) with $\gamma = 0.5$ (green curves), **(C, F, I)**: scenario (III) with $\gamma = 1.0$ (magenta curves). **(A, B, C)**: birth and death rates estimated from 100-trajectory ensembles. **(D, E, F)**: a representative trajectory without drug and two representative trajectories treated with drugs. The red and curves trajectories have the same mean-field behavior but the drug mechanisms are different. **(G, H, I)**: estimated birth and death rates throughout time corresponding to the trajectories in **(D, E, F)**.

Figure 3.5 illustrates the effects of -cidal versus -static drugs in scenario (I) in the first column (panels **A, D, G**), scenario (II) in the second column (panels **B, E, H**), and scenario (III) in the third column (panels **C, F, I**). As is evident in Figure 3.5 **(D, E, F)**, the observed cell number dynamics can be very similar in each scenario. However, Figure 3.5 panels **(A, B, C)** and **(G, F, H)** show that the underlying birth and death processes that give rise to the dynamics can be very different. Specifically, in **(D, E, F)**, we see that the red and blue curves are almost indistinguishable. Thus, these scenarios could not easily be distinguished from the general shape of the growth curve alone. However, to obtain the red curves, we keep the *per capita* birth rates the same and increase the *per capita* death rates, and to obtain the blue curves, it is the other way around–as illustrated in panels **(A, B, C)**. The time-dependent birth and death rates in panels **(G, H, I)** also show significant differences. In particular, the *per capita* birth rates under the "-static" drug treatment (blue curves) are monotonically increasing in scenario (I) (density-dependent death rate, as shown in **(G)**), but show a pronounced increase and then decrease in scenario (III), as shown in **(I)**. Thus, by extracting birth and death rates separately from cell number time series, we are able to disambiguate underlying drug mechanisms.

3.4.3 Stage 3: Drug Resistance

After having been treated with drugs that reduce their carrying capacities as described in Section 3.4.2, cell populations can overcome the drug effects and revert to their original carrying capacities. We refer to this phenomenon as drug resistance. In this section, we demonstrate different mechanisms through which cell populations might develop drug resistance against a -cidal drug (Figure 3.6) and against a -static drug (Figure 3.7), for the

three scenarios (I), (II), and (III) described in Section 3.4.1. In simulating the scenarios for these two cases, we set the original carrying capacity to be $K = 10^3$, and keep the other original parameters to be the same as in Table A.1. On this scale, the fluctuations are readily apparent in the traces; the method works robustly for larger values of K as well. Throughout, "original" means "wild-type" and "before drug treatment". We consider the case where the "-cidal" and "-static" drugs reduce the carrying capacity by a factor of 2.

The effect of a drug and the cell population's resistance mechanism can be captured in part by a change in its carrying capacity, in part by a change in the distribution of density-dependent effects, described by γ, and in part by a change in the *per capita* intrinsic/low-density birth and death rates, b_0 and d_0. Figure 3.6 illustrates different mechanisms of drug resistance to the "-cidal" effect described in Section 3.4.2.

- In scenario (I) as shown in Figure 3.6 **(A, B)**, the cell population can develop resistance either by decreasing its *per capita* death rate back to the original rate:

$$(d_N)_{\text{cidal, resistant}} = (d_N)_{\text{original}} = d_0 + \frac{r}{2(K/2)}N, \tag{3.61}$$

$$(b_N)_{\text{cidal, resistant}} = (b_N)_{\text{cidal}} = (b_N)_{\text{original}} = b_0, \tag{3.62}$$

  or by increasing its *per capita* intrinsic birth rate to $b_0 + r = 2r + d_0$:

$$(b_N)_{\text{cidal, resistant}} = b_0 + r = 2r + d_0, \tag{3.63}$$

$$(d_N)_{\text{cidal, resistant}} = (d_N)_{\text{cidal}} = d_0 + \frac{r}{(K/2)}N = d_0 + \frac{2r}{K}N, \tag{3.64}$$

  which leads to the *per capita* intrinsic net growth rate r increasing to $2r$. Such an increase in the intrinsic cell division rate could potentially arise through mutation.

(Why such a mutation would not already have been exploited in the wild-type cell line is a question beyond the scope of this chapter.) In both drug resistance scenarios, the density dependence parameter γ remains 0, which suggests no significant change in the cell-to-cell interaction modality.

- In scenario (II) as shown in Figure 3.6 **(C, D)**, the cell population can develop resistance by either decreasing its *per capita* death rate back to the original rate:

$$(d_N)_{\text{cidal, resistant}} = (d_N)_{\text{original}} = d_0 + 0.5\frac{r}{2(K/2)}N, \tag{3.65}$$

$$(b_N)_{\text{cidal, resistant}} = (b_N)_{\text{cidal}} = (b_N)_{\text{original}} = \max\left\{b_0 - 0.5\frac{r}{K}N, 0\right\}, \tag{3.66}$$

or by increasing its *per capita* intrinsic birth rate to $b_0 + 2r = 3r + d_0$:

$$(b_N)_{\text{cidal, resistant}} = \max\left\{3r + d_0 - 0.5\frac{3r}{K}N, 0\right\}, \tag{3.67}$$

$$(d_N)_{\text{cidal, resistant}} = (d_N)_{\text{cidal}} = d_0 + 1.5\frac{r}{2(K/2)}N = d_0 + 0.5\frac{3r}{K}N, \tag{3.68}$$

which shows that the *per capita* intrinsic net growth rate r would have to increase to $3r$. Such an increase in the intrinsic cell division rate could potentially arise through mutation. In both drug resistance scenarios, the density dependence parameter γ changes from 0.25 back to 0.5, which suggests a change in cell-to-cell interaction modality. Note that the drug resistance mechanism through death in this scenario is different from scenario (I), because in scenario (I), the *per capita* death rate decreases only due to increased carrying capacity, while in this scenario, the *per capita* death rate decreases also due to decreased density dependence of death (i.e. $(1-\gamma)$ changes from 0.75 to 0.5).

- In scenario (III) as shown in Figure 3.6 **(E, F)**, the cell population can become drug resistant either by decreasing its *per capita* death rate back to the original rate:

$$(d_N)_{\text{cidal, resistant}} = (d_N)_{\text{original}} = d_0, \tag{3.69}$$

$$(b_N)_{\text{cidal, resistant}} = (b_N)_{\text{cidal}} = (b_N)_{\text{original}} = \max\left\{b_0 - \frac{r}{K}N, 0\right\}, \tag{3.70}$$

  or by increasing its *per capita* birth rate to

$$(b_N)_{\text{cidal, resistant}} = b_0, \tag{3.71}$$

$$(d_N)_{\text{cidal, resistant}} = (d_N)_{\text{cidal}} = d_0 + \frac{r}{2(K/2)}N. \tag{3.72}$$

  In the first scenario (drug resistance mechanism via modified death rate), the density dependence parameter γ changes from 0.5 back to 1, while in the drug resistance mechanism through birth, the density dependence parameter γ changes from 0.5 to 0. Both of these scenarios suggest changes in the cell-to-cell interaction modalities. The latter suggests a significant change from full density dependence in birth (before drug treatment) to full density dependence in death (after "-cidal" drug treatment and resistance). Note that the *per capita* intrinsic rates, b_0 and d_0, remain the same.

Figure 3.6 shows that having been treated with a "-cidal" drug, the cell population can develop resistance either by reverting to its original dynamics–the red curves change back to the black, green, and magenta curves for scenarios (I), (II), and (III) respectively in the figure, or by increasing its *per capita* birth rate as illustrated by the cyan curves. We may call the latter drug resistance mechanism "enhanced fecundity" or "hyper-birth." Without computing the birth and death rates explicitly, we observe from cell number time series that if the resistant cell population (cyan curves) reaches its original carrying capacity earlier

than the wild-type population (black, green, magenta curves) as in Figure 3.6 **(B, D)** or if the typical fluctuations around the mean population size are visibly larger than the fluctuations of the wild-type as in Figure 3.6 **(F)**, we may hypothesize that the population has developed drug resistance through the "hyper-birth" mechanism.

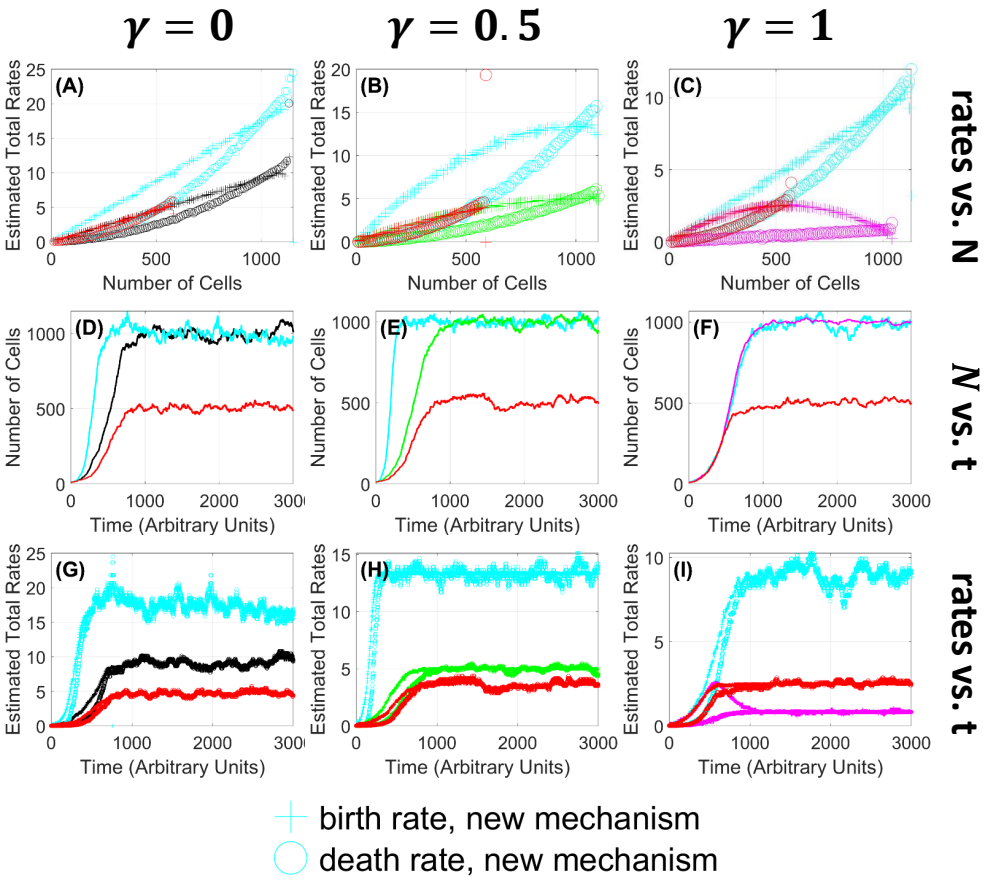

Figure 3.6: **Resolving separate birth vs. death rates distinguishes different underlying mechanisms of resistance to -cidal drugs.** In each of the three density-dependent scenarios (I), (II), (III), a cell population can restore its carrying capacity after a "-cidal" drug treatment via different mechanisms: decreasing death rate to return to the original dynamics (shown in the black, green, magenta curves) or increasing birth rate (shown in the cyan curves). **(A, B, C)**: estimated birth and death rates using an ensemble of 100 cell number trajectories. Plus signs $(+)$ denote estimated birth rates; circles $(\circ)$ denote estimated death rates. **(D, E, F)**: selected cell number trajectories. **(G, H, I)**: estimated birth and death rates corresponding to the cell number trajectories in **(D, E, F)**. **(A, D, G)**: scenario (I) where the density dependence $\gamma = 0$. **(B, E, H)**: scenario (II) where the density dependence $\gamma = 0.5$. **(C, F, I)**: scenario (III) where the density dependence $\gamma = 1$. The red curves represent the case where the cell population has been treated with a "-cidal" drug. The black, green, and magenta curves represent the case where the cell population develops resistance by decreasing its *per capita* death rate and returns to the original dynamics for the scenarios (I), (II), and (III) introduced in Section 3.4.1. The cyan curves represent the case in which the cell population develops resistance by increasing its *per capita* birth rate.

Figure 3.7 illustrates different mechanisms of drug resistance to the "-static" effect described in Section 3.4.2.

- In scenario (I) as shown in Figure 3.7 **(A, B)**, the cell population can become drug resistant either by increasing its *per capita* birth rate back to the original rate:

$$(b_N)_{\text{static, resistant}} = (b_N)_{\text{original}} = b_0, \tag{3.73}$$

$$(d_N)_{\text{static, resistant}} = (d_N)_{\text{static}} = (d_N)_{\text{original}} = d_0 + \frac{r}{K}N, \tag{3.74}$$

  or by decreasing its *per capita* death rate:

$$(d_N)_{\text{static, resistant}} = d_0, \tag{3.75}$$

$$(b_N)_{\text{static, resistant}} = (b_N)_{\text{static}} = b_0 - 0.5\frac{r}{(K/2)}N = \max\left\{b_0 - \frac{r}{K}N, 0\right\}. \tag{3.76}$$

  In the drug resistance mechanism through birth, the density dependence parameter γ changes from 0.5 back to 0, while in the drug resistance mechanism through death, the density dependence parameter γ changes from 0.5 to 1. Both of these scenarios suggest changes in the cell-to-cell interaction modalities. The latter suggests a significant change from full density dependence in death (before drug treatment) to full density dependence in birth (after "-static" drug treatment and resistance).

- In scenario (II) as shown in Figure 3.7 **(C, D)**, the cell population can develop

resistance by either increasing its *per capita* birth rate back to the original rate:

$$(b_N)_{\text{static, resistant}} = (b_N)_{\text{original}} = b_0 - 0.5\frac{r}{K}N, \tag{3.77}$$

$$(d_N)_{\text{static, resistant}} = (d_N)_{\text{static}} = (d_N)_{\text{original}} = d_0 + 0.25\frac{r}{(K/2)}N = d_0 + 0.5\frac{r}{K}N, \tag{3.78}$$

or by decreasing its *per capita* intrinsic death rate to $d_0 - 2r$:

$$(d_N)_{\text{static, resistant}} = d_0 - 2r + 0.5\frac{3r}{K}N, \tag{3.79}$$

$$(b_N)_{\text{static, resistant}} = (b_N)_{\text{static}} = \max\left\{b_0 - 0.5\frac{3r}{K}N, 0\right\}. \tag{3.80}$$

In the drug resistance mechanism through birth, the density dependence parameter γ changes from 0.75 back to 0.5, which suggests a change in the cell-to-cell interaction modality. In the drug resistance mechanism through death, the new *per capita* intrinsic death rate, $d_0 - 2r$, can be negative, which is not biologically meaningful.

- In scenario (III) as shown in Figure 3.7 **(E, F)**, the cell population can develop resistance by either increasing its *per capita* birth rate back to the original rate:

$$(b_N)_{\text{static, resistant}} = (b_N)_{\text{original}} = \max\left\{b_0 - \frac{r}{K}N\right\}, \tag{3.81}$$

$$(d_N)_{\text{static, resistant}} = (d_N)_{\text{static}} = (d_N)_{\text{original}} = d_0, \tag{3.82}$$

or by decreasing its *per capita* intrinsic death rate to $d_0 - r$:

$$(d_N)_{\text{static, resistant}} = d_0 - r, \tag{3.83}$$

$$(b_N)_{\text{static, resistant}} = (b_N)_{\text{static}} = \max\left\{b_0 - \frac{2r}{K}N, 0\right\}. \tag{3.84}$$

In the drug resistance mechanism through birth, the density dependence parameter γ remains 1, which suggests no significant change in the cell-to-cell interaction modality. In the drug resistance mechanism through death, the new *per capita* intrinsic death rate, $d_0 - r$, can be negative, which is not biologically meaningful.

Figure 3.7 shows that having been treated with a "-static" drug, the cell population can develop resistance either by reverting to its original dynamics–the blue curves change back to the black, green, and magenta curves for scenarios (I), (II), and (III) respectively in the figure–or by decreasing its *per capita* death rate as illustrated by the cyan curves. We may call the latter drug resistance mechanism "reduced mortality" or "hypo-death." We note that the decreased *per capita* death rate can become algebraically negative and not biologically meaningful, as seen in Equations (3.79) and (3.83), which is consistent with the fact that drug resistance has previously been considered mainly for "-cidal" drugs, not "-static" drugs, in the literature, cf. [12]. However, in contrast to some recent literature [12], in this chapter, we propose the possibility of mechanisms through which cell populations can overcome the "-static" effect (birth inhibition) of drugs–that is, increasing the *per capita* birth rates back to the original rates, as seen in the black, green, and magenta curves in Figure 3.7. For instance if, through preexisting genetic variation, the cell population contained a mutant with an alternative sequence for the protein by which the drug targets the cell, then as this variant propagated in favor of the principal variant, the cell line could develop resistance to the "-static" drug. It is interesting to observe in Figure 3.7 **(G)** that

even after being with a "-static" drug that inhibits birth, the cell population can develop resistance by reducing birth rates *throughout time*–as we can see the cyan curves are lower than the blue curves as time increases.

We note that for scenario (I) where $\gamma = 0$, we observe a second possible drug resistance mechanism, in which the cell population decreases its *per capita* death rate without making it negative. In this scenario, the cell population also changes its density dependence parameter from $\gamma = 0$ to $\gamma = 1$ as it becomes resistant to the "-static" drug.

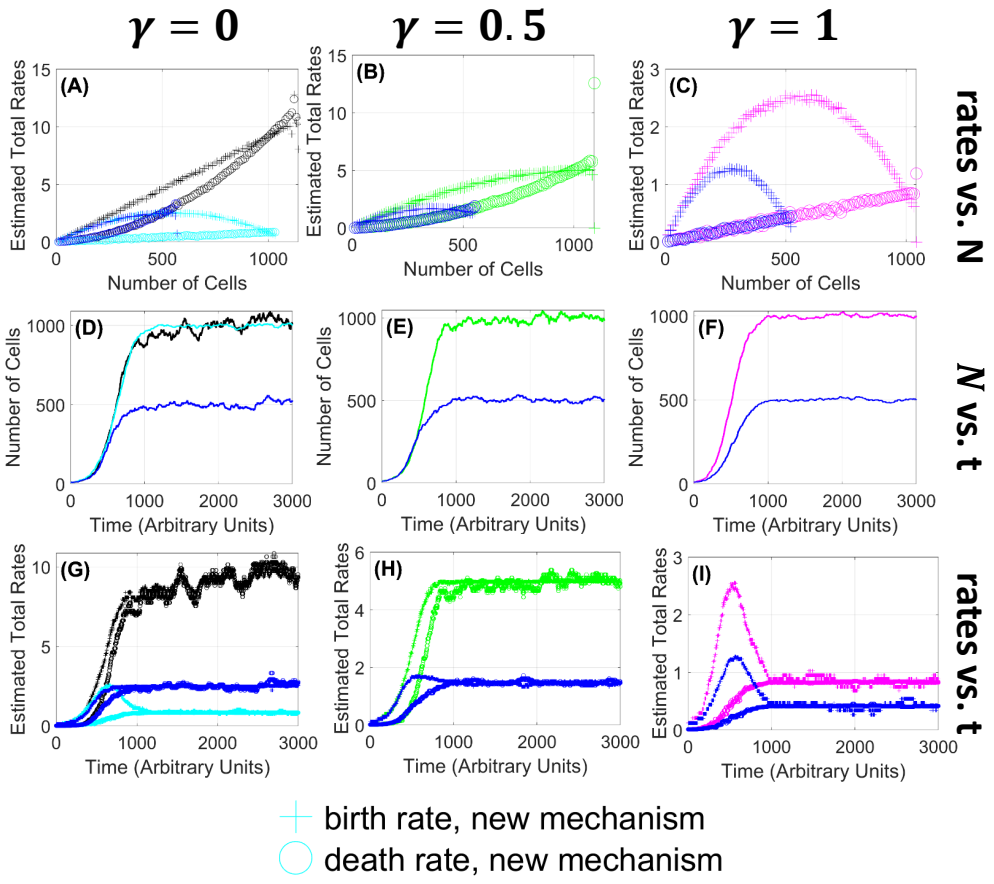

Figure 3.7: **Resolving separate birth vs. death rates distinguishes different underlying mechanisms of resistance to -static drugs.** A cell population can restore its carrying capacity after a "-static" drug treatment via different mechanisms: increasing birth rate to return to the original dynamics (shown in the black, green, magenta curves) or decreasing death rate (shown in the cyan curves). The latter can happen only for scenario (I) where originally, the density dependence is fully in the death rate. **(A, B, C)**: estimated birth and death rates using an ensemble of 100 cell number trajectories. Plus signs $(+)$ denote estimated birth rates; circles $(\circ)$ denote estimated death rates. **(D, E, F)**: selected cell number trajectories. **(G, H, I)**: estimated birth and death rates corresponding to the cell number trajectories in **(D, E, F)**: selected cell number trajectories. **(A, D, G)**: scenario (I) where the density dependence $\gamma = 0$. **(B, E, H)**: scenario (II) where the density dependence $\gamma = 0.5$. **(C, F, I)**: scenario (III) where the density dependence $\gamma = 1$. The blue curves represent the case where the cell population has been treated with a "-static" drug. The black, green, and magenta curves represent the case where the cell population develops resistance by decreasing its *per capita* death rate and returns to the original dynamics for the scenarios (I), (II), and (III) introduced in Section 3.4.1. The cyan curves represent the case in which the cell population develops resistance by increasing its *per capita* birth rate.

3.5 Conclusion and Discussion

In order to infer density-dependent population dynamics mechanisms from data, we separately identify density-dependent *per capita* birth and death rates from net growth rates using the method described in Section 3.3 and infer whether density dependence is manifest in the birth process, death process, or some combination of the two. Our method involves directly estimating the mean and variance of cell number increments, as functions of population size, and expressing birth and death rates in terms of these two statistics. In order to obtain the mean and variance with tolerable accuracy, we compute them from an ensemble of cell number time series (e.g. multiple experiments). We analyze the accuracy of this method and derive analytical expressions for the theoretical expected errors and variance of errors in estimating birth and death rates as functions of the bin size (details are in Appendix A.2). We discover that small bin sizes do not necessarily result in small errors in estimating birth and death rates, due to small sample sizes. In fact, we find that intermediate bin sizes are optimal. Our error analysis also shows that if the intrinsic *per capita* net growth rate r is large relative to the carrying capacity K, then the expected error in estimating the mean cell number increment is high, as shown in Equation (A.33), which suggests that the estimation is not as good for fast-producing cell types.

Our method is distinct from other methods in the literature. It provides a novel perspective on the problem of stochastic parameter identification. Existing methods typically require numerical solution of a high-dimensional optimization problem, e.g. in a Bayesian inverse problem setting [14] or a likelihood function maximization framework. [19] constructed an expectation-maximization algorithm to identify birth and death rates for general birth-death processes. This method enjoys fast convergence and benefits from an elegant formulation of

conditional expectations in terms of convolutions of transition probabilities. Their approach results from solving a maximum likelihood problem. In contrast, we suggest a simple direct estimation approach that accurately extracts birth and death rates from the conditional first and second moments of the cell number time series data. Aside from [19], to the best of our knowledge, other work addressing disambiguation of birth and death rates has been confined to *linear* birth-death processes. For example, [58] used a Bayesian approach to parameter estimation for linear birth-death models in order to quantify the effects of changing drug concentrations. Here, we also consider different drug treatment scenarios, but in the context of nonlinear, logistic population models rather than linear growth models. [28] estimated birth and death rates as functions of time for a continuous-time branching process. Their method applies to multi-type cell populations and is illustrated with density-independent *per capita* birth and death rates. In contrast, our framework encompasses density-dependent *per capita* rates.

Our direct estimation method is a data-hungry approach. As an alternative, for small sample sizes, we also present a maximum likelihood approach in Section 4.2, in which we evaluate the log-likelihood function and maximize it over the density dependence parameter $\gamma \in [0, 1]$. This approach, which involves solving a constrained nonlinear optimization problem, is limited to the assumption that the other system parameters are known.

The significance of both approaches is the application in studying treatments of pathogens and their resistance to the treatments. Specifically, in Section 3.4, we consider the scenario where a homogeneous cell population goes through three stages: (1) grows naturally to its carrying capacity, (2) is treated with a drug that reduces its carrying capacity, and (3) overcomes the drug effect to gain back its carrying capacity. Our method allows us to identify

whether each stage happens through the birth process, death process, or some combination of the two. Our analysis contributes to disambiguating underlying mechanisms such as exploitation vs. interference competition in ecology, bacteriostatic vs. bactericidal antibiotics in clinical treatments, and enhanced fecundity vs. reduced mortality in pathogens' defense against drug treatments, which we may define as drug resistance. The mechanisms shown in this chapter can help explain biological phenomena and may suggest novel approaches for engineering synthetic biological systems. More microscopic mechanisms within the birth process or death process, such as inactivating mutations of the gene for p53 protein [4], are beyond the scope of the model in this chapter.

In Section 3.4.2, we show how to apply our method to distinguish the action of "-static" (birth-inhibiting) versus "-cidal" (death-promoting) drugs. However, the classification of drugs as being "-static" or "-cidal" is complicated by potentially stochastic factors such as external growth conditions [71]. For bacterial infections in a clinical setting, the "-static/-cidal" distinction is defined in terms of drug concentrations–specifically in terms of the ratio between Minimum Inhibitory Concentration (MIC) and Minimum Bactericidal Concentration (MBC). The Minimum Inhibitory Concentration (MIC) is defined as the lowest drug concentration that prevents visible growth, and the Minimum Bactericidal Concentration (MBC) is defined as the lowest drug concentration that results in a 99.99% decrease in the initial population size over a fixed period of time. Bacteriostatic drugs have been defined as those for which the ratio of the MBC to the MIC is larger than 4. Bactericidal drugs are those for which the ratio is ≤ 4 [94]. Including the differential effects of drugs at larger or smaller concentrations will be an interesting direction for expanding our birth/death rate analysis in future work.

In Section 3.4.3, we use our direct estimation method to disambiguate different drug resistance mechanisms. In our chapter, we define "drug resistance" as the cell population's ability to overcome the drug effect and gain back its original carrying capacity. However, the term "drug resistance" is used to mean different things in the research literature. For example, in Davison et al. 2000 [21], drug resistance is defined in terms of the drug concentration needed to inhibit growth or kill the pathogen. Brauner et al. 2016 [12] quantify cell populations overcoming drug effects in terms of MIC and the minimum time needed to kill the pathogens (MDK). Based on these two measures, MIC and MDK, the pathogens' defense against the drug can be called drug tolerance, persistence, or resistance. For future work, we will look into different definitions of "drug resistance".

For the present study, we confine our investigation to simulated data because of several factors. First, generating large ensembles of cell population trajectories is expensive, although high-throughput methods continue to accelerate the pace of data generation [36]. In a typical bioreactor, the data available are optical density time series, rather than direct cell number measurements. In theory, the relation between optical density and cell count is expected to be linear. Unfortunately, that is not always the case. McClure et al. 1993 [63] show that it can be second order and Stephens et al. 1997 [81] show that it can be third order. Moreover, Stevenson et al. 2016 [82] show that the relation between cell count and optical density varies for different cell sizes and shapes, as well as other properties such as the index of refraction of the media. Some experimental calibration techniques have been developed to overcome these discrepancies, such as Francois et al. 2005 [30] and Beal et al. 2020 [6]. Finally, experimental data may include measurement noise that obscures finite population driven density fluctuations. Swain et al. 2016 [85] attempts to estimate net growth rates from optical density data using a Gaussian process framework.

In contrast, we would like disambiguate net growth rate into separate birth and death rates. Extending our method to take into account the mapping from cell number to noisy optical density measurements is an interesting subject for future work.

In the derivation of the relations between birth/death rates and the mean/variance of cell number increments in Section 3.3, we assume that the birth and death processes are approximately independent. However, the birth and death processes cannot be truly independent, as a cell cannot both divide and die at the same time. Suppose we do not assume independence in data, which means our data simulation include the nonzero covariance between the number of cells going through division and the number of cells going through death. The covariance is computed as follows. Let $I_+^{(p)}$ denotes a binary-valued random variable indicating whether the pth cell going through division after a time period Δt and is equal to 1 for yes, and 0 for no. Similarly, let $I_-^{(q)}$ denotes a binary-valued random variable indicating whether the qth cell going through death after a time period Δt and is equal to 1 for yes, and 0 for no. Then,

$$\Delta N_+ = \sum_{p=1}^{N} I_+^{(p)} \quad \text{and} \quad \Delta N_+ = \sum_{q=1}^{N} I_-^{(q)}, \tag{3.85}$$

and

$$\text{Cov}\Big(\Delta N_+|N, \Delta N_-|N\Big) = \text{Cov}\Big(\sum_{p=1}^{N} I_+^{(p)}, \sum_{q=1}^{N} I_-^{(q)}\Big) \tag{3.86}$$

$$= \sum_{p=1}^{N}\sum_{q=1}^{N} \text{Cov}\Big(I_+^{(p)}, I_-^{(q)}\Big) \tag{3.87}$$

$$= \sum_{p=q}\sum_{q=1}^{N} \text{Cov}\Big(I_+^{(p)}, I_-^{(q)}\Big) + \sum_{p=1, p\neq q}^{N}\sum_{q=1}^{N} \text{Cov}\Big(I_+^{(p)}, I_-^{(q)}\Big) \tag{3.88}$$

$$= 0 + \sum_{q=1}^{N} \text{Cov}\Big(I_+^{(q)}, I_-^{(q)}\Big), \text{ as } p \neq q \text{ represents two different cells, so they are independent.} \tag{3.89}$$

$$= \sum_{q=1}^{N} \mathbb{E}\Big[I_+^{(q)} I_-^{(q)}\Big] - \sum_{q=1}^{N} \mathbb{E}\Big[I_+^{(q)}\Big]\mathbb{E}\Big[I_-^{(q)}\Big] \tag{3.90}$$

$$= 0 - \sum_{q=1}^{N} \mathbb{E}\Big[I_+^{(q)}\Big]\mathbb{E}\Big[I_-^{(q)}\Big], \text{ as } I_+^{(q)} I_-^{(q)} = 0, \text{ since a cell cannot divide and die at the same time.} \tag{3.91}$$

$$= -N\Big(b_N \Delta t\Big)\Big(d_N \Delta t\Big) \tag{3.92}$$

$$= -N b_N d_N \Delta t^2, \tag{3.93}$$

which we could include in our data simulation using Gaussian approximation and τ-leaping simulation as follows:

$$X(t+\Delta t) - X(t) = \Big(b_{X(t)} - d_{X(t)}\Big)X(t)\Delta t + \sqrt{\Big(b_{X(t)} + d_{X(t)}\Big)X(t) - \underbrace{b_{X(t)} d_{X(t)} X(t)\Delta t}_{\text{from nonzero covariance}}}\, \Delta W(t). \tag{3.94}$$

We simulate cell number trajectories using Equation (3.95) and estimate birth and death rates with the independence assumption, as in Equations (3.15) with the approximation $\mathbb{V}[\Delta X|X] \approx \Big(b_X + d_X\Big)X\Delta t$. Although the data do not assume independence between the birth and death processes and our estimation does, we see that the estimated and true birth and death rates are well-aligned, showing our method described in Section 3.3 with the independence assumption is robust, as shown in Figure 3.8.

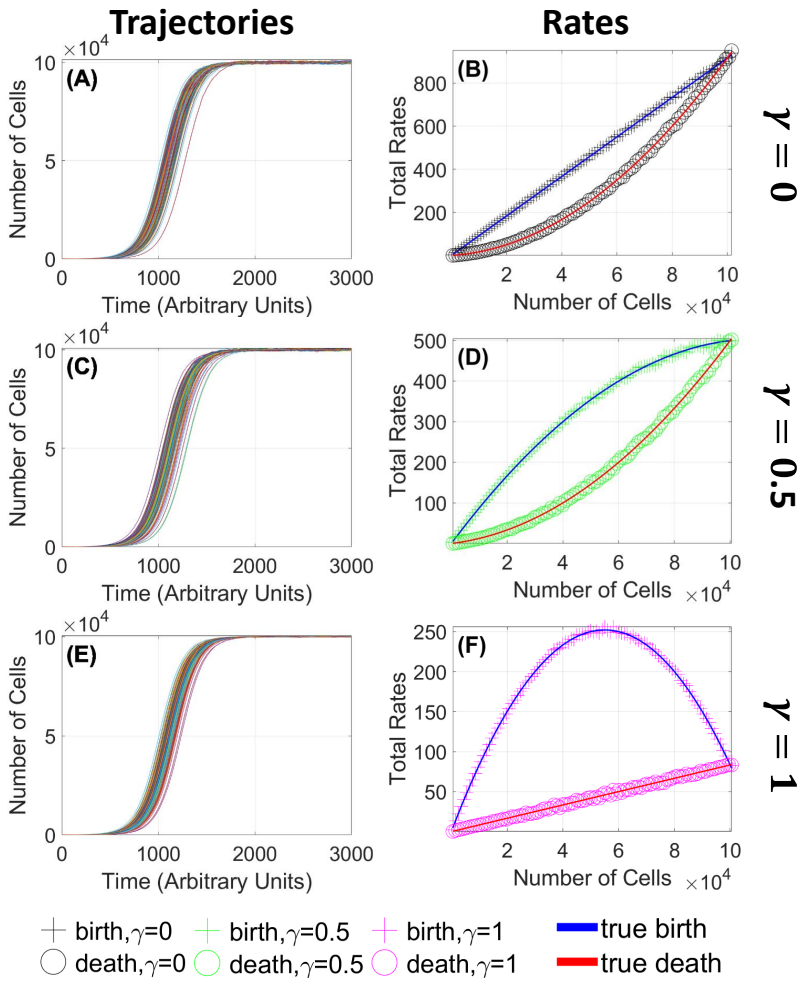

Figure 3.8: **The estimation method works well even without the independence assumption in the simulated data. (A, C, E)**: Time series ensembles simulated using the τ-leaping approximation for the cases $\gamma = 0$ **(A)**, $\gamma = 0.5$ **(C)**, and $\gamma = 1$ **(E)** respectively. In the simulation, we do not assume independence between the birth and death processes, and include the nonzero covariance between the two processes. Each figure shows $S = 100$ trials. The estimated rates are computed using a bin size of $\eta = 10^3$. Carrying capacity $K = 10^5$ cells; low-density growth rate $r = 1/120$ (arbitrary time units); for other parameters see Table A.1. **(B, D, F)**: Estimated and true birth and death rates, as functions of population size. Blue line: true birth rate. Red line: true death rate. Plus signs $(+)$ denote estimated birth rates; circles $(\circ)$ denote estimated death rates. Throughout the we will use distinct colors to denote values of γ. **(B)** Black: $\gamma = 0$; **(D)** Green: $\gamma = 0.5$; **(E)** Magenta: $\gamma = 1.0$. We observe that the estimated birth and death rates are well-aligned with the true birth and death rates used to simulate the trajectories in **(A)**, **(C)**, and **(E)**.

In this thesis, we consider the stochastic dynamics where the waiting time to the next event (i.e. birth or death) is exponentially distributed. However, in real life, this distribution does not have to be exponential and the variation from cell to cell can be significant. For example, [57] considers the case where the generation time (i.e. the time from cell birth to division) follows a general probability distribution, and expresses the long-term net growth rate in terms of the distribution. Similarly, the book by Kimmel and Axelrod [47] provides theories in branching processes to analyze and make inferences when the cell division process includes transitions among multiple stages. As pointed out in [57], many factors can give rise to cell-to-cell variability such as asymmetries in cell division. It is an interesting problem to consider extending our analysis to systems with a richer description of the cell division process. For future work, we would like take the variations between cells into consideration, and infer the dynamics accordingly.

As mentioned in the Introduction (Section 3.1), throughout the chapter, we interpret the density dependence term (interaction between individuals) as competition, which either reduces birth rates or increases death rates. However, in some situations, interactions among individuals can be cooperative, and increase the birth rate or reduce death rate with increasing population size [9]. To address this possibility, in future work one might introduce to a cooperation parameter $c \geq 0$:

$$\frac{d\phi}{dt} = r\phi - \frac{r}{K}\phi^2 = r\phi + \underbrace{c\frac{r}{K}\phi^2}_{\text{cooperation}} \underbrace{-(1+c)\frac{r}{K}\phi^2}_{\text{competition}}. \tag{3.95}$$

One may interpret the cooperation term $c\frac{r}{K}\phi^2$ as a positive interaction between individuals that increases cell population growth. One could parameterize this term with parameter γ_c, to quantify how much of the cooperation increases birth and how much of the coop-

eration decreases death. Similarly, one may interpret the competition term $(1+c)\frac{r}{K}\phi^2$ as a negative interaction between individuals that reduces cell population growth. One could parameterize the competition term with parameter $\gamma_{\sim c}$ to quantify how much of the competition decreases birth and how much of the competition increases death:

$$\frac{d\phi}{dt} = r\phi + \underbrace{\gamma_c c\frac{r}{K}\phi^2}_{\text{cooperation}} + \underbrace{(1-\gamma_c)c\frac{r}{K}\phi^2}_{\text{cooperation}} - \underbrace{\gamma_{\sim c}(1+c)\frac{r}{K}\phi^2}_{\text{competition}} - \underbrace{(1-\gamma_{\sim c})(1+c)\frac{r}{K}\phi^2}_{\text{competition}} \tag{3.96}$$

$$= \underbrace{\left(b_0\phi + \gamma_c c\frac{r}{K}\phi^2 - \gamma_{\sim c}(1+c)\frac{r}{K}\phi^2\right)}_{\text{birth}} \tag{3.97}$$

$$- \underbrace{\left(d_0\phi - (1-\gamma_c)c\frac{r}{K}\phi^2 + (1-\gamma_{\sim c})(1+c)\frac{r}{K}\phi^2\right)}_{\text{death}}. \tag{3.98}$$

Exploring these and other extensions provide interesting directions for future investigation.

Chapter 4

Likelihood-Based Inference: Density-Dependent Dynamics Disambiguation

4.1 Motivation

Separately identifying birth and death rates using the direct estimation method in Chapter 3 requires sufficiently large sample sizes. However, large datasets can be costly in terms of time and resources. In this chapter, we consider alternative approaches for small sample sizes. Most statistical inference methods are for *linear* birth-death-processes, in which the *per capita* birth and death rates do not depend on cell numbers [87], while we are interested in density dependence. Since we want to perform inference in the context of our logistic birth-death process model described in Chapter 3, Section 3.2, we adopt one of the classical methods: the likelihood approach [76]. The key aspects of our model are the *per capita*

birth and death rates defined in Equations (3.6) and (3.7). For convenience, we are going to recall those definitions here:

$$b_N = \max\{b_0 - \gamma \frac{r}{K} N, 0\}, \tag{4.1}$$

$$d_N = d_0 + (1 - \gamma) \frac{r}{K} N. \tag{4.2}$$

Our model is characterized by the parameter set $\Theta := \{b_0, d_0, K, \gamma\}$. Although there are four parameters in our model, we can easily estimate the carrying capacity K, and we can estimate the intrinsic net growth rate $r = b_0 - d_0$ using methods such as [85]. Therefore, we just need to estimate γ and either b_0 or d_0. In Section 4.2, we assume that we know b_0, d_0, K, and solve a one-dimensional maximum likelihood problem with γ being the variable. Moving from one-dimension to two-dimension introduces computational complexity. Moreover, even for the one-dimensional problem, it is hard to theoretically verify that the numerical solution is indeed a maximizer. In particular, as shown in Section 4.2, we can only verify numerically that the second derivative of the log-likelihood at the numerical solution is negative; it remains very difficult to prove the second-order optimality condition in the context of our problem. Therefore, instead of solving a two-dimensional optimization problem for γ and b_0 (or d_0), we take a different perspective. We compute the mean and variance of the log-likelihood, and we use these two statistics to answer the "more likely" question. By "more likely", we mean *For which scenario is the log-likelihood higher?* We are hoping that $\mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_1) > \mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_2)$, from which we can infer that scenario Θ_1 is "more likely" than scenario Θ_2. Since $\mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_1)$ and $\mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_2)$ are random variables, we need to assess the following probability:

$$P\left(\mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_1) > \mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_2)\right), \tag{4.3}$$

Define $Z_1 := \mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_1) - \mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_2)$. Then,

$$P\Bigg(\mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_1) > \mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_2)\Bigg) = P\Bigg(Z_{\text{data}} > 0\Bigg), \tag{4.4}$$

and

$$\mathbb{E}[Z_1] = \mathbb{E}\Big[\mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_1)\Big] - \mathbb{E}\Big[\mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_2)\Big] \tag{4.5}$$

$$\mathbb{V}[Z_1] = \mathbb{V}\Big[\mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_1)\Big] + \mathbb{V}\Big[\mathcal{L}(\boldsymbol{X}_{\Theta_1}|\Theta_2)\Big]. \tag{4.6}$$

We may then use $\mathbb{E}[Z_1]$ and $\mathbb{V}[Z_1]$ to compute $P\Bigg(Z_{\text{data}} > 0\Bigg)$. Computing the mean and variance of the log-likelihood in Section 4.4 and Section 4.5 requires the probability of the state in our model (i.e. number of cells), which we compute in Section 4.3.

4.2 Maximization over Density Dependence Parameter

In the instance that there is only a single cell number trajectory, we would like to infer the density dependence scenario from which the data are generated. This question leads us to consider a maximum likelihood approach.

Let a cell number time series $\boldsymbol{X}_{\text{data}} = [x_0, x_1, \ldots, x_T]$ be a realization for the normally distributed random variable $\boldsymbol{X} = [X(t_0), X(t_1), \ldots, X(t_T)]$, which approximates the discrete random variable $\boldsymbol{N} = [N(t_0), N(t_1), \ldots, N(t_T)]$ as discussed in Section 3.3.2. For clarity, we denote x_j as $x_{j,\text{data}}, j = 0, \ldots, T$. Recall that $X(t)$ follow a Gaussian birth-death process[1] characterized by the parameter set $\Theta_{\text{data}} = \{b_{0,\text{data}}, d_{0,\text{data}}, \gamma_{\text{data}}, K_{\text{data}}\}$. These

[1]This approximation requires sufficiently large summed birth and death rates.

parameters determine the birth and death rates of the birth-death process from which the time series is generated. In particular, the *per capita* birth and death rates are defined as follows:

$$b_{X_{j,\text{data}}} = \max\left\{b_{0,\text{data}} - \gamma_{\text{data}}\frac{r_{\text{data}}}{K_{\text{data}}}X_{j,\text{data}},\ 0\right\}, \tag{4.7}$$

$$d_{X_{j,\text{data}}} = d_{0,\text{data}} + (1-\gamma_{\text{data}})\frac{r_{\text{data}}}{K_{\text{data}}}X_{j,\text{data}}, \tag{4.8}$$

where $r_{\text{data}} = b_{0,\text{data}} - d_{0,\text{data}}$. To test whether a given time series $\boldsymbol{X}_{\text{data}}$ belongs to a scenario characterized by the parameter set $\Theta_{\text{test}} = \{b_{0,\text{test}}, d_{0,\text{test}}, \gamma_{\text{test}}, K_{\text{test}}\}$, we evaluate the log likelihood function at the time series:

$$\mathcal{L}(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}}) = \underbrace{\ln\left(P(x_{0,\text{data}})\right)}_{=\ln(1)} + \sum_{j=1}^{T-1}\ln\left(P(x_{j+1,\text{data}}|x_{j,\text{data}},\Theta_{\text{test}})\right) \tag{4.9}$$

$$= \sum_{j=1}^{T-1}\frac{1}{2}\ln\left(\frac{1}{2\pi x_{j,\text{data}}(b_{\text{test},x_{j,\text{data}}} + d_{\text{test},x_{j,\text{data}}})\Delta t}\right) \tag{4.10}$$

$$-\frac{1}{2}\frac{\left(x_{j+1,\text{data}} - x_{j,\text{data}} - x_{j,\text{data}}(b_{\text{test},x_{j,\text{data}}} - d_{\text{test},x_{j,\text{data}}})\Delta t\right)^2}{x_{j,\text{data}}(b_{\text{test},x_{j,\text{data}}} + d_{\text{test},x_{j,\text{data}}})\Delta t}, \tag{4.11}$$

where

$$b_{\text{test},x_{j,\text{data}}} = \max\left\{b_{0,\text{test}} - \gamma_{\text{test}}\frac{r_{\text{test}}}{K_{\text{test}}}x_{j,\text{data}},\ 0\right\}, \tag{4.12}$$

$$d_{\text{test},x_{j,\text{data}}} = d_{0,\text{test}} + (1-\gamma_{\text{test}})\frac{r_{\text{test}}}{K_{\text{test}}}x_{j,\text{data}}, \tag{4.13}$$

$$r_{\text{test}} = b_{0,\text{test}} - d_{0,\text{test}}. \tag{4.14}$$

Suppose we know $b_{0,\text{data}}, d_{0,\text{data}}$, and K_{data}. That is, suppose that $b_{0,\text{test}} = b_{0,\text{data}}$, $d_{0,\text{test}} = d_{0,\text{data}}$, and $K_{\text{test}} = K_{\text{data}}$. Given one cell number time series, to infer which density

dependence scenario the data mostly likely belongs, we treat the log-likelihood function as a function f of $\gamma := \gamma_{\text{test}}$, and find $\gamma \in [0, 1]$ that maximizes $f(\gamma)$. We thus formulate a constrained nonlinear optimization problem as follows:

$$\max_{\gamma} f(\gamma) = \mathcal{L}(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}}) \quad \text{subject to} \quad 0 \leq \gamma \leq 1. \tag{4.15}$$

For shorter notation, here we denote $b_{0,\text{test}}$ and $b_{0,\text{data}}$ as b_0, $d_{0,\text{test}}$ and $d_{0,\text{data}}$ as d_0, K_{test} and K_{data} as K, and $x_{j,\text{data}}$ as x_j. We calculate the first derivative $df/d\gamma$ and find critical points by solving $\dfrac{df}{d\gamma} = 0, \gamma \in [0, 1]$ numerically using the Bisection method on the interval $[0 - \Delta\gamma, 1 + \Delta\gamma], \Delta\gamma = 0.5 > 0$. We use the interval $[-0.5, 1.5]$, which is wider than the domain of γ, so that we can find the end points $\gamma = 0$ and $\gamma = 1$ using the Bisection method.

Given multiple samples of cell number time series (e.g. from multiple experimental trials), we obtain an empirical distribution of solutions γ to the optimization problem (4.15). In Figure 4.1, for each of the three scenarios (I) $\gamma_{\text{data}} = 0$, (II) $\gamma_{\text{data}} = 0.5$, and (III) $\gamma_{\text{data}} = 1$, we plot the results upon solving the optimization problem 100 times for 100 independent time series, and obtain a distribution of estimated γ parameters. In addition we obtain a distribution of the estimation error, which we define as the absolute difference $(\gamma_{\text{data}} - \gamma_{\text{estimated}})$, where $\gamma_{\text{estimated}}$ is the numerical solution to the optimization problem (4.15). The values of the parameters b_0, d_0, and K used in data simulation are the same as in Table A.1. The empirical means and variances of the estimated γ values and estimation errors for the three scenarios (I), (II), and (III) are listed in Table 4.1.

Table 4.1: Numerical solution to the optimization problem (4.15) for a time series of length $T = 90,000$ points (timestep $\Delta t = 1/30$, total time 3000 arbitrary units).

True γ Value	Mean Estimated γ	Variance of Estimated γ	Mean Error	Error Variance
0	0.0010	1.4858×10^{-4}	-0.0010	1.4858×10^{-4}
0.5	0.4996	8.4605×10^{-5}	2.7218×10^{-5}	8.4605×10^{-5}
1	1.0000	3.6839×10^{-5}	-2.8675×10^{-5}	3.6839×10^{-5}

We note that the mean values of γ for the three scenarios (I), (II), and (III) are separated by margins that are an order of magnitude larger than the standard errors of the estimates. Thus, for the data generated by our birth/death simulations, the distribution the density-dependent effects can clearly be distinguished in terms of fully a birth-rate effect, fully a death-rate effect, or an evenly mixed effect.

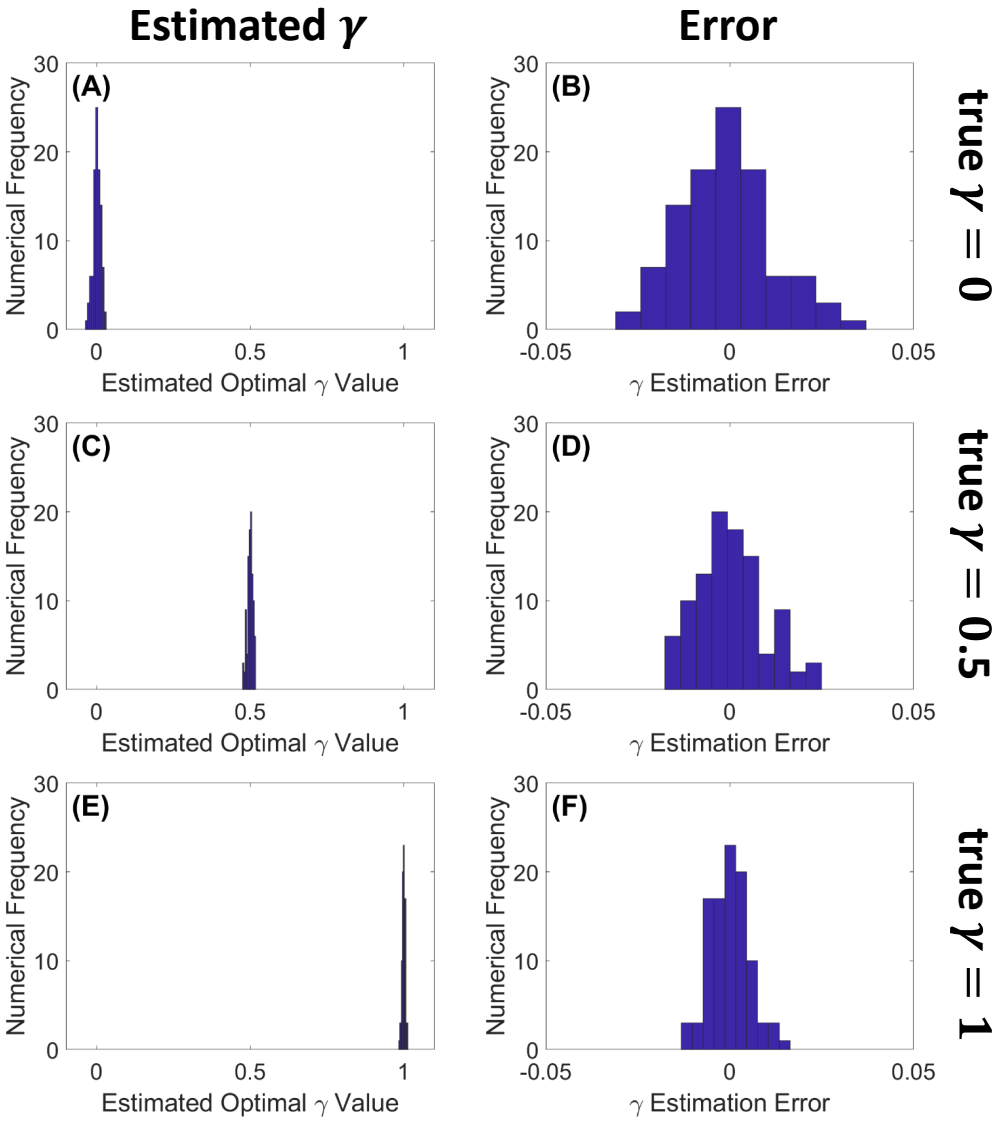

Figure 4.1: **Numerical solutions to the optimization problem for different density dependence scenarios are clearly separated.** We plot empirical distributions of estimated γ values, and corresponding errors, for an ensemble of individual cell number time series. **(A, B)**: cell number time series simulated with $\gamma = 0$; **(C, D)**: cell number time series simulated with $\gamma = 0.5$; **(E, F)**: cell number time series simulated with $\gamma = 1$. **(A, C, E)**: distributions of estimated γ; **(B, D, F)**: distributions of the corresponding estimation errors, defined as $(\gamma_{\text{data}} - \gamma_{\text{estimated}})$. We observe that the distributions of the estimated γ have small variances, and the errors are approximately normally distributed. Here we used time series of length $T = 90,000$ points (timestep $\Delta t = 1/30$, total time 3000).

We calculate the first and second derivatives of the log-likelihood function $f(\gamma)$ (4.11) for a single trajectory as a function of the density-dependence parameter γ. Let $\Delta x_j = x_{j+1} - x_j$.

$$f(\gamma) = \sum_{j=1}^{T-1} \frac{1}{2} \ln\left(\frac{1}{2\pi \mathbb{V}[\Delta x_j]}\right) - \frac{1}{2} \frac{\Big(\Delta x_j - \mathbb{E}[\Delta x_j]\Big)^2}{\mathbb{V}[\Delta x_j]} \tag{4.16}$$

$$= \sum_{j=1}^{T-1} -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln\Big(\mathbb{V}[\Delta x_j]\Big) - \frac{1}{2} \frac{\Big(\Delta x_j - \mathbb{E}[\Delta x_j]\Big)^2}{\mathbb{V}[\Delta x_j]}, \tag{4.17}$$

where

$$\mathbb{E}[\Delta x_j] = x_j \Delta t (b_{x_j} - d_{x_j}) = x_j \Delta t \tag{4.18}$$

$$\mathbb{V}[\Delta x_j] = x_j \Delta t (b_{x_j} + d_{x_j}) \tag{4.19}$$

We observe that $\mathbb{E}[\Delta x_j]$ is a piecewise linear function of γ, i.e. $\mathbb{E}[\Delta x_j]$ has the form $c_1^j + c_2^j \gamma$, where

$$c_1^j = \begin{cases} x_j \Delta t \Big(b_0 - d_0 - \dfrac{r}{K} x_j\Big), & \text{for } \Big(b_0 - \gamma \dfrac{r}{K} x_j\Big) > 0 \\ -x_j \Delta t \Big(d_0 + \dfrac{r}{K} x_j\Big), & \text{for } \Big(b_0 - \gamma \dfrac{r}{K} x_j^2\Big) = 0 \end{cases} \tag{4.20}$$

and

$$c_2^j = \begin{cases} 0, & \text{for } \Big(b_0 - \gamma \dfrac{r}{K} x_j\Big) > 0 \\ \Delta t \dfrac{r}{K} x_j, & \text{for } \Big(b_0 - \gamma \dfrac{r}{K} x_j^2\Big) = 0 \end{cases} \tag{4.21}$$

The variance $\mathbb{V}[\Delta x_j]$ is also a linear function of γ, i.e. $\mathbb{V}[\Delta x_j]$ has the form $c_3^j - c_4^j\gamma$ with

$$c_3^j = \begin{cases} x_j \Delta t\Big(b_0 + d_0 + \dfrac{r}{K}x_j\Big), & \text{for } \Big(b_0 - \gamma\dfrac{r}{K}x_j\Big) > 0 \\ x_j \Delta t\Big(d_0 + \dfrac{r}{K}x_j\Big), & \text{for } \Big(b_0 - \gamma\dfrac{r}{K}x_j^2\Big) = 0 \end{cases} \tag{4.22}$$

and

$$c_4^j = \begin{cases} 2\Delta t\dfrac{r}{K}x_j^2, & \text{for } \Big(b_0 - \gamma\dfrac{r}{K}x_j\Big) > 0 \\ \Delta t\dfrac{r}{K}x_j^2, & \text{for } \Big(b_0 - \gamma\dfrac{r}{K}x_j^2\Big) = 0 \end{cases} \tag{4.23}$$

Therefore,

$$f(\gamma) = \sum_{j=1}^{T-1} -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\Big(c_3^j - c_4^j\gamma\Big) - \frac{1}{2}\frac{\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)^2}{c_3^j - c_4^j\gamma} \tag{4.24}$$

Denote $v_j = \dfrac{1}{c_3^j - c_4^j\gamma} \Rightarrow v_j > 0$ and $\dfrac{dv_j}{d\gamma} = \dfrac{c_4^j}{(c_3^j - c_4^j\gamma)^2} = c_4^j v_j^2$. We have

$$f(\gamma) = \sum_{j=1}^{T-1} -\frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln(v_j) - \frac{1}{2}\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)^2 v_j \tag{4.25}$$

If $b_0 - \gamma(r/K)x_j > 0$, then $c_2^j = 0$ and $\mathbb{E}[\Delta x_j] = c_1^j$ is independent of γ. Hence,

$$\frac{df}{d\gamma} = \sum_{j=1}^{T-1} \frac{1}{2}\frac{1}{v_j}\frac{dv_j}{d\gamma} - \frac{1}{2}\Big(\Delta x_j - c_1^j\Big)^2 \frac{dv_j}{d\gamma} = \sum_{j=1}^{T-1} \frac{1}{2}c_4^j v_j - \frac{1}{2}\Big(\Delta x_j - c_1^j\Big)^2 c_4^j v_j^2 \quad (4.26)$$

$$\Rightarrow \frac{d^2 f}{d\gamma^2} = \sum_{j=1}^{T-1} \frac{1}{2}c_4^j \frac{dv_j}{d\gamma} - \frac{1}{2}\Big(\Delta x_j - c_1^j\Big)^2 c_4^j 2v_j \frac{dv_j}{d\gamma} \quad (4.27)$$

$$= \sum_{j=1}^{T-1} \frac{1}{2}(c_4^j)^2 v_j^2 - \Big(\Delta x_j - c_1^j\Big)^2 (c_4^j)^2 v_j^3 \quad (4.28)$$

$$= \sum_{j=1}^{T-1} (c_4^j)^2 v_j^2 \left(\frac{1}{2} - \Big(\Delta x_j - c_1^j)\Big)^2 v_j\right) \quad (4.29)$$

$$= \sum_{j=1}^{T-1} (c_4^j)^2 v_j^2 \left(\frac{1}{2} - \frac{\big(\Delta x_j - \mathbb{E}[\Delta x_j]\big)^2}{\mathbb{V}[\Delta x_j]}\right) \quad (4.30)$$

In general,

$$\frac{df}{d\gamma} = \sum_{j=1}^{T-1} \frac{1}{2}\frac{1}{v_j}\frac{dv_j}{d\gamma} - \frac{1}{2}\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)^2 \frac{dv_j}{d\gamma} + c_2^j\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)v_j \quad (4.31)$$

$$= \sum_{j=1}^{T-1} \frac{1}{2}c_4^j v_j - \frac{1}{2}\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)^2 c_4^j v_j^2 + c_2^j\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)v_j \quad (4.32)$$

$$\Rightarrow \frac{d^2 f}{d\gamma^2} = \sum_{j=1}^{T-1} \frac{1}{2}c_4^j \frac{dv_j}{d\gamma} + \frac{1}{2}2c_2^j\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)c_4^j v_j^2 - \frac{1}{2}\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)^2 c_4^j 2v_j \frac{dv_j}{d\gamma} \quad (4.33)$$

$$+ \sum_{j=1}^{T-1} c_2^j\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)\frac{dv_j}{d\gamma} \quad (4.34)$$

$$= \sum_{j=1}^{T-1} \frac{1}{2}(c_4^j)^2 v_j^2 + 2c_2^j c_4^j\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)v_j^2 - \Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)^2 (c_4^j)^2 v_j^3 \quad (4.35)$$

$$= \sum_{j=1}^{T-1} c_4^j v_j^2 \left(\frac{1}{2}c_4^j + 2c_2^j\Big(\Delta x_j - c_1^j - c_2^j\gamma\Big) - \Big(\Delta x_j - c_1^j - c_2^j\gamma\Big)^2 c_4^j v_j\right) \quad (4.36)$$

Figure 4.2 shows the frequency $\frac{d^2 f}{d\gamma^2}$ evaluated at the numerical root of $\frac{df}{d\gamma}$ on $[0, 1]$ is negative among 100 times of solving the optimization problem (4.15). We observe that the second derivatives are negative for all of the cases, which implies that the numerical root is reasonably a maximum.

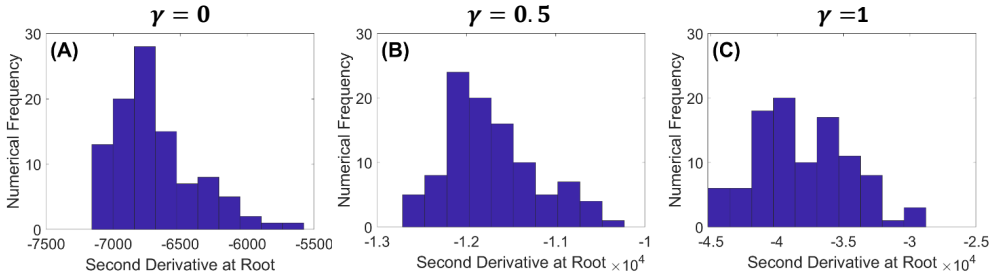

Figure 4.2: **The second derivative of the log likelihood is confirmed empirically to be negative at the numerical solution of the optimization problem.** We plot empirical distributions of the second derivative of the log-likelihood function $f(\gamma)$ evaluated at the numerical root on $[0, 1]$ of the first derivative of $f(\gamma)$ for 100 cell number trajectories. **(A), (B), (C)** correspond to three different scenarios of density dependence $\gamma = 0$, $\gamma = 0.5$, $\gamma = 1$ respectively. The distributions are obtained from maximizing the log-likelihood function $f(\gamma)$ 100 times for each of the three γ scenarios. We observe that the second derivatives are negative for all of the cases.

We explicitly calculate the first derivative of $f(\gamma)$ below to find critical points:

$$\frac{df}{d\gamma} = \sum_{j=1}^{T-1} -\frac{1}{2}\frac{1}{x_j(b_{x_j}+d_{x_j})}x_j\frac{d}{d\gamma}(b_{x_j}+d_{x_j}) \tag{4.37}$$

$$-\sum_{j=1}^{T-1}\frac{1}{2}2\Big(\Delta x_j - x_j(b_{x_j}-d_{x_j})\Delta t\Big)\Delta t(-x_j)\frac{d}{d\gamma}(b_{x_j}-d_{x_j})\frac{1}{x_j(b_{x_j}+d_{x_j})}\frac{1}{\Delta t}$$

(4.38)

$$-\sum_{j=1}^{T-1}\frac{1}{2}(\Delta x_j - x_j(b_{X_j}-d_{x_j})\Delta t)^2\frac{-1}{\Delta t x_j^2(b_{x_j}+d_{x_j})^2}x_j\frac{d}{d\gamma}(b_{x_j}+d_{x_j}) \tag{4.39}$$

$$=\sum_{j=1}^{T-1} -\frac{1}{2}\frac{1}{(b_{x_j}+d_{x_j})}\frac{d}{d\gamma}(b_{x_j}+d_{x_j}) \tag{4.40}$$

$$+\sum_{j=1}^{T-1}\Big(\Delta x_j - x_j(b_{x_j}-d_{x_j})\Delta t\Big)\frac{1}{(b_{x_j}+d_{x_j})}\frac{d}{d\gamma}(b_{x_j}-d_{x_j}) \tag{4.41}$$

$$+\sum_{j=1}^{T-1}\frac{1}{2}(\Delta x_j - x_j(b_{X_j}-d_{x_j})\Delta t)^2\frac{1}{\Delta t x_j(b_{x_j}+d_{x_j})^2}\frac{d}{d\gamma}(b_{x_j}+d_{x_j}), \tag{4.42}$$

where

$$b_{x_j} + d_{x_j} = \frac{b_0 - \gamma(r/K)x_j + \left|b_0 - \gamma(r/K)x_j\right|}{2} + d_0 + (1-\gamma)(r/K)x_j \tag{4.43}$$

$$\Rightarrow \frac{d}{d\gamma}\Big(b_{x_j} + d_{x_j}\Big) = -\frac{(r/K)x_j}{2} - (r/K)x_j - \frac{1}{2}(r/K)x_j\frac{\left|b_0 - \gamma(r/K)x_j\right|}{b_0 - \gamma(r/K)x_j} \tag{4.44}$$

$$= -\frac{3}{2}(r/K)x_j - \frac{1}{2}(r/K)x_j\frac{\left|b_0 - \gamma(r/K)x_j\right|}{b_0 - \gamma(r/K)x_j} \tag{4.45}$$

and

$$b_{x_j} - d_{x_j} = \frac{b_0 - \gamma(r/K)x_j + \left|b_0 - \gamma(r/K)x_j\right|}{2} - d_0 - (1-\gamma)(r/K)x_j \tag{4.46}$$

$$\Rightarrow \frac{d}{d\gamma}\Big(b_{x_j} - d_{x_j}\Big) = -\frac{(r/K)x_j}{2} + (r/K)x_j - \frac{1}{2}(r/K)x_j\frac{\left|b_0 - \gamma(r/K)x_j\right|}{b_0 - \gamma(r/K)x_j} \tag{4.47}$$

$$= \frac{1}{2}(r/K)x_j - \frac{1}{2}(r/K)x_j\frac{\left|b_0 - \gamma(r/K)x_j\right|}{b_0 - \gamma(r/K)x_j}. \tag{4.48}$$

4.3 Probability Distribution of the State in the Logistic Birth-Death Process Model

In this section we develop a semianalytic Gaussian approximation to the probability distribution for the population size as it evolves over time.

Let $\phi(t)$ denote the deterministic cell number at time t, normalized to the carrying capacity K; that is, $K\phi(t)$ represents the total population. Then $\phi(t)$ satisfies the following logistic differential equation:

$$\frac{d(K\phi)}{dt} = (K\phi)r\Big(1 - \frac{K\phi}{K}\Big) = \underbrace{b_{K\phi}K\phi}_{\text{birth}} - \underbrace{d_{K\phi}K\phi}_{\text{death}} \tag{4.49}$$

$$\Rightarrow \frac{d\phi}{dt} = r\phi(1-\phi) = b_{K\phi}\phi - d_{K\phi}\phi, \tag{4.50}$$

with a predetermined initial condition $\phi(0) = \phi_0$ and $0 \leq \phi \leq 1$. As in Section 3.2, the parameter r represents the intrinsic, low-density net growth rate, and the density-dependent *per capita* birth and death rates for population size n are b_n and death rate d_n defined by Equations (3.6) and (3.7). However, for our analysis in this section, we drop the $\max$ function in the birth rate, since the birth rate is positive in all of our data simulations. In Equation (4.50), $n = K\phi$. Solving Equation (4.50) with a fixed (deterministic) initial condition $\phi(0) = \phi_0$, we obtain an explicit expression for $\phi(t)$:

$$\phi(t) = \frac{1}{1 + Ae^{-rt}}, \text{ with } A = \frac{1-\phi_0}{\phi_0}. \tag{4.51}$$

We represent the stochastic total population size $X(t)$ in our logistic birth-death process model as the deterministic solution plus a (stochastic) variation, using the Ansatz

$$X(t) = K\phi(t) + \sqrt{K}U(t), \tag{4.52}$$

where $\mathbb{E}[U(t)] = \mathcal{O}(1)$. Scaling out the carrying capacity we have

$$\frac{X(t)}{K} = \phi(t) + \varepsilon U(t), \tag{4.53}$$

where we introduce the small parameter $\varepsilon := K^{-1/2}$. As discussed in Section 3.3.2 we use a Langevin equation to approximate trajectories of the total population, whence $U(t)$ and $X(t)$ are Gaussian stochastic processes. To be consistent with the notations used in the log-likelihood function, we define $X_j := X(t_j)$ to mean the continuous random variable representing the number of cells at time t_j, and denote x_j as a realization of the random variable X_j. We adopt a finite time step Δt such that $t_j = t_0 + j\Delta t$. Similarly, we define $U_j := U(t_j)$, with realization u_j. We have:

$$\mathbb{E}[X_j] = K\phi_j + \varepsilon K\mathbb{E}[U_j] = K\phi_j + \sqrt{K}\,\mathbb{E}[U_j], \tag{4.54}$$

$$\mathbb{V}[X_j] = \varepsilon^2 K^2 \mathbb{V}[U_j] = K\mathbb{V}[U_j]. \tag{4.55}$$

To find $\mathbb{E}[X_j]$ and $\mathbb{V}[X_j]$, it suffices to find $\mathbb{E}[U_j]$ and $\mathbb{V}[U_j]$. Using the Ansatz (4.52), we have

$$dX(t) = Kd\phi(t) + \varepsilon K\, dU(t) \tag{4.56}$$

$$= \Big(b_{K\phi(t)} - d_{K\phi(t)}\Big) K\phi(t)\, dt + \varepsilon K\, dU(t). \tag{4.57}$$

At the same time, from Equation (3.21) we have

$$dX(t) = \Big(b_{X(t)} - d_{X(t)}\Big)X(t)dt + \sqrt{\Big(b_{X(t)} + d_{X(t)}\Big)X(t)}\, dW(t) \tag{4.58}$$

$$= \Big(b_{K\phi(t)+\varepsilon KU(t)} - d_{K\phi(t)+\varepsilon KU(t)}\Big)\Big(K\phi(t) + \varepsilon KU(t)\Big)dt \tag{4.59}$$

$$+ \sqrt{\Big(b_{K\phi(t)+\varepsilon KU(t)} + d_{K\phi(t)+\varepsilon KU(t)}\Big)\Big(K\phi(t) + \varepsilon KU(t)\Big)}\, dW(t),$$

since $X(t) = K\phi(t) + \varepsilon KU(t)$. Noting that $b_x = b_0 - \gamma\frac{r}{K}x$ and $d_x = d_0 + (1-\gamma)\frac{r}{K}x$ are linear in x, we have

$$b_{K\phi+\varepsilon KU} = b_0 - \gamma\frac{r}{K}\Big(K\phi + \varepsilon KU\Big) = b_0 - \gamma\frac{r}{K\phi}K\phi - \gamma\varepsilon U = b_{K\phi} - \gamma\varepsilon rU \tag{4.60}$$

$$d_{K\phi+\varepsilon KU} = d_0 + (1-\gamma)\frac{r}{K}\Big(K\phi + \varepsilon KU\Big) = d_{K\phi} + (1-\gamma)\varepsilon rU, \tag{4.61}$$

which implies

$$b_{K\phi+\varepsilon KU} - d_{K\phi+\varepsilon KU} = b_{K\phi} - d_{K\phi} - \varepsilon rU, \tag{4.62}$$

$$b_{K\phi+\varepsilon KU} + d_{K\phi+\varepsilon KU} = b_{K\phi} + d_{K\phi} + (1-2\gamma)\varepsilon rU. \tag{4.63}$$

Substituting these expressions into Equation (4.59) gives

$$
\begin{aligned}
& dX && (4.64)\\
=&\Big(b_{K\phi}-d_{K\phi}-\varepsilon rU\Big)\Big(K\phi+\varepsilon KU\Big)dt && (4.65)\\
&+\sqrt{\Big(b_{K\phi}+d_{K\phi}+(1-2\gamma)\varepsilon rU\Big)\Big(K\phi+\varepsilon KU\Big)}\,dW\\
=&\Big(b_{K\phi}-d_{K\phi}\Big)K\phi dt+\Big(b_{K\phi}-d_{K\phi}\Big)\varepsilon KUdt-\varepsilon K\phi rUdt+\underbrace{\varepsilon K\Big(\varepsilon U\Big)rUdt}_{\text{drop}} && (4.66)\\
&+\sqrt{K}\sqrt{\Big(b_{K\phi}+d_{K\phi}\Big)\phi+\underbrace{\Big(b_{K\phi}+d_{K\phi}\Big)\varepsilon U+(1-2\gamma)r\varepsilon U\Big(\phi+\varepsilon U\Big)}_{\text{drop}}}\,dW\\
\approx&\Big(b_{K\phi}-d_{K\phi}\Big)K\phi dt+\varepsilon K\Big(b_{K\phi}-d_{K\phi}\Big)Udt-\varepsilon Kr\phi Udt && (4.67)\\
&+\varepsilon K\sqrt{\Big(b_{K\phi}+d_{K\phi}\Big)\phi}\,dW+\mathcal{O}(\sqrt[4]{K}).
\end{aligned}
$$

Note that we drop the term $\varepsilon K\Big(\varepsilon U\Big)rUdt=r\Big(U\Big)^2dt$ because it is small relatively to the terms $\Big(b_{K\phi}-d_{K\phi}\Big)K\phi dt$ and $\Big(b_{K\phi}-d_{K\phi}\Big)\varepsilon KUdt$ in the same equation. Specifically, the "small" term is $\mathcal{O}(1)$ while the other terms are $\mathcal{O}(K)$ or $\mathcal{O}(\sqrt{K})$. Similarly, we also drop the term $\Big(b_{K\phi}+d_{K\phi}\Big)\varepsilon U+(1-2\gamma)r\varepsilon U\Big(\phi+\varepsilon U\Big)$ in the square root, because it is of smaller order than the leading term under the radical sign.

Setting Equation (4.57) equal to Equation (4.67), and neglecting terms of order $\mathcal{O}(\sqrt[4]{K})$

and smaller, we require

$$dU = \Big(b_{K\phi} - d_{K\phi} - r\phi\Big)U dt + \sqrt{\Big(b_{K\phi} + d_{K\phi}\Big)\phi}\, dW \tag{4.68}$$

$$= \Big(b_0 - d_0 - \frac{r}{K}K\phi - r\phi\Big)U dt + \sqrt{\Big(b_0 + d_0 + (1-2\gamma)r\phi\Big)\phi}\, dW \tag{4.69}$$

$$= \Big(r - r\phi - r\phi\Big)U dt + \sqrt{\Big(b_0 + d_0 + (1-2\gamma)r\phi\Big)\phi}\, dW \tag{4.70}$$

$$= \underbrace{r\Big(1-2\phi\Big)}_{\text{deterministic}} U dt + \sqrt{\underbrace{\Big(b_0 + d_0 + (1-2\gamma)r\phi\Big)}_{\text{deterministic}}\phi}\, dW. \tag{4.71}$$

From Equation (4.71), we have

$$d\mathbb{E}[U] = r\Big(1-2\phi\Big)\mathbb{E}[U]dt + \sqrt{\Big(b_0 + d_0 + (1-2\gamma)r\phi\Big)\phi}\,\underbrace{\mathbb{E}[dW]}_{=0} \tag{4.72}$$

$$\Rightarrow \frac{d\mathbb{E}[U]}{dt} = r\Big(1-2\phi\Big)\mathbb{E}[U]. \tag{4.73}$$

We solve equation (4.73) using integration by parts. The integrating factor is

$$I(t) = e^{\left(\int r\Big(2\phi(t)-1\Big)dt\right)} = e^{\left(2r\int \frac{1}{1+Ae^{-rt}}dt - \int r dt\right)} \tag{4.74}$$

$$= e^{\left(2r\frac{1}{r}\int \frac{1}{z(1-z)}du - rt + C_1\right)}, \quad \text{with } z = 1 + Ae^{-rt} \tag{4.75}$$

$$= e^{\left(2\int \frac{1}{z} - \frac{1}{z-1}dz - rt + C_1\right)} \tag{4.76}$$

$$= e^{\left(2\Big(\ln(z)\ln(z-1)\Big) - rt + C_1\right)} \tag{4.77}$$

$$= e^{\left(2\ln\left|\frac{z}{z-1}\right| - rt + C_1\right)} = e^{\left(2\ln\left|\frac{1+Ae^{-rt}}{Ae^{-rt}}\right| - rt + C_1\right)} \tag{4.78}$$

$$= e^{\left(\ln\left|\frac{1}{A}e^{rt}+1\right|^2 - rt + C_1\right)} \tag{4.79}$$

$$= C_2\left|\frac{e^{rt}+A}{A}\right|^2 e^{-rt} = C(e^{rt}+A)^2 e^{-rt}, \quad \text{for some constant } C. \tag{4.80}$$

We have

$$(e^{rt}+A)^2 e^{-rt}\frac{d}{dt}\mathbb{E}[U(t)] - (e^{rt}+A)^2 e^{-rt} r\Big(2\phi(t)-1\Big)\mathbb{E}[U(t)] = 0 \tag{4.81}$$

$$\Rightarrow \left((e^{rt}+A)^2 e^{-rt}\mathbb{E}[U(t)]\right) = C \tag{4.82}$$

$$\Rightarrow \mathbb{E}[U(t)] = \frac{Ce^{rt}}{(e^{rt}+A)^2}. \tag{4.83}$$

If we assume $\mathbb{E}[U(0)] = \dfrac{C}{(1+A)^2} = 0$, then $C = 0$, which implies $\mathbb{E}[U(t)] = 0\ \forall t$. If $\mathbb{E}[U(0)] = u_0 \neq 0$, then $C = u_0(1+A)^2 \neq 0$.

Let $\mathbb{V}[U(t)] = \mathbb{E}\Big[(U(t) - \mathbb{E}[U(t)])^2\Big]$ be the variance of $U(t)$. We have

$$\frac{d}{dt}\mathbb{V}[U(t)] = \frac{d}{dt}\mathbb{E}\Big[(U(t) - \mathbb{E}[U(t)])^2\Big] = \frac{d}{dt}\mathbb{E}[(U(t))^2] - \frac{d}{dt}\Big(\mathbb{E}[U(t)]\Big)^2. \tag{4.84}$$

As discussed above, if we choose $\mathbb{E}[U(0)] = 0$, then $\mathbb{E}[U(t)] = 0\ \forall t$, and

$$\mathbb{V}[U(t)] = \mathbb{E}[(U(t))^2] \Rightarrow \frac{d}{dt}\mathbb{V}[U(t)] = \frac{d}{dt}\mathbb{E}[(U(t))^2]. \tag{4.85}$$

Now, we find $\frac{d}{dt}\mathbb{E}[(U(t))^2]$. Using Ito's formula, we have

$$d\left[U(t)^2\right] = \Big(2r(1-2\phi(t))U(t)^2 + \Big(b_{K\phi(t)} + d_{K\phi(t)}\Big)\phi(t)\Big)dt \tag{4.86}$$

$$+ 2\sqrt{\Big(b_{K\phi(t)} + d_{K\phi(t)}\Big)\phi(t)}\,U(t)\,dW(t). \tag{4.87}$$

Taking expectations of both sides, and recalling that $U(t)$ and $dW(t)$ are independent, we have

$$\mathbb{E}\Big[d\left(U(t)^2\right)\Big] = \Big(2r(1-2\phi(t))\Big)\mathbb{E}[U(t)^2]dt + \Big(b_{K\phi(t)} + d_{K\phi(t)}\Big)\phi(t)dt. \tag{4.88}$$

It follows that the rate of change of the second moment is given by

$$\frac{d}{dt}\mathbb{E}\Big[U(t)^2\Big] = \Big(2r(1-2\phi(t))\Big)\mathbb{E}[U(t)^2] + \Big(b_{K\phi(t)} + d_{K\phi(t)}\Big)\phi(t). \tag{4.89}$$

As mentioned above, we choose $\mathbb{E}[U(0)] = 0$, so $\mathbb{V}[U(t)] = \mathbb{E}\Big[U(t)^2\Big]$. Therefore,

$$\frac{d}{dt}\mathbb{V}\Big[U(t)\Big] = \Big(2r(1-2\phi(t))\Big)\mathbb{V}\Big[U(t)\Big] + \Big(b_{K\phi(t)} + d_{K\phi(t)}\Big)\phi(t). \tag{4.90}$$

Thus the variance of $U(t)$ spontaneously grows during the initial growth phase of the population, when $\phi(t) < 1/2$, and tends to decay during the approach to equilibrium, when $\phi(t) > 1/2$. Meanwhile the sum of the *per capita* birth and death rates act as a source term driving the variance all along the trajectory.

Let $D(t) = \dfrac{\Big(b_{K\phi(t)} + d_{K\phi(t)}\Big)\phi(t)}{2}$. For notational convenience, we will write U for $U(t)$ in the following computation. Similarly, we are going to omit the notation for time dependence in the other variables.

$$\frac{1}{2}\frac{d}{dt}\mathbb{V}[U] = r\Big(1-2\phi\Big)\mathbb{V}[U] + D \tag{4.91}$$

$$\Rightarrow\left((e^{rt}+A)^2e^{-rt}\mathbb{V}[U]\right) = \int D(e^{rt}+A)^2e^{-rt}dt \tag{4.92}$$

$$= \int\Big(\frac{b_0+d_0}{2} + \frac{(1-2\gamma)r}{2}\frac{1}{1+Ae^{-rt}}\Big)\frac{1}{1+Ae^{-rt}}(e^{rt}+A)^2e^{-rt}dt \tag{4.93}$$

$$= \int\Big(\frac{b_0+d_0}{2} + \frac{(1-2\gamma)r}{2}\frac{1}{1+Ae^{-rt}}\Big)\frac{1}{1+Ae^{-rt}}(1+Ae^{-rt})^2e^{2rt}e^{-rt}dt \tag{4.94}$$

$$= \int\Big(\frac{b_0+d_0}{2} + \frac{(1-2\gamma)r}{2}\frac{1}{1+Ae^{-rt}}\Big)\frac{1}{1+Ae^{-rt}}(1+Ae^{-rt})^2e^{rt}dt \tag{4.95}$$

$$= \int\frac{b_0+d_0}{2}(1+Ae^{-rt})e^{rt}dt + \int\frac{(1-2\gamma)r}{2}e^{rt}dt \tag{4.96}$$

$$= \int\frac{b_0+d_0}{2}e^{rt} + A\frac{b_0+d_0}{2}dt + \int\frac{(1-2\gamma)r}{2}e^{rt}dt \tag{4.97}$$

$$= \frac{1}{r}\frac{b_0+d_0}{2}e^{rt} + A\frac{b_0+d_0}{2}t + \frac{(1-2\gamma)}{2}e^{rt} + C_{\mathbb{V}[U]}, \tag{4.98}$$

which implies

$$\mathbb{V}[U(t)] = (e^{rt}+A)^{-2}e^{rt}\left(\frac{b_0+d_0}{r}e^{rt} + A(b_0+d_0)t + (1-2\gamma)e^{rt} + C_{\mathbb{V}[U]}\right), \tag{4.99}$$

for some constant $C_{\mathbb{V}[U]}$. Assuming the initial condition $\mathbb{V}[U(0)] = 0$, then

$$C_{\mathbb{V}[U]} = -\frac{b_0 + d_0}{r} - (1 - 2\gamma). \tag{4.100}$$

To visualize and check our analytical computation, in Figure 4.3 **(A)**, we plot 10 empirical, simulated cell number trajectories (multicolor stochastic curves), the theoretical mean $\mathbb{E}[X(t)]$ (solid black curve), and the theoretical mean plus/minus two theoretical standard deviations, i.e. $\mathbb{E}[U(t)] \pm 2\sqrt{\mathbb{V}[U(t)]}$, (dashed black curves). In Figure 4.3 **(B)**, we plot the theoretical variance $\mathbb{V}[U(t)]$ of the stochastic fluctuation $U(t)$ (recall that we assume $X(t)$ is sum of the deterministic solution $K\phi(t)$ and stochastic fluctuation). We verify that $\mathbb{E}[U(t)] = 0$, but it is not necessary to plot it. We observe that nearly all empirical, simulated cell number trajectories are bounded between $\mathbb{E}[U(t)] \pm 2\sqrt{\mathbb{V}[U(t)]}$, which is consistent with our Gaussian distribution assumption. We also see that variance of the stochastic fluctuation $U(t)$ increases exponentially, and then decreases to a stabilized value.

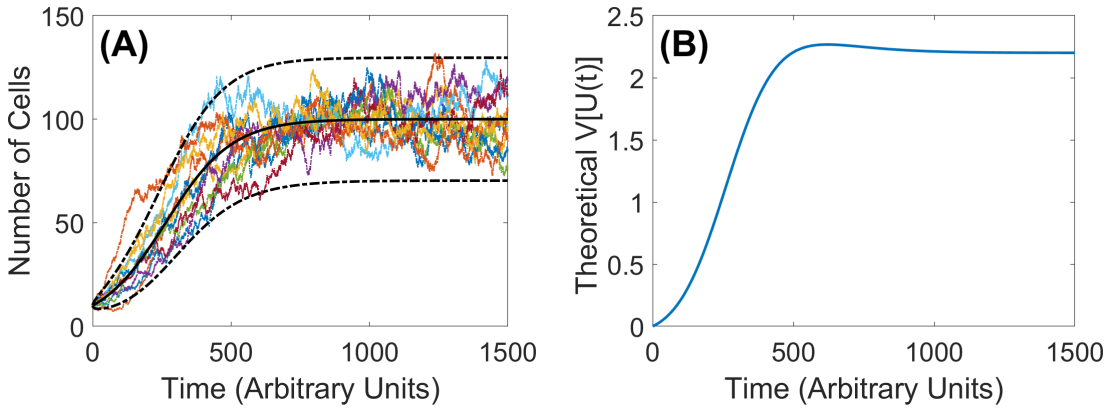

Figure 4.3: **Analytically derived mean and variance agree with simulated trajectories.** Note that nearly all simulated trajectories are bounded between $\mathbb{E}[X(t)] \pm 2\sqrt{\mathbb{V}[X(t)]}$. In **(A)**, we plot 10 empirical, simulated cell number trajectories (multicolor stochastic curves), the theoretical mean $\mathbb{E}[X(t)]$ (solid black curve), and the theoretical mean plus and minus two theoretical standard deviations (dash black curves). In **(B)**, we plot the theoretical variance $\mathbb{V}[U(t)]$ of the stochastic fluctuation $U(t)$. We observe that variance of the stochastic fluctuation $U(t)$ increases exponentially, and then decreases to a stabilized value.

4.4 Mean of Log-Likelihood

In this section, we apply the results of Section 4.3 to calculate the expected value of the log-likelihood function for a single trajectory. By the Markov property of our model, we

have $P(\boldsymbol{X}_{\text{data}}) = P(x_0)P(x_1|x_0)\dots P(x_N|x_{N-1})$ and

$$\mathbb{E}\Big[\mathcal{L}(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}})\Big] \tag{4.101}$$

$$= \int_{\boldsymbol{X}_{\text{data}}} \ln\Big(\underbrace{P(\boldsymbol{X}_{\text{data}})}_{\Theta_{\text{test}}}\Big)\,\underbrace{P(\boldsymbol{X}_{\text{data}})}_{\Theta_{\text{data}}}\,d\boldsymbol{X}_{\text{data}} \tag{4.102}$$

$$= \int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} \ln\Big(\underbrace{P(x_0)}_{\Theta_{\text{test}}}\Big)\,\underbrace{P(x_0)P(x_1|x_0)\dots P(x_N|x_{N-1})}_{\Theta_{\text{data}}}\,dx_N\dots dx_0 \tag{4.103}$$

$$+\sum_{j=1}^{T-1}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} \ln\Big(\underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{test}}}\Big)\,\underbrace{P(x_0)P(x_1|x_0)\dots P(x_N|x_{N-1})}_{\Theta_{\text{data}}}\,dx_N\dots dx_0. \tag{4.104}$$

Note that the notation $\underbrace{P(\cdot)}_{\Theta_{\text{test}}}$ means the probability distribution is characterized by the parameter set Θ_{test}. Similarly, the notation $\underbrace{P(\cdot)}_{\Theta_{\text{data}}}$ means the probability distribution is characterized by the parameter set Θ_{data}. All the x_j in this section denote realizations from the distribution characterized by the parameter set Θ_{data}. The notations $b_{x_j,\text{test}}$ and $d_{x_j,\text{test}}$ mean the *per capita* birth and death rates are defined with the parameter set Θ_{test} as follows

$$b_{x_j,\text{test}} = b_{0,\text{test}} - \gamma_{\text{test}}\frac{r_{\text{test}}}{K_{\text{test}}}x_j \tag{4.105}$$

$$d_{x_j,\text{test}} = d_{0,\text{test}} + (1-\gamma_{\text{test}})\frac{r_{\text{test}}}{K_{\text{test}}}x_j. \tag{4.106}$$

Continuing with the computation, we have

$$\mathbb{E}\Big[\mathcal{L}(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}})\Big] \tag{4.107}$$

$$= \int_{-\infty}^{\infty} \underbrace{\ln\Big(\underbrace{P(x_0)}_{\Theta_{data}}\Big)}_{=\ln(1)} \underbrace{P(x_0)}_{\Theta_{\text{data}}} dx_0 \tag{4.108}$$

$$+ \sum_{j=1}^{T-1} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \ln\Big(\underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{test}}}\Big) \underbrace{P(x_0)P(x_1|x_0)\dots P(x_{j+1}|x_j)}_{\Theta_{\text{data}}} dx_{j+1}dx_j \dots dx_0 \tag{4.109}$$

$$= \sum_{j=1}^{T-1} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \ln\Big(\underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{test}}}\Big) \underbrace{P(x_2|x_1)\dots P(x_{j+1}|x_j)}_{\Theta_{\text{data}}} \tag{4.110}$$

$$\left(\int_{-\infty}^{\infty} \underbrace{P(x_0)P(x_1|x_0)}_{\Theta_{\text{data}}} dx_0\right) dx_1 \dots dx_{j+1}$$

$$= \sum_{j=1}^{T-1} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \ln\Big(\underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{test}}}\Big) \underbrace{P(x_3|x_2)\dots P(x_{j+1}|x_j)}_{\Theta_{\text{data}}} \tag{4.111}$$

$$\left(\int_{-\infty}^{\infty} \underbrace{P(x_1)P(x_2|x_1)}_{\Theta_{\text{data}}} dx_1\right) dx_2 \dots dx_{j+1}$$

$$= \sum_{j=1}^{T-1} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \ln\Big(\underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{test}}}\Big) \underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{data}}} dx_{j+1}\right) \underbrace{P(x_j)}_{\Theta_{\text{data}}} dx_j \tag{4.112}$$

$$= \sum_{j=1}^{T-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{1}{2}\ln\Big(\frac{1}{2\pi\mathbb{V}_{\text{test}}[X(t_{j+1})|X(t_j)=x_j]}\Big) \underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{data}}} dx_{j+1}\right) \underbrace{P(x_j)}_{\Theta_{\text{data}}} dx_j \tag{4.113}$$

$$- \sum_{j=1}^{T-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{(x_{j+1}-x_j-\mathbb{E}_{\text{test}}[X(t_{j+1})|X(t_j)=x_j])^2}{2\mathbb{V}_{\text{test}}[X(t_{j+1})|X(t_j)=x_j]} \underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{data}}} dx_{j+1}\right) \underbrace{P(x_j)}_{\Theta_{\text{data}}} dx_j \tag{4.114}$$

$$= \sum_{j=1}^{T-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{1}{2} \ln \left(\frac{1}{2\pi x_j (b_{x_j\text{test}} + d_{x_j\text{test}}) \Delta t} \right) \underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{data}}} dx_{j+1} \right) \underbrace{P(x_j)}_{\Theta_{\text{data}}} dx_j \quad (4.115)$$

$$- \sum_{j=1}^{T-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{(x_{j+1} - x_j - x_j(b_{x_j,\text{test}} - d_{x_j,\text{test}})\Delta t)^2}{2x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Delta t} \underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{data}}} dx_{j+1} \right) \underbrace{P(x_j)}_{\Theta_{\text{data}}} dx_j. \quad (4.116)$$

Observing that $P(x_{j+1}|x_j)$ may be computed from $P(x_{j+1} - x_j|x_j)$, we define the random variable $\Delta X_j := X(t_{j+1}) - X(t_j)$ and a realization $\Delta x_j := (x_{j+1} - x_j)$ of ΔX_j. We have

$$g_1(\Delta x_j) := \frac{\left(\Delta x_j - x_j(b_{x_j,\text{test}} - d_{x_j,\text{test}})\Delta t \right)^2}{2x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Delta t}. \quad (4.117)$$

With these definitions and the observation that integration with respect to the infinitesimal $d\Delta x_j$ and dx_{j+1} mean the same thing, we have

$$\mathbb{E}\Big[\mathcal{L}(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}})\Big] \tag{4.118}$$

$$=\sum_{j=1}^{T-1}\int_{-\infty}^{\infty}\frac{1}{2}\ln\Big(\frac{1}{2\pi x_j(b_{x_j,\text{test}}+d_{x_j,\text{test}})\Delta t}\Big)\left(\int_{-\infty}^{\infty}\underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{data}}}dx_{j+1}\right)\underbrace{P(x_j)}_{\Theta_{\text{data}}}dx_j \tag{4.119}$$

$$-\sum_{j=1}^{T-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big(g_1(\Delta x_j)\Big)\underbrace{P(\Delta x_j|x_j)}_{\Theta_{\text{data}}}d\Delta x_j\Bigg)\underbrace{P(x_j)}_{\Theta_{\text{data}}}dx_j \tag{4.120}$$

$$=\sum_{j=1}^{T-1}\int_{-\infty}^{\infty}\frac{1}{2}\ln\Big(\frac{1}{2\pi x_j(b_{x_j,\text{test}}+d_{x_j,\text{test}})\Delta t}\Big)\underbrace{P(x_j)}_{\Theta_{\text{data}}}dx_j \tag{4.121}$$

$$-\sum_{j=1}^{T-1}\int_{-\infty}^{\infty}\mathbb{E}_{\text{data}}\Big[g_1(\Delta X_j)\Big|X_j)\Big]\underbrace{P(x_j)}_{\Theta_{\text{data}}}dx_j \tag{4.122}$$

$$=-\sum_{j=1}^{T-1}\frac{1}{2}\ln(2\pi\Delta t)-\sum_{j=1}^{T-1}\frac{1}{2}\mathbb{E}_{\text{data}}[\ln(X_j)]-\sum_{j=1}^{T-1}\frac{1}{2}\mathbb{E}_{\text{data}}[\ln(b_{X_j,\text{test}}+d_{X_j,\text{test}})] \tag{4.123}$$

$$-\sum_{j=1}^{T-1}\int_{-\infty}^{\infty}\mathbb{E}_{\text{data}}\Big[g_1(\Delta X_j)\Big|X_j\Big]\underbrace{P(x_j)}_{\Theta_{\text{data}}}dx_j \tag{4.124}$$

$$=\sum_{j=1}^{T-1}\left(-\frac{1}{2}\ln(2\pi\Delta t)-\frac{1}{2}\mathbb{E}_{\text{data}}[\ln(X_j)]-\frac{1}{2}\mathbb{E}_{\text{data}}[\ln(b_{X_j,\text{test}}+d_{X_j,\text{test}})]\right) \tag{4.125}$$

$$-\sum_{j=1}^{T-1}\mathbb{E}_{\text{data}}\left[\mathbb{E}_{\text{data}}\Big[g_1(\Delta X_j)\Big|X_j\Big]\right]. \tag{4.126}$$

We approximate the expected values $\mathbb{E}_{\text{data}}[\ln(X_j)]$, $\mathbb{E}_{\text{data}}[\ln(b_{X_j,\text{test}}+d_{X_j,\text{test}})]$, $\mathbb{E}_{\text{data}}\Big[g_1(\Delta X_j)|X_j\Big]$, and $\mathbb{E}_{\text{data}}\left[\mathbb{E}_{\text{data}}\Big[g_1(\Delta X_j)|X_j\Big]\right]$ as follows. In general, given a random variable Y and a function $g(Y)$, we consider the Taylor expansion of $g(Y)$ around the expected value $\mathbb{E}[Y]$

of the random variable Y, and we write

$$g(Y) \approx g(y)\Big|_{y=\mathbb{E}[Y]} + \frac{dg}{dy}\Big|_{y=\mathbb{E}[Y]}\Big(Y - \mathbb{E}[Y]\Big) + \frac{1}{2}\frac{d^2g}{dy^2}\Big|_{y=\mathbb{E}[Y]}\Big(Y - \mathbb{E}[Y]\Big)^2 \tag{4.127}$$

$$\Rightarrow \mathbb{E}[g(Y)] \approx g(y)\Big|_{y=\mathbb{E}[Y]} + \frac{1}{2}\frac{d^2g}{dy^2}\Big|_{y=\mathbb{E}[Y]}\mathbb{V}[Y]. \tag{4.128}$$

Applying the approximation in Equation (4.128), we have

$$\mathbb{E}_{\text{data}}[\ln(X_j)] \approx \ln(\mathbb{E}_{\text{data}}[X_j]) - \frac{1}{2}\frac{1}{\mathbb{E}_{\text{data}}[X_j]^2}\mathbb{V}_{\text{data}}[X_j] \tag{4.129}$$

$$\mathbb{E}_{\text{data}}[\ln(b_{X_j,\text{test}} + d_{X_j,\text{test}})] \approx \ln(b_{\mathbb{E}_{\text{data}}[X_j],\text{test}} + d_{\mathbb{E}_{\text{data}}[X_j],\text{test}}) \tag{4.130}$$

$$- \frac{1}{2}\frac{\left(\frac{d(b_{\mathbb{E}_{\text{data}}[X_j],\text{test}} + d_{\mathbb{E}_{\text{data}}[X_j],\text{test}})}{x_j}\right)^2}{(b_{\mathbb{E}_{\text{data}}[X_j],\text{test}} + d_{\mathbb{E}_{\text{data}}[X_j],\text{test}})^2}, \tag{4.131}$$

where

$$\frac{d(b_{x_j,\text{test}} + d_{x_j,\text{test}})}{dx_j} = \begin{cases} (1 - 2\gamma_{\text{test}}) & \text{if } \left(b_{0,\text{test}} - \gamma_{\text{test}}\frac{r_{\text{test}}}{K_{\text{test}}}x_j\right) > 0 \\ (1 - \gamma_{\text{test}}) & \text{otherwise} \end{cases}. \tag{4.132}$$

Similarly,

$$\frac{d(b_{x_j,\text{data}} + d_{x_j,\text{data}})}{dx_j} = \begin{cases} (1 - 2\gamma_{\text{data}}) & \text{if } \left(b_{0,\text{data}} - \gamma_{\text{data}}\frac{r_{\text{data}}}{K_{\text{data}}}x_j\right) > 0 \\ (1 - \gamma_{\text{data}}) & \text{otherwise.} \end{cases} \tag{4.133}$$

We apply Equation (4.128) and have

$$\mathbb{E}_{\text{data}}\Big[g_1(\Delta X_j)|X_j\Big] \approx g_1(\Delta x_j)\Big|_{\Delta x_j=\mathbb{E}_{\text{data}}[\Delta X_j|X_j]} + \frac{1}{2}\frac{d^2 g_1}{d\Delta x_j^2}\Big|_{\Delta x_j=\mathbb{E}_{\text{data}}[\Delta X_j|X_j]} \mathbb{V}_{\text{data}}[\Delta X_j|X_j] \tag{4.134}$$

$$= \frac{\Big(X_j(b_{X_j,\text{data}} - d_{X_j,\text{data}})\Delta t - X_j(b_{X_j,\text{test}} - d_{X_j,\text{test}})\Delta t\Big)^2}{2X_j(b_{X_j,\text{test}} + d_{X_j,\text{test}})\Delta t} \tag{4.135}$$

$$+ \frac{1}{2}\frac{X_j(b_{X_j,\text{data}} + d_{X_j,\text{data}})\Delta t}{X_j(b_{X_j,\text{test}} + d_{X_j,\text{test}})\Delta t} \tag{4.136}$$

$$= \frac{\Big(X_j(b_{X_j,\text{data}} - d_{X_j,\text{data}})\Delta t - X_j(b_{X_j,\text{test}} - d_{X_j,\text{test}})\Delta t\Big)^2}{2X_j(b_{X_j,\text{test}} + d_{X_j,\text{test}})\Delta t} \tag{4.137}$$

$$+ \frac{1}{2}\frac{b_{X_j,\text{data}} + d_{X_j,\text{data}}}{b_{X_j,\text{test}} + d_{X_j,\text{test}}}. \tag{4.138}$$

In the context of our research problem, we are interested in cases where the net growth rates are the same. Hence, it is reasonable to assume $b_{X_j,\text{data}} - d_{X_j,\text{data}} = b_{X_j,\text{test}} - d_{X_j,\text{test}}$, which leads to

$$\mathbb{E}_{\text{data}}\Big[g_1(\Delta X_j)|X_j\Big] \approx \frac{1}{2}\frac{b_{X_j,\text{data}} + d_{X_j,\text{data}}}{b_{X_j,\text{test}} + d_{X_j,\text{test}}} \tag{4.139}$$

$$\Rightarrow \mathbb{E}_{\text{data}}\left[\mathbb{E}_{\text{data}}\Big[g_1(\Delta X_j)|X_j\Big]\right] = \frac{1}{2}\mathbb{E}_{\text{data}}\left[\frac{b_{X_j,\text{data}} + d_{X_j,\text{data}}}{b_{X_j,\text{test}} + d_{X_j,\text{test}}}\right]. \tag{4.140}$$

Using Equation (4.128), we have

$$\frac{d}{dx_j}\frac{b_{x_j,\text{data}}+d_{x_j,\text{data}}}{b_{x_j,\text{test}}+d_{x_j,\text{test}}} = \frac{\frac{d(b_{x_j,\text{data}}+d_{x_j,\text{data}})}{dx_j}}{b_{x_j,\text{test}}+d_{x_j,\text{test}}} \tag{4.141}$$

$$- \Big(b_{x_j,\text{data}}+d_{x_j,\text{data}}\Big)\Big(b_{x_j,\text{test}}+d_{x_j,\text{test}}\Big)^{-2}\frac{d(b_{x_j,\text{test}}+d_{x_j,\text{test}})}{dx_j} \tag{4.142}$$

$$\frac{d^2}{dx_j^2}\frac{b_{x_j,\text{data}}+d_{x_j,\text{data}}}{b_{x_j,\text{test}}+d_{x_j,\text{test}}} = -\frac{\frac{d(b_{x_j,\text{data}}+d_{x_j,\text{data}})}{dx_j}\frac{d(b_{x_j,\text{test}}+d_{x_j,\text{test}})}{dx_j}}{\Big(b_{x_j,\text{test}}+d_{x_j,\text{test}}\Big)^2} \tag{4.143}$$

$$-\frac{\frac{d(b_{x_j,\text{data}}+d_{x_j,\text{data}})}{dx_j}\frac{d(b_{x_j,\text{test}}+d_{x_j,\text{test}})}{dx_j}}{\Big(b_{x_j,\text{test}}+d_{x_j,\text{test}}\Big)^2} \tag{4.144}$$

$$+\frac{2\Big(b_{x_j,\text{data}}+d_{x_j,\text{data}}\Big)}{\Big(b_{x_j,\text{test}}+d_{x_j,\text{test}}\Big)^3}\Big(\frac{d(b_{x_j,\text{test}}+d_{x_j,\text{test}})}{dx_j}\Big)^2 \tag{4.145}$$

$$\mathbb{E}_{\text{data}}\left[\frac{b_{X_j,\text{data}}+d_{X_j,\text{data}}}{b_{X_j,\text{test}}+d_{X_j,\text{test}}}\right] \approx \frac{b_{\mathbb{E}_{\text{data}}[X_j],\text{data}}+d_{\mathbb{E}_{\text{data}}[X_j],\text{data}}}{b_{\mathbb{E}_{\text{data}}[X_j],\text{test}}+d_{\mathbb{E}_{\text{data}}[X_j],\text{test}}} \tag{4.146}$$

$$+\frac{1}{2}\frac{d^2}{dx_j^2}\frac{b_{x_j,\text{data}}+d_{x_j,\text{data}}}{b_{x_j,\text{test}}+d_{x_j,\text{test}}}\bigg|_{x_j=\mathbb{E}_{\text{data}}[X_j]}\mathbb{V}_{\text{data}}[X_j]. \tag{4.147}$$

Note that when $\Theta_{\text{data}} = \Theta_{\text{test}}$, we have

$$\frac{d^2}{dx_j^2}\frac{b_{x_j,\text{data}} + d_{x_j,\text{data}}}{b_{x_j,\text{test}} + d_{x_j,\text{test}}} = -\Big(b_{x_j,\text{test}} + d_{x_j,\text{test}}\Big)\frac{\Big(\frac{d(b_{x_j,\text{test}} + d_{x_j,\text{test}})}{dx_j}\Big)^2}{\Big(b_{x_j,\text{test}} + d_{x_j,\text{test}}\Big)^3} \tag{4.148}$$

$$-\Big(b_{x_j,\text{test}} + d_{x_j,\text{test}}\Big)\frac{\Big(\frac{d(b_{x_j,\text{test}} + d_{x_j,\text{test}})}{dx_j}\Big)^2}{\Big(b_{x_j,\text{test}} + d_{x_j,\text{test}}\Big)^3} \tag{4.149}$$

$$+\frac{2\Big(b_{x_j,\text{test}} + d_{x_j,\text{test}}\Big)}{\Big(b_{x_j,\text{test}} + d_{x_j,\text{test}}\Big)^3}\Big(\frac{d(b_{x_j,\text{test}} + d_{x_j,\text{test}})}{dx_j}\Big)^2 \tag{4.150}$$

$$=0 \tag{4.151}$$

and $\mathbb{E}_{\text{data}}\Big[\mathbb{E}_{\text{data}}\Big[g_1(\Delta X_j)|X_j\Big]\Big] = \frac{1}{2}$, which is consistent with

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{(x_{j+1} - x_j - \mathbb{E}_{\text{test}}[X(t_{j+1})|X(t_j) = x_j])^2}{2\mathbb{V}_{\text{test}}[X(t_{j+1})|X(t_j) = x_j]}\underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{data}}}dx_{j+1}\right)\underbrace{P(x_j)}_{\Theta_{\text{data}}}dx_j \tag{4.152}$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{(x_{j+1} - x_j - \mathbb{E}_{\text{data}}[X(t_{j+1})|X(t_j) = x_j])^2}{2\mathbb{V}_{\text{data}}[X(t_{j+1})|X(t_j) = x_j]}\underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{data}}}dx_{j+1}\right)\underbrace{P(x_j)}_{\Theta_{\text{data}}}dx_j \tag{4.153}$$

$$=\int_{-\infty}^{\infty}\frac{1}{2\mathbb{V}_{\text{data}}[X(t_{j+1})|X(t_j) = x_j]}\mathbb{V}_{\text{data}}[X(t_{j+1})|X(t_j) = x_j]\underbrace{P(x_j)}_{\Theta_{\text{data}}}dx_j \tag{4.154}$$

$$=\frac{1}{2}. \tag{4.155}$$

In Figure 4.4, we compare the theoretical mean with empirical mean for various combinations of γ_{data} and γ_{test} listed in Table 4.2. We see that the two means are well-aligned.

We compute the theoretical mean using the approximation described above, except for $\mathbb{E}_{\text{data}}\left[\frac{b_{X_j,\text{data}} + d_{X_j,\text{data}}}{b_{X_j,\text{test}} + d_{X_j,\text{test}}}\right]$; it only works when we use first-order Taylor approximation for this mean. For the other expected values, it works even when we use second-order Taylor approximation.

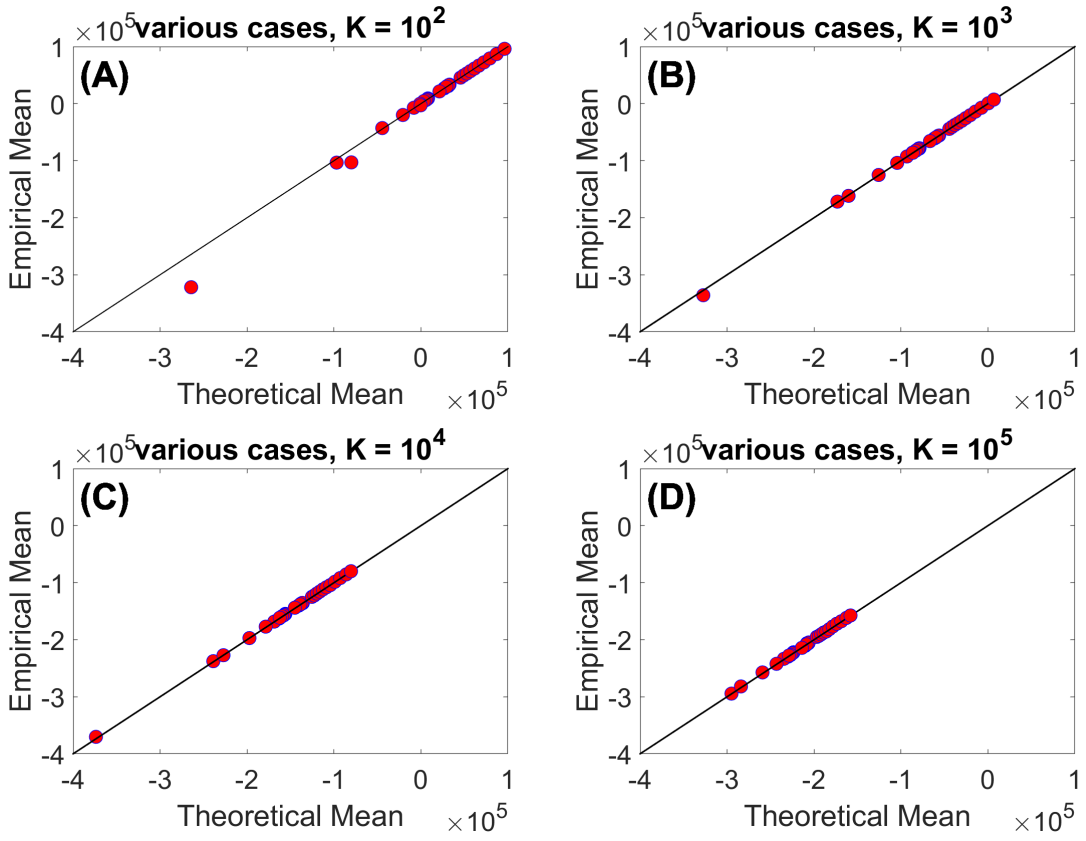

Figure 4.4: **Theoretical mean and empirical mean of the log-likelihood are well-aligned.** In this figure, we compare the theoretical mean and the empirical mean of the log-likelihood for various combinations of γ_{data} and γ_{test} listed in Table 4.2 and different carrying capacities: **(A)**: $K = 10^2$, **(B)**: $K = 10^3$, **(C)**: $K = 10^3$, **(D)**: $K = 10^4$.

γ_{data}	γ_{test}	γ_{data}	γ_{test}	γ_{data}	γ_{test}
0	0	0.5	0	1	0
0	0.1	0.5	0.1	1	0.1
0	0.2	0.5	0.2	1	0.2
0	0.3	0.5	0.3	1	0.3
0	0.4	0.5	0.4	1	0.4
0	0.5	0.5	0.5	1	0.5
0	0.6	0.5	0.6	1	0.6
0	0.7	0.5	0.7	1	0.7
0	0.8	0.5	0.8	1	0.8
0	0.9	0.5	0.9	1	0.9
0	1	0.5	1	1	1

Table 4.2: Various Combinations of γ_{data} and γ_{test} used in Figure 4.4.

4.5 Variance of Log-Likelihood

In this section we apply the results of Section 4.3 to attempt to find the variance of the log-likelihood function, using a linear Taylor approximation for the log-likelihood $\mathcal{L}\Big(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}}\Big)$.

$$\mathbb{V}\Big[\mathcal{L}\Big(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}}\Big)\Big] = \mathbb{E}\Big[\mathcal{L}^2\Big(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}}\Big)\Big] - \left(\mathbb{E}\Big[\mathcal{L}\Big(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}}\Big)\Big]\right)^2, \tag{4.156}$$

where

$$\mathbb{E}\Big[\mathcal{L}^2\Big(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}}\Big)\Big] = \mathbb{E}\Big[\ln^2\Big(\underbrace{P(\boldsymbol{X}_{\text{data}})}_{\Theta_{\text{test}}}\Big)\Big] \tag{4.157}$$

$$= \mathbb{E}\left[\left(\ln\Big(\underbrace{P(x_{0,\text{data}})}_{\Theta_{\text{test}}}\Big) + \sum_{j=1}^{N-1}\ln\Big(\underbrace{P(x_{j+1}|x_j)}_{\Theta_{\text{test}}}\Big)\right)^2\right], \tag{4.158}$$

which is hard to compute. Therefore, we consider the linear Taylor expansion of $g_{\mathcal{L}}\Big(\boldsymbol{X}_{\text{data}}\Big) := \mathcal{L}\Big(\boldsymbol{X}_{\text{data}}\Big|\Theta_{\text{test}}\Big)$ around $\mathbb{E}\Big[\boldsymbol{X}_{\text{data}}\Big]$ as follows:

$$g_{\mathcal{L}}\Big(\boldsymbol{X}_{\text{data}}\Big) \approx \mathcal{L}\Big(\mathbb{E}\Big[\boldsymbol{X}_{\text{data}}\Big]\Big|\Theta_{\text{test}}\Big) + \left(\nabla g_{\mathcal{L}}\Big|_{\mathbb{E}[\boldsymbol{X}_{\text{data}}]}\right)^T \left(\boldsymbol{X}_{\text{data}} - \mathbb{E}\Big[\boldsymbol{X}_{\text{data}}\Big]\right) \tag{4.159}$$

$$\Rightarrow \mathbb{V}\Big[g_{\mathcal{L}}\Big(\boldsymbol{X}_{\text{data}}\Big)\Big] \approx \left((\nabla g_{\mathcal{L}})^2\Big|_{\mathbb{E}[\boldsymbol{X}_{\text{data}}]}\right)^T \mathbb{V}\Big[\boldsymbol{X}_{\text{data}}\Big], \tag{4.160}$$

where $(\nabla g_{\mathcal{L}})^2$ denotes the vector of the squares of the elements of $\nabla g_{\mathcal{L}}$. For $j = 2, \ldots, T-1$,

$$(\nabla g_{\mathcal{L}})_j = -\frac{1}{2}\frac{\dfrac{d\Big(x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Big)}{dx_j}}{x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})} \tag{4.161}$$

$$- \frac{2\Big(x_{j+1} - x_j - x_j(b_{x_j,\text{test}} - d_{x_j,\text{test}})\Delta t\Big)}{2x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Delta t}\Big(-1 - \Delta t\frac{d\Big(x_j(b_{x_j,\text{test}} - d_{x_j,\text{test}})\Big)}{dx_j}\Big) \tag{4.162}$$

$$+ \frac{\Big(x_{j+1} - x_j - x_j(b_{x_j,\text{test}} - d_{x_j,\text{test}})\Delta t\Big)^2}{\Big(2x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Delta t\Big)^2} 2\Delta t\frac{d\Big(x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Big)}{dx_j} \tag{4.163}$$

$$- \frac{2\Big(x_j - x_{j-1} - x_{j-1}(b_{x_{j-1},\text{test}} - d_{x_{j-1},\text{test}})\Delta t\Big)}{2x_{j-1}(b_{x_{j-1},\text{test}} + d_{x_{j-1},\text{test}})\Delta t} \tag{4.164}$$

For $j = 1$,

$$(\nabla g_{\mathcal{L}})_j = -\frac{1}{2}\frac{\dfrac{d\Big(x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Big)}{dx_j}}{x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})} \tag{4.165}$$

$$- \frac{2\Big(x_{j+1} - x_j - x_j(b_{x_j,\text{test}} - d_{x_j,\text{test}})\Delta t\Big)}{2x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Delta t}\Big(- 1 - \Delta t\frac{d\Big(x_j(b_{x_j,\text{test}} - d_{x_j,\text{test}})\Big)}{dx_j}\Big) \tag{4.166}$$

$$+ \frac{\Big(x_{j+1} - x_j - x_j(b_{x_j,\text{test}} - d_{x_j,\text{test}})\Delta t\Big)^2}{\Big(2x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Delta t\Big)^2}2\Delta t\frac{d\Big(x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Big)}{dx_j}, \tag{4.167}$$

where

$$\frac{d\Big(x_j(b_{x_j,\text{test}} + d_{x_j,\text{test}})\Big)}{dx_j} = \begin{cases} b_{0,\text{test}} + d_{0,\text{test}} + (1 - 2\gamma_{\text{test}})\dfrac{r_{\text{test}}}{K_{\text{test}}}x_j & \text{if } b_{0,\text{test}} - \gamma_{\text{test}}\dfrac{r_{\text{test}}}{K_{\text{test}}}x_j > 0 \\ d_{0,\text{test}} + (1 - \gamma_{\text{test}})\dfrac{r_{\text{test}}}{K_{\text{test}}}x_j & \text{if } b_{0,\text{test}} - \gamma_{\text{test}}\dfrac{r_{\text{test}}}{K_{\text{test}}}x_j \leq 0 \end{cases} \tag{4.168}$$

and

$$\frac{d\Big(x_j(b_{x_j,\text{test}} - d_{x_j,\text{test}})\Big)}{dx_j} = \begin{cases} b_{0,\text{test}} - d_{0,\text{test}} - \dfrac{r_{\text{test}}}{K_{\text{test}}}x_j & \text{if } b_{0,\text{test}} - \gamma_{\text{test}}\dfrac{r_{\text{test}}}{K_{\text{test}}}x_j > 0 \\ -d_{0,\text{test}} - (1 - \gamma_{\text{test}})\dfrac{r_{\text{test}}}{K_{\text{test}}}x_j & \text{if } b_{0,\text{test}} - \gamma_{\text{test}}\dfrac{r_{\text{test}}}{K_{\text{test}}}x_j \leq 0 \end{cases} \tag{4.169}$$

Unfortulatey, we do not consider this approximation to be successful. In our estimation, the failure of the approxmation most likely results because here we approximate $\mathcal{L}\Big(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}}\Big)$ with a linear function, but this function is too highly nonlinear for the

approximation to be accurate.

4.6 Conclusion and Discussion

In this chapter, we use the log-likelihood approach to disambiguate density-dependent dynamics. In Section 4.2, we maximize the log-likelihood function to find the most likely density dependence parameter $\gamma \in [0, 1]$ for a given single cell number time series, with the assumption that we know the other parameters b_0, d_0, and K.

Solving numerical optimization problems can be challenging as we move to higher dimensions (i.e. add more unknown parameters to the system). In Sections 4.3, 4.4, and 4.5, we use a probabilistic/statistical approach to disambiguate density-dependent dynamics. In particular, we have observed empirically that the log-likelihood value is higher when γ_{test} matches with γ_{data} in Figure **??** and **??**, which is also well-known analytically [76]. Therefore, instead of asking the question of "most likely" as in Section 4.2, we ask the question of "more likely" in Sections 4.3, 4.4, and 4.5. Instead of considering a continuous spectrum of density dependence candidates (i.e. γ on the whole interval $[0, 1]$) as in Section 4.2, we consider a discrete number of candidate γ values, and we check which one gives the highest value of when plugged into the log-likelihood, then we may say that the value is "more likely" than the others on the discrete list. In order to justify this approach, we need to analyze the theoretical distributions of the log-likelihood evaluated at different values of γ. Therefore, we compute the theoretical mean and variance of the log-likelihood. In order to compute these two statistics, we must know the theoretical mean and variance of the cell number $X(t)$. In section 4.3, we derive analytical expressions for the theoretical mean and variance of $X(t)$, by considering $X(t)$ as a stochastic deviation from the deter-

ministic solution $\phi(t)$ to the classical logistic differential equation [93]. We assume that the stochastic deviation $U(t)$ added to the deterministic solution follows a time-dependent Ornstein–Uhlenbeck process [32] and compute the theoretical mean and variance of $U(t)$. We use those theoretical statistics to derive the theoretical mean and variance of the log-likelihood in Section 4.4 and Section 4.5. The main idea of computing the theoretical mean of the log-likelihood in Section 4.4 is to use properties of Markov process to have a formula for $P(\boldsymbol{X}_{\text{data}}|\Theta_{\text{test}})$, and approximate functions of random variables with second-order Taylor polynomials [8, 39]. The second-order approximation works well, except when computing the mean $\mathbb{E}_{\text{data}}\left[\dfrac{b_{X_j,\text{data}}+d_{X_j,\text{data}}}{b_{X_j,\text{test}}+d_{X_j,\text{test}}}\right]$, which would equal to 1 if the parameter set $\Theta_{\text{test}}=\Theta_{\text{data}}$. For this mean, we find a linear Taylor approximation works well. We confirm our analytical computation by comparing the theoretical means with the empirical means for various combinations of γ_{data} and γ_{test}, in Figure 4.4.

Finally, in Section 4.5, we attempt to compute the variance of the log-likelihood. However, the computation involves the square of the sum of many random variables, many of which may not be independent. Moreover, second-order Taylor approximation requires a covariance matrix, which we do not know. When we use linear Taylor approximation, the theoretical variance and empirical variance do not match well, as the log-likelihood is highly nonlinear. This problem of computing the theoretical variance remains an open problem for future work. As for future directions, we will also explore other approaches such as Bayesian inference and Gibbs sampling. [14, 54].

Chapter 5

Future Directions

At the end of each of Chapter 2, Chapter 3, and Chapter 4, we summarize our results and discuss, in detail, future work related to the project discussed in the chapter. In this last chapter of the thesis, we give a big picture of future directions.

A big-picture takeaway of this thesis is that stochastic fluctuations allow us to separately infer birth and death rates from cell number time series data, which is significant in disambiguating different mechanisms underlying the same *net* growth rates.

Mathematically, we are broadly interested in roles of stochasticity in biological dynamics and parameter identification. For future directions, we will consider different stochastic models for biological systems and analyze the roles of noise in the analysis of the models. We will also explore how we can further utilize stochasticity to uncover underlying mechanisms at different scales–more microscopic (e.g. within the birth process or within the death process) and more macroscopic (e.g. cell populations moving in space). Biologically,

we are broadly interested studying ecological and evolutionary processes on different time scales and the relation between these processes and drug resistance. Combing mathematics and biology, we would like to develop translational optimal treatments that can be effectively applied in clinical settings.

Appendix A

Appendix for Chapter 3

A.1 Model Parameters Used in Simulation

Table A.1: Model Parameters Used in Simulation

Parameter	Value	Unit
b_0	1.1/120	1/time
d_0	0.1/120	1/time
r	1/120	1/time
K	10^5	Dimensionless

A.2 Error Analysis of the Direct Estimation Method

As described in Section 3.3.3, we discretize all the values of cell number across the whole ensemble of trajectories into bins. Denote the bin size as η. The left end point N_k of the kth bin $[N_k, N_k + \eta)$ with $k = 1, 2, \ldots, k_{\max}$ is equal to $N_k := N_{\min} + (k-1)\eta$, where $N_{\min}$ is the smallest value of cell number across the whole ensemble of trajectories. In many instances, $N_{\min} = N(t_0)$, the initial population size. The total number of bins $k_{\max} \in \mathbb{Z}^+$ is

equal to $\left\lceil \frac{N_{\max} - N_{\min}}{\eta} \right\rceil$, where $N_{\max}$ is the largest value of cell number across the whole ensemble of trajectories, and $\lceil n \rceil$ is the smallest integer not less than n. The ith cell number element to have landed in the kth bin $[N_k, N_k + \eta)$ is equal to $N_k + \eta_i$. For simplicity, we make the approximation that for each bin, the random variables η_i are i.i.d. and uniformly distributed on $[0, \eta)$. We expect this approximation to be reasonably accurate when the bin size η is small enough that a given trajectory is unlikely to land in any particular bin twice in succession; the approximation may become inaccurate for excessively large bin sizes. In light of this uniform distribution assumption, we use the midpoint $N_k + \frac{\eta}{2}$ to represent the kth bin $[N_k, N_k + \eta)$.

We approximate the theoretical mean $\mathbb{E}\left[\Delta N \middle| N = N_k + \frac{\eta}{2}\right]$ with the empirical mean $\left\langle \Delta N \middle| N = N_k + \eta_i, 0 \leq \eta_i < \eta, \hat{S}_k \right\rangle$ and the theoretical variance $\mathbb{V}\left[\Delta N \middle| N = N_k + \frac{\eta}{2}\right]$ the empirical variance $\sigma^2\left[\Delta N \middle| N = N_k + \eta_i, 0 \leq \eta_i < \eta, \hat{S}_k\right]$ obtained from simulation of S cell number trajectories. Recall that $\hat{S}_k$ denotes the number of population size $N_k + \eta_i$ landing in bin k. These sample sizes $\hat{S}_k, k = 1, 2, \ldots, k_{\max}$, are not necessarily equal to each other or equal to the number of cell number trajectories S, which is pre-determined and independent of the bin size η. Different bin sizes η result in different sets of $\hat{S}_k, k = 1, 2, \ldots, k_{\max}$. With the same bin size η, different simulations may also result in different sets of cell number values and hence different sets of $\hat{S}_k, k = 1, 2, \ldots, k_{\max}$. It is well-known that as the larger the sample size $\hat{S}_k$, the smaller the estimation errors [28].

In this section, we analyze how the bin size influences distributions of estimation errors of birth and death rates. In particular, we compute the theoretical means and variances of errors as functions of bin size η. We use the notation N for cell number to be consistent with the mathematical model discussed in Section 3.2. A summary of notations can be

found in Section A.2.3.

A.2.1 Theoretical Mean and Variance of Cell Number Increment as Functions of Bin Size

As mentioned above, our estimation of the birth and death rates corresponding to $N = N_k + \frac{\eta}{2}$ uses the empirical mean $\left\langle \Delta N \middle| N = N_k + \eta_i, 0 \leq \eta_i < \eta, \hat{S}_k \right\rangle$ and empirical variance $\sigma^2\Big[\Delta N | N = N_k + \eta_i] \Big| 0 \leq \eta_i < \eta, \hat{S}_k\Big]$. The theoretical means and variances of the estimation errors involves the theoretical mean $\mathbb{E}\Big[\Delta N \Big| N = N_k + U, U \sim \text{Unif}[0, \eta)\Big]$ and theoretical variance $\mathbb{V}\Big[\Delta N \Big| N = N_k + U, U \sim \text{Unif}[0, \eta)\Big]$, as shown in Section 3.3.4. In this subsection, we analyze how the bin size η influences these theoretical mean and variance. We present the analysis for nonnegative birth rates, that is, in which we can drop the $\max$ function in Equation (3.6), as the birth rates are always positive in our simulated datasets.

Theoretical mean:

$$\mathbb{E}\Big[\Delta N \Big| N = N_k + U, U \sim \text{Unif}[0, \eta)\Big] = \mathbb{E}\Big[\mathbb{E}[\Delta N | N = N_k + U] \Big| U \sim \text{Unif}[0, \eta)\Big] \tag{A.1}$$

$$= \mathbb{E}\Big[(b_{N_k+U} - d_{N_k+U})(N_k + U)\Delta t \Big| U \sim \text{Unif}[0, \eta)\Big] \tag{A.2}$$

$$= \mathbb{E}\Big[(r - \frac{r}{K}N_k - \frac{r}{K}U)(N_k + U)\Delta t \Big| U \sim \text{Unif}[0, \eta)\Big] \tag{A.3}$$

$$= \mathbb{E}\Big[(r - \frac{r}{K}N_k)N_k\Delta t \Big| U \sim \text{Unif}[0, \eta)\Big] - \frac{r}{K}N_k\Delta t\mathbb{E}\Big[U \Big| U \sim \text{Unif}[0, \eta)\Big] \tag{A.4}$$

$$+ (r - \frac{r}{K}N_k)\Delta t\mathbb{E}\Big[U \Big| U \sim \text{Unif}[0, \eta)\Big] - \frac{r}{K}\Delta t\mathbb{E}\Big[U^2 \Big| U \sim \text{Unif}[0, \eta)\Big]$$

$$= \mathbb{E}[\Delta N \mid N = N_k] + (r - 2\frac{r}{K}N_k)\Delta t\frac{\eta}{2} - \frac{r}{K}\Delta t\frac{\eta^2}{3}. \tag{A.5}$$

Theoretical variance:

$$\begin{aligned}&\mathbb{V}\Big[\Delta N\Big|N=N_k+U,U\sim\text{Unif}[0,\eta)\Big]\\=&\mathbb{E}\Big[\Delta N^2\Big|N=N_k+U,U\sim\text{Unif}[0,\eta)\Big]-\Big(\mathbb{E}\Big[\Delta N\Big|N=N_k+U,U\sim\text{Unif}[0,\eta)\Big]\Big)^2,\end{aligned}\tag{A.6}$$

where

$$\mathbb{E}\Big[\Delta N^2\Big|N=N_k+U,U\sim\text{Unif}[0,\eta)\Big]=\mathbb{E}\Big[\mathbb{E}[\Delta N^2|N=N_k+U]\Big|U\sim\text{Unif}[0,\eta)\Big]$$

$$=\mathbb{E}\left[\mathbb{V}[\Delta N|N=N_k+U]+\Big(\mathbb{E}[\Delta N|N=N_k+U]\Big)^2\Big|U\sim\text{Unif}[0,\eta)\right]\tag{A.7}$$

$$=\mathbb{E}\Big[\mathbb{V}[\Delta N|N=N_k+U]\Big|U\sim\text{Unif}[0,\eta)\Big]+\mathbb{E}\Big[\Big(\mathbb{E}[\Delta N|N=N_k+U]\Big)^2\Big|U\sim\text{Unif}[0,\eta)\Big],\tag{A.8}$$

and

$$\mathbb{E}\Big[\mathbb{V}[\Delta N|N = N_k + U]\Big|U \sim \text{Unif}[0,\eta]\Big]$$

$$= \mathbb{E}\Big[(b_{N_k+U} + d_{N_k+U})(N_k + U)\Delta t\Big|U \sim \text{Unif}[0,\eta)\Big] \tag{A.9}$$

$$= \mathbb{E}\Big[(b_0 + d_0 + (1-2\gamma)\frac{r}{K}N_k + (1-2\gamma)\frac{r}{K}U)(N_k + U)\Delta t\Big|U \sim \text{Unif}[0,\eta)\Big] \tag{A.10}$$

$$= \mathbb{E}\Big[(b_0 + d_0 + (1-2\gamma)\frac{r}{K}N_k)N_k\Delta t\Big|U \sim \text{Unif}[0,\eta)\Big] \tag{A.11}$$

$$+ \mathbb{E}\Big[(b_0 + d_0 + (1-2\gamma)\frac{r}{K}N_k)U\Delta t\Big|U \sim \text{Unif}[0,\eta)\Big] \tag{A.12}$$

$$+ \mathbb{E}\Big[(1-2\gamma)\frac{r}{K}N_kU\Delta t\Big|U \sim \text{Unif}[0,\eta)\Big] + \mathbb{E}\Big[(1-2\gamma)\frac{r}{K}U^2\Delta t\Big|U \sim \text{Unif}[0,\eta)\Big]$$

$$= \mathbb{V}[\Delta N_k] + \Big(b_0 + d_0 + 2(1-2\gamma)\frac{r}{K}N_k\Big)\Delta t\mathbb{E}[U|U \sim \text{Unif}[0,\eta)] \tag{A.13}$$

$$+ (1-2\gamma)\frac{r}{K}\Delta t\mathbb{E}[U^2|U \sim \text{Unif}[0,\eta)] \tag{A.14}$$

$$= \mathbb{V}[\Delta N_k] + \Big(b_0 + d_0 + 2(1-2\gamma)\frac{r}{K}N_k\Big)\Delta t\frac{\eta}{2} + (1-2\gamma)\frac{r}{K}\Delta t\frac{\eta^2}{3}, \tag{A.15}$$

and

$$\mathbb{E}\Big[(b_{N_k+U} - d_{N_k+U})^2(N_k + U)^2\Delta t^2\Big|U \sim \text{Unif}[0,\eta)\Big]$$

$$= \Delta t^2\mathbb{E}\Big[\Big((b_0 - d_0)(N_k + U) - \frac{r}{K}(N_k + U)^2\Big)^2\Big|U \sim \text{Unif}[0,\eta)\Big] \tag{A.16}$$

$$= \Delta t^2\mathbb{E}\Big[r^2(N_k + U)^2 - 2\frac{r^2}{K}(N_k + U)^3 + \frac{r^2}{K^2}(N_k + U)^4\Big|U \sim \text{Unif}[0,\eta)\Big] \tag{A.17}$$

$$= \Delta t^2r^2\mathbb{E}\Big[(N_k + U)^2\Big|U \sim \text{Unif}[0,\eta)\Big] - 2\frac{r^2}{K}\Delta t^2\mathbb{E}\Big[(N_k + U)^3\Big|U \sim \text{Unif}[0,\eta)\Big] \tag{A.18}$$

$$+ \frac{r^2}{K^2}\Delta t^2\mathbb{E}\Big[(N_k + U)^4\Big|U \sim \text{Unif}[0,\eta)\Big]$$

$$= \Delta t^2r^2\frac{(N_k + \eta)^3 - N_k^3}{3\eta} - 2\frac{r^2}{K}\Delta t^2\frac{(N_k + \eta)^4 - N_k^4}{4\eta} + \frac{r^2}{K^2}\Delta t^2\frac{(N_k + \eta)^5 - N_k^5}{5\eta}. \tag{A.19}$$

Therefore,

$$
\begin{aligned}
&\mathbb{V}\Big[\Delta N \Big| N = N_k + U, U \sim \text{Unif}[0,\eta)\Big] \\
=&\mathbb{V}[\Delta N_k] + \Big(b_0 + d_0 + 2(1-2\gamma)\frac{r}{K}N_k\Big)\Delta t\frac{\eta}{2} + (1-2\gamma)\frac{r}{K}\Delta t\frac{\eta^2}{3} \\
&+ \Delta t^2 r^2 \frac{(N_k+\eta)^3 - N_k^3}{3\eta} - 2\frac{r^2}{K}\Delta t^2\frac{(N_k+\eta)^4 - N_k^4}{4\eta} + \frac{r^2}{K^2}\Delta t^2\frac{(N_k+\eta)^5 - N_k^5}{5\eta} \\
&- \left(\mathbb{E}[\Delta N \mid N = N_k] + (r - 2\frac{r}{K}N_k)\Delta t\frac{\eta}{2} - \frac{r}{K}\Delta t\frac{\eta^2}{3}\right)^2.
\end{aligned}
\tag{A.20}
$$

In Figure A.1, we compare the theoretical mean $\mathbb{E}\Big[\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0,\eta)\Big]$ that we just computed with the theoretical mean $\mathbb{E}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]$ and the empirical mean $\Big\langle \Delta N\Big|N = N_k + U, U \sim \text{Unif}[0,\eta)\Big\rangle$ using data from a simulation of $S = 100$ cell number trajectories. Similarly, we also compare the population variance $\mathbb{V}\Big[\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0,\eta)\Big]$ that we just computed with the theoretical variance $\mathbb{V}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]$ and the empirical variance $\sigma^2\Big[\Delta N\Big|N = N_k + \eta_i, 0 \le \eta_i < \eta\Big]$ using data from a simulation of $S = 100$ cell number trajectories.

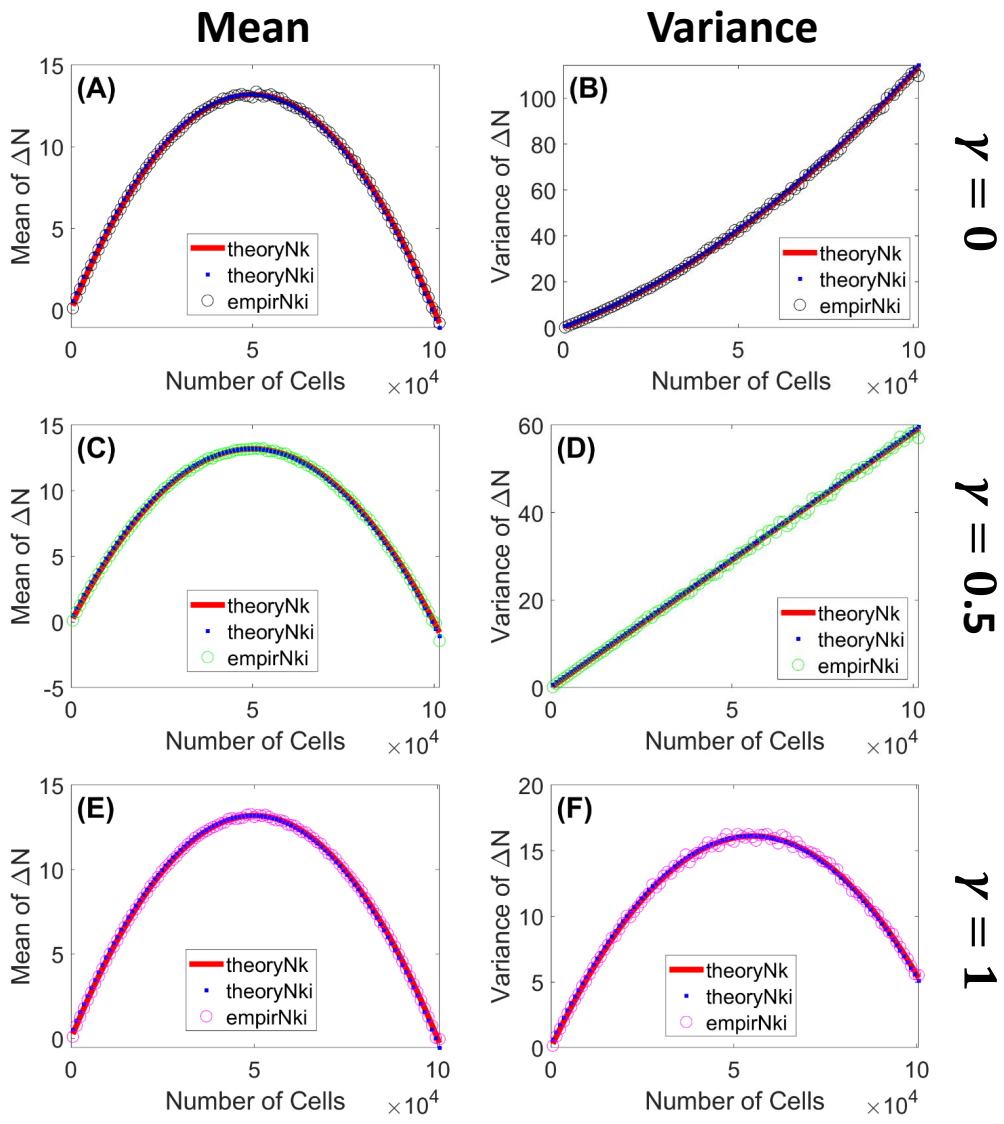

Figure A.1: **Theoretical mean and variance of cell number increments ΔN as functions of population size are well-aligned with empirical mean and variance**. The statistics are computed using carrying capacity $K = 10^5$ and bin size $\eta = 10^3$. In **(A, C, E)**, we compare the theoretical mean $\mathbb{E}\left[\Delta N \middle| N = N_k + U, U \sim \text{Unif}[0, \eta)\right]$ with the theoretical mean $\mathbb{E}\left[\Delta N \middle| N = N_k + \frac{\eta}{2}\right]$ and the empirical mean $\left\langle \Delta N \middle| N = N_k + \eta_i, 0 \leq \eta_i < \eta \right\rangle$ using data from a simulation of $S = 100$ cell number trajectories. In **(B, E, F)**, we compare the theoretical variance $\mathbb{V}\left[\Delta N \middle| N = N_k + U, U \sim \text{Unif}[0, \eta)\right]$ with theoretical variance $\mathbb{V}\left[\Delta N \middle| N = N_k + \frac{\eta}{2}\right]$ and the empirical variance $\sigma^2\left[\Delta N \middle| N = N_k + \eta_i, 0 \leq \eta_i < \eta\right]$ using data from a simulation of $S = 100$ cell number trajectories. **(A, B)**: $\gamma = 0$ (black color); **(C, D)**: $\gamma = 0.5$ (green color); **(E, F)**: $\gamma = 0.5$ (magenta color). Red lines (-) denote theoretical statistics (i.e. mean and variance) for $N = N_k + \frac{\eta}{2}$; blue squares (□) denote theoretical statistics for $N = N_k + U, U \sim \text{Unif}[0, \eta)$; circles (○) denote empirical statistics for $N = N_k + \eta_i$ with $i = 1, \ldots, \hat{S}_k$.

A.2.2 Errors of Birth and Death Rate Estimation as Functions of Bin Size

In this section we consider the effect of bin size on the accuracy with which we can estimate the birth and death rates. Thus we compare the theoretical mean and variance of the population increment, given that a point of the trajectory lies within a given bin, versus the empirical mean and variance obtained from simulation with a finite sample size. We use $\mathcal{E}$ to represent expected differences in these errors. Define

$$\mathcal{E}_{k\text{mean}} := \mathbb{E}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big] - \Big\langle \Delta N \Big| N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k \Big\rangle, \tag{A.21}$$

$$\mathcal{E}_{k\text{var}} := \mathbb{V}\Big[\Delta N \Big| N = N_k + \frac{\eta}{2}\Big] - \sigma^2\Big[\Delta N \Big| N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k\Big]. \tag{A.22}$$

The errors in estimating the birth and death rates corresponding to $N = N_k + \dfrac{\eta}{2}$ are

$$\mathcal{E}_{k\text{birth}} = \frac{\mathcal{E}_{k\text{var}} + \mathcal{E}_{k\text{mean}}}{2\Delta t} \quad \text{and} \quad \mathcal{E}_{k\text{death}} = \frac{\mathcal{E}_{k\text{var}} - \mathcal{E}_{k\text{mean}}}{2\Delta t}. \tag{A.23}$$

The theoretical means of the errors over *all* realizations η_i of the iid uniform random variable $U \sim \text{Unif}[0, \eta)$ are

$$\mathbb{E}\Big[\mathcal{E}_{k\text{birth}}\Big] = \frac{\mathbb{E}\Big[\mathcal{E}_{k\text{var}}\Big] + \mathbb{E}\Big[\mathcal{E}_{k\text{mean}}\Big]}{2\Delta t} \quad \text{and} \quad \mathbb{E}\Big[\mathcal{E}_{k\text{death}}\Big] = \frac{\mathbb{E}\Big[\mathcal{E}_{k\text{var}}\Big] - \mathbb{E}\Big[\mathcal{E}_{k\text{mean}}\Big]}{2\Delta t}. \tag{A.24}$$

The theoretical variances of the errors over *all* realizations η_i of U are

$$\mathbb{V}\Big[\mathcal{E}_{k\text{birth}}\Big] = \mathbb{V}\Big[\mathcal{E}_{k\text{death}}\Big] = \frac{\mathbb{V}\Big[\mathcal{E}_{k\text{var}}\Big] + \mathbb{V}\Big[\mathcal{E}_{k\text{mean}}\Big]}{4\Delta t^2}. \tag{A.25}$$

We analyze how the bin size η influences these analytical expected values and variances of errors $\mathbb{E}\Big[\mathcal{E}_{k\text{mean}}\Big]$, $\mathbb{E}\Big[\mathcal{E}_{k\text{var}}\Big]$, $\mathbb{V}\Big[\mathcal{E}_{k\text{var}}\Big]$, and $\mathbb{V}\Big[\mathcal{E}_{k\text{mean}}\Big]$.

Treating the samples of $\Big(\Delta N\Big|N = N_k + U, U \sim \text{Unif}[), \eta)\Big)$ as if they were identically and independently distributed, the expected value of the sample mean is equal to the theoretical mean. Therefore,

$$\mathbb{E}\Big[\mathcal{E}_{k\text{mean}}\Big] = \mathbb{E}\Bigg[\underbrace{\mathbb{E}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]}_{\text{independent of } \eta_i}\Bigg] - \mathbb{E}\Bigg[\Big\langle \Delta N\Big|N = N_k + \eta_i, 0 \leq \eta_i < \eta, \hat{S}_k\Big\rangle\Bigg] \tag{A.26}$$

$$= \mathbb{E}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big] - \mathbb{E}\Big[\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0, \eta)\Big], \tag{A.27}$$

where

$$\begin{aligned}
&\mathbb{E}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big] \\
&= \Big(b_{N_k+(\eta/2)} - d_{N_k+(\eta/2)}\Big)\Big(N_k + \frac{\eta}{2}\Big)\Delta t && \text{(A.28)} \\
&= \Big(r - \frac{r}{K}N_k - \frac{r}{K}\frac{\eta}{2}\Big)\Big(N_k + \frac{\eta}{2}\Big)\Delta t && \text{(A.29)} \\
&= \Big(r - \frac{r}{K}N_k\Big)N_k\Delta t + \Big(r - \frac{r}{K}N_k\Big)\Delta t\frac{\eta}{2} - \frac{r}{K}N_k\Delta t\frac{\eta}{2} - \frac{r}{K}\frac{\eta^2}{4}\Delta t && \text{(A.30)} \\
&= \mathbb{E}[\Delta N|N = N_k] + \Big(r - 2\frac{r}{K}N_k\Big)\Delta t\frac{\eta}{2} - \frac{r}{K}\Delta t\frac{\eta^2}{4}. && \text{(A.31)}
\end{aligned}$$

Hence,

$$\mathbb{E}\Big[\mathcal{E}_{k\text{mean}}\Big] = \mathbb{E}[\Delta N | N = N_k] + \Big(r - 2\frac{r}{K}N_k\Big)\Delta t\frac{\eta}{2} - \frac{r}{K}\Delta t\frac{\eta^2}{4} \tag{A.32}$$

$$- \mathbb{E}[\Delta N \mid N = N_k] - (r - 2\frac{r}{K}N_k)\Delta t\frac{\eta}{2} + \frac{r}{K}\Delta t\frac{\eta^2}{3}$$

$$= \frac{1}{12}\frac{r}{K}\Delta t\eta^2. \tag{A.33}$$

We observe that the expected error $\mathbb{E}\Big[\mathcal{E}_{k\text{mean}}\Big]$ in approximating the true mean $\mathbb{E}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]$ for each bin k is independent of k and is increasing quadratically for $\eta > 0$. If we write the expected error $\mathbb{E}\Big[\mathcal{E}_{k\text{mean}}\Big]$ as $\frac{1}{12}\Big(r\Delta t\Big)\Big(\frac{\eta}{K}\Big)\eta$, then we see that the expected error depends on the ratio $\Big(\frac{\eta}{K}\Big)$, which shows how big the bin size is relative to the system size (i.e. carrying capacity K), and also depends on the product $r\Delta t$, which can be interpreted roughly as the *per capita* change in cell number $\Big(\frac{\Delta N}{N}\Big)$ after Δt. The higher these ratios are, the higher expected error is. Looking from a different angle, the expected error $\mathbb{E}\Big[\mathcal{E}_{k\text{mean}}\Big]$ can be written as $\Big(\frac{r\Delta t}{K}\Big)\Big(\frac{1}{12}\eta^2\Big)$. This shows that the expected error depends on $\Big(\frac{1}{12}\eta^2\Big)$, which is the variance of the random variable η_i, and how big the *per capita* change in cell number $r\Delta t$ after Δt is relative to the system size K. This observation suggests that it may be harder to estimate the cell number increments with high accuracy for fast-reproducing cell types. Further analysis on the relation between r and K would be interesting for future work, since existing work such as [73] shows that the product rK can influence the evolution of antibiotic-resistant bacterial genomes.

We assume that the samples of $\Big(\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0, \eta)\Big)$ are independently and identically distributed, so the expected value of the sample variance is equal to the

population variance. Therefore,

$$\mathbb{E}\Big[\mathcal{E}_{k\text{var}}\Big] = \mathbb{E}\Bigg[\underbrace{\mathbb{V}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]}_{\text{independent of } \eta_i}\Bigg] - \mathbb{E}\Bigg[\sigma^2\Big[\Delta N\Big|N = N_k + \eta_i, 0 \le \eta_i < \eta, \hat{S}_k\Big]\Bigg] \tag{A.34}$$

$$= \mathbb{V}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big] - \mathbb{V}\Big[\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0,\eta)\Big], \tag{A.35}$$

where

$$\begin{aligned}
&\mathbb{V}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big] \\
=&\Big(b_0 + d_0 + (1-2\gamma)\frac{r}{K}N_k + (1-2\gamma)\frac{r}{K}\frac{\eta}{2}\Big)\Big(N_k + \frac{\eta}{2}\Big)\Delta t &\text{(A.36)}\\
=&\mathbb{V}[\Delta N_k] + \Big(b_0 + d_0 + 2(1-2\gamma)\frac{r}{K}N_k\Big)\Delta t\frac{\eta}{2} + (1-2\gamma)\frac{r}{K}\Delta t\frac{\eta^2}{4}. &\text{(A.37)}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}\Big[\mathcal{E}_{k\text{var}}\Big] =& -(1-2\gamma)\frac{r}{K}\Delta t\frac{\eta^2}{12} &\text{(A.38)}\\
&+ \Delta t^2 r^2\frac{(N_k+\eta)^3 - N_k^3}{3\eta} - 2\frac{r^2}{K}\Delta t^2\frac{(N_k+\eta)^4 - N_k^4}{4\eta} + \frac{r^2}{K^2}\Delta t^2\frac{(N_k+\eta)^5 - N_k^5}{5\eta}\\
&- \left(\mathbb{E}[\Delta N \mid N = N_k] + (r - 2\frac{r}{K}N_k)\Delta t\frac{\eta}{2} - \frac{r}{K}\Delta t\frac{\eta^2}{3}\right)^2.
\end{aligned}$$

Now, we compute the theoretical variances $\mathbb{V}\Big[\mathcal{E}_{k\text{mean}}\Big]$ and $\mathbb{V}\Big[\mathcal{E}_{k\text{var}}\Big]$ over *all* realizations of η_i. We assume the samples of $\Big(\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0,\eta)\Big)$ are identically distributed, so the variance of the sample mean is equal to the population variance divided

by the sample size. Therefore,

$$\mathbb{V}\Big[\mathcal{E}_{k\text{mean}}\Big] = \mathbb{V}\Bigg[\underbrace{\mathbb{E}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]}_{\text{independent of } \eta_i}\Bigg] + \mathbb{V}\Bigg[\Big\langle \Delta N\Big|N = N_k + U, 0 \leq \eta_i < \eta, \hat{S}_k\Big\rangle\Bigg] \tag{A.39}$$

$$= \mathbb{V}\Bigg[\Big\langle \Delta N\Big|N = N_k + \eta_i, 0 \leq \eta_i < \eta, \hat{S}_k\Big\rangle\Bigg] \tag{A.40}$$

$$= \frac{\mathbb{V}\Big[\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0, \eta)\Big]}{\hat{S}_k}. \tag{A.41}$$

As mentioned above, the samples of $\Big(\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0, \eta)\Big)$ are independently and identically distributed. For computation convenience here, we approximate the binomial distribution of these samples with the Gaussian distribution with the empirical mean and variance as discussed in Section 3.3.2. We still use the notation N instead of X here to be consistent with the other statistics computed above. With this approximation, the theoretical variance of the empirical variance is equal to two times the theoretical variance squared divided by the sample size minus one. Therefore,

$$\mathbb{V}\Big[\mathcal{E}_{k\text{var}}\Big] = \mathbb{V}\Bigg[\underbrace{\mathbb{V}\Big[\Delta N\Big|N = N_k + \frac{\eta}{2}\Big]}_{\text{independent of } \eta_i}\Bigg] + \mathbb{V}\Bigg[\sigma^2\Big[\Delta N\Big|N = N_k + \eta_i, 0 \leq \eta_i < \eta, \hat{S}_k\Big]\Bigg] \tag{A.42}$$

$$= \mathbb{V}\Bigg[\sigma^2\Big[\Delta N\Big|N = N_k + \eta_i, 0 \leq \eta_i < \eta, \hat{S}_k\Big]\Bigg] \tag{A.43}$$

$$= \frac{2\Big(\mathbb{V}\Big[\Delta N\Big|N = N_k + U, U \sim \text{Unif}[0, \eta)\Big]\Big)^2}{\hat{S}_k - 1}. \tag{A.44}$$

The theoretical variance $\mathbb{V}\left[\Delta N \middle| N = N_k + U, U \sim \text{Unif}[0, \eta)\right]$ is given by Equation (A.20).

Using the $\mathbb{E}\left[\mathcal{E}_{k\text{mean}}\right]$, $\mathbb{E}\left[\mathcal{E}_{k\text{var}}\right]$, $\mathbb{V}\left[\mathcal{E}_{k\text{mean}}\right]$, and $\mathbb{V}\left[\mathcal{E}_{k\text{var}}\right]$ that we just computed, we obtain the theoretical means and variances of the errors in estimating birth and death rates corresponding to $N = N_k + \frac{\eta}{2}$ for all $k = 1, 2, \ldots, k_{\max}$ using Equations (A.23) and (A.24).

In Figure 3.3, we compare the l_2-norm of the theoretical means and variances of the errors and compare them with the l_2-norm of the empirical errors (i.e. realizations of the error random variables) computed using data from a simulation of $S = 100$ cell number trajectories. To computed the theoretical variances of the errors shown in Figure 3.3, we use the empirical sample sizes $\hat{S}_k, k = 1, 2, \ldots, k_{\max}$, from the same data simulation.

We observe that as the bin size η increases, the theoretical means of the errors increase, the theoretical variances (or standard deviations) of the errors decreases, and the empirical errors balance between the theoretical means and variances (or standard deviations) and have convex quadratic shapes. The theoretical means of the errors reflect the differences between ΔN at one point $\left(N = N_k + \frac{\eta}{2}\right)$ and ΔN at multiple points $\left(N = N_k + \eta_i, 0 \leq \eta_i < \eta\right)$; the smaller the bin size, the closer multiple points are to one point, so the error is smaller (for example, Equation (A.33) shows that the expected errors in estimating the mean of cell number increments are $(r\Delta t/12K)\eta^2$). However, if the bin is too small, then there are not enough samples to estimate theoretical statistics with empirical statistics with accuracy. The theoretical variances of errors involves sample sizes; the bigger the bin size, the more samples we have. These two competing effects of bin size result in the empirical errors being intermediate values between the two theoretical statistics (means and variances) of the estimation errors. The optimal bin size reflects a balancing of these two

effects. When the bin size is smaller than the optimal bin size, the sample error coincides with the sum of the expected error and the standard deviation of the error. When the bin size is bigger than the optimal bin size, this relationship breaks down, which may reflect growing inaccuracy of our approximation that the trajectory points are uniformly and i.i.d. within each bin.

A.2.3 Notation

N	:	discrete cell number random variable
t_0 and t_T	:	deterministic initial and final times respectively
η	:	deterministic bin size
k	:	bin index, $k = 1, 2, \ldots, k_{\max}$
U	:	uniformly distributed random variable, $\Big(N = N_k + U\Big) \in [N_k, N_k + \eta)$
η_i	:	realization of the random variable U
S	:	number of cell number trajectories/time series
$\hat{S}_k$	:	number of samples of $\Delta N := N(t + \Delta t) - N(t)$ in bin $[N_k, N_k + \eta)$
$\mathbb{E}[\cdot]$	:	theoretical mean
$\langle \cdot \rangle$	:	empirical mean
$\mathbb{V}[\cdot]$	:	theoretical variance
$\sigma^2[\cdot]$	:	empirical variance
$\mathcal{E}_{[\cdot]}$	:	error

Bibliography

[1] WC Allee and Edith S Bowen. Studies in animal aggregations: Mass protection against colloidal silver among goldfishes. *Journal of Experimental Zoology*, 61(2):185–207, 1932.

[2] Linda Allen. *An Introduction to Stochastic Processes with Applications to Biology*. CRC Press, 2010.

[3] Nicolas Bacaër. *A Short History of Mathematical Population Dynamics*. Springer Science & Business Media, 2011.

[4] Bruce C Baguley. Multiple drug resistance mechanisms in cancer. *Molecular Biotechnology*, 46(3):308–316, 2010.

[5] Norman TJ Bailey. *The Elements of Stochastic Processes with Applications to the Natural Sciences*, volume 25. John Wiley & Sons, 1991.

[6] Jacob Beal, Natalie G. Farny, Traci Haddock-Angelli, Vinoo Selvarajah, Geoff S. Baldwin, Russell Buckley-Taylor, Markus Gershater, Daisuke Kiga, John Marken, Vishal Sanchania, Abigail Sison, Christopher T. Workman, and the iGEM Interlab Study Contributors. Robust estimation of bacterial cell count from optical density. *Communications Biology*, 3(512), 2020.

[7] Richard Bellman. A Markovian Decision Process. *Journal of Mathematics and Mechanics*, 6(5):679–684, 1957.

[8] Haym Benaroya, Seon Mi Han, and Mark Nagurka. *Probability Models in Engineering and Science*, volume 192. CRC press, 2005.

[9] Amiya Ranjan Bhowmick, Bapi Saha, Joydev Chattopadhyay, Santanu Ray, and Sabyasachi Bhattacharya. Cooperation in species: Interplay of population regulation and extinction through global population dynamics database. *Ecological Modelling*, 312:150–165, 2015.

[10] Michael B Bonsall and Alan Hastings. Demographic and environmental stochasticity in predator–prey metapopulation dynamics. *Journal of Animal Ecology*, 73(6):1043–1055, 2004.

[11] Rebecca K Borchering and Scott A McKinley. Continuum approximation of invasion probabilities. *Multiscale Modeling & Simulation*, 16(2):551–582, 2018.

[12] Asher Brauner, Ofer Fridman, Orit Gefen, and Nathalie Q Balaban. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nature Reviews Microbiology*, 14(5):320–330, 2016.

[13] Catherine Calder, Michael Lavine, Peter Müller, and James S Clark. Incorporating multiple sources of stochasticity into dynamic population models. *Ecology*, 84(6):1395–1402, 2003.

[14] Daniela Calvetti and Erkki Somersalo. *An Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing*, volume 2. Springer Science & Business Media, 2007.

[15] Scott P Carroll, Andrew P Hendry, David N Reznick, and Charles W Fox. Evolution on ecological time scales, 2007.

[16] George Casella and Roger L Berger. *Statistical Inference*. Cengage Learning, 2021.

[17] Jessica Coates, Bo Ryoung Park, Dai Le, Emrah Şimşek, Waqas Chaudhry, and Minsu Kim. Antibiotic-induced population fluctuations and stochastic clearance of bacteria. *eLife*, 7:e32976, 2018.

[18] Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An Introduction to Statistical Modeling of Extreme Values*, volume 208. Springer, 2001.

[19] Forrest W. Crawford, Vladimir N. Minin, and Marc A. Suchard. Estimation for general birth-death processes. *Journal of the American Statistical Association*, 109:730–747, 2014.

[20] Sophie E Darch, Stuart A West, Klaus Winzer, and Stephen P Diggle. Density-dependent fitness benefits in quorum-sensing bacterial populations. *Proceedings of the National Academy of Sciences*, 109(21):8259–8263, 2012.

[21] Helen C. Davison, Mark E.J Woolhouse, and J. Chris Low. What is antibiotic resistance and how can we measure it? *Trends in Microbiology*, 8(12):554–559, 2000.

[22] Michael Doebeli, Yaroslav Ispolatov, and Burt Simon. Towards a mechanistic foundation of evolutionary theory. *eLife*, 2017.

[23] John M. Drake and Andrew M. Kramer. Allee effects. *Nature Education Knowledge*, 3:2, 2011.

[24] Rena Emond, Jason I Griffiths, Vince Kornél Grolmusz, Rachel S Sousa, Andrea H Bild, and Frederick R Adler. Ecological interactions in breast cancer: Cell facilitation promotes growth and survival under drug pressure. *bioRxiv*, 2021.

[25] Steinar Engen, Øyvind Bakke, and Aminul Islam. Demographic and environmental stochasticity-concepts and definitions. *Biometrics*, pages 840–846, 1998.

[26] Steinar Engen and Bernt-Erik Sæther. Stochastic population models: Some concepts, definitions and results. *Oikos*, pages 345–352, 1998.

[27] Nathan Farrokhian, Jeff Maltas, Mina Dinh, Arda Durmaz, Patrick Ellsworth, Masahiro Hitomi, Erin McClure, Andriy Marusyk, Artem Kaznatcheev, and Jacob G Scott. Measuring competitive exclusion in non-small cell lung cancer. *bioRxiv*, 2022.

[28] Jeremy Ferlic. *Quantitative Approaches to Cancer and Cellular Differentiation*. PhD thesis, Harvard University, Graduate School of Arts and Sciences, 2019.

[29] Jasmine Foo and Franziska Michor. Evolution of resistance to anti-cancer therapy during general dosing schedules. *Journal of Theoretical Biology*, 263(2):179–188, 2010.

[30] K. Francois, F. Devlieghere, A.R. Standaert, A.H. Geeraerd, I. Cools, J.F. Van Impe, and J. J. Debevere. Environmental factors influencing the relationship between optical density and cell count for listeria monocytogenes. *Journal of Applied Microbiology*, 99:503–1515, 2005.

[31] Antoine Frenoy and Sebastian Bonhoeffer. Death and population dynamics affect mutation rate estimates and evolvability under stress in bacteria. *Plos Biology*, 16(5), 2018.

[32] Crispin Gardiner. *Stochastic Methods*, volume 4. Springer Berlin, 2009.

[33] Philip Gerlee. The model muddle: in search of tumor growth laws. *Cancer Research*, 73(8):2407–2411, 2013.

[34] Philip Gerlee. Weak selection and the separation of eco-evo time scales using perturbation analysis. *Bulletin of Mathematical Biology*, 84(5):1–19, 2022.

[35] Daniel T Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, 115(4):1716–1733, 2001.

[36] Vishhvaan Gopalakrishnan, Nikhil P Krishnan, Erin McClure, Julia Pelesko, Dena Crozier, Kyle J Card, Drew FK Williamson, Nathan Webster, Daniel Nichol, Soumyajit Mandal, et al. A low-cost, open source, self-contained bacterial evolutionary bioreactor (eve). *bioRxiv*, page 729434, 2020.

[37] Alastair R Hall. *Generalized Method of Moments*. Oxford University Press, 2004.

[38] Magnus J Haughey, Aleix Bassolas, Sandro Sousa, Ann-Marie Baker, Trevor A Graham, Vincenzo Nicosia, and Weini Huang. First passage time analysis of spatial mutation patterns reveals evolutionary dynamics of pre-existing resistance in colorectal cancer. *bioRxiv*, 2022.

[39] Gustaf Hendeby and Fredrik Gustafsson. On nonlinear transformations of Gaussian distributions. *Technical Report from Automatic Control*, 2007.

[40] Mark A Hixon and Darren W Johnson. *Density Dependence and Independence. In: Encyclopedia of Life Sciences (ELS)*. John Wiley and Sons, Ltd: Chichester, 2009.

[41] Linh Huynh, Jacob G. Scott, and Peter J. Thomas. Inferring density-dependent population dynamics mechanisms through rate disambiguation for logistic birth-death processes. Submitted, 2022.

[42] Yoh Iwasa, Franziska Michor, and Martin A Nowak. Evolutionary dynamics of escape from biomedical intervention. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1533):2573–2578, 2003.

[43] A.L. Jesen. Simple models for exploitive and inference competition. *Ecological Modelling*, 35:113–121, 1987.

[44] Andrew R. Kanarek and Colleen T. Webb. Allee effects, adaptive evolution, and invasion success. *Evolutionary Applications*, 3:122–135, 2010.

[45] Jason Karslake, Jeff Maltas, Peter Brumm, and Kevin B Wood. Population density modulates drug inhibition and gives rise to potential bistability of treatment outcomes for bacterial infections. *PLoS Computational Biology*, 12(10):e1005098, 2016.

[46] Artem Kaznatcheev, Jeffrey Peacock, David Basanta, Andriy Marusyk, and Jacob G Scott. Fibroblasts and alectinib switch the evolutionary games played by non-small cell lung cancer. *Nature Ecology & Evolution*, 3(3):450–456, 2019.

[47] Marek Kimmel and David Axelrod. *Branching Processes in Biology*. Interdisciplinary Applied Mathematics, 2015.

[48] Ron Kohavi, David H Wolpert, et al. Bias plus variance decomposition for zero-one loss functions. In *ICML*, volume 96, pages 275–83, 1996.

[49] Natalia Komarova. Stochastic modeling of drug resistance in cancer. *Journal of Theoretical Biology*, 239:351–366, 2006.

[50] Russell Lande, Steinar Engen, Bernt-Erik Saether, et al. *Stochastic population dynamics in ecology and conservation*. Oxford University Press, 2003.

[51] Orit Lavi, Michael M Gottesman, and Doron Levy. The dynamics of drug resistance: a mathematical perspective. *Drug Resistance Updates*, 15(1-2):90–97, 2012.

[52] Sean D Lawley. Extreme first-passage times for random walks on networks. *Physical Review E*, 102(6):062118, 2020.

[53] Sean D Lawley. Extreme first passage times of piecewise deterministic Markov processes. *Nonlinearity*, 34(5):2750, 2021.

[54] Se Yoon Lee. Gibbs sampler and coordinate ascent variational inference: A set-theoretical review. *Communications in Statistics-Theory and Methods*, 51(6):1549–1568, 2022.

[55] X. Lei, W. Tian, H. Zhu, T. Chen, and P. Ao. Biological sources of intrinsic and extrinsic noise in cI expression of Lysogenic Phage Lambda. *Scientific Reports*, 5:1–12, 2015.

[56] Thomas Lenormand, Denis Roze, and François Rousset. Stochasticity in evolution. *Trends in Ecology & Evolution*, 24(3):157–165, 2009.

[57] Ethan Levien, Trevor GrandPre, and Ariel Amir. Large deviation principle linking lineage statistics to fitness in microbial populations. *Physical Review Letters*, 125:048102, 2020.

[58] Yiyi Liu and Forrest W. Crawford. Estimating dose-specific cell division and apoptosis rates from chemo-sensitivity experiments. *Scientific Reports*, 8:2705, 2018.

[59] M.A. Lobritz, P. Belenky, C.B.M Porter, A. Gutierrez, J.H. Yang, and E.G. Schwarz. Antibiotic efficacy is linked to bacterial cellular respiration. *Proceedings of the National Academy of Sciences of the United States of America*, 112:8173–8180, 2015.

[60] Alfred James Lotka. *Elements of Physical Biology*. Williams & Wilkins, 1925.

[61] Jeff Maltas and Kevin B. Wood. Pervasive and diverse collateral sensitivity profiles inform optimal strategies to limit antibiotic resistance. *PloS Biology*, 17:e3000515, 2019.

[62] GS Mani and Bryan Campbell Clarke. Mutational order: A major stochastic process in evolution. *Proceedings of the Royal Society of London. B. Biological Sciences*, 240(1297):29–37, 1990.

[63] P.J. McClure, B.M. Cole, K.W. Davies, and W.A. Anderson. The use of automated turbidimetric data for the construction of kinetic models. *Journal of Industrial Microbiology*, 12:277–285, 1993.

[64] Alexander S Moffett, Peter J Thomas, Michael Hinczewski, and Andrew W Eckford. Cheater suppression and spite through quorum sensing. *bioRxiv*, 2021.

[65] Sean Nee. Birth-death models in macroevolution. *Annu. Rev. Ecol. Evol. Syst.*, 37:1–17, 2006.

[66] Martha I Nelson, Lone Simonsen, Cecile Viboud, Mark A Miller, Jill Taylor, Kirsten St George, Sara B Griesemer, Elodie Ghedin, Naomi A Sengamalay, David J Spiro, et al. Stochastic processes are key determinants of short-term evolution in influenza A virus. *PLoS Pathogens*, 2(12):e125, 2006.

[67] Daniel Nichol, Joseph Rutter, Christopher Bryant, Andrea M Hujer, Sai Lek, Mark D Adams, Peter Jeavons, Alexander RA Anderson, Robert A Bonomo, and Jacob G Scott. Antibiotic collateral sensitivity is contingent on the repeatability of evolution. *Nature Communications*, 10(1):1–10, 2019.

[68] James R Norris. *Markov Chains*. Number 2. Cambridge university press, 1998.

[69] Otso Ovaskainen and Stephen J Cornell. Space and stochasticity in population dynamics. *Proceedings of the National Academy of Sciences*, 103(34):12781–12786, 2006.

[70] Marcin Paczkowski, Warren W Kretzschmar, Bostjan Markelc, Stanley K Liu, Leoni A Kunz-Schughart, Adrian L Harris, Mike Partridge, Helen M Byrne, and Pavitra Kannan. Reciprocal interactions between tumour cell populations enhance growth and reduce radiation sensitivity in prostate cancer. *Communications Biology*, 4(1):1–13, 2021.

[71] G.A. Pankey and L.D. Sabath. Clinical relevance of bacteriostatic versus bactericidal mechanisms of action in the treatment of gram-positive bacterial infections. *Clinical Infectious Diseases*, 38:864–870, 2004.

[72] G. Ch. Pflug. Stochastic optimization and statistical inference. *Handbooks in Operations Research and Management Science*, 10:427–482, 2003.

[73] Carlos Reding-Roman, Mark Hewlett, Sarah Duxbury, Fabio Gori, Ivana Gudelj, and Robert Beardmore. The unconstrained evolution of fast and efficient antibiotic-resistant bacterial genomes. *Nature Ecology & Evolution*, 1(3):1–11, 2017.

[74] Sidney Redner et al. *A Guide to First-Passage Processes*. Cambridge University Press, 2001.

[75] Adin Ross-Gillespie, Andy Gardner, Angus Buckling, Stuart A West, and Ashleigh S Griffin. Density dependence and cooperation: Theory and a test with bacteria. *Evolution: International Journal of Organic Evolution*, 63(9):2315–2325, 2009.

[76] Richard J Rossi. *Mathematical Statistics: an Introduction to Likelihood Based Inference*. John Wiley & Sons, 2018.

[77] J.A. Scarborough, M.C. Tom, M.W. Kattan, and J.G. Scott. Revisiting a null hypothesis: Exploring the parameters of oligometastasis treatment. *International Journal of Radiation Oncology, Biology, Physics*, pages 1–11, 2021.

[78] Jessica A Scarborough, Erin McClure, Peter Anderson, Andrew Dhawan, Arda Durmaz, Stephen L Lessnick, Masahiro Hitomi, and Jacob G Scott. Identifying states of collateral sensitivity during the evolution of therapeutic resistance in Ewing's sarcoma. *iScience*, 23(7):101293, 2020.

[79] Lauren G Shoemaker, Lauren L Sullivan, Ian Donohue, Juliano S Cabral, Ryan J Williams, Margaret M Mayfield, Jonathan M Chase, Chengjin Chu, W Stanley Harpole, Andreas Huth, et al. Integrating the underlying structure of stochasticity into community ecology. *Ecology*, 101(2):e02922, 2020.

[80] Isabel M Smallegange, Jaap Van Der Meer, and Ralf HJM Kurvers. Disentangling interference competition from exploitative competition in a crab–bivalve system using a novel experimental approach. *Oikos*, 113(1):157–167, 2006.

[81] P.J. Stephens, J.A. Joynson, K.W. Davies, R. Holbrook, H.M. Lappin-Scott, and T.J. Humphrey. The use of an automated growth analyser to measure recovery times of single heat-injured salmonella cells. *Journal of Applied Microbiology*, 83:445–455, 1997.

[82] K. Stevenson, A. F. Alexander F. McVey, I.B.N. Clark, P. Swain, and T. Pilizota. General calibration of microbial growth in microplate readers. *Scientific Reports*, 6(38828), 2016.

[83] Alex G. Strang, Karen C. Abbott, and Peter J. Thomas. How to avoid an extinction time paradox. *Theoretical Ecology*, 12:467–487, 2019.

[84] Zachary Susswein, Surojeet Sengupta, Robert Clarke, and Shweta Bansal. Borrowing ecological theory to infer interactions between sensitive and resistant breast cancer cell populations. *bioRxiv*, 2022.

[85] Peter S. Swain, Keiran Stevenson, Allen Leary, Luis F. Montano-Gutierrez, Ivan B.N. Clark1, Jackie Vogel, and Teuta Pilizota1. Inferring time derivatives including cell growth rates using Gaussian processes. *Nature Communications*, 2016.

[86] Jyoti Tanwar, Shrayanee Das, Zeeshan Fatima, and Saif Hameed. Multidrug resistance: An emerging crisis. *Interdisciplinary Perspectives on Infectious Diseases*, 2014, 2014.

[87] Simone Tavare. The linear birth–death process: An inferential retrospective. *Advances in Applied Probability*, 50:253–269, 2018.

[88] Hamid Teimouri and Anatoly B Kolomeisky. Theoretical investigation of stochastic clearance of bacteria: First-passage analysis. *Journal of the Royal Society Interface*, 16(152):20180765, 2019.

[89] John N Thompson. Rapid evolution as an ecological process. *Trends in Ecology & Evolution*, 13(8):329–332, 1998.

[90] G Upton and I Cook. *A Dictionary of Statistics*. Oxford University Press, 2008.

[91] Mark Vellend, Diane S Srivastava, Kathryn M Anderson, Carissa D Brown, Jill E Jankowski, Elizabeth J Kleynhans, Nathan JB Kraft, Alathea D Letaw, A Andrew M Macdonald, Janet E Maclean, et al. Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos*, 123(12):1420–1430, 2014.

[92] P Verhulst. Notice sur la loi que la population suit dans son accroissement. *Correspondance mathematique et physique*, 10:113–121, 1838.

[93] Vito Volterra. *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*. C. Ferrari Venezia, 1927.

[94] N. Wald-Dickler, P. Paul Holtom, and B. Spellberg. Busting the myth of "static vs cidal": A systemic literature review. *Clinical Infectious Diseases*, 66:1470–1474, 2018.

[95] Cassandra Willyard. The drug-resistant bacteria that pose the greatest health threats. *Nature News*, 543(7643):15, 2017.

[96] Nara Yoon, Nikhil Krishnan, and Jacob Scott. Theoretical modeling of collaterally sensitive drug cycles: shaping heterogeneity to allow adaptive therapy. *Journal of Mathematical Biology*, 83(5):1–29, 2021.

[97] Nara Yoon, Robert V. Veld, Andriy Marusyk, and Jacob Scott. Optimal therapy scheduling based on a pair of collaterally sensitive drugs. *Bulletin of Mathematical Biology*, 80, 2018.