DISSERTATION

zur Erlangung der Doktorwürde
der Naturwissenschaftlich-Mathematischen Gesamtfakultät
der Ruprecht-Karls-Universität Heidelberg

vorgelegt von

Dipl.-Math. Frederik Ziebell

aus Jena

Tag der mündlichen Prüfung: .....

Mathematical Modeling of Neural Stem Cell Dynamics in the Adult Hippocampus

Gutachter: Prof. Dr. Anna Marciniak-Czochra

.....

Die Mathematiker sind eine Art Franzosen:
redet man zu ihnen, so übersetzen sie es in ihre Sprache,
und dann ist es alsobald ganz etwas anderes.

Mathematicians are a sort of Frenchmen:
if one talks to them, they translate it into their own language,
and then it is something completely different.

Johann Wolfgang von Goethe

Abstract

The mammalian hippocampus is a brain region in which neural stem cells (NSCs) continuously generate new neurons and astrocytes during adulthood. The production of the former cell type is known as adult neurogenesis. This new progeny, neurons and also astrocytes, is crucial for cognitive tasks such as learning and memory. Understanding the mechanisms that allow for adult neuron generation forms the basis of new clinical applications. Due to the complexity of adult neurogenesis, mathematical modeling is needed in order to identify its dynamics and to evaluate experimental data.

In this thesis, we model the dynamics of neural stem cells and downstream cell compartments using systems of ordinary differential equations (ODEs). To study the role of sudden changes in neural stem cell dynamics, we develop a method to compute the sign of the derivative of the ODE solution with respect to model parameters. Moreover, we compare different hypotheses about NSC dynamics by analyzing the quasi steady-state to which the respective system converges and propose a compromise model. The quasi steady state approach is also used to reveal age-related changes of neural stem cell dynamics. In the case of evaluating single-cell level data, we additionally employ stochastic simulations utilizing the Gillespie algorithm [26].

Finally, to achieve the original aim of this modeling project, we apply our developed model to evaluate the knockout experiment of the Dkk1 gene, a study published by the group of our collaborator Prof. Martin-Villalba (DKFZ) [51]. We demonstrate that our model is a suitable description of adult hippocampal neurogenesis and give a data-driven identification of the parameters most probably changing to explain the observed effects upon Dkk1 deletion.

Zusammenfassung

Der Hippocampus vieler Säugetierarten beherbergt neuronale Stammzellen, welche im Erwachsenenalter kontinuierlich neue Nervenzellen und Astrozyten bilden. Diese Produktion von neuen Neuronen (Neurogenese) und ebenso Astrozyten ist für kognitive Funktionen wie das Lernen oder die Gedächtnisleistung von wesentlicher Bedeutung. Ein Verständnis der zugrundeliegenden Mechanismen, welche die Produktion neuer Neuronen im Erwachsenenalter ermöglichen, bildet zudem die Basis neuer klinischer Anwendungen. Aufgrund der Komplexität des Produktionsprozesses von neuen Neuronen werden mathematische Modelle benötigt um dessen Ablauf zu entschlüsseln und um experimentelle Daten auszuwerten.

In dieser Dissertation modellieren wir die Dynamik von neuronalen Stammzellen und deren nachgeschalteten Kompartimenten mit Systemen gewöhnlicher Differentialgleichungen (DGL). Um eine abrupte Änderung der Stammzelldynamik zu verstehen, entwickeln wir eine neue Methode, die es erlaubt das Vorzeichen der partiellen Ableitung der DGL Lösung nach Parametern des Modells zu berechnen. Außerdem vergleichen wir verschiedene Hypothesen zur Dynamik neuronaler Stammzellen indem wir die zugehörigen Quasigleichgewichte, gegen welche die jeweilige Dynamik konvergiert, analysieren. Außerdem schlagen wir ein Konsensusmodell vor, welches sich aus den verschiedenen Hypothesen zusammensetzt. Die Analyse von Quasigleichgewichten nutzen wir auch, um altersbedingte Veränderungen der neuronalen Stammzelldynamik aufzudecken. Um auch Daten zu verstehen, die aus einigen wenigen Zellen bestehen, nutzen wir zudem den Gillespie-Algorithmus [26], um entsprechende stochastische Simulationen durchzuführen.

Abschließend widmen wir uns der ursprünglichen Zielsetzung dieses Dissertationsprojektes und analysieren ein Knockoutexperiment des Dkk1 Gens, welches in der Gruppe unseres Kollaborationspartners Prof. Martin-Villalba (DKFZ) durchgeführt wurde [51]. Wir zeigen, dass das von uns etablierte Modell Neurogenese im Hippocampus zutreffend beschreibt. Außerdem treffen wir eine datenbasierte Aus-

sage darüber, wie sich die Dynamik neuronaler Stammzellen aufgrund des Dkk1 Knockouts aller Wahrscheinlichkeit nach ändert.

Acknowledgements

I would like to thank my two supervisors Anna Marciniak-Czochra and Ana Martin-Villalba for giving me the opportunity to work on this interesting topic. They always believed in my abilities to understand biology with mathematics, even when I didn't. Furthermore, I would like to thank them for their constant support throughout the whole project.

I am thankful to all present and former members of the Marciniak-Czochra and Martin-Villalba group for sharing their insights during many fruitful discussions. In particular, I want to thank the members of the former group for their help with modeling issues and the members of the latter group for explaining biology to me in a way a mathematician can understand.

Last but not least, I want to thank my friends and family for their encouragement and idealistic support.

Contents

Chapter 1

Introduction

In almost all mammals including humans, neural stem cells (NSCs) continuously give rise to new neurons during adulthood [30, 45, 54]. This process, called adult neurogenesis, predominantly occurs in two regions of the brain: The subgranular zone of the hippocampal dentate gyrus and the subventricular zone of the lateral ventricles [3, 25, 41]. The focus of this work is on identifying the dynamic behavior of the neuron production system in the hippocampus with respect to cell cycle kinetics, self-renewal properties and apoptosis. Moreover, we aim to understand how the dynamics of neurogenesis change upon inducing a stem cell targeting knockout (KO) of the Dkk1 gene. The latter application is a follow-up on a study which has been published by the group of our collaborator Prof. Martin-Villalba [51].

These aims are being approached using mathematical modeling methods. We employ two kinds of models to account for the different types of data that have been generated. Population level data, which originate from observing a large amount of cells, admit inferences about the collective behavior of a whole cell population. In this case, we use ODEs to describe mean properties of the cell system. In contrast, clonal data collect single-cell level behavior by tracing the progeny of individual cells. Here, stochastic effects emerging from the biological variability among cells need to be taken into account. Accordingly, we use the Gillespie method [26] to perform stochastic simulations corresponding to the ODE models.

1.1 Motivation

Since the first evidence of adult hippocampal neurogenesis by Altman and Das [3], extensive research has been carried out to gain further insights into the topic. In particular, progress has been made towards establishing links between adult hippocampal neurogenesis and several diseases. For instance, it has been shown that Alzheimer's disease (AD) and Parkinson's disease are correlated with impaired hippocampal neurogenesis [18, 27, 44, 60] and that a breakdown of the blood brain barrier, also occurring in AD patients, begins in the hippocampus [42]. Moreover, hippocampal neurogenesis is involved in the development of chronic seizures [15] and mouse models of human epilepsy report an increase of hippocampal volume [47]. Apart from being crucial for cognitive tasks such as learning and memory [55], the hippocampus has been shown to be involved in buffering stress responses and depressive behavior [53]. Taken together, these studies identify the adult hippocampus as potential target for many therapeutic applications. Understanding the dynamic behavior and regulation of hippocampal neurogenesis is therefore of great clinical relevance.

1.2 Biological Background

1.2.1 Adult Hippocampal Neurogenesis

In the adult hippocampus, neurogenesis takes place in the subgranular zone of the dentate gyrus. Here, neuron production originates from neural stem cells, which constitute the top of a hierarchical cell system, leading to mature neurons at the end of the differentiation cascade [25]. Recent data obtained from single-cell level analysis demonstrate that neural stem cells may undergo four different types of events in order to produce progeny: Symmetric divisions by dividing into two stem cells, two types of asymmetric divisions by either dividing into a stem cell and an astrocyte or a stem cell and a neural progenitor cell, and astrocytic transformation, the direct conversion of a stem cell into an astrocyte [9]. Originating from stem cells, neural progenitors are capable of expanding their own pool by symmetric divisions and produce immature neurons called neuroblasts. After a transient period in which

a large majority of neuroblasts die, the remaining cells mature to become granule cell neurons [9, 52].

On a population level, it has been shown that the number of neural stem cells, progenitors and immature neurons decreases during aging [7, 21, 31] with an accompanying decrease of the number of newborn neurons [37]. Although hippocampal neurogenesis is in a steady decline during the aging process, the fraction of dividing stem cells as well as the total number of granule neurons in the dentate gyrus remains constant during aging [7, 21].

1.2.2 Knockout Experiments

A gene-knockout is a procedure that eliminates a certain gene (a DNA sequence encoding a protein) from an organisms' DNA. Thus, the corresponding protein is not synthesized and it is said to be "knocked out". In order to study the protein-driven regulatory mechanisms involved in the dynamics of adult neural stem cells, such knockout experiments have been conducted in the past. A particular version of knockout experiment is the inducible knockout: Cells with a pre-marked gene-sequence react to the administration of Tamoxifen, a chemical that is injected in the animal. The reaction leads to the activation of a cutting enzyme (Cre-recombinase) that excises the marked sequence. Since it is possible to generate an animal in which this cutting enzyme is only present in cells expressing a certain protein, for instance the NSC characteristic protein Nestin, one can selectively knock out the gene of interest in NSCs. The advantage of an inducible knockout is the control over the time point of Tamoxifen injection. In this way, it is possible to influence that the knockout affects neurogenesis only during adulthood, hence eliminating possible developmental effects.

Over the last few years, knockout experiments have been used to study the regulatory mechanisms governing adult hippocampal neurogenesis [1, 9, 17, 24, 29, 51]. Evaluating the results of such experiments is a non-trivial task due to the multifactorial nature of the neurogenesis process. These complex dynamics severely limit intuitive interpretation of experimental data and call for tools such as mathematical modeling and analysis.

1.3 State-of-art in Modeling of Neurogenesis

Mathematical and computational models have been applied before to study adult neurogenesis. Ashbourn et al. [4] provide a system of partial differential equations to model the migration of immature neurons from the subventricular zone of the lateral ventricles to the olfactory bulb and investigate parameters that lead to biologically plausible solutions. Aimone et al. [2] model the functional integration of new neurons to the hippocampus as an artificial neural network. To our knowledge, there exists no model addressing the cellular dynamics of adult hippocampal neurogenesis.

Our proposed model is a neurogenesis adjusted modification of the model of hematopoiesis investigated by Marciniak-Czochra et al. [40] and Stiehl and Marciniak-Czochra [56]. Dynamics of hierarchical cell production systems, which maintain a continuous supply of differentiated functional cells to various parts of a living organism, have attracted the attention of biologists and mathematicians since a long time in the context of blood cell production [59]. Besides common elements that can be found in all cell production systems, there are significant differences depending on the type of cells considered. To model the hierarchical structure of neurogenesis we apply a system of ODEs, each of which describes a discrete differentiation stage. In such models the pace of commitment is dictated by successive divisions. However, in the case of neurogenesis there are indications that stem cell differentiation also involves direct (continuous) transitions. Our proposed model also accounts for these observations.

1.4 Outline of the Thesis

The thesis is organized as follows:

In **Chapter 2** we construct an ODE model of adult hippocampal neurogenesis, incorporating the existing knowledge about this process, which has been outlined in the previous section. The model is applied to simulate the effects of knockout experiments by computing the derivatives of the ODE solution with respect to stem cell parameters. Using this approach, we show that the effects of an inducible knockout experiment are time-dependent in the sense that there exist distinct time-intervals in which cell counts are either increased or decreased. The method of computing parameter-derivatives of the ODE solution allows to describe knockout

effects only qualitatively. This is due to the fact that the constructed model cannot be quantified using the available experimental data. In order to fully understand knockout experiments, it is important to develop a quantitative model that makes it possible to estimate parameters based on experimental data.

Chapter 3 is devoted to build such a quantitative model, departing from the hypotheses of Bonaguidi et al. [9] and Encinas et al. [21]. In both studies, a model for the dynamic behavior of NSCs has been suggested based on their experimental findings. To differentiate between the two hypotheses, we translate both scenarios into ODE models and estimate parameters based on population level data generated by Sascha Dehler in the Martin-Villalba laboratory. In addition, we translate the ODE models into stochastic counterparts using the Gillespie method [26] to also take single-cell level data into account, which has been published by Bonaguidi et al. [9]. The key point for estimating parameters is that the two hypotheses of Bonaguidi et al. [9] and Encinas et al. [21] distinguish between cycling and quiescent (non-cycling) stem cells, a feature that was missing in the model proposed in Chapter 2.

We then demonstrate that the model based on the hypothesis of Bonaguidi et al. [9] provides a better fit considering all the data, since it allows to maintain a constant fraction of proliferating stem cells, a feature that is observed in vivo. Additionally, we reveal age-related changes in the dynamics of neural stem cells and provide experimentally testable predictions.

Chapter 4 considers the dynamic contribution of progenitor cells to the neuron production process. Using the data published in ref. [21], we analyze the hypothesis that progenitors, being generated from stem cells, perform several rounds of symmetric divisions followed by either apoptosis or differentiation into a neuroblast. We show that under a steady state assumption, these dynamics do not agree with the data. However, the data can be explained under the assumption of a non-steady initial distribution, indicating that the steady state was not attained at the time point of the experiment.

In **Chapter 5**, we couple the stem cell model of Chapter 3 and the progenitor model of Chapter 4 to investigate the Dkk1 experiment of Seib et al. [51]. The very good fit to the wild type data independently confirms that we have developed a suitable model of adult hippocampal neurogenesis. Moreover, we demonstrate that

the Dkk1 knockout can best be explained by an increased activation of NSCs, rather than the hypothesized increased self-renewal of NSCs.

Chapter 6 includes a summary and discussion of the findings of this thesis as well as an outlook for further studies.

Moreover, we provide an **Appendix** to give an introduction into the foundations of nonlinear regression. This theory forms the basis of estimating parameters for the developed ODE models.

The results of Chapter 2 have already been published in ref. [62].

Chapter 2

Modeling of Knockout Experiments

In this chapter, we propose a basic ODE model of adult hippocampal neurogenesis based on experimental data. To analyze the results of knockout (KO) experiments, we investigate how changes of cell properties, reflected by model parameters, influence the dynamics of cell counts and of the experimentally observed counts of cells labeled by the cell division marker bromodeoxyuridine (BrdU).

We find that changing cell proliferation rates or the fraction of self-renewal, reflecting the balance between symmetric and asymmetric cell divisions, may result in multiple time phases in the response of the system, such as an initial increase in cell counts followed by a decrease. Furthermore, these phases may be qualitatively different in cells at different differentiation stages and even between BrdU labeled cells and all cells existing in the system.

2.1 Derivation of a Multicompartmental Model

Based on the experimental evidence outlined in Chapter 1, we propose a model of adult neurogenesis consisting of five cellular compartments: Stem cells (c_1), (neural) progenitors (c_2), neuroblasts (c_3), mature neurons (c_4) and astrocytes (c_5):

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}c_1(t) &= (2a_1\theta_1 - 1)p_1c_1(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_2(t) &= \theta_1 2(1-a_1)\kappa p_1c_1(t) + \big((2a_2\theta_2 - 1)p_2 - d_2\big)c_2(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_3(t) &= \big((1+\theta_2) - 2a_2\theta_2\big)p_2c_2(t) - p_3c_3(t) - d_3c_3(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_4(t) &= p_3c_3(t) - d_4c_4(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_5(t) &= \big(\theta_1 2(1-a_1)(1-\kappa) + 1 - \theta_1\big)p_1c_1(t) - d_5c_5(t).
\end{aligned}
\tag{2.1.1}
$$

The model equations describe the following events: Stem cells can either divide with probability θ_1 or perform an astrocytic transformation with probability $1-\theta_1$. The rate at which stem cells undergo these events is given by the parameter p_1, called further as the proliferation rate. The balance between symmetric self-renewal and asymmetric divisions is reflected by the fraction of self-renewal $a_1 \geq 1/2$, which is the probability of a daughter cell to have the same fate as the mother cell. The relation

$$a_1 = 1\,(2a_1 - 1) + \frac{1}{2}\,(1 - (2a_1 - 1)) \tag{2.1.2}$$

shows that $2a_1 - 1$ is the corresponding probability of a symmetric division and $1-(2a_1-1) = 2(1-a_1)$ the probability of an asymmetric division, since the fraction of progeny cells that are stem cells is 1 in a symmetric and $1/2$ in an asymmetric division. It follows that the expected net change of the number of stem cells after one stem cell event (division or transformation) is given by

$$\theta_1(2a_1 - 1) + (1-\theta_1)(-1) = 2a_1\theta_1 - 1.$$

Asymmetric cell divisions may lead to two types of differentiated cells. The non-stem daughter cell is assumed to be either a neural progenitor with probability κ or an astrocyte with probability $1-\kappa$ (see Figure 2.1 for the diagram showing possible scenarios followed by a stem cell).

For the proliferative capacity of progenitors, we again assume two possible modes of generating progeny: Division, which occurs with probability θ_2, or direct transformation to a neuroblast with probability $1-\theta_2$. Analogously to stem cells, progenitors have a proliferation rate p_2 and a fraction of self-renewal a_2. Neuroblasts mature

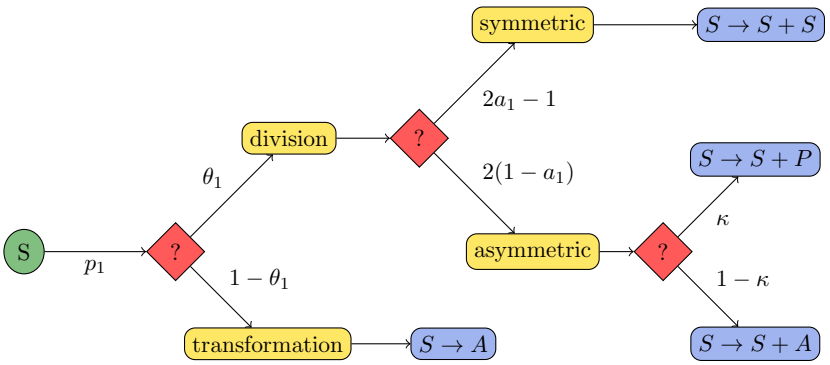


Figure 2.1: Proliferation diagram of a stem cell. Red nodes indicate events with stochastic outcome (e.g. division or transformation; symmetric or asymmetric division), blue nodes describe the outcome of particular events using chemical reaction notation (S: stem cell, P: neural progenitor, A: astrocyte). θ_1 denotes the probability of stem cell division, p_1 the proliferation rate; a_1 reflects the probability that a daughter cell has the same fate as its parent cell (self-renewal takes place) and κ is the probability that a neural progenitor is produced in an asymmetric division rather than an astrocyte.

by transforming into a neuron at the rate p_3. Furthermore, we assume that all cell types except stem cells are subject to apoptosis, modeled by the parameters d_i corresponding to the compartment i.

The independent time variable t is used in two contexts. In section 2.2, we analyze age-related properties of the neurogenesis system and we use time t for the adult age of the animal, i.e. the time point $t = 0$ refers to the beginning of adult age and the initial data consists of the number of cells present at $t = 0$ for each compartment. In section 2.3 and 2.4, we investigate the effect of altered stem cell dynamics on cell counts and on the number of BrdU labeled cells in the framework of an inducible knockout experiment, which imposes altered dynamics upon administration of a certain chemical. In this context, the variable t is used as the time after the knockout, i.e. $t = 0$ refers to the start at which the neurogenesis system operates under the altered dynamics.

2.2 Decline of Stem Cell and Progenitor Counts

Since it was observed that the number of stem and progenitor cells declines with age [21], we first derive parameter conditions that account for this effect.

Lemma 2.2.1. *The solutions $c_1(t)$ and $c_2(t)$ of (2.1.1) are monotonically decreasing for all $t \geq 0$ if and only if*

$$a_1\theta_1 < \frac{1}{2} \quad \textit{and} \quad \frac{c_1(0)}{c_2(0)} < \frac{d_2 - (2a_2\theta_2 - 1)p_2}{\kappa\theta_1 2(1-a_1)p_1}.$$

Furthermore

$$\lim_{t\to\infty} \frac{c_1(t)}{c_2(t)} \begin{cases} > 0 & (2a_1\theta_1 - 1)p_1 > (2a_2\theta_2 - 1)p_2 - d_2, \\ = 0 & \textit{otherwise.} \end{cases}$$

Proof. It holds $\frac{\mathrm{d}}{\mathrm{d}t}c_1(t) < 0$ if and only if $a_1\theta_1 < \frac{1}{2}$. To analyze the decay conditions of the progenitor compartment c_2, we introduce the abbreviations

$$\begin{aligned} p &:= (2a_1\theta_1 - 1)p_1, \\ q &:= \kappa\theta_1 2(1-a_1)p_1, \\ r &:= d_2 - (2a_2\theta_2 - 1)p_2. \end{aligned}$$

Firstly, let us note that $p < 0$ and $r > 0$ is a necessary though not sufficient condition for $\frac{\mathrm{d}}{\mathrm{d}t}c_2(t) < 0$. Furthermore, it can be seen that the quantity $\frac{c_1}{c_2}$ satisfies a Riccati differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{c_1(t)}{c_2(t)}\right) = \frac{c_1(t)}{c_2(t)}\left(p + r - q\frac{c_1(t)}{c_2(t)}\right). \tag{2.2.1}$$

Since $q > 0$, the right-hand side of (2.2.1) describes a downward opening parabola with roots 0 and $\frac{p+r}{q}$. If $p + r > 0$, the state $\frac{p+r}{q}$ is globally asymptotically stable. Otherwise, 0 is globally asymptotically stable. In both cases, the steady state is approached monotonically. Furthermore, $p < 0$ and $q, r > 0$ imply

$$\frac{p+r}{q} < \frac{r}{q}$$

and hence

$$\frac{c_1(0)}{c_2(0)} < \frac{r}{q} \text{ implies } \frac{c_1(t)}{c_2(t)} < \frac{r}{q} \text{ for all } t \geq 0. \tag{2.2.2}$$

Note from (2.1.1) that $\frac{\mathrm{d}}{\mathrm{d}t}c_2(t) < 0$ holds if and only if

$$\frac{c_1(t)}{c_2(t)} < \frac{r}{q}.$$

Taken together with (2.2.2), we see that $\frac{\mathrm{d}}{\mathrm{d}t}c_2(t) < 0$ for all $t \geq 0$ if and only if

$$\frac{c_1(0)}{c_2(0)} < \frac{r}{q}.$$

□

Biological Interpretation. Lemma 2.2.1 states that the depletion of the stem cell pool takes place if and only if symmetric stem cell divisions, accompanied by a gain of stem cells, are less likely than astrocytic transformations with the resulting loss of the stem cell. The second part states that the ratio of the number of stem cells to the number of progenitors converges to zero, if the net depletion rate of stem cells is higher than the one of progenitors. Otherwise, it converges to a positive value. Furthermore, the positive steady state is achieved monotonically, either increasing or decreasing. Interestingly, both behaviors have been observed experimentally. In ref. [21] it was reported that the ratio of the number of stem cells to the number of progenitors is monotonically decreasing while [31] reports an increasing progression. The discrepancy in both observations might thus result from different labeling procedures and measurements of different sub-populations.

In [21, Supplementary Table 2], a time series for the age-related decline of the stem cell and progenitor count was presented. Fitting those data to the solution of (2.1.1) indicates that the parameters of our model satisfy the relations $(2a_2\theta_2 - 1)p_2 - d_2 - (2a_1\theta_1 - 1)p_1 < 0$ and $a_1\theta_1 < 1/2$ (see section 2.4.1). Hence, the data are consistent with the scenario in which the net depletion rate of progenitors is higher than the net depletion rate of stem cells. This results in the following assumption of our subsequent mathematical analysis in order to stay consistent with the experimental data of [21]:

Assumption 2.2.2. The parameters of model (2.1.1) have the properties

$$(2a_2\theta_2 - 1)p_2 - d_2 - (2a_1\theta_1 - 1)p_1 < 0$$

and

$$a_1\theta_1 < \frac{1}{2}\,.$$

2.3 Analysis of Altered Stem Cell Parameters

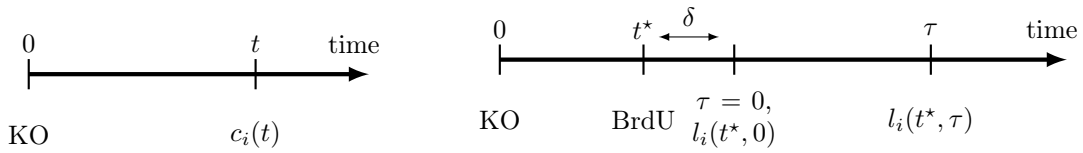
2.3.1 Preliminaries

As already described in Chapter 1, KO experiments have been used to manipulate the dynamics of neural stem cells in the adult hippocampus. If the knockout of a certain gene results in a difference between knockout and wild-type (non-knockout) individuals regarding the number of counted cells, the question arises of which stem cell parameter was affected by the knockout and caused the observed difference. To treat this question in a general way, we examine the effects of alterations of the stem cell parameters a_1 (fraction of self-renewal), θ_1 (division probability) and p_1 (proliferation rate) on the number of existing cells and the number of cells labeled by BrdU, a chemical that is incorporated in cell's DNA after they performed a division and are in the stage of DNA synthesis.

Basic notation: We occasionally write $x(t;p)$ to emphasize the dependence of the solution $x(t)$ of a differential equation on a parameter p and $\operatorname{sgn}(a)$ denotes the sign of a real number a.

Modeling of Two Experimental Scenarios

We consider two scenarios related to knockout experiments, for which we analyze the effect of altered stem cell parameters. **Scenario (i)**, (Figure 2.2a): Starting from a time point zero, which corresponds to the fixed age of the studied animal at which the knockout is conducted, the number of cells of compartment i is analyzed at t time units after time point zero. **Scenario (ii)**, (Figure 2.2b): At time $t^\star$ after the initial time point of the knockout, BrdU is administered and is present in the organism for a duration δ, thus labeling DNA synthesizing cells during that



(a) Cell counts $c_i(t)$ are evaluated at time t after the knockout

(b) The number of BrdU labeled cells is evaluated at time τ after the end of the labeling period ($\tau = 0$) and BrdU was given at time $t^\star$ after knockout

Figure 2.2: Graphical representation of the analyzed knockout scenarios

period. At τ time units after the labeling has ended, the number of BrdU labeled cells ($l_i(t^\star, \tau)$) is examined.

To evaluate in both scenarios the effects of a change of a parameter $p \in \{a_1, p_1, \theta_1\}$ from a value $\hat{p}$ to a value $\hat{p} + \Delta p$ ($\Delta p > 0$), we analyze the sign of the derivative $\partial_p c_i(t; \hat{p})$, respectively $\partial_p l_i(t^\star, \tau; \hat{p})$, with respect to the parameter p. Thus, we assume that Δp is so small that

$$\operatorname{sgn}(c_i(t; \hat{p} + \Delta p) - c_i(t; \hat{p})) = \operatorname{sgn}(\partial_p c_i(t; \hat{p}))$$

and

$$\operatorname{sgn}(l_i(t^\star, \tau; \hat{p} + \Delta p) - l_i(t^\star, \tau; \hat{p})) = \operatorname{sgn}(\partial_p l_i(t^\star, \tau; \hat{p})).$$

To model scenario (i), we use model (2.1.1) together with initial data $c_i(0)$ corresponding to the number of cells of compartment i, which are present at the time point of the knockout. For scenario (ii), the initial data is $l_i(t^\star, 0)$, the number of cells that have incorporated BrdU at the end of the labeling period, where BrdU was given at time point $t^\star$ and the independent variable in this scenario is τ, the time that passed since the end of the labeling period.

Initial Data for BrdU Labeled Cells

It is known from the theory of branching processes that in a model of proliferation in which particles (cells) have exponentially distributed lifetimes, mean counts of particles (cells) follow a system of ordinary differential equations [34, Chapter 4]. Conversely, if the population of cells is described by a system of ODEs, then the interpretation is that the cells have exponentially distributed lifetimes. This is

a simple and widely used model (see relevant discussion of the cell proliferation models in Kimmel and Axelrod [34]), despite the fact that it has been known that cell lifetime distributions are not exponentially distributed [28]. Thus, we apply the relationship between ODEs and the exponential distribution in order to derive equations for the number of cells belonging to type i and have incorporated BrdU ($l_i(t^\star, 0)$).

Recall that BrdU is a chemical which is incorporated in cells after they performed a division and are in the stage of DNA synthesis. Thus, in order to be labeled by BrdU, a cell must be in S-phase during a time interval of length δ in which BrdU is present in the animal. Assuming that the fraction of dividing cells equals the fraction of DNA-synthesizing cells during any time interval of fixed length, it follows that the number of cells that have incorporated BrdU at the end of the labeling period, starting with labeling at time $t^\star$, is given by

$$
\begin{aligned}
l_1(t^\star, 0) &= 2a_1\theta_1(1 - e^{-p_1\delta})c_1(t^\star), \\
l_2(t^\star, 0) &= \kappa\theta_1 2(1 - a_1)(1 - e^{-p_1\delta})c_1(t^\star) + \theta_2(2a_2 - 1)(1 - e^{-p_2\delta})c_2(t^\star), \\
l_3(t^\star, 0) &= \theta_2 2(1 - a_2)(1 - e^{-p_2\delta})c_2(t^\star), \\
l_4(t^\star, 0) &= 0, \\
l_5(t^\star, 0) &= (1 - \kappa)\theta_1 2(1 - a_1)(1 - e^{-p_1\delta})c_1(t^\star)
\end{aligned}
$$

where $c_i(t^\star)$ is the number of cells of compartment i, present at time $t^\star$ of BrdU injection. More specifically, a random variable X, exponentially distributed with parameter λ, has the property

$$\Pr(\{X \in [t, t + \delta]\}|\{X > t\}) = 1 - e^{-\lambda\delta}.$$

Thus, the fraction of stem cells performing a transformation or division during marker exposure is $1 - e^{-p_1\delta}$ and a fraction θ_1 of them divides with each division contributing, on average, $2a_1$ stem cells. Analogous considerations lead to the equations for all other cell compartments. Note that neurons are assumed to be the result of a transformation from neuroblasts rather than a division. Thus, there are no BrdU labeled neurons right after the BrdU labeling period has ended.

2.3.2 Effects of Altered Stem Cell Parameters

We model the scenarios (i) and (ii) of section 2.3.1 using the system of ODEs that follows. The equations for the quantities $l_i(t^\star, \tau)$ have been derived based on the assumption that labeled cells of type i follow the same dynamics as their corresponding compartment c_i. Thus, the initial data for the labeled cells, $l_i(t^\star, 0)$, depend on $c_i(t^\star)$, the number of cells present at the time point $t = t^\star$ of BrdU injection. We obtain the following system of equations.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}c_1(t) &= (2a_1\theta_1 - 1)p_1c_1(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_2(t) &= \theta_1 2(1-a_1)\kappa p_1c_1(t) + \big((2a_2\theta_2 - 1)p_2 - d_2\big)c_2(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_3(t) &= \big((1+\theta_2) - 2a_2\theta_2\big)p_2c_2(t) - d_3c_3(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_4(t) &= p_3c_3(t) - d_4c_4(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_5(t) &= \big(\theta_1 2(1-a_1)(1-\kappa) + 1 - \theta_1\big)p_1c_1(t) - d_5c_5(t),\\
c_i(0) &= n_i,\\
\frac{\mathrm{d}}{\mathrm{d}\tau}l_1(t^\star,\tau) &= (2a_1\theta_1 - 1)p_1l_1(t^\star,\tau),\\
\frac{\mathrm{d}}{\mathrm{d}\tau}l_2(t^\star,\tau) &= \theta_1 2(1-a_1)\kappa p_1l_1(t^\star,\tau) + \big((2a_2\theta_2 - 1)p_2 - d_2\big)l_2(t^\star,\tau),\\
\frac{\mathrm{d}}{\mathrm{d}\tau}l_3(t^\star,\tau) &= \big((1+\theta_2) - 2a_2\theta_2\big)p_2l_2(t^\star,\tau) - d_3l_3(t^\star,\tau),\\
\frac{\mathrm{d}}{\mathrm{d}\tau}l_4(t^\star,\tau) &= p_3l_3(t^\star,\tau) - d_4l_4(t^\star,\tau),\\
\frac{\mathrm{d}}{\mathrm{d}\tau}l_5(t^\star,\tau) &= \big(\theta_1 2(1-a_1)(1-\kappa) + 1 - \theta_1\big)p_1l_1(t^\star,\tau) - d_5l_5(t^\star,\tau),\\
l_1(t^\star,0) &= 2a_1\theta_1(1-e^{-p_1\delta})c_1(t^\star),\\
l_2(t^\star,0) &= \kappa\theta_1 2(1-a_1)(1-e^{-p_1\delta})c_1(t^\star) + 2a_2\theta_2(1-e^{-p_2\delta})c_2(t^\star),\\
l_3(t^\star,0) &= \theta_2 2(1-a_2)(1-e^{-p_2\delta})c_2(t^\star),\\
l_4(t^\star,0) &= 0,\\
l_5(t^\star,0) &= (1-\kappa)\theta_1 2(1-a_1)(1-e^{-p_1\delta})c_1(t^\star),
\end{aligned}
\tag{2.3.1}
$$

where n_i is the number of cells of compartment i at the fixed age of the animal where the knockout is induced. We will analyze the derivatives of $c_1(t)$, $c_2(t)$ and $l_1(t,0)$, $l_2(t,0)$ and $l_5(t,0)$ with respect to the parameter $p \in \{a_1, \theta_1, p_1\}$, thus evaluating the implications of altered stem cell parameters on the number of existing stem cells and progenitors and on the number of BrdU labeled stem cells, progenitors and astrocytes at the end of the labeling period. The parameter-derivatives of the number of neurons and BrdU labeled neurons ($c_4(t)$ and $l_i(t,\tau)$) cannot be analyzed in full generality for arbitrary τ. They are investigated numerically in section 2.4. For notational convenience, we define

$$l_i^0(t^\star) := l_i(t^\star, 0)$$

and denote this quantity as the number of BrdU incorporating cells of the compartment i.

Proposition 2.3.1. *It holds*

$$\begin{aligned} \partial_{a_1} c_1(t; a_1) &= 2\theta_1 p_1 t c_1(t) \\ \partial_{p_1} c_1(t; p_1) &= (2a_1\theta_1 - 1) t c_1(t) \\ \partial_{\theta_1} c_1(t; \theta_1) &= 2a_1 p_1 t c_1(t) \end{aligned} \tag{2.3.2}$$

and

$$\partial_p c_i(t; p) = \frac{(e^{\alpha t} - 1)(\alpha\beta + \gamma) - \alpha\gamma t}{\alpha^2} c_1(t) \tag{2.3.3}$$

for $i \in \{2, 5\}$ and $p \in \{a_1, \theta_1, p_1\}$, where α, β and γ depend on both, i and p (cf. Table 2.3).

p	$i = 2$	$i = 5$
a_1	$\alpha = (2a_2\theta_2 - 1)p_2 - (2a_1\theta_1 - 1)p_1 - d_2$	$\alpha = -(d_4 + (2a_1\theta_1 - 1)p_1)$
	$\beta = -2\kappa\theta_1 p_1$	$\beta = 2\theta_1(\kappa - 1)p_1$
	$\gamma = 4\kappa\theta_1^2 p_1^2(1 - a_1)$	$\gamma = 2\theta_1 p_1^2(\theta_1(1-\kappa)2(1-a_1) + 1 - \theta_1)$
p_1	$\alpha = (2a_2\theta_2 - 1)p_2 - (2a_1\theta_1 - 1)p_1 - d_2$	$\alpha = -(d_4 + (2a_1\theta_1 - 1)p_1)$
	$\beta = \kappa\theta_1 2(1 - a_1)$	$\beta = \theta_1(1-\kappa)2(1-a_1) + 1 - \theta_1$
	$\gamma = \kappa\theta_1 2(1 - a_1)(2a_1\theta_1 - 1)p_1$	$\gamma = (2a_1\theta_1 - 1)p_1(\theta_1(1-\kappa)2(1-a_1) + 1 - \theta_1)$
θ_1	$\alpha = (2a_2\theta_2 - 1)p_2 - (2a_1\theta_1 - 1)p_1 - d_2$	$\alpha = -(d_4 + (2a_1\theta_1 - 1)p_1)$
	$\beta = \kappa 2(1 - a_1)p_1$	$\beta = ((1-\kappa)2(1-a_1) - 1)p_1$
	$\gamma = \kappa\theta_1 4(1 - a_1)a_1 p_1^2$	$\gamma = 2a_1 p_1^2(\theta_1(1-\kappa)2(1-a_1) + 1 - \theta_1)$

Table 2.3: Coefficients of $\partial_p c_i(t; p)$ for $p \in \{a_1, p_1, \theta_1\}$

For the proof of Proposition 2.3.1, we need a technical lemma.

Lemma 2.3.2. *Let $p, h \in \mathbb{R}$, f, g real-valued non-zero functions and consider the system*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}x(t) &= f(p)x(t),\\ \frac{\mathrm{d}}{\mathrm{d}t}y(t) &= g(p)x(t) + hy(t)\end{aligned}$$

together with non-zero initial conditions $x(0)$ and $y(0)$, which are independent of p. The solution then satisfies

$$\partial_p x(t;p) = f'(p)tx(t)$$

and

$$\partial_p y(t;p) = \frac{(e^{\alpha t}-1)(\alpha\beta+\gamma)-\alpha\gamma t}{\alpha^2}x(t),$$

where

$$\alpha = h - f(p), \quad \beta = g'(p) \quad \text{and} \quad \gamma = f'(p)g(p). \tag{2.3.4}$$

Proof. Define $x_p(t) := \partial_p x(t;p)$ and $y_p(t) := \partial_p y(t;p)$. Since $x(0)$ and $y(0)$ are independent of p, we have

$$x_p(0) = y_p(0) = 0 \tag{2.3.5}$$

and it follows from symmetry of second partial derivatives that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}x_p(t) &= f'(p)x(t) + f(p)x_p(t),\\ \frac{\mathrm{d}}{\mathrm{d}t}y_p(t) &= g'(p)x(t) + g(p)x_p(t) + hy_p(t).\end{aligned}$$

Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{x_p(t)}{x(t)}\right) = f'(p) + f(p)\left(\frac{x_p(t)}{x(t)}\right) - x_p(t)\frac{\frac{\mathrm{d}}{\mathrm{d}t}x(t)}{x(t)^2} = f'(p).$$

The assumption $x(0) \neq 0$ leads to $\left(\frac{x_p}{x}\right)(0) = 0$ by (2.3.5) and solving the above differential equation yields

$$x_p(t) = f'(p)tx(t).$$

Furthermore,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{y_p(t)}{x(t)}\right) = g'(p) + g(p)f'(p)t + (h - f(p))\left(\frac{y_p(t)}{x(t)}\right)$$

and

$$\frac{y_p(0)}{x(0)} = 0$$

imply

$$\frac{y_p(t)}{x(t)} = \frac{(e^{\alpha t} - 1)(\alpha\beta + \gamma) - \alpha\gamma t}{\alpha^2},$$

with α, β and γ as stated in (2.3.4), since a differential equation of the form

$$\begin{aligned} x'(t) &= \alpha x(t) + \beta + \gamma t, \\ x(0) &= 0 \end{aligned}$$

has the solution

$$x(t) = \frac{(e^{\alpha t} - 1)(\alpha\beta + \gamma) - \alpha\gamma t}{\alpha^2}.$$

□

Proof of Proposition 2.3.1. We apply Lemma 2.3.2 to (2.3.1) with $x = c_1$, $y \in \{c_2, c_4\}$ and $p \in \{a_1, p_1, \theta_1\}$. Thus, the derivatives appearing in (2.3.4) have to be understood as partial derivatives with respect to the considered parameter $p \in \{a_1, p_1, \theta_1\}$. It follows that

$$f = (2a_1\theta_1 - 1)p_1.$$

If $y = c_2$, we have

$$\begin{aligned} g &= \kappa\theta_1 2(1 - a_1)p_1, \\ h &= (2a_2\theta_2 - 1)p_2 - d_2 \end{aligned}$$

and if $y = c_4$, it holds

$$\begin{aligned} g &= (\theta_1(1 - \kappa)2(1 - a_1) + 1 - \theta_1)p_1, \\ h &= -d_4. \end{aligned}$$

The explicit values of α, β and γ appearing in (2.3.4) are summarized in Table 2.3. □

In order to evaluate (2.3.3), we need a second technical lemma.

Lemma 2.3.3. *For $t \geq 0$, $\alpha < 0$ and $\beta, \gamma \in \mathbb{R}$, consider the function*

$$f(t) = \frac{(e^{\alpha t} - 1)(\alpha\beta + \gamma) - \alpha\gamma t}{\alpha^2}.$$

It holds $f(0) = 0$, $f(t) \in \mathcal{O}(t)$ and f has the following properties:

(P1) *If $\beta > 0$ and $\gamma > 0$,*

$$f \text{ is monotonically increasing.}$$

(P2) *If $\beta > 0$ and $\gamma < 0$, there exists a unique $t_0 > 0$ such that for all $t > 0$*

$$\operatorname{sgn}(f(t)) = -\operatorname{sgn}(t - t_0).$$

(P3) *If $\beta < 0$ and $\gamma > 0$, there exists a unique $t_0 > 0$ such that for all $t > 0$*

$$\operatorname{sgn}(f(t)) = \operatorname{sgn}(t - t_0).$$

Proof. Follows from evaluating $f'(t) = \frac{1}{\alpha}((\alpha\beta + \gamma)e^{\alpha t} - \gamma)$ for every single case. □

Furthermore, we introduce a notion for the *sign-sequence* of a real-valued function:

Definition 2.3.4 (Sign-sequence)**.** Let $f\colon [0, \infty) \to \mathbb{R}$ be a function. The sign-sequence $\sigma(f)$ is defined as the sequence of distinct signs of $f(t)$ that are encountered by traversing the domain of f from zero to infinity. For instance, $\sigma((t-1)(t-2)) = (+, 0, -, 0, +)$.

Altered Fraction of Self-Renewal

In this section, we describe the effect of increasing the fraction of self-renewal of stem cells. Recall that the fraction a_1 is defined as the probability that a progeny cell, which resulted from a stem cell division, becomes a stem cell itself. Thus, increasing the fraction of self-renewal increases the proportion of symmetric stem cell divisions that give rise to two stem cells at the expense of asymmetric stem cell divisions.

Lemma 2.3.5. *The solution of* (2.3.1) *satisfies*

$$\begin{aligned}
&\sigma(\partial_{a_1} c_1(t;a_1)) = (0,+),\\
&\sigma\big(\partial_{a_1} l_1^0(t;a_1)\big) = (+),\\
&\sigma(\partial_{a_1} c_2(t;a_1)) = (0,-,0,+),\\
&\sigma\big(\partial_{a_1} l_2^0(t;a_1)\big) = (-,0,+),\\
&\sigma\big(\partial_{a_1} l_5^0(t;a_1)\big) = (-,0,+).
\end{aligned}$$

Proof. From (2.3.1) and (2.3.2), it follows that

$$\begin{aligned}
\partial_{a_1} c_1(t;a_1) &= 2\theta_1 p_1 t c_1(t),\\
\partial_{a_1} l_1^0(t;a_1) &= 2\theta_1(1-e^{-p_1\delta})(1+2a_1\theta_1 p_1 t)c_1(t),\\
\partial_{a_1} l_2^0(t;a_1) &= \kappa\theta_1(1-e^{-p_1\delta})(-2+2(1-a_1)2\theta_1 p_1 t)c_1(t)\\
&\quad + 2a_2\theta_2(1-e^{-p_2\delta})\partial_{a_1} c_2(t;a_1),\\
\partial_{a_1} l_5^0(t;a_1) &= (1-\kappa)\theta_1(1-e^{-p_1\delta})(-2+2(1-a_1)2\theta_1 p_1 t)c_1(t).
\end{aligned} \tag{2.3.6}$$

Thus, $\partial_{a_1} c_1(t;a_1)$ and $\partial_{a_1} c_1(0;a_1)$ are positive for all positive t and $\partial_{a_1} l_1^0(0;a_1) = 0$. Consider now the quantity $\partial_{a_1} c_2(t;a_1)$. Assumption 2.2.2 together with Table 2.3 imply that the first factor of $\partial_{a_1} c_2(t;a_1)$ in (2.3.3) has the property (P3) in Lemma 2.3.3. Consequently, $\partial_{a_1} c_2(t;a_1)$ is negative on $(0,t_0)$ and positive on (t_0,∞) for some positive t_0. From (2.3.6), we deduce that the same is true for $\partial_{a_1} l_2^0(t;a_1)$ and $\partial_{a_1} l_5^0(t;a_1)$, since the term $-2+2(1-a_1)2\theta_1 p_1 t$ is negative for $t=0$ and positive for sufficiently large t. □

Biological Interpretation. Lemma 2.3.5 shows that increasing stem cells self-renewal increases the stem cell count and the number of BrdU incorporating stem cells at any time point after the increase was performed. Conversely, the increased self-renewal initially decreases the progenitor count and the number of BrdU incorporating progenitors and astrocytes. However, the decrease is reversed and turns into an increase after the initial period. Furthermore, the effect of altered self-renewal is instantaneous on the number of BrdU labeled cells in the sense that the corresponding parameter-derivative is non-zero at time zero. Figure 2.4 depicts a simulation of the time-dependent responses of an increased fraction of self-renewal. The simulation is consistent with Lemma 2.3.5.

The two-phase progression on progenitors and astrocytes can be explained intuitively as follows: The increased number of symmetric stem cell divisions at the expense of asymmetric divisions reduces the proportion of events at which progenitors are born. Thus, a decreased number of progenitors is observed initially. At the same time, the increased number of symmetric stem cell divisions, which result in an enlarged stem cell pool, benefits the progenitor count in the long run: Although a reduced fraction of stem cells generates progenitors via asymmetric divisions, the increased number of stem cells dominates this effect, meaning that the total number of asymmetric stem cell divisions is elevated. The immediate effect on BrdU incorporating progenitors and astrocytes can be explained by the observation that changing a parameter that affects division instantaneously changes the output of the division and that the mitotic marker BrdU exactly labels this output. Thus, the number of cells labeled by BrdU is faster influenced by a parameter change than the actual cell count.

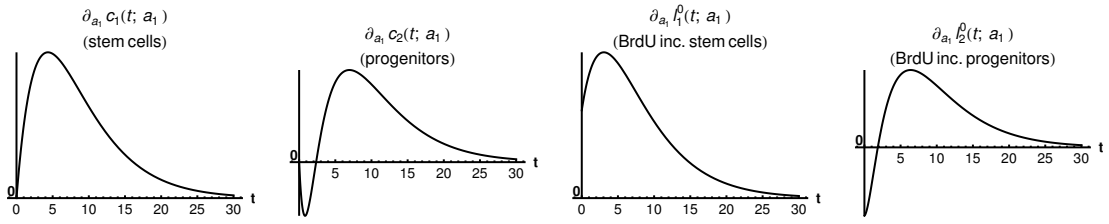


Figure 2.4: Simulated responses to an infinitesimal increase in stem cells fraction of self-renewal a_1 of the number of stem cells (c_1), progenitors (c_2), BrdU incorporating stem cells (l_1^0) and BrdU incorporating progenitors (l_2^0) respectively at time t after the increase. The parameter set of section 2.4.2 was used.

Altered Proliferation Rate

The proliferation rate p_1 is the rate at which stem cells undergo division or transformation events. Increasing this rate shortens the waiting time between successive events of a given stem cell.

Lemma 2.3.6. *The solution of* (2.3.1) *satisfies*

$$\begin{aligned}
\sigma(\partial_{p_1} c_1(t;p_1)) &= (0,-),\\
\sigma\big(\partial_{p_1} l_1^0(t;p_1)\big) &= (+,0,-),\\
\sigma(\partial_{p_1} c_2(t;p_1)) &= (0,+,0,-),\\
\sigma\big(\partial_{p_1} l_2^0(t;p_1)\big) &= (+,0,-),\\
\sigma\big(\partial_{p_1} l_5^0(t;p_1)\big) &= (+,0,-).
\end{aligned}$$

Proof. (2.3.1) and (2.3.2) imply

$$\begin{aligned}
\partial_{p_1} c_1(t;p_1) &= (2a_1\theta_1 - 1)tc_1(t),\\
\partial_{p_1} l_1^0(t;p_1) &= 2a_1\theta_1(\delta e^{-p_1\delta} + (1-e^{-p_1\delta})(2a_1\theta_1-1)t)c_1(t),\\
\partial_{p_1} l_2^0(t;p_1) &= \kappa\theta_1 2(1-a_1)(\delta e^{-p_1\delta} + (1-e^{-p_1\delta})(2a_1\theta_1-1)t)c_1(t)\\
&\quad + 2a_2\theta_2(1-e^{-p_2\delta})\partial_{p_1} c_2(t;p_1),\\
\partial_{p_1} l_5^0(t;p_1) &= (1-\kappa)\theta_1 2(1-a_1)(\delta e^{-p_1\delta} + (1-e^{-p_1\delta})(2a_1\theta_1-1)t)c_1(t).
\end{aligned}$$

Because of Assumption 2.2.2, it holds $2a_1\theta_1 - 1 < 0$. Thus, $\partial_{p_1} c_1(t;p_1)$ is negative on $(0,\infty)$ and $\partial_{p_1} l_1^0(t;p_1)$ is positive on $(0,t_0)$ and negative on (t_0,∞) for some positive t_0. The same holds true for $\partial_{p_1} l_5^0(t;p_1)$, but with different t_0. Consider now $\partial_{p_1} c_2(t;p_1)$. The first factor of $\partial_{p_1} c_2(t;p_1)$ in formula (2.3.3) has the property (P2) in Lemma 2.3.3. Hence, $\partial_{p_1} c_2(t;p_1)$ and $\partial_{p_1} l_2^0(t;p_1)$ show the same qualitative progression as stated for $\partial_{p_1} l_1^0(t;p_1)$. Analogously to the considerations of an altered fraction of self-renewal, the BrdU incorporating quantities $\partial_{p_1} l_1^0(t;p_1), \partial_{p_1} l_2^0(t;p_1)$ and $\partial_{p_1} l_5^0(t;p_1)$ are positive at time zero whereas $\partial_{p_1} c_1^0(0;p_1) = \partial_{p_1} c_2^0(0;p_1) = 0$. □

Biological Interpretation. The above lemma shows that an increased proliferation rate of stem cells decreases the stem cell count while the number of BrdU incorporating stem cells initially increases and later on decreases. This effect can be explained by the observation that an increased proliferation rate causes more stem cell divisions over any given time interval, resulting initially in more BrdU incorporating stem cells. Furthermore, the increased proliferation rate also causes a higher rate of astrocytic transformations, the events responsible for the depletion of the stem cell pool. As time progresses, the increased decay rate of stem cells compensates for the higher proportion of BrdU incorporating cells and results thus in a net

decrease of labeled stem cells. See Figure 2.5 for a corresponding simulation, which is consistent with the above lemma. It is not surprising that the number of progenitors and BrdU labeled progenitors and astrocytes display the same qualitative trend as labeled stem cells, since these quantities also depend on stem cell divisions.

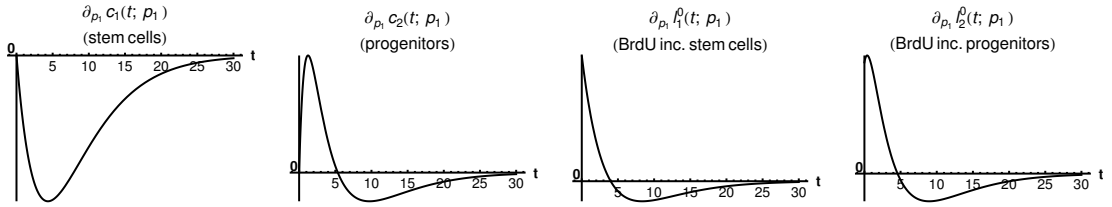


Figure 2.5: Simulated responses to an infinitesimal increase in stem cells proliferation rate p_1 of the number of stem cells (c_1), progenitors (c_2), BrdU incorporating stem cells (l_1^0) and BrdU incorporating progenitors (l_2^0) respectively at time t after the increase. The parameter set of section 2.4.2 was used.

Altered Division Probability

The division probability θ_1 of a stem cell is the probability that the next event a stem cell undergoes is a division rather than a transformation. Consequently, increasing the division probability causes more division and fewer transformation events.

Lemma 2.3.7. *The solution of* (2.3.1) *satisfies*

$$\begin{aligned}
\sigma(\partial_{\theta_1} c_1(t;\theta_1)) &= (0,+),\\
\sigma\big(\partial_{\theta_1} l_1^0(t;\theta_1)\big) &= (+),\\
\sigma(\partial_{\theta_1} c_2(t;\theta_1)) &= (0,+),\\
\sigma\big(\partial_{\theta_1} l_2^0(t;\theta_1)\big) &= (+),\\
\sigma\big(\partial_{\theta_1} l_5^0(t;\theta_1)\big) &= (+).
\end{aligned}$$

Proof. It holds

$$\begin{aligned}
\partial_{\theta_1} c_1(t;\theta_1) &= 2a_1p_1tc_1(t),\\
\partial_{\theta_1} l_1^0(t;\theta_1) &= 2a_1(1-e^{-p_1\delta})(1+\theta_1 2a_1p_1t)c_1(t),\\
\partial_{\theta_1} l_2^0(t;\theta_1) &= \kappa 2(1-a_1)(1-e^{-p_1\delta})(1+\theta_1 2a_1p_1t)c_1(t)\\
&\quad + 2a_2\theta_2(1-e^{-p_2\delta})\partial_{\theta_1} c_2(t;\theta_1),\\
\partial_{\theta_1} l_5^0(t;\theta_1) &= (1-\kappa)2(1-a_1)(1-e^{-p_1\delta})(1+\theta_1 2a_1p_1t)c_1(t).
\end{aligned}$$

Thus, $\partial_{\theta_1} c_1(t;\theta_1)$, $\partial_{\theta_1} l_1^0(t;\theta_1)$ and $\partial_{\theta_1} l_5^0(t;\theta_1)$ are positive for positive t. Furthermore, the first factor of $\partial_{\theta_1} c_2(t;\theta_1)$ in (2.3.3) has the property (P1) in Lemma 2.3.3. Hence, $\partial_{\theta_1} c_2(t;\theta_1)$ and $\partial_{\theta_1} l_2^0(t;\theta_1)$ are also positive for positive t. □

Biological Interpretation. From the above lemma we conclude that increasing the stem cell division probability causes an increase in cell counts and in the number of BrdU incorporating cells for all considered compartments. There is no two-phase progression displaying increased or decreased cell numbers in distinct phases after the change in stem cell dynamics. The effect on the progenitor count is qualitatively the same as on the stem cell count; the same holds true for BrdU incorporating stem cells and progenitors. Figure 2.6 illustrates a corresponding simulation of the discussed quantities.

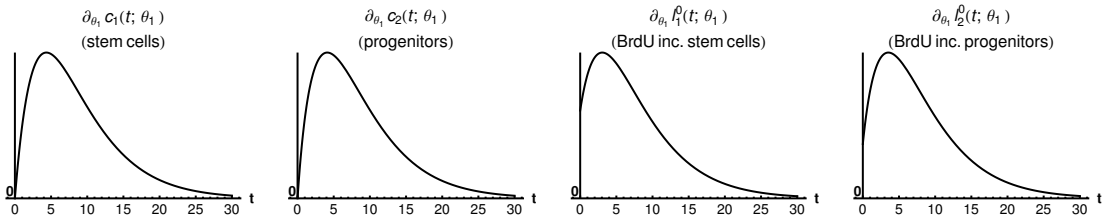


Figure 2.6: Simulated responses to an infinitesimal increase in stem cells division probability θ_1 of the number of stem cells (c_1), progenitors (c_2), BrdU incorporating stem cells (l_1^0) and BrdU incorporating progenitors (l_2^0) respectively at time t after the increase. The parameter set of section 2.4.2 was used.

2.3.3 Conclusion of the Analysis

The above considerations show that all derivatives of the quantities c_1, c_2, l_1^0, l_2^0 and l_5^0 with respect to the stem cell parameters fraction of self-renewal, proliferation rate and division probability are products of the exponentially decreasing function $c_1(t)$ and a factor that can be bounded by an affine linear function of t. Thus, the effects of an altered stem cell parameter on the number of existing cells and the number of BrdU incorporating cells weakens with time due to the decline of the stem cell compartment. The effect of altered stem cell parameters on the number of existing astrocytes, $\partial_p c_5(t;p)$, cannot be analyzed due to no available data on the sign of $\alpha\beta+\gamma$ with α, β and γ as stated in Table 2.3.

The sign sequences of the parameter-derivatives of the five quantities stem cell count, progenitor count and BrdU incorporating stem cells, progenitors and astrocytes with respect to the three considered parameters a_1, p_1 and θ_1 are summarized in Table 2.7.

p	$\sigma(\partial_p c_1(t;p))$ (stem cells)	$\sigma(\partial_p l_1^0(t;p))$ (Brdu inc. stem cells)	$\sigma(\partial_p c_2(t;p))$ (progenitors)	$\sigma(\partial_p l_2^0(t;p))$ (Brdu inc. progenitors)	$\sigma(\partial_p l_5^0(t;p))$ (BrdU inc. astrocytes)
a_1	$(0,+)$	$(+)$	$(0,-,0,+)$	$(-,0,+)$	$(-,0,+)$
p_1	$(0,-)$	$(+,0,-)$	$(0,+,0,-)$	$(+,0,-)$	$(+,0,-)$
θ_1	$(0,+)$	$(+)$	$(0,+)$	$(+)$	$(+)$

Table 2.7: Time-dependent responses of stem cells (c_1), BrdU incorporating stem cells (l_1^0), progenitors (c_2), BrdU incorporating progenitors (l_2^0) and BrdU incorporating astrocytes (l_5^0) respectively to an infinitesimal increase of the respective parameter p

2.4 Numerical Investigations

2.4.1 Parameter Estimations

Our neurogenesis model (2.1.1) involves 18 parameters, including 5 parameters for the initial data of each compartment. We first analyze the parameter region that is consistent with the data presented in [21, Supplementary Table 2]. For this purpose, we simultaneously fit the analytical solution of (2.1.1) either for c_1 and $\frac{c_1}{c_2}$

(the number of stem cells and the ratio of the number of stem cells to the number of progenitors), depicted in Figure 2.8, or for c_1 and c_2 (the number of stem cells and the number of progenitors), shown in Figure 2.9.

Throughout this thesis, fitting is performed using the `NonlinearModelFit` procedure of *Mathematica* 9 to numerically minimize the weighted sum of squared residuals (cf. equation (A.2.8) in the appendix). The weights σ_i were chosen as the data points standard error of the mean as documented in the Mathematica reference for fitting data involving measurement errors. For numerical minimization, the random search method is chosen, which results in the highest coefficient of determination (R^2) value among all available minimization methods.

Fitting the solution for $c_1(t)$ and $\frac{c_1(t)}{c_2(t)}$ of (2.1.1) to the stated data results in the values

$$\begin{aligned} (2a_1\theta_1 - 1)p_1 &= -5.16 \times 10^{-4}\,\mathrm{h}^{-1}, \\ \kappa\theta_1 2(1-a_1)p_1 &= 1.71 \times 10^{-4}\,\mathrm{h}^{-1}, \\ d_2 - (2a_2\theta_2 - 1)p_2 &= 5.41 \times 10^{-4}\,\mathrm{h}^{-1}. \end{aligned} \tag{2.4.1}$$

Although this parameter set agrees with the data of [21] for the age related decay of the stem cell compartment and the dynamics on the ratio of the number of stem cells and the number of progenitors, the fit underestimates the number of progenitors at late time points (Figure 2.8). Conversely, fitting the solution for $c_1(t)$ and $c_2(t)$ to the data of [21, Supplementary Table 2] displays agreement with the stem cell and the progenitor compartment, but not with their ratio (Figure 2.9).

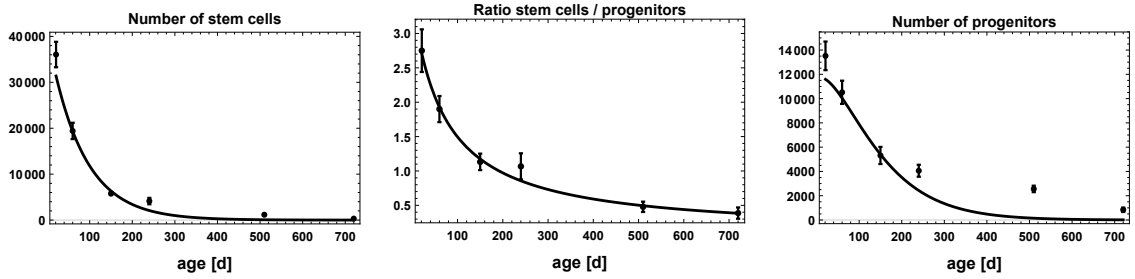


Figure 2.8: Fitting model (2.1.1) to the number of stem cells and the ratio of the number of stem cells to the number of progenitors in [21, Supplementary Table 2] results in a poor agreement with the number of progenitors.

It appears that the decay of the stem cell compartment involves a saturation effect for late time points, which cannot be reproduced by our linear model. More

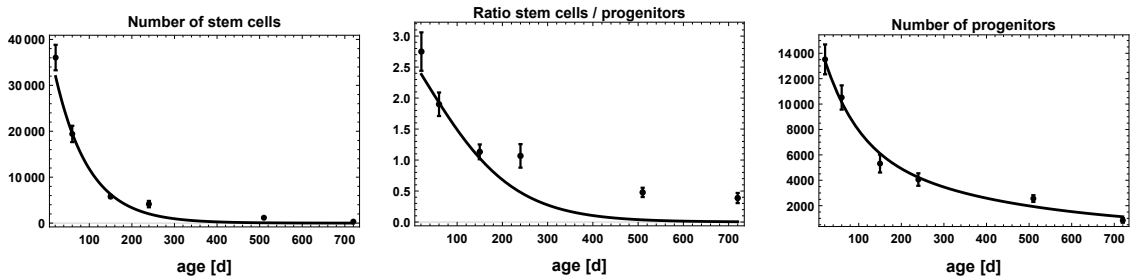


Figure 2.9: Fitting model (2.1.1) to the number of stem cells and the number of progenitors in [21, Supplementary Table 2] results in a poor agreement with the ratio of the number of stem cells to the number of progenitors.

precisely, the solution of (2.1.1) for the number of stem cells (c_1) is an exponentially declining curve and fitting this curve to the data estimates 5 stem cells remaining at 2 years of age, while this number was measured to be 320. We hypothesize that this saturation is caused by either a feedback mechanism on stem cells that induces their quiescence with increasing age or by the existence of a mixture of two populations with one population performing adult neurogenesis and a quiescent one. Moreover, physiological conditions of a_1, p_1, θ_1 and κ, i.e. $a_1 \approx 0.5$, $1/p_1 < 2$ years, $\theta_1 > 0.1$ and $\kappa > 0.5$ contradict the restrictions imposed by (2.4.1), potentially because of the missing saturation effect that cannot be explained with our model. Thus, our current model indicates that there are some novel aspects in adult neurogenesis required to explain the experimental data. The proposed explanation of how saturation could be achieved should be the subject of future experimental validation.

2.4.2 Simulations

The mathematical analysis conducted in section 2.3 depends on Assumption 2.2.2. Since the purpose of this section is to extend our analysis to the effects of altered stem cell parameters on the number of neurons (c_4) and the number of BrdU labeled cells ($l_i(t,\tau)$) for $\tau > 0$, we utilize a parameter set satisfying Assumption 2.2.2 for our numerical investigations. Unless otherwise stated, we set $a_1 = 0.55$, $\theta_1 = 0.7$, $p_1 = 1$, $\kappa = 0.6$, $a_2 = 0.7$, $\theta_2 = 0.4$, $p_2 = 2.5$, $d_2 = 0.1$, $p_3 = 1.5$, $d_3 = 0.4$, $\hat{t} = 10$, $d_4 = 0.05$, $d_5 = 0.05$, $c_1(0) = 10000$, $c_2(0) = 5000$, $c_3(0) = 15000$, $c_4(0) = 350000$, $c_5(0) = 100000$ to numerically solve (2.1.1) by employing the `NDSolve` framework of Mathematica. Note that the chosen parameter set does not include any time units,

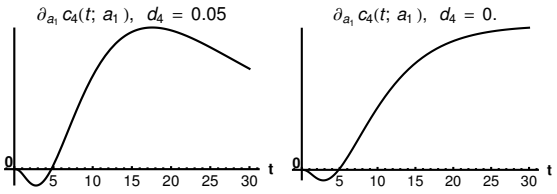
since a choice of physiological parameters is not feasible as explained in section 2.4.1. Thus, there are no time units in any figure using this parameter set.

At first, we investigate the effect of altered stem cell parameters on the number of existing neurons at time t after the KO, i.e. $\partial_p c_4(t;p)$ for $p \in \{a_1, p_1, \theta_1\}$.

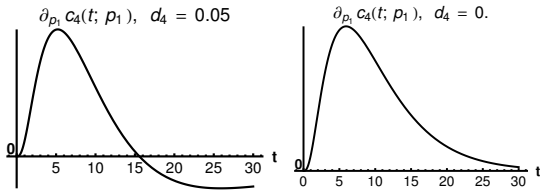
An increase of a_1, which increases the proportion of symmetric stem cell divisions at the expense of asymmetric divisions, displays the same qualitative progression on the number of neurons as the effect of an increase in a_1 on the number of progenitors: Initially, the neuron count is decreased, followed by an always ongoing period displaying an increase. Moreover, simulations suggest that the magnitude of the increase weakens with time, i.e. $\lim_{t\to\infty} \partial_{a_1} c_4(t;a_1) = 0$. Interestingly, the weakening effect depends on the decay rate of neurons: If there is no decay ($d_4 = 0$), we find that $\partial_{a_1} c_4(t;a_1)$ converges to a positive value (Figure 2.10a). Hence, the magnitude of the increase will not decline with time if neurons do not decay.

Increasing the stem cell proliferation rate p_1 affects the number of neurons in the same way as the number of progenitors. Consequently, the neuron count is increased in the initial period while it is decreased afterwards. The magnitude of the increase declines with time. Interestingly, assuming that neurons do not decay, the two-phase progression is lost and the number of neurons always increases but the magnitude of the increase weakens with time (Figure 2.10b).

Increasing the probability of stem cell divisions θ_1 yields an increase in the number of neurons. If $d_4 > 0$, our simulations indicate that the magnitude of the increase converges to zero and if $d_4 = 0$ that it converges to a positive value (no figure shown).



(a) Effect of an increased fraction of self-renewal with and without death of neurons respectively



(b) Effect of an increased proliferation rate with and without death of neurons respectively

Figure 2.10: Responses to an infinitesimal increase in stem cells fraction of self-renewal and proliferation rate respectively on the number of neurons

Our analysis indicates that the effect of altered stem cell dynamics on the number of neurons depends on the death rate of neurons. Thus, investigating whether

neurons have the ability to decay and if so determining their decay rate is vital in order to understand the short- and long-term impact of altered stem cell dynamics on the number of neurons in the dentate gyrus.

Next, we turn our investigations to the quantity $l_i(t^\star, \tau)$, the number of BrdU labeled cells of cellular compartment i, where BrdU was applied at a particular time point $t = t^\star$ and $t = 0$ corresponds to the time point of the knockout. The effect of an increase of stem cells fraction of self-renewal on BrdU labeled cells depends on the time point $t^\star$. If BrdU was given directly after the fraction of self-renewal was increased, i.e. $t^\star = 0$, the number of BrdU labeled progenitors $l_2(0, \tau)$ shows a two-phase progression with an initial decrease and subsequent increase. The same holds true for the number of BrdU labeled neurons $l_4(0, \tau)$ (Figure 2.11, first row). As the time $t^\star$ between the change of dynamics and the BrdU administration increases, the first phase that shows a decrease in the number of progenitors and neurons becomes shorter (Figure 2.11, second row) until there remains only a one-phase progression, with increased numbers of BrdU labeled progenitors and neurons at every time point τ after BrdU was given (Figure 2.11, third row).

Similarly to an increased fraction of self-renewal, the effect of an increased stem cell proliferation rate depends on the time $t^\star$ between the change of dynamics an BrdU administration. If $t^\star = 0$, BrdU labeled progenitors and neurons display a two-phase progression with increased numbers in the initial phase and decreased numbers subsequently (Figure 2.12, first row). Interestingly, the two-phase progression after BrdU labeling gains an additional phase as $t^\star$ increases, displaying now an initial decrease, a subsequent increase and again a later on decrease in the number of BrdU labeled progenitors and neurons (Figure 2.12, second row). Increasing $t^\star$ further, the three-phase progression will be lost and then the number of labeled progenitors and neurons is decreased at any time τ after BrdU was given (Figure 2.12, third row).

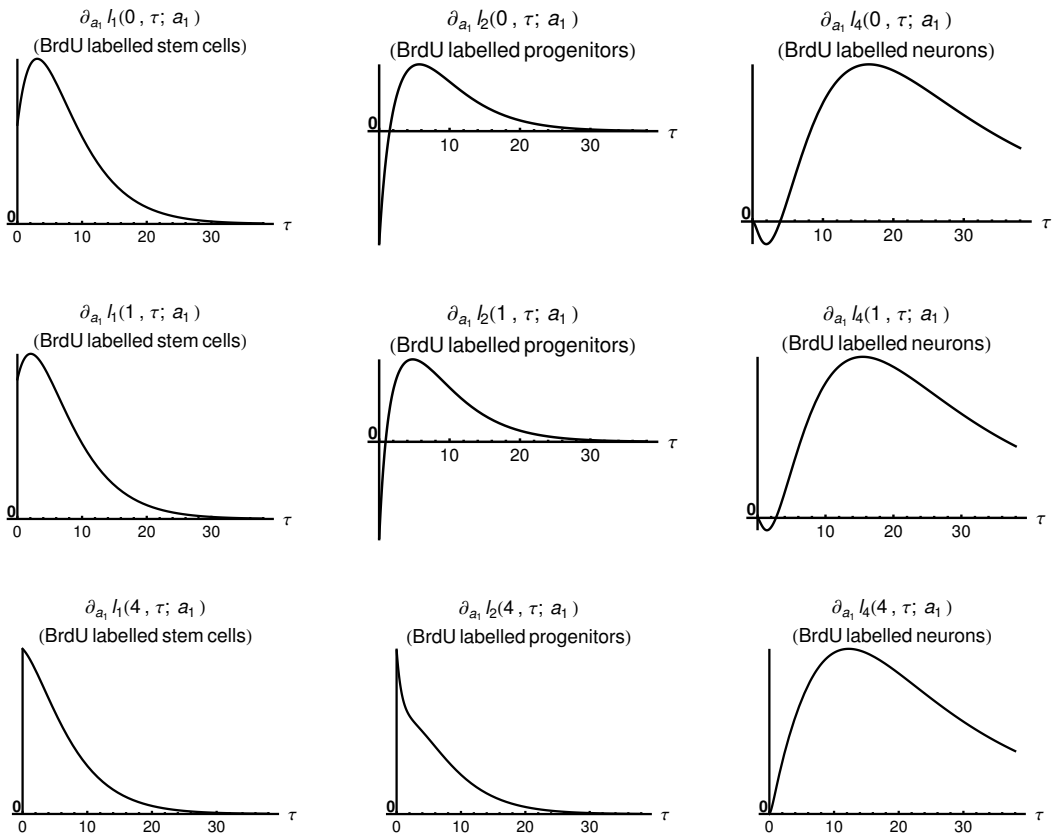


Figure 2.11: Effect of an infinitesimally increased stem cell self-renewal on BrdU labeled stem cells, progenitors and neurons at time τ after BrdU labeling, where BrdU was given at time $t = 0$ (first row), $t = 1$ (second row) and $t = 4$ (third row)

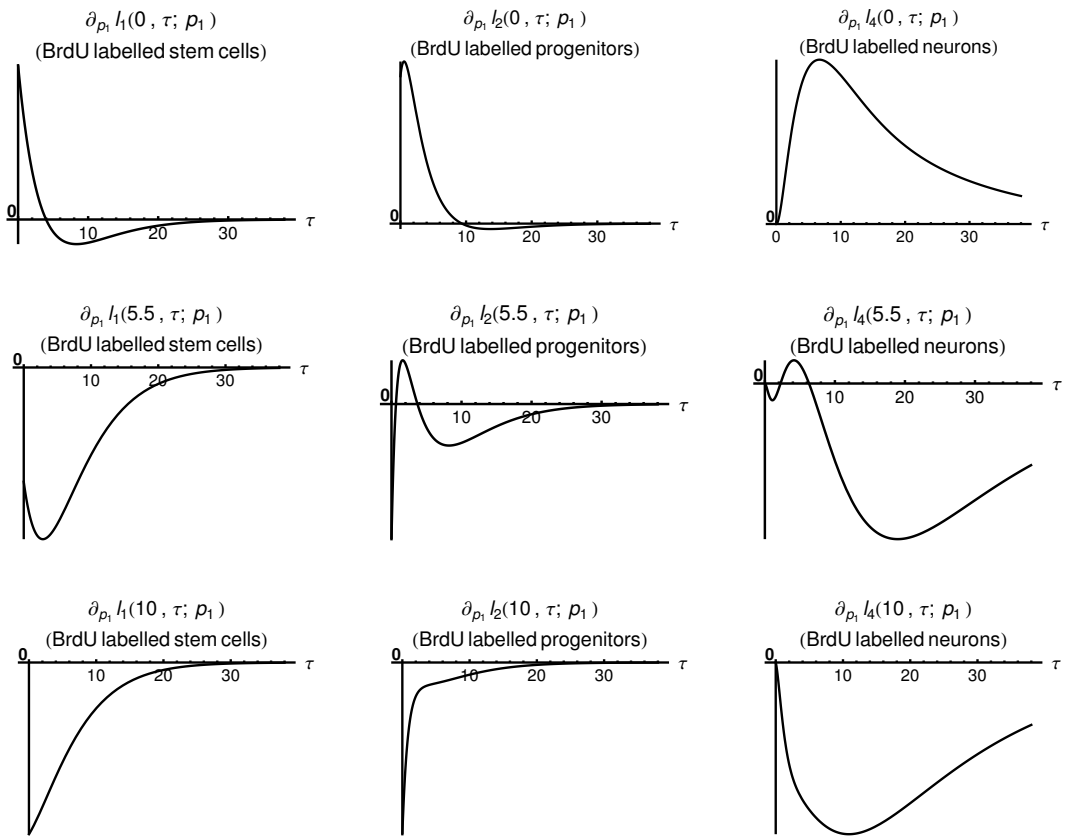


Figure 2.12: Effect of an infinitesimally increased stem cell proliferation rate on BrdU labeled stem cells, progenitors and neurons at time τ after BrdU labeling, where BrdU was given at time $t = 0$ (first row), $t = 5.5$ (second row) and $t = 10$ (third row)

2.5 Summary

We have established a mathematical model of adult hippocampal neurogenesis based on experimental data. Although we consider a basic model not accounting for any feedback mechanisms or a spatial component, we demonstrate that modifying the dynamics of adult neural stem cells, which corresponds to inducing a stem cell targeting knockout, exhibits a rich variety of effects due to the high complexity of the hippocampal neurogenic niche.

Our investigation shows that observed differences in cell numbers due to altered stem cell dynamics depend not only on the alteration that was induced by a particular knockout but also on the time at which cell counts were measured. Therefore, it is necessary to perform measurements at multiple time points in order to draw conclusions from knockout experiments. Moreover, we find that cell numbers of cells at different differentiation stages may respond qualitatively different to altered stem cell dynamics. Additionally, this response may also be different between the total number of cells and the number of BrdU labeled cells of a given cellular compartment and also depends on the time point of BrdU administration after the knockout. Thus, labeling cells with BrdU does not generate a subset of the neurogenesis system that is sufficient to perform an analysis of the impact of altered stem cell dynamics. The reason for this is that changing stem cell dynamics influences either the number of cell divisions or the ratio of symmetric to asymmetric divisions, which in turn affects the initial distribution of cell types among all BrdU labeled cells. Furthermore, our reasoning proves that under the assumptions of our model, the effect of altered stem cell dynamics declines as time passes and that this decline is a result of the depletion of the stem cell pool. Evaluating differences in cell counts at late time points after a knockout can therefore be used to test the notion of a declining stem cell population.

Chapter 3

A Refined Model of Stem Cells

Although multiple studies have been conducted in the past to identify qualitative features of neural stem cells (NSCs) such as multipotency or the age-related decline of the NSC pool, a quantitative understanding of the dynamics of adult neurogenesis is still missing. This lack of quantification is mainly due to sparse data and diverse labeling approaches used by different studies in order to observe NSCs. In particular, the two landmark studies of Bonaguidi et al. [9] and Encinas et al. [21] propose different hypotheses about the dynamics of neural stem cells.

In this chapter, we examine both hypotheses by formulating them as ODE models. We find that the Bonaguidi hypothesis explains a wider range of data than the one of Encinas, provided that it subscribes to Encinas' theory of NSC depletion. Moreover, our analysis shows that repeated activation is an important feature of NSCs in order to maintain a constant fraction of proliferating stem cells. In addition, we provide new data to test Encinas' theory that NSCs deplete by transforming into astrocytes and find that these transformation events can account for at least 40% of the NSC decline. Finally, we develop novel experimentally testable predictions and reveal possible age-related changes during adult hippocampal neurogenesis.

3.1 Problem Formulation

Understanding the dynamic behavior of adult hippocampal NSCs in the process of neuron production has been subject to many studies in recent years [9, 10, 11, 21, 39, 52]. Moreover, much of current research has been focused in discovering age-related changes during hippocampal neurogenesis [7, 24, 31, 32, 33, 51, 58].

The two landmark studies on neural stem cell dynamics have been conducted by Bonaguidi et al. [9] and Encinas et al. [21]. Although both studies focus on identifying the dynamics of NSCs, they arrive at different conclusions. Bonaguidi et al. [9] performed clonal analysis and concluded that NSCs can get activated multiple times from their quiescence in order to produce offspring. Upon division, NSCs return to quiescence and have the ability to get activated again. Additionally, they infer that NSCs are multipotent and can give rise to multiple cell types such as astrocytes, transit amplifying progenitors and other stem cells. In contrast, Encinas et al. [21] carried out population level analysis and reasoned that NSCs get activated only once, enter a series of asymmetric divisions by producing additional progenitor cells and finally vanish by transforming, i.e. directly differentiating, into an astrocyte. We will refer to the reasoning of Bonaguidi et al. [9] as *repeated activation hypothesis* and to the one of Encinas et al. [21] as *one-time activation hypothesis*.

We show that the repeated activation hypothesis has a greater capability of collectively explaining the data, since repeated activation is a robust construct in order to maintain the observed constant fraction of proliferating stem cells. Moreover, the repeated activation dynamics needs to be accompanied by the ability of NSCs to perform depletion in general and astrocytic transformation in particular. Otherwise, the age-related decline of NSC numbers in conjunction with an accumulation of astrocytes cannot be explained.

3.2 Modeling Data

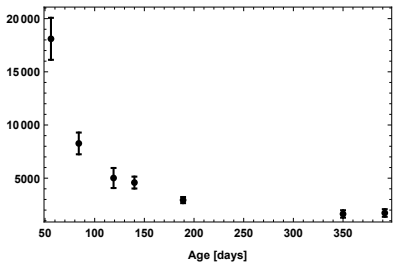
In order to investigate the one-time and repeated activation hypothesis, it is important to examine how well both theories can explain experimental data. We consider two data sets. The first, which we refer to as *population level data* was generated by Sascha Dehler (Martin-Villaba laboratory) as an extension of the data set of Encinas et al. [21]. In the experiment, the number of NSCs, the fraction of BrdU incorporating NSCs and the number of astrocytes was measured at several time points during adulthood (Figures 3.1a-3.1c).

The new data set confirms the results of ref. [21], showing that the number of NSCs declines during aging (Figure 3.1a) and that at the same time the fraction of BrdU incorporating NSCs remains constant at about 1% (Figure 3.1b). Interestingly, we

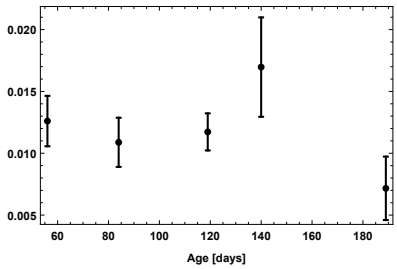
also find an accumulation of the number of astrocytes during aging, an observation pointing towards the in ref. [21] suggested relationship between NSC decline and the production of new astrocytes (Figure 3.1c).

The second data set, which we refer to as *clonal data*, has been published by Bonaguidi et al. [9]. In the study, individual NSCs were labeled at an age of 8-12 weeks and the clonal progeny of these cells was evaluated at one month, two months and one year after labeling. The authors then classified these NSC clones in the categories *quiescent*, i.e. only one NSC is present, *active*, i.e one NSC and at least one additional cell is present and *depleted*, i.e. no NSC is present (Figure 3.1d).

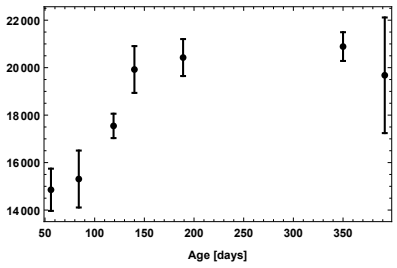
The discussed clonal data set complements the population level data and the aim of our subsequent analysis of the one-time and repeated activation model is to evaluate whether the two hypotheses can explain both data sets.



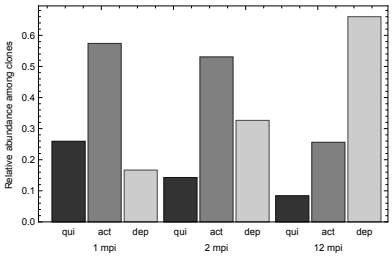
(a) Number of NSCs per mm^3



(b) Fraction of BrdU incorporating NSCs



(c) Number of astrocytes per mm^3



(d) Plot of the clonal data published by Bonaguidi et al. [9]. *Qui* denotes quiescent, *act* active and *dep* depleted clones.

Figure 3.1: Data used during modeling. The population level data of Figures a-c were generated in the lab of Prof. Martin-Villalba. The clonal data displayed in Figure d has been published in ref. [9].

3.3 Analysis of the One-Time Activation Model

We implement the one-time activation hypothesis of Encinas et al. [21] with the equations

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}c_0(t) &= -rc_0(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_1^n(t) &= rc_0(t) - pc_1^n(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_1^k(t) &= pc_1^{k+1}(t) - pc_1^k(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}c_1^0(t) &= pc_1^1(t) - qc_1^0(t),
\end{aligned} \tag{3.3.1}$$

where $n \in \mathbb{N}$ is the maximum number of NSC divisions. Moreover, c_0 denotes quiescent NSCs and c_1^k $(1 \leq k \leq n)$ cycling NSCs with k divisions remaining. The variable r represents the activation rate of quiescent NSCs, p the division rate of proliferating NSCs and q is the rate of NSCs with no remaining divisions to transform into astrocytes.

Figure 3.2 depicts a graphical representation of system (3.3.1). For subsequent simulations, we consider the case of $n = 3$ asymmetric divisions, which has been suggested by Encinas et al. [21]. Assuming a different number of asymmetric divisions leads to similar results and does not impact the conclusions drawn.

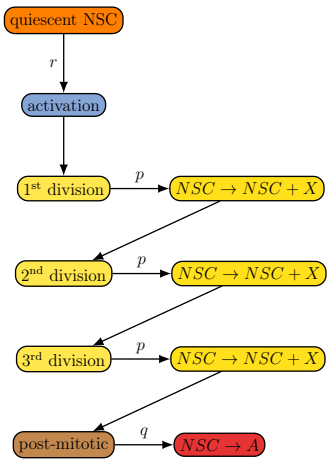


Figure 3.2: Graphical representation of system (3.3.1). Quiescent NSCs can get activated to enter the cell cycle, subsequently perform a series of (three) asymmetric divisions by producing a NSC and another cell (X) before entering the post-mitotic stage and transforming into an astrocyte (A).

The dynamics of system (3.3.1) is given by

Theorem 3.3.1. *Let $n \in \mathbb{N}$ and $p > q > r > 0$. Let $C(t) = (c_0(t), c_1^n(t), \ldots, c_1^0(t))^\intercal$ be the solution of* (3.3.1) *and denote with $\mathrm{e}^\intercal C(t)$ the sum of all components of $C(t)$. Then*

$$\lim_{t\to\infty} \frac{C(t)}{\mathrm{e}^\intercal C(t)} = \bar{c} > 0$$

if $c_0(0) > 0$. Moreover, if $C(0) = \bar{c}$, then

$$\mathrm{e}^\intercal C(t) = e^{-rt} \tag{3.3.2}$$

and

$$\frac{c_1^1(t) + \ldots + c_1^n(t)}{\mathrm{e}^\intercal C(t)} = \left(1 - \frac{r}{q}\right)\left(1 - \left(1 - \frac{r}{p}\right)^n\right). \tag{3.3.3}$$

To proof the above Theorem, we need an auxiliary lemma and a corollary.

Lemma 3.3.2. *Let A be a matrix and let $x(t)$ be the solution of*

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}x(t) &= Ax(t), \\ x(0) &= x_0 \end{aligned}$$

where x_0 can be written as linear combination of normalized eigenvectors of A, i.e.

$$x_0 = \sum_i \mu_i v_i$$

with $Av_i = \lambda_i v_i$ and $\mathrm{e}^\intercal v_i = 1$. Let

$$\lambda_{max} = \max_i \lambda_i.$$

Then

$$\lim_{t\to\infty} \frac{x(t)}{\mathrm{e}^\intercal x(t)} = \left(\sum_{\{j|\lambda_j=\lambda_{max}\}} \mu_j v_j\right) \Bigg/ \left(\sum_{\{j|\lambda_j=\lambda_{max}\}} \mu_j\right)$$

Proof. We use the fact that if v is an eigenvector of A with eigenvalue λ, v is also an eigenvector of $\exp(tA)$ with eigenvalue $e^{\lambda t}$. It holds

$$\frac{x(t)}{\mathrm{e}^\intercal x(t)} = \frac{\exp(tA)x_0}{\mathrm{e}^\intercal \exp(tA)x_0} = \frac{\exp(tA)\sum_i \mu_i v_i}{\mathrm{e}^\intercal \exp(tA)\sum_i \mu_i v_i} = \frac{\sum_i \mu_i e^{\lambda_i t} v_i}{\sum_i \mu_i e^{\lambda_i t}}.$$

Hence

$$\lim_{t\to\infty} \frac{x(t)}{\mathrm{e}^\intercal x(t)} = \lim_{t\to\infty} \frac{e^{-\lambda_{max} t}\sum_i \mu_i e^{\lambda_i t} v_i}{e^{-\lambda_{max} t}\sum_i \mu_i e^{\lambda_i t}} = \left(\sum_{\{j|\lambda_j=\lambda_{max}\}} \mu_j v_j\right) \Bigg/ \left(\sum_{\{j|\lambda_j=\lambda_{max}\}} \mu_j\right).$$

□

Corollary 3.3.3. *Let A be a matrix with the largest eigenvalue $\lambda^\star$ having geometric multiplicity 1. Let $x(t)$ be the solution of*

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = Ax(t),$$
$$x(0) = x_0$$

where x_0 has an eigenvector decomposition

$$x_0 = \mu^\star v^\star + \sum_i \mu_i v_i$$

with $\mu^\star \neq 0$, $Av^\star = \lambda^\star v^\star$, $Av_i = \lambda_i v_i$, $\lambda^\star > \max_i \lambda_i$, $\mathrm{e}^\intercal v^\star = 1$ and $\mathrm{e}^\intercal v_i = 1$. Then

$$\lim_{t\to\infty} \frac{x(t)}{\mathrm{e}^\intercal x(t)} = v^\star.$$

Moreover, if $x_0 = v^\star$,

$$x(t) = e^{\lambda^\star t} v^\star.$$

Proof. Lemma 3.3.2 implies

$$\lim_{t\to\infty} \frac{x(t)}{\mathrm{e}^\intercal x(t)} = (\mu^\star v^\star)/\mu^\star = v^\star.$$

Since $v^\star$ is an eigenvector of A, $x_0 = v^\star$ implies

$$x(t) = \exp(tA)v^\star = e^{\lambda^\star t}v^\star.$$

□

Proof of Theorem 3.3.1. System (3.3.1) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}C(t) = (-p\mathbb{I}_{n+2} + A)C(t)$$

with

$$A = \begin{pmatrix} p-r & & & & & & \\ r & 0 & & & & & \\ & p & \ddots & & & & \\ & & \ddots & \ddots & & & \\ & & & \ddots & \ddots & & \\ & 0 & & & \ddots & 0 & \\ & & & & & p & p-q \end{pmatrix} \in \mathbb{R}^{n+2\times n+2}.$$

The solution $C(t)$ satisfies

$$C(t) = \exp(t(-p\mathbb{I}_{n+2} + A))C(0) = e^{-pt}\exp(tA)C(0).$$

The matrix A has characteristic polynomial

$$\chi_A(x) = (-x)^n(p-q-x)(p-r-x)$$

and the condition $p > q > r > 0$ implies that

$$\lambda^\star = p - r$$

is the largest eigenvalue of A and has geometric multiplicity 1. Thus, Corollary 3.3.3 implies that $x(t) := \exp(tA)C(0)$ being the solution of

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}x(t) &= Ax(t), \\ x(0) &= C(0) \end{aligned}$$

satisfies

$$\lim_{t\to\infty}\frac{C(t)}{\mathrm{e}^{\intercal}C(t)}=\lim_{t\to\infty}\frac{x(t)}{\mathrm{e}^{\intercal}x(t)}=\bar{c}$$

where $\bar{c}$ is the eigenvector of A with eigenvalue $\lambda^{\star}$ satisfying

$$\mathrm{e}^{\intercal}\bar{c}=1.$$

Moreover, Corollary 3.3.3 implies that if $C(0)=\bar{c}$,

$$C(t)=e^{-pt}x(t)=e^{-pt}e^{\lambda^{\star}t}\bar{c}=e^{-rt}\bar{c}.$$

Thus

$$\mathrm{e}^{\intercal}C(t)=e^{-rt}$$

and

$$\frac{c_1^1(t)+\ldots+c_1^n(t)}{\mathrm{e}^{\intercal}C(t)}=\frac{\sum_{i=2}^{n+1}C(t)^{(i)}}{\mathrm{e}^{\intercal}C(t)}=\sum_{i=2}^{n+1}\bar{c}^{(i)}$$

where the superscript $^{(i)}$ denotes the i-th component of a vector. The vector

$$v=\left(\frac{(p-r)^n}{rp^n},\frac{(p-r)^{n-1}}{p^n},\ldots,\frac{(p-r)^0}{p},\frac{1}{q-r}\right)^{\intercal}$$

is an eigenvector of A with eigenvalue $p-r$. Hence,

$$\bar{c}=\frac{v}{\mathrm{e}^{\intercal}v}.$$

We have

$$\sum_{i=2}^{n+1}v^{(i)}=\frac{1}{p}\sum_{j=0}^{n-1}\left(\frac{p-r}{p}\right)^j=\frac{1-\left(\frac{p-r}{p}\right)^n}{r}$$

and

$$\mathrm{e}^{\intercal}v=\frac{(p-r)^n}{rp^n}+\frac{1-\left(\frac{p-r}{p}\right)^n}{r}+\frac{1}{q-r}=\frac{q}{r(q-r)}.$$

Thus

$$\frac{c_1^1(t)+\ldots+c_1^n(t)}{\mathrm{e}^{\intercal}C(t)}=\sum_{i=2}^{n+1}\bar{c}^{(i)}=\frac{\sum_{i=2}^{n+1}v^{(i)}}{\mathrm{e}^{\intercal}v}=\left(1-\frac{r}{q}\right)\left(1-\left(1-\frac{r}{p}\right)^n\right).$$

□

To estimate the parameters of model (3.3.1), we assume that at $t = 0$, the begin of adult age, the system is already in the steady state $\bar{c}$ mentioned in Theorem 3.3.1. We justify this assumption with the observation of a constant fraction of BrdU incorporating stem cells throughout life (Figure 3.1b). The value of the proliferation rate p can be inferred from the literature [11], since the cell cycle length of NSCs was measured as

$$T_c = 22.8\,\mathrm{h}$$

and we can interpret this value as the doubling-time of an exponential growth process, leading to

$$p = \log(2)/T_c.$$

The remaining two parameters q and r can be estimated from the population level data using Theorem 3.3.1, since they appear in the equation for the total number of NSCs (3.3.2) and the equation for relative amount of cycling NSCs on all NSCs (3.3.3). To estimate q and r, the decline of the number of NSCs (Figure 3.1a) is fitted to a function of the form

$$ne^{-\alpha t}$$

and the fraction of BrdU incorporating NSCs (Figure 3.1b) is fitted to a constant value c. The resulting equations for q and r are thus given by

$$\alpha = r$$

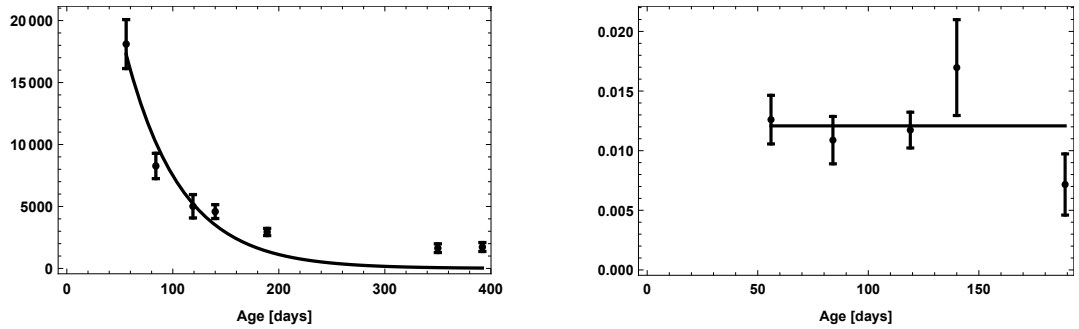
and

$$c = \left(1 - \frac{r}{q}\right)\left(1 - \left(1 - \frac{r}{p}\right)^n\right)\frac{T_s}{T_c}.$$

Here, we use that the fraction of BrdU incorporating NSCs is the product of the fraction of cycling NSCs on all NSCs times the relative length of the S-phase in the cell cycle, i.e. T_s/T_c, with $T_s = 9.7\,\mathrm{h}$ [11].

As can be seen from Figure 3.3, the one-time activation model (3.3.1) can be fit to our population level data of the decline of the NSC count and the fraction of BrdU incorporating NSCs. To compare the dynamics of this fit to the clonal data of Bonaguidi et al. [9], we use the Gillespie method [26] to translate (3.3.1) into a

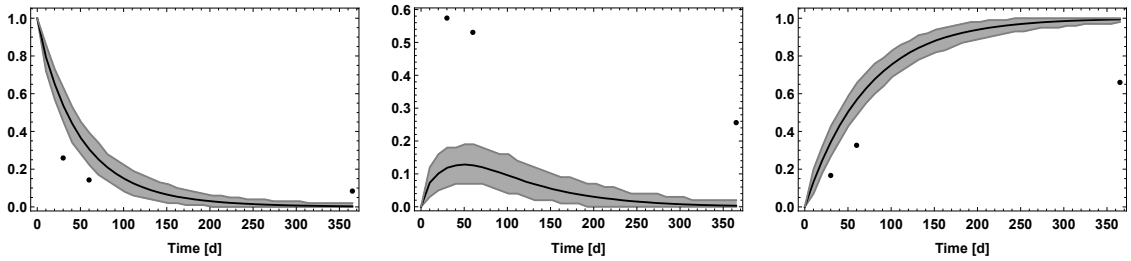
corresponding stochastic processes. However, comparing the dynamics derived from the population level fit to the clonal data, we find that the clonal behavior of NSCs cannot be predicted accurately (Figure 3.4). The reason for this is on the one hand the rapid rise of the fraction of activated clones within the first two months after labeling (Figure 3.4b) and on the other hand the persistence of a fraction of $\sim 8\%$ of quiescent clones over a one year period (Figure 3.4a).



(a) Model fit to the number of NSCs

(b) Model fit to the fraction of BrdU incorporating NSCs

Figure 3.3: Fit of the one-time activation model (system (3.3.1)) to population level data (Figures 3.1a and 3.1b).



(a) Fraction of quiescent clones

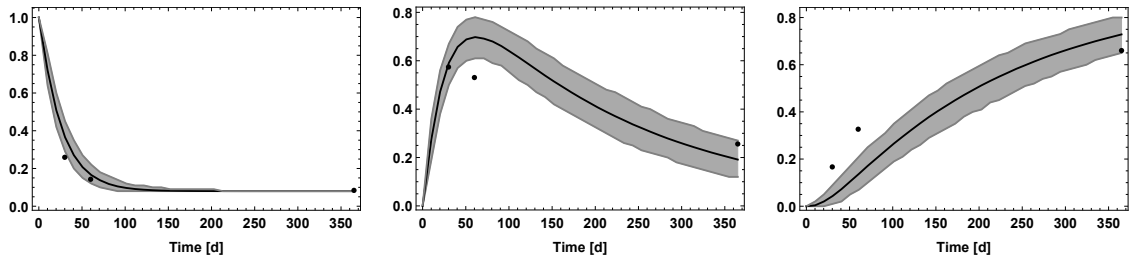
(b) Fraction of activated clones

(c) Fraction of depleted clones

Figure 3.4: Simulation of the clonal evolution of the one-time activation model with parameters estimated from population level data (Figure 3.3). Results are obtained by simulating 100 NSC clones for 1000 times. Simulation data is represented as mean (solid black line) and (gray) band containing 95% of all simulated trajectories. Black dots correspond to the data of Figure 3.1d.

Conversely, the one-time activation model can be fitted to the clonal data of Bonaguidi et al. [9] by estimating model parameters based on this clonal data set

(Figure 3.5). Interestingly, in order to explain the clonal data with the one-time activation model, it is necessary to introduce an additional population of *resilient* NSCs, cells that never deplete or get activated. Without this extra population, which makes up $\sim 8\%$ of all NSCs in a ten weeks old mouse, the fraction of quiescent clones would decrease very rapidly to zero and the one year time point of the fraction of quiescent clones could not be matched.



(a) Fraction of quiescent clones

(b) Fraction of activated clones

(c) Fraction of depleted clones

Figure 3.5: Fit of the one-time activation model to the clonal data of Figure 3.1d.

Although the one-time activation model can be fit to the clonal data of Bonaguidi et al. [9], the dynamics necessary to obtain this fit is biologically implausible (Figure 3.6): Assuming an initial composition of the NSC pool with 8% of resilient NSCs and the remaining cells being in quiescent phase, then, under the dynamics of the fit of Figure 3.5, within less than three months the fraction of cycling NSCs would initially grow and thereafter decline to make up less than 1% of all NSCs. However, a dynamics causing cycling NSCs to make up less than 1% of the whole stem cell population within only three months contradicts the observation of 1% of all NSCs being in S-phase and incorporating BrdU constantly throughout aging ([21] and Figure 3.1b).

In conclusion, the one-time activation model can be fit to the population level data (Figure 3.3), but then it fails to explain the clonal data (Figure 3.4). Conversely, the model can be fit to the clonal data (Figure 3.5), but the dynamics necessary to obtain this fit are accompanied by a rapid extinction of the pool of cycling NSCs (Figure 3.6), contradicting the observation of dividing NSCs in old-age animals.

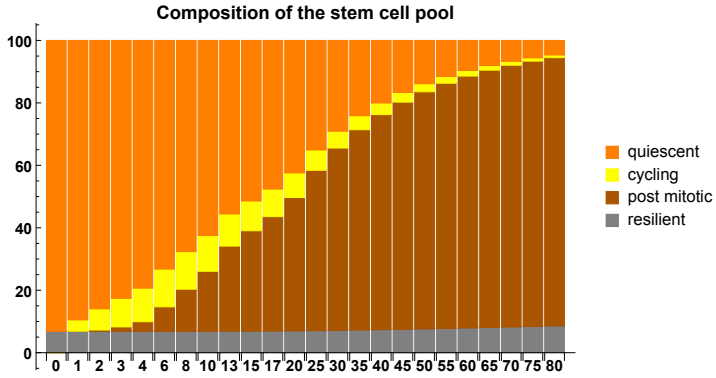


Figure 3.6: Time course of the composition of the NSC pool in the one-time activation model with the dynamics of the clonal fit (Figure 3.5). X-axis represents time in days, Y-axis relative amount in per cent.

3.4 Analysis of the Repeated Activation Model

The repeated activation hypothesis of Bonaguidi et al. [9] is implemented with the equations

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}c_0 &= -(r+q)c_0 + 2apc_1, \\ \frac{\mathrm{d}}{\mathrm{d}t}c_1 &= rc_0 - pc_1, \end{aligned} \tag{3.4.1}$$

where c_0 represents quiescent NSCs and c_1 cycling NSCs (Figure 3.7). Quiescent NSCs are assumed to either enter the cell cycle with rate r or deplete with rate q. The justification for the depletion process is that the data of Bonaguidi et al. [9] display an increase of the number of NSC depleted clones (Figure 3.1d) and the reasoning of the authors that there exist NSC depletion events. Moreover, cycling NSCs are assumed to proliferate with rate p. Since the repeated activation hypothesis also considers symmetric NSC divisions, we introduce the parameter a as the fraction of self-renewal, which is the probability of a progeny cell to have the same fate as the mother cell [40]. Moreover, since the hypothesis of Bonaguidi et al. [9] assumes that cycling NSCs return to quiescence after division, the resulting $2a$ units of progeny cells that stay stem cells are assumed to return to the compartment c_0.

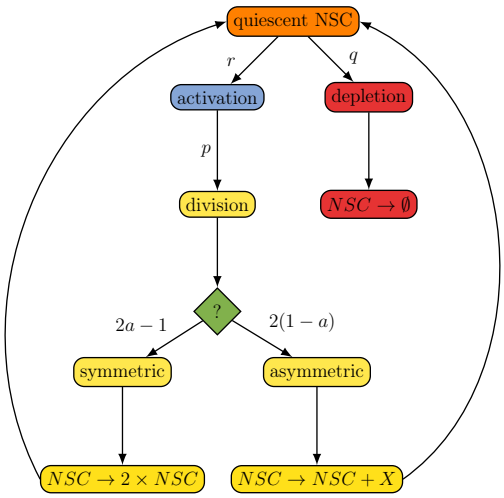


Figure 3.7: Graphical representation of system (3.4.1). Quiescent NSCs are either activated to enter the cell cycle and subsequently perform a symmetric or asymmetric division, or vanish from the pool of NSCs by performing a depletion event. Moreover, cycling NSCs re-enter the quiescent phase after division. Thus, stem cells have the potential get activated repeatedly.

The dynamics of system (3.4.1) is given by

Theorem 3.4.1. *Define* $s := \sqrt{(p-r-q)^2 + 8apr}$, *then the solution of* (3.4.1) *satisfies*

$$\lim_{t\to\infty} c_0(t)/c_1(t) = \varphi := \frac{p-r-q+s}{2r} \tag{3.4.2}$$

and if $c_0(0)/c_1(0) = \varphi$, *it holds that*

$$c_1(t)/(c_0(t)+c_1(t)) = 1/(1+\varphi) \tag{3.4.3}$$

and

$$c_0(t) + c_1(t) = \exp\left(-\frac{t}{2}(p+q+r-s)\right). \tag{3.4.4}$$

Proof. Rewriting (3.4.1) as

$$\frac{\mathrm{d}}{\mathrm{d}t}c(t) = Ac(t),$$

where $c(t) = (c_0(t), c_1(t))^\intercal$, and

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{pmatrix} -(r+q) & 2ap \\ r & -p \end{pmatrix}, \tag{3.4.5}$$

it follows that

$$s = \sqrt{(a-d)^2 + 4bc}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{c_0}{c_1}\right) = b + \left(\frac{c_0}{c_1}\right)\left(a - d - c\left(\frac{c_0}{c_1}\right)\right). \tag{3.4.6}$$

Since $c = r > 0$, the right-hand side of the above equation defines a downward-opened parabola with roots

$$x_{1/2} = \frac{a-d \pm s}{2c}.$$

Since $bc = 2apr > 0$,

$$s > |a-d|$$

and there is exactly one positive root

$$\varphi := \frac{a-d+s}{2c}, \tag{3.4.7}$$

which is the globally asymptotically stable steady state of (3.4.6) thus showing (3.4.2).

The solution of (3.4.1) is given by

$$c(t) = \exp(tA)c(0) = e^{\delta}\left(\cosh(\Delta)\mathbb{I} + \frac{\sinh(\Delta)}{\Delta}\begin{pmatrix}\gamma & tb\\ tc & -\gamma\end{pmatrix}\right)\begin{pmatrix}c_0(0)\\ c_1(0)\end{pmatrix}$$

with

$$\delta = \frac{t}{2}(a+d), \quad \gamma = \frac{t}{2}(a-d), \quad \Delta = \sqrt{\gamma^2 + t^2bc}. \tag{3.4.8}$$

Here, we have used a general formula for the two-dimensional matrix exponential [57] and $\mathbb{I}$ denotes the two-dimensional identity matrix. It holds $c_0(0) = \varphi/(\varphi+1)$ and $c_1(0) = 1/(\varphi+1)$. Hence,

$$\begin{aligned} c_0(t) + c_1(t) &= \frac{e^{\delta}}{\varphi+1}\left(\cosh(\Delta)(\varphi+1) + \frac{\sinh(\Delta)}{\Delta}(\gamma\varphi + tb + tc\varphi - \gamma)\right)\\ &= e^{\delta}\left(e^{\Delta} - \sinh(\Delta) + \frac{\sinh(\Delta)}{\Delta(\varphi+1)}(\gamma\varphi + tb + tc\varphi - \gamma)\right). \end{aligned}$$

From (3.4.7) and (3.4.8), it follows that $\varphi = (\gamma + \Delta)/(tc)$. Thus, $tc\varphi - \gamma = \Delta$ and

$$\gamma\varphi + tb = \frac{\gamma\Delta + \gamma^2}{tc} + tb = \frac{\gamma\Delta + \Delta^2}{tc} = \Delta\varphi.$$

Taken together, it holds

$$\frac{\gamma\varphi + tb + tc\varphi - \gamma}{\Delta(\varphi + 1)} = \frac{\Delta\varphi + \Delta}{\Delta(\varphi + 1)} = 1,$$

and consequently

$$c_0(t) + c_1(t) = e^{\delta+\Delta}.$$

Substituting back using (3.4.8) and then (3.4.5) establishes (3.4.4). □

To estimate parameters for model (3.4.1), we assume for the fraction of self-renewal a that $2a - 1$, the corresponding probability of a NSC division being symmetric, is 5% (cf. equation (2.1.2)). We justify this choice of a with the suggestion of Bonaguidi et al. [9] that the fraction of symmetric NSC divisions on all divisions is relatively small. A different value of a does not affect the drawn conclusions. Next, we follow the same parameter estimation procedure for system (3.4.1) as already conducted for system (3.3.1) using steady state initial data: Equation (3.4.4) implies that the total number of NSCs is given by an exponentially declining function $ne^{-\alpha t}$ and equation (3.4.3) implies that the ratio of the number of cycling NSCs to all NSCs is given by a constant value c. The resulting equations for q and r are thus given by

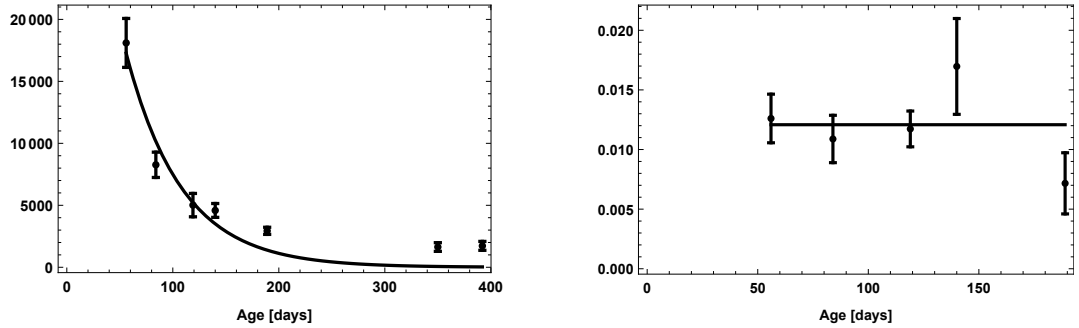
$$\alpha = \frac{p + q + r - s}{2}$$

and

$$c = \frac{1}{1 + \varphi} \cdot \frac{T_s}{T_c}.$$

Theorems 3.3.1 and 3.4.1 show that the temporal progression of the total number of NSCs and the ratio of cycling NSCs on all NSCs are identical for model (3.3.1) and model (3.4.1) under steady state initial conditions (function of the form $ne^{-\alpha t}$ respectively as constant value c). Accordingly, the fit of model (3.4.1) to the population level data is identical to the one of model (3.3.1) (Figure 3.8). This indicates that the one-time activation hypothesis of Encinas et al. [21] and the repeated acti-

vation hypothesis of Bonaguidi et al. [9] are indistinguishable solely with population level data.

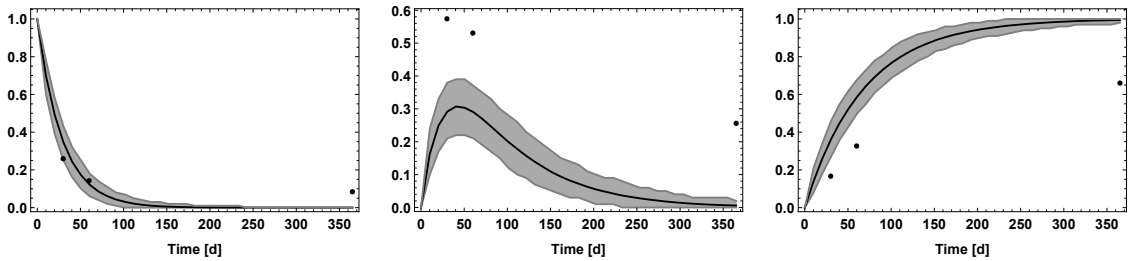


(a) Model fit to the number of NSCs

(b) Model fit to the fraction of BrdU incorporating NSCs

Figure 3.8: Fit of the repeated activation model (system (3.4.1)) to the population level data (Figures 3.1a and 3.1b).

Comparing the dynamics derived from the population level fit of the repeated activation model to the clonal data shows a lack of fit for the same reasons as in the one-time activation case (Figure 3.9). Moreover, the repeated activation model can be fit to the clonal data as well, assuming a population of resilient NSCs (Figure 3.10). An important aspect of this fit is that it does not lead to a biologically implausible contradiction as is the case for the one-time activation model: Under the dynamics of the fit of Figure 3.10, the fraction of cycling NSCs stays relatively constant over a two year period (Figure 3.11), a feature which is necessary to explain the constant fraction of BrdU incorporating NSCs during aging.

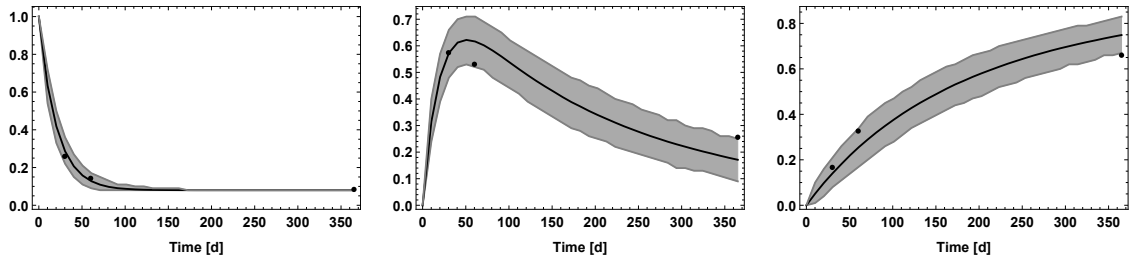


(a) Fraction of quiescent clones

(b) Fraction of activated clones

(c) Fraction of depleted clones

Figure 3.9: Simulation of the clonal evolution of the repeated activation model with parameters estimated from population level data (Figure 3.8).



(a) Fraction of quiescent clones (b) Fraction of activated clones (c) Fraction of depleted clones

Figure 3.10: Fit of the repeated activation model to the clonal data of Figure 3.1d.

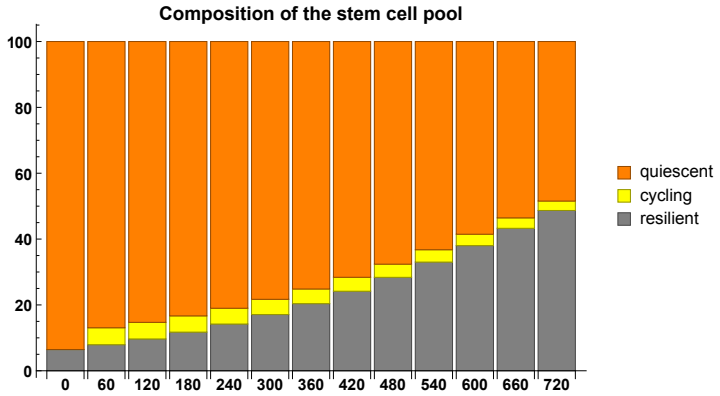


Figure 3.11: Time course of the composition of the NSC pool in the repeated activation model with the dynamics of the clonal fit (Figure 3.10). X-axis represents time in days, Y-axis relative amount in per cent. As can be seen, the relative amount of cycling NSCs stays constant throughout aging.

To summarize, both, the one-time activation model (3.3.1) and the repeated activation model (3.4.1), can be fit to the population level and to the clonal data, but not simultaneously. The difference between the two models is that the repeated activation model displays a wider range of flexibility to explain the data, since it does not have biologically implausible implications when fitted to the clonal data set. Because of this flexibility, we further investigate the repeated activation model by analyzing age-related changes of the dynamics of NSCs and evaluating the hypothesis that NSCs deplete by transforming into astrocytes. The values of all estimated parameters for the one-time activation model and the repeated activation model are summarized in Table 3.12.

Figure #	Model Type	Data Type	Parameters
3.3	One-time activation	Population level	$q = 0.0265\,\mathrm{d}^{-1}$ $r = 0.0140\,\mathrm{d}^{-1}$ $\rho = 0$
3.5	One-time activation	Clonal	$q = 0.00476\,\mathrm{d}^{-1}$ $r = 0.0402\,\mathrm{d}^{-1}$ $\rho = 0.08$
3.8	Repeated activation	Population level	$q = 0.0154\,\mathrm{d}^{-1}$ $r = 0.0195\,\mathrm{d}^{-1}$ $\rho = 0$
3.10	Repeated activation	Clonal	$q = 0.00601\,\mathrm{d}^{-1}$ $r = 0.0429\,\mathrm{d}^{-1}$ $\rho = 0.08$

Table 3.12: Parameters estimated during the comparison between the one-time and repeated activation model. Figure # displays the number of the corresponding figure in the main text. Data type indicates whether the model was fitted to the population level data (Figures 3.1a and 3.1b) or to the clonal data (Figure 3.1d). The parameter ρ is the fraction of resilient NSCs on all NSCs at the start of the experiment.

3.5 Age-Dependent Changes of Neural Stem Cell Dynamics

The age-related decline of the NSC pool is the central characteristic of the neuron production process in the adult hippocampus. A related question is whether this decline occurs uniformly throughout aging or whether the dynamics of NSCs change during adulthood.

Exponential decline is one of the most widely observed decay processes in nature. Thus, it is natural to expect that the decline of NSC numbers is also exponential, implying a constant decline rate of NSC throughout aging. However, as outlined in our previous study [62], the decline of NSC numbers saturates during aging (Figure 3.13), indicating a non-constant decline. To discuss possible mechanisms that can account for this saturation, we modify system (3.4.1) in several ways. Our subsequent evaluation consists of six possible mechanisms, of which only the first one is biologically plausible. The reason for this is that in addition to explaining the

saturation pattern of NSC's decline, the mechanism also needs to predict the time course of BrdU incorporating NSCs (Figure 3.1b).

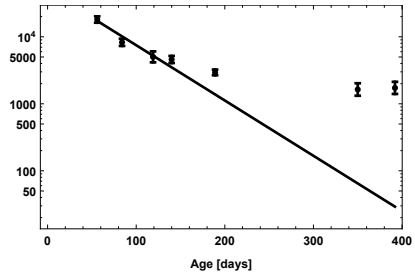


Figure 3.13: Logarithmic plot of the time course of NSC numbers (data of Figure 3.1a) together with the fit of an exponentially declining function. At late time points, the number of NSCs is higher than expected based on the exponential decline process, indicating a saturation effect.

3.5.1 Discussion of Possible Scenarios

Decreasing Depletion

The repeated activation model contains two competing processes associated with quiescent NSCs. Activation causes NSCs to enter the cell cycle and leads to a maintenance or expansion of the NSC pool, depending on whether an asymmetric or symmetric division occurred. In contrast, depletion causes NSCs to disappear. The relative frequency of the two processes determines whether the pool of NSCs expands or declines over time.

The depletion process is the integral part of the NSC decline. Without it, the number of NSCs would not decrease. One possibility to explain the saturation of NSC's decline is that the depletion rate of NSCs declines during aging, leading to a decreased fraction of depleting stem cells at old age. Accordingly, we modify (3.4.1) by assuming

$$q = q(t) = q_{\max} e^{-\beta_q t}. \tag{3.5.1}$$

As can be seen from Figure 3.14, the discussed mechanism is a very good explanation for the population level data.

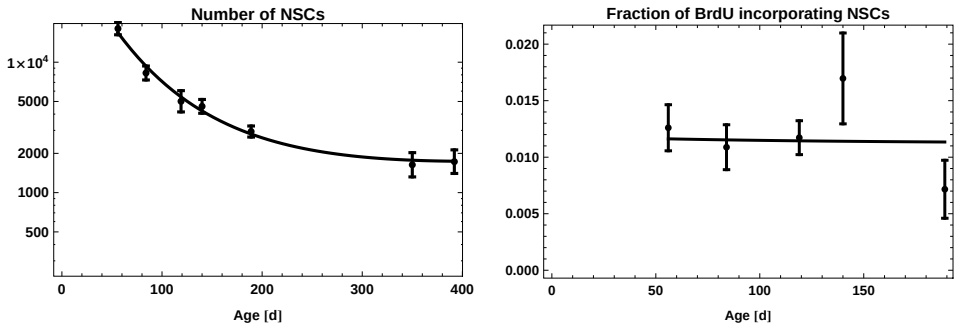


Figure 3.14: Fit of the repeated activation model to population level data, assuming a decreasing fraction of depleting NSCs during aging.

Increasing Activation

Another possibility to explain the saturation of the NSC decline is to assume that the fraction of quiescent NSCs that get activated per time unit increases during aging. Thus, activation will counteract the decline of NSC numbers at old age. The increasing activation scenario is implemented with

$$r = r(t) = r_{\min} + \alpha_r t.$$

However, comparing the above mechanism to our population level data shows that an increase of activation fails to explain the saturation of NSC's decline (Figure 3.15).

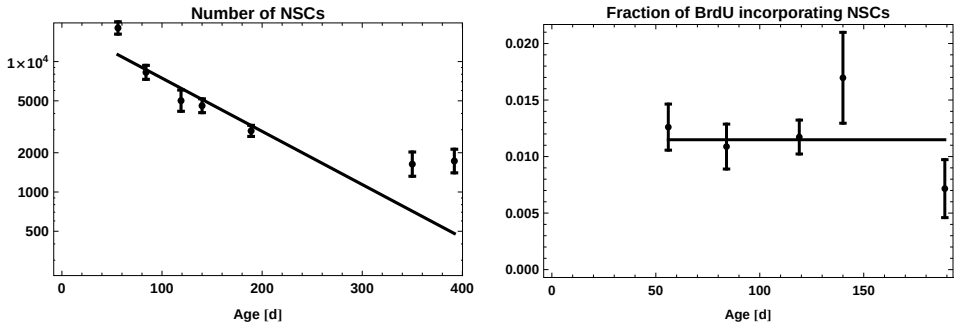


Figure 3.15: Fit of the repeated activation model to population level data, assuming that the fraction of activated stem cells increases during aging.

Increasing Quiescence

An increase of NSC's quiescence, corresponding to an age-related lengthening of the G_0 phase, could also explain the decline pattern of NSCs. Since leaving the quiescent phase is associated with a higher tendency to deplete than to maintain or expand the pool of NSCs—otherwise it could not be explained why this pool declines—remaining in quiescence would neutralize the decline. The modification of system (3.4.1) takes the form

$$q = q(t) = q_{\max} e^{-\beta_{qr} t}$$

and

$$r = r(t) = r_{\max} e^{-\beta_{qr} t}.$$

The justification for these equations is that leaving the quiescent phase c_0 is driven by a joint decay process consisting of activation (rate $r(t)$) and depletion (rate $q(t)$) [38]. The mean time of a NSC to sojourn in quiescence is thus given by

$$\frac{1}{q(t) + r(t)} = \frac{1}{q_{\max} + r_{\max}} e^{\beta_{qr} t}.$$

However, in order to explain the saturation of NSC's decline in this scenario, the strong increase of quiescence reduces the fraction of cycling NSCs and alongside the fraction of BrdU incorporating stem cells, contradicting what is observed in the data (Figure 3.16).

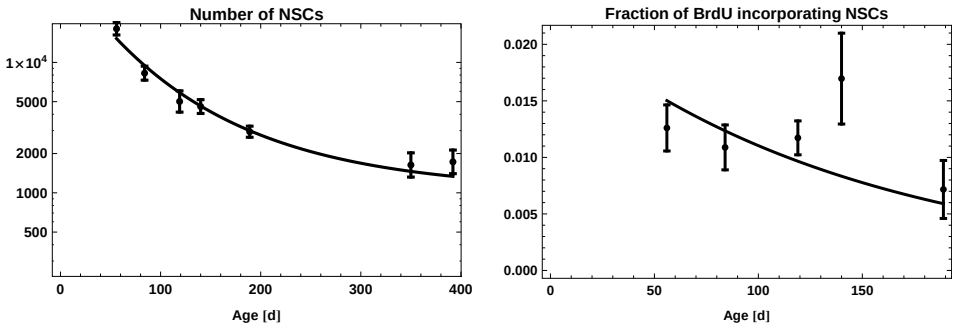


Figure 3.16: Fit of the repeated activation model to population level data, assuming that stem cells stay progressively longer in quiescence during aging.

Increasing Self-Renewal

The saturation of the NSC decline could also indicate that NSCs increase their self-renewal to counteract the depletion. The corresponding modification of (3.4.1) takes the form

$$a = a(t) = a_{\min} + \alpha_a t.$$

An analysis of this scenario shows that in order to explain our population level data, NSC numbers would start to increase after about 1 year of age (Figure 3.17). However, this implication contradicts the fact that NSCs and other downstream compartments such as neural progenitors and immature neurons decline in numbers even at later time points than one year of age [21, 58].

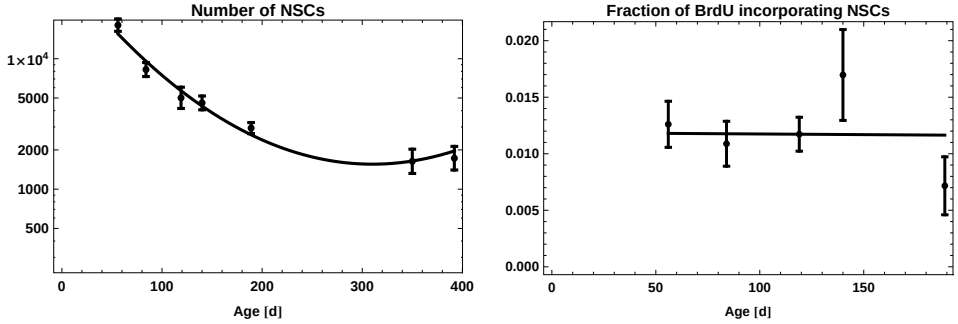


Figure 3.17: Fit of the repeated activation model to population level data, assuming that stem cells increase their self-renewal during aging.

Lengthening of the Cell Cycle

If NSCs take progressively longer during aging to complete the cell cycle, the number of stem cells entering the quiescent phase declines with time, which in turn leads to a decreasing net depletion of stem cells residing in quiescence. Accordingly, we assume a decline of the proliferation rate given by

$$p = p(t) = p_{\max} e^{-\beta_p t}.$$

A comparison of the suggested mechanism with the data again shows a lack of fit (Figure 3.18).

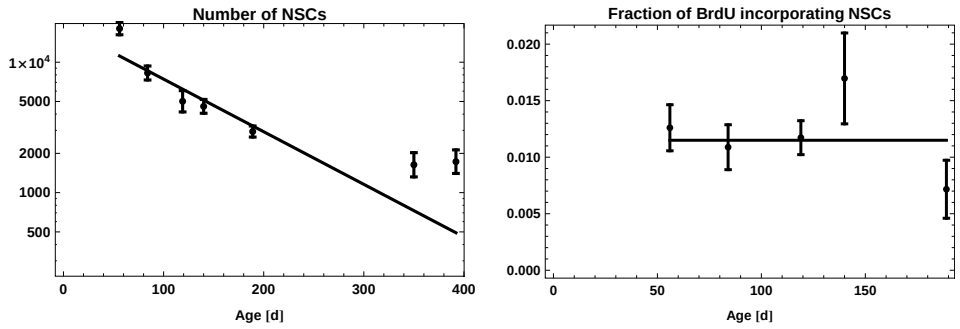


Figure 3.18: Fit of the repeated activation model to population level data, assuming an age-related lengthening of NSCs cell cycle.

Existence of Resilient NSCs

Our preceding analysis of the clonal data set of Bonaguidi et al. [9] points towards a possible second population of resilient NSCs, cells that can neither get activated nor deplete. The additional population c_{res} is implemented with

$$\frac{\mathrm{d}}{\mathrm{d}t}c_{\mathrm{res}}(t) = 0,$$
$$\frac{c_{\mathrm{res}}(0)}{c_0(0) + c_1(0)} = \frac{\rho}{1-\rho},$$

where ρ is the fraction of resilient NSCs on all NSCs at the start of adulthood. However, the existence of such a population in combination with the decline of NSCs leads to a decrease of the fraction of cycling NSCs (Figure 3.19), contradicting the observed constant fraction of BrdU incorporating NSCs [21].

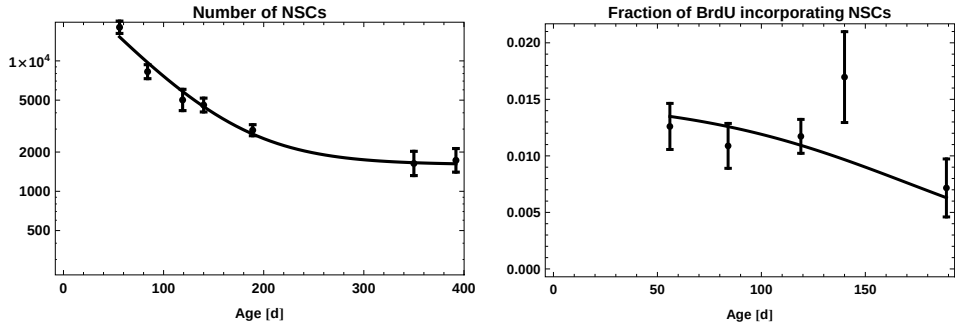


Figure 3.19: Fit of the repeated activation model to population level data, assuming an additional population of resilient NSCs.

3.5.2 Model Selection

To further evaluate the plausibility of the different scenarios of age-dependent NSC dynamics, we can compare their corresponding Akaike weights [12]. Table 3.20 displays the weight of each scenario together with the estimated values of the corresponding parameters. The recommendation is that the level of empirical support of a certain model i is substantial, if $0 \leq \Delta_i \leq 2$, considerably less, if $4 \leq \Delta_i \leq 7$ and essentially none, if $\Delta_i > 10$ [12, Section 2.6]. Thus, the only two considerable mechanisms besides the decreasing depletion mechanism are the scenario of increasing self-renewal and the existence of a population of resilient NSCs. However, as outlined previously, the former scenario contradicts the observation of a decline of NSCs and downstream cell types at old age [58], while the latter contradicts the observation of a constant fraction of BrdU incorporating NSCs [21].

Figure #	Mechanism	Parameters	R^2	AICc	Δ_i
3.14	Decreasing depletion	$q_{\max} = 0.0431\,\mathrm{d}^{-1}$ $r = 0.0198\,\mathrm{d}^{-1}$ $\beta_q = 0.00921\mathrm{d}^{-1}$	0.915	242.2	0
3.15	Increasing activation	$q = 0.0107\,\mathrm{d}^{-1}$ $r_{\min} = 0.0200\,\mathrm{d}^{-1}$ $\alpha_r = 0\,\mathrm{d}^{-1}$	0.865	267.5	25.3
3.16	Increasing quiescence	$q_{\max} = 0.0307\,\mathrm{d}^{-1}$ $r_{\max} = 0.0383\,\mathrm{d}^{-1}$ $\beta_{qr} = 0.00707\,\mathrm{d}^{-1}$	0.903	249.6	7.4
3.17	Increasing self-renewal	$q = 0.0241\,\mathrm{d}^{-1}$ $r = 0.0203\,\mathrm{d}^{-1}$ $\alpha_a = 0.00183\,\mathrm{d}^{-1}$	0.910	245.7	3.5
3.18	Cell cycle lengthening	$q = 0.0106\,\mathrm{d}^{-1}$ $r = 0.0200\,\mathrm{d}^{-1}$ $\beta_p = 0\,\mathrm{d}^{-1}$	0.865	267.5	25.3
3.19	Resilient population	$q = 0.0206\,\mathrm{d}^{-1}$ $r = 0.0261\,\mathrm{d}^{-1}$ $\rho = 0.04$	0.910	245.6	3.4

Table 3.20: Parameters estimated during the analysis of different mechanisms to explain the saturation of the NSC decline. AICc is the small sample size corrected Akaike information criterion and Δ_i the corresponding Akaike weight [12].

3.6 Neural Stem Cell Depletion by Astrocytic Transformation

In order to explain the decline of NSC numbers during aging, Encinas et al. [21] suggested that NSCs deplete by transforming into astrocytes. To test this theory, we calculate the expected number of astrocytes based on the age-related decline of NSC numbers and compare this prediction to our newly generated data. To analyze this hypothesis, we take model (3.4.1) together with our best explanation of the saturation of the NSC decline (equation (3.5.1)) and add a new compartment c_2 of astrocytes satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}c_2 = \theta q(t)c_0(t).$$

Here, $\theta \in [0,1]$ is the fraction of NSC depletion events where depletion occurs via astrocytic transformation. An estimation of this parameter yields $\theta = 0.395$, indicating that astrocytic transformation can account for about 40% of the NSC decline (Figure 3.21).

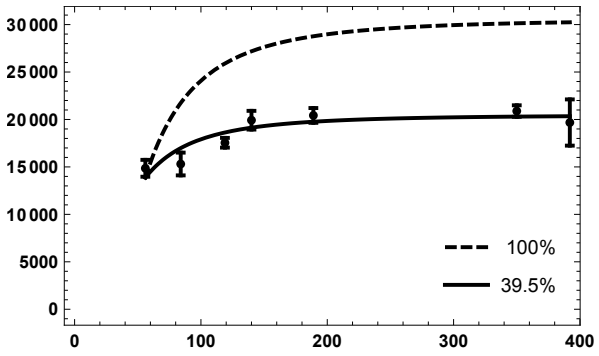


Figure 3.21: Age-related accumulation of astrocyte numbers. Dashed line represents expected number of astrocytes, assuming that 100% of the NSC decline is caused by astrocytic transformation. Solid line results from fitting the fraction of transformation events to the data of Figure 3.1c, indicating that $\sim 40\%$ of the NSC decline is caused by astrocytic transformation.

3.7 Predictions

One central element in mathematical modeling of natural phenomena is to develop model-based predictions, which can be tested experimentally by future studies. This part of model validation is an important step within the scientific method [46].

In the previous section, we discussed possible mechanisms to explain the saturation pattern of the decline of NSC numbers. Our best explanation is that of a declining fraction of depleting NSCs during aging. In order to quantify this time-dependent depletion process given by equation (3.5.1), recall that leaving the quiescent phase c_0 is a joint decay process consisting of activation (rate r) and depletion (rate $q(t)$). Thus, the fraction of stem cells leaving c_0 within a period δ is $1 - e^{-(q(t)+r)\delta}$ and a fraction $q(t)/(q(t)+r)$ of those cells undergoes depletion. The fraction ψ of stem cells that deplete at age t within a period δ is thus given by

$$\psi(t,\delta) = \frac{q_{\max} e^{-\beta_q t}}{q_{\max} e^{-\beta_q t} + r}\Big(1 - e^{-(q_{\max} e^{-\beta_q t}+r)\delta}\Big).$$

Using the corresponding parameter values from Tables 3.12 and 3.20, the quantification predicts that the fraction of NSCs that deplete within $\delta = 1$ week decreases from 15% at the age of two months via 1% at one year of age to 0.2% at one and a half years of age (Figure 3.22).

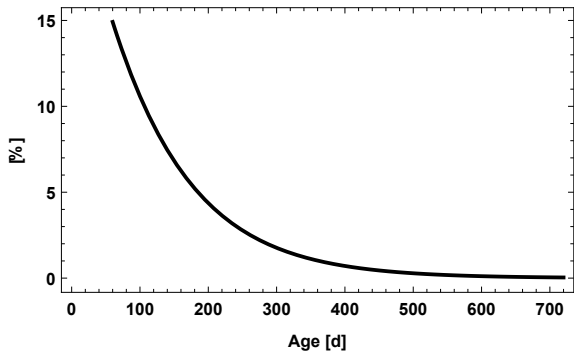


Figure 3.22: Predicted time course of the fraction $\psi(t,\delta)$ of NSCs that undergo depletion within $\delta = 1$ week.

Another prediction concerns the total number of divisions a certain stem cell performs. Entering the cell cycle multiple times is the key element of the repeated activation model. Thus, we can ask how the number of already performed divisions is

distributed in the whole pool of existing NSCs along aging. To answer this question, we have simulated 20000 stem cells starting at 2 months of age, which is the total number of existing NSCs at that age, and then evaluated the number of performed divisions at different time points during aging (Figure 3.23). As can be seen, the resulting time course of the distribution of performed divisions depends on whether the dynamics of the best clonal fit (Figure 3.10) or the best population level fit (Figure 3.14) is assumed. In the clonal case, the fraction of NSCs with no divisions initially drops and later on increases. This dynamics can be explained with the postulated existence of resilient NSCs. Initially, quiescent NSCs that enter the cell cycle and perform divisions cause a drop of the fraction of cells with no divisions. As time progresses, the fraction of resilient NSCs starts to grow owing to the decline of non-resilient NSCs. In contrast, the fraction of NSCs with no divisions constantly declines under the population level dynamics. Moreover, the difference between the dynamics of the clonal and the population level fit is that stem cells with divisions have performed on average almost twice as many divisions in the clonal case than in the population level one.

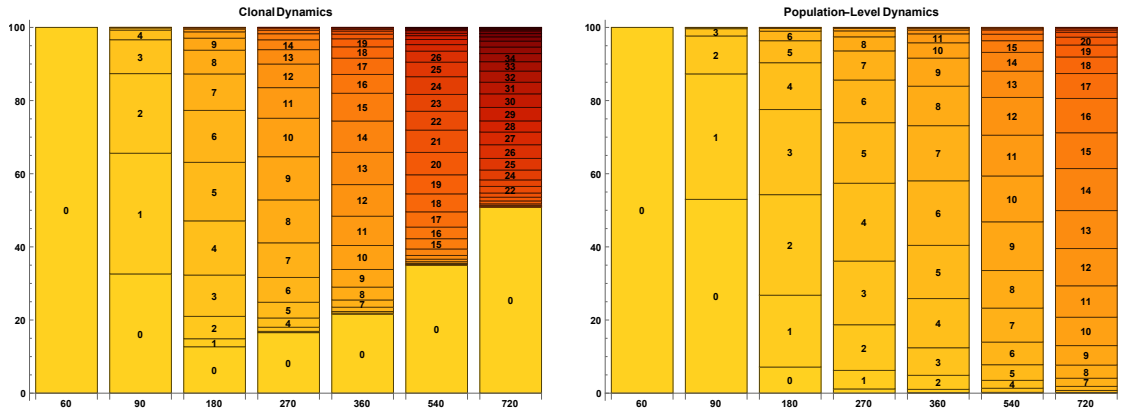


Figure 3.23: Composition of the pool of existing NSCs regarding the number of divisions performed since the age of 2 months. Left panel assumes the dynamics of the best fit to the clonal data (Figure 3.10). Right panel assumes the dynamics of the best fit to the population level data (Figure 3.14). To obtain the respective distributions, 20000 stem cells were simulated. X-axis represents time in days, Y-axis relative amount in per cent.

3.8 Summary

The studies of Bonaguidi et al. [9] and Encinas et al. [21] arrived at different theories about the dynamics of neural stem cells in order to explain their respective data. Although both hypotheses seem to be fundamentally different, we were able to use mathematical modeling in order to identify features of both models which are plausible in the context of our own as well as their published data.

In a first step, we have seen that both models are able to explain our population level data of the age-related decline of NSC numbers as well as the constant fraction of BrdU incorporating stem cells. However, only the repeated activation model can predict the clonal data of Bonaguidi et al. [9]. The reason for this is that the one-time activation model causes NSCs to accumulate in the post-mitotic phase, leading to a rapid extinction of the pool of cycling stem cells. In contrast, repeated activation allows NSCs to revisit the cell cycle, thus keeping a constant fraction of cycling cells. Moreover, we found that the clonal data not only allows to distinguish between one-time and repeated activation, but also points towards a subpopulation of resilient NSCs, cells that do not have the ability to either enter the cell cycle or to deplete. The biological implication of such a population remains speculative, but one possible explanation could be that these cells constitute a backup of non-resilient NSCs and could be activated in case of demand.

We then asked for possible mechanisms that can explain the saturation pattern of the age-related decline of NSC numbers, an observation we were able to reproduce from the study of Encinas et al. [21]. After evaluating multiple mechanisms, the only explanation left is that of a declining fraction of stem cells performing depletion events. Moreover, a quantification of the decline needed in order to reproduce the saturation has led to a new prediction.

The second major postulation of the one-time activation hypothesis is that NSCs deplete by transforming into astrocytes. Here, we were able to partially confirm this hypothesis by observing an age-related accumulation of astrocyte numbers and our mathematical analysis showed that these transformation events can account for about 40% of the NSC decline. Taken together, our study shows that the best model predicting the considered data constitutes a compromise: Repeated activation together with astrocytic depletion.

Although the repeated activation model can predict the population level as well as the clonal data, it should be noted that both datasets require a different dynamics, i.e. different parameters, of the model in order to be explained. To work out this difference, we simulated a high number of NSCs with the respective dynamics to predict the relative abundance of the number of performed divisions in the pool of existing NSCs. This prediction could be used by future studies in order to differentiate between both dynamics.

An interesting aspect of the repeated activation model is that leaving the quiescent phase is a stochastic event with the two random outcomes being activation or depletion. In the case of depletion, the stem cell vanishes, whereas in the case of activation, the stem cell progresses through the cell cycle, returns to quiescence and the two possible outcomes are again randomly determined. Such stochastic behavior of single cells has been found in other tissues using mathematical modeling [16, 19, 20, 35] and also motivated theoretical work on the subject [48].

Our study shows that mathematical modeling plays an important role in understanding the dynamics of a complex cell system like the neurogenic niche of the hippocampus. Moreover, we were able to derive experimentally testable predictions that can be addressed by future studies in order to validate the suggested compromise model of repeated activation together with astrocytic transformation.

Chapter 4

The Dynamics of Progenitor Cells

Neural progenitors are generated from stem cells via asymmetric divisions [9]. It is believed that the role of these cells is a rapid expansion of the pool of undifferentiated cells via symmetric divisions, followed by a subsequent differentiation into immature neurons (neuroblasts) [21]. Understanding the dynamic behavior of progenitors is important, since the magnitude of the expansion influences the net output of a stem cell to generate new neurons.

In this chapter, we model the dynamics of neural progenitors. We propose the concept of a capacity, which is inherent in all progenitor cells and guides their potential to divide multiple times into two progenitor cells with lower capacity or to directly transform into a neuroblast. In the simplest case, the capacity can be thought of measure for the number of already performed divisions.

We find that this concept is consistent with experimental data of BrdU labeled progenitors [21], provided that the cell cycle of these cells lengthens with the number of performed divisions or that the distribution of the capacity among progenitors is *not* in a steady state at the time of the experiment. In addition, we can show that two alternative scenarios where either the rate of apoptosis is capacity dependent or where progenitors divide asymmetrically by giving rise to a progenitor with equal and one with lower capacity cannot explain the experimental data.

4.1 Problem Formulation

In the study of Encinas et al. [21], the authors observed an interesting behavior of neural progenitor cells: After labeling a cohort of dividing cells with the S-phase

marker BrdU, the number of labeled progenitors increases rapidly within 48 hours and decreases thereafter (Figure 4.1). The authors claim that this dynamics is consistent with a model stating that progenitors are born from stem cells via asymmetric divisions, perform a series of on average 2.3 symmetric divisions and subsequently transform into neuroblasts.

In the following sections, we evaluate the biological plausibility for several scenarios of neural progenitor dynamics. In section 4.2, we formulate a generalization of the dynamics suggested in ref. [21] and proof that it fails to reproduce the experimental data under steady state conditions. In section 4.3, we provide an alternative scenario—division-coupled lengthening of the cell cycle—which can explain the data with steady state assumptions. Sections 4.4 and 4.5 address two other possibilities to explain the BrdU data, but we show that both scenarios fail to reproduce the experimental observations. Section 4.6 is devoted to conclude our findings.

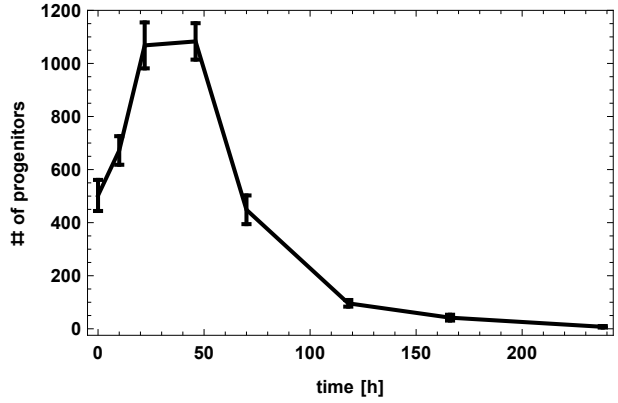


Figure 4.1: Time course of the dynamics of neural progenitor cells. Two months old mice were injected with 150 mg/kg BrdU and sacrificed at several time points after injection. Depicted is the number of BrdU positive progenitor cells. Data is reproduced from the publication of Encinas et al. [21].

4.2 A Division Capacity-Structured Model of Progenitor Dynamics

To model the proposed progenitor dynamics of Encinas et al. [21], we make the following assumptions.

Assumption 4.2.1 (Dynamics of progenitors)**.**

(**A1**) Progenitors are born from stem cells via asymmetric divisions.

(**A2**) Each progenitor has a capacity $x \in [0, 1]$.

(**A3**) A progenitor which is born from a stem cell has capacity 1.

(**A4**) A progenitor with positive capacity can either divide symmetrically by giving rise to two progenitors with lower capacity or transforms into a neuroblast.

(**A5**) A progenitor with capacity 0 always transforms into a neuroblast.

(**A6**) The rate at which division or transformation events occur is independent of the capacity.

In order to model the dynamics of progenitor cells with ordinary differential equations, we assume that the capacity x is discrete, i.e. we are given an $N \in \mathbb{N}$ such that

$$x \in \left\{0, \frac{1}{N}, \ldots, \frac{N-1}{N}, 1\right\}.$$

This assumption implies that N is the maximum number of divisions a progenitor can perform. Moreover, we assume that the division probability of a progenitor with capacity x is a $[0, 1]$-valued function $f(x)$ with $f(0) = 0$.

4.2.1 Model Formulation

Given the preceding considerations, the dynamics of progenitors is given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}P_k(t) &= 2f\left(\frac{k+1}{N}\right)pP_{k+1}(t) - pP_k(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}P_N(t) &= -pP_N(t) + b \end{aligned} \tag{4.2.1}$$

for $k = 0, \ldots, N-1$. Here, P_i is the number of progenitor cells with i remaining divisions having capacity $\frac{i}{N}$, $p > 0$ is the rate at which progenitors perform a symmetric division or transformation event and $b \geq 0$ is the rate at which progenitors of capacity 1 are born from stem cells via asymmetric divisions.

We use system (4.2.1) in two contexts: At first, we calculate the steady state distribution of the capacity among all progenitor cells by calculating the steady

state of (4.2.1) . For this, we assume a constant input into the compartment of capacity 1 progenitors, i.e. $b > 0$ and the independent time variable t means the adult life time of the animal, i.e. $t = 0$ refers to the begin of adult age. Second, we analyze the temporal progression of progenitor cells that were labeled with the S-phase marker BrdU. In this context, t refers to the time after labeling and we need to specify $P_i(0)$, i.e. the initial number of BrdU labeled progenitors having capacity of i/N. Here, we assume that the distribution of the capacity among progenitors is in a steady state at the time of BrdU administration. Furthermore, we have $b = 0$ in this scenario, since we only analyze labeled cells and progenitors that are born from stem cells after BrdU labeling stopped are not captured.

4.2.2 Analysis of BrdU-labeled Progenitors

Initial Data

As mentioned in the previous section, we assume a steady state of the capacity distribution among progenitors at the time point of BrdU labeling. Setting the right-hand side of (4.2.1) to zero results in

$$\bar{P}_N = \frac{b}{p}$$

and

$$\bar{P}_k = 2f\left(\frac{k+1}{N}\right)\bar{P}_{k+1}$$

for $k = 0, \ldots, N-1$. Consequently

$$\bar{P}_k = \frac{b}{p}2^{N-k}\prod_{i=k+1}^{N} f\left(\frac{i}{N}\right) \tag{4.2.2}$$

for all $k = 0, \ldots, N$. The number of BrdU labeled progenitors having capacity k/N is then given by

$$L_k = \rho\bar{P}_k.$$

with

$$\rho = \min\left(\frac{t_s + \delta}{t_c}, 1\right), \tag{4.2.3}$$

where t_c is the cell cycle length of progenitors, t_s the length of the S-phase and δ the length of BrdU bioavailability in the organism. To see this, consider a periodic interval of length t_c corresponding to the cell cycle (Figure 4.2). We assume that cells are uniformly distributed within the cell cycle at time point of BrdU injection. Moreover, we assume that a cell receives the BrdU label if and only if the time interval of BrdU bioavailability intersects with the S-phase. Thus, a cell is labeled with BrdU either if the time point of BrdU injection falls within the S-phase or if BrdU injection occurs outside of the S-phase and during the time of BrdU bioavailability, the S-phase is reached. The probability of the first case is the relative length of the S-phase within the cell cycle, i.e. t_s/t_c. The second case occurs if the cell is outside the cell cycle and within a duration δ the S-phase is reached, thus the corresponding probability is δ/t_c. Furthermore, we assume that the case $t_s + \delta > t_c$ is biologically implausible, since the length of the cell cycle is several hours, the S-phase makes up about half of the length of the cell cycle and BrdU bioavailability is about 1/4 hour for the data shown in Figure 4.1 [11, 21, 23, 39].

Taken together, the initial data for BrdU labeled progenitors is given by

$$P_k(0) = \frac{t_s + \delta}{t_c} \cdot \frac{b}{p} \cdot 2^{N-k} \prod_{i=k+1}^{N} f\left(\frac{i}{N}\right)$$

for $k = 0, \ldots, N$.

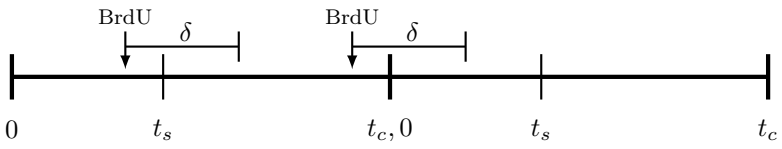


Figure 4.2: Graphical representation of the two scenarios leading to a BrdU labeled cell. In the first case, BrdU injection takes places during the S-phase, in the second, injection takes place outside the S-phase but within the time interval of BrdU bioavailability (length δ), the S-phase is reached. The time point 0 corresponds to the start of the S-phase.

Temporal Progression of BrdU-labeled Progenitors

Our previous considerations have shown that the dynamics of BrdU labeled progenitors is given by the system

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}P_k(t) &= 2f\left(\frac{k+1}{N}\right)pP_{k+1}(t) - pP_k(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}P_N(t) &= -pP_N(t), \\ P_l(0) &= \frac{t_s+\delta}{t_c}\cdot\frac{b}{p}\cdot 2^{N-l}\prod_{i=l+1}^{N} f\left(\frac{i}{N}\right) \end{aligned} \tag{4.2.4}$$

for $k = 0, \ldots, N-1$ and $l = 0, \ldots, N$. In order to compare this dynamics with the data depicted in Figure 4.1, we need to consider the total number of progenitors, i.e.

$$\mathrm{e}^{\intercal}P(t) = \sum_{k=0}^{N} P_k(t), \tag{4.2.5}$$

since it is not possible to distinguish progenitors by the number of remaining divisions based on the BrdU data.

However, the total number of progenitors of system (4.2.4) is monotonically declining, thus showing that the dynamics of (4.2.1) does *not* correspond to the data of Figure 4.1 under steady state conditions. In order to prove this result, we rescale time in (4.2.4) by setting $p = 1$. Moreover, the temporal progression of (4.2.5) does not change by setting

$$\frac{t_s+\delta}{t_c}\cdot\frac{b}{p} = 1.$$

We now apply the subsequent theorem to show that the total number of progenitors in declining under the dynamics of (4.2.4).

Theorem 4.2.2. *Let $N \in \mathbb{N}$ and consider the system*

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}x_k(t) &= 2f\left(\frac{k+1}{N}\right)x_{k+1}(t) - x_k(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}x_N(t) &= -x_N(t), \\ x_l(0) &= 2^{N-l}\prod_{i=l+1}^{N} f\left(\frac{i}{N}\right), \end{aligned} \tag{4.2.6}$$

for $k = 0, \dots, N-1$ and $l = 0, \dots, N$, where f is a positive function defined on the interval $[0, 1]$. The solution

$$x(t) = (x_0(t), \dots, x_N(t))^\intercal$$

has the property that the function

$$\mathrm{e}^\intercal x(t) = \sum_{k=0}^{N} x_k(t)$$

is monotonically decreasing on $(0, \infty)$.

Before proofing the above theorem, we need a technical lemma.

Lemma 4.2.3. *Let $n \in \mathbb{N}$ and consider the $n+1 \times n+1$ matrix*

$$X = \begin{pmatrix} 0 & x_1 & & & \\ & 0 & x_2 & & \\ & & \ddots & & \\ & & & 0 & x_n \\ & & & & 0 \end{pmatrix}.$$

Then

$$\exp(X) = \begin{pmatrix} 1 & x_1 & \frac{1}{2!}x_1x_2 & \frac{1}{3!}x_1x_2x_3 & \cdots & \frac{1}{n!}x_1 \cdot \ldots \cdot x_n \\ & 1 & x_2 & \frac{1}{2!}x_2x_3 & \cdots & \frac{1}{(n-1)!}x_2 \cdot \ldots \cdot x_n \\ & & 1 & x_3 & \cdots & \frac{1}{(n-2)!}x_3 \cdot \ldots \cdot x_n \\ & & & \ddots & & \\ & & & & 1 & x_n \\ & & & & & 1 \end{pmatrix}.$$

Proof. Since X is nilpotent, i.e. $X^{n+1} = 0$, $\exp(X) = \sum_{k=0}^{n} \frac{X^k}{k!}$. The statement thus follows from a direct calculation. □

Proof of Theorem 4.2.2. For $A = -\mathbb{I}_{N+1} + B$ and

$$B = \begin{pmatrix} 0 & 2f\left(\frac{1}{N}\right) & & & \\ & 0 & 2f\left(\frac{2}{N}\right) & & 0 \\ & & \ddots & & \\ & 0 & & 0 & 2f(1) \\ & & & & 0 \end{pmatrix},$$

equation (4.2.6) takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = Ax(t). \tag{4.2.7}$$

The solution of (4.2.7) is given by

$$x(t) = \exp(tA)x(0).$$

The initial data $x(0)$ of (4.2.6) is the steady state of the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}y(t) = Ay(t) + e_{N+1},$$

i.e.

$$x(0) = -A^{-1}e_{N+1},$$

where e_{N+1} is the $(N+1)$-st unit vector. Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = -\exp(tA)e_{N+1}.$$

Since A is the sum of a scalar multiple of the identity matrix and the nilpotent matrix B, i.e. $B^{N+1} = 0$, it holds

$$\exp(tA) = e^{-t}\exp(tB).$$

Using Lemma 4.2.3, we have

$$\exp(tB) = \begin{pmatrix} 1 & 2tf\left(\frac{1}{N}\right) & \frac{(2t)^2}{2!}f\left(\frac{1}{N}\right)f\left(\frac{2}{N}\right) & \cdots & \frac{(2t)^N}{N!}\prod_{i=1}^{N} f\left(\frac{i}{N}\right) \\ & 1 & 2tf\left(\frac{2}{N}\right) & \cdots & \frac{(2t)^{N-1}}{(N-1)!}\prod_{i=2}^{N} f\left(\frac{i}{N}\right) \\ & & \ddots & & \\ & \mathit{0} & & 1 & 2tf\left(\frac{N}{N}\right) \\ & & & & 1 \end{pmatrix}.$$

Thus

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}x(t) &= -\exp(tA)e_{N+1} \\ &= -e^{-t}\exp(tB)e_{N+1} \\ &= -e^{-t}\sum_{j=0}^{N}\frac{(2t)^{N-j}}{(N-j)!}\prod_{i=j+1}^{N} f\left(\frac{i}{N}\right) \\ &< 0 \end{aligned}$$

for $t > 0$. □

Convergence to the Steady State Distribution

Model (4.2.4) assumes that the distribution of the capacity among progenitors is in a steady state at the time point of BrdU injection. However, since BrdU was administered at the start of adulthood in the experiment of Figure 4.1, adult neurogenesis was only active for a short amount of time. Thus, it is necessary to quantify the convergence speed of model (4.2.1) to its steady state in order to justify the steady state assumption.

For quantification purposes, we assume that there are no progenitors at the start of adulthood, i.e. we consider the dynamics of (4.2.1) together with the initial data

$$P_k(0) = 0$$

for $k = 0, \ldots, N$. The distribution of progenitors regarding their capacity is given by $P(t)/(\mathrm{e}^{\intercal}P(t))$ where $\mathrm{e}^{\intercal}P(t)$ denotes the sum of all components of the solution vector $P(t) = (P_0(t), \ldots, P_N(t))^{\intercal}$ of (4.2.1). Denoting with $\bar{P} = (\bar{P}_0, \ldots, \bar{P}_N)^{\intercal}$ the

steady state given by (4.2.2), we consider the total variation distance

$$\Delta(t) = \mathrm{d_{TV}}\left(\frac{P(t)}{\mathrm{e}^{\intercal}P(t)}, \frac{\bar{P}}{\mathrm{e}^{\intercal}\bar{P}}\right) = \frac{1}{2}\left\|\frac{P(t)}{\mathrm{e}^{\intercal}P(t)} - \frac{\bar{P}}{\mathrm{e}^{\intercal}\bar{P}}\right\|_1$$

as measurement for the distance between the distribution of the solution and the steady state distribution, where $\|\cdot\|_1$ denotes the l_1-norm. Figure 4.3 shows the time course of $\Delta(t)$ for different division probabilities f and a biologically plausible cell cycle length of $\log(2)/p = 20\,\mathrm{h}$ [11, 21]. For all considered f, $\Delta(t)$ is within a 10% distance to the steady state after 5 days, indicating a rapid convergence to the equilibrium distribution and that the steady state assumption is valid.

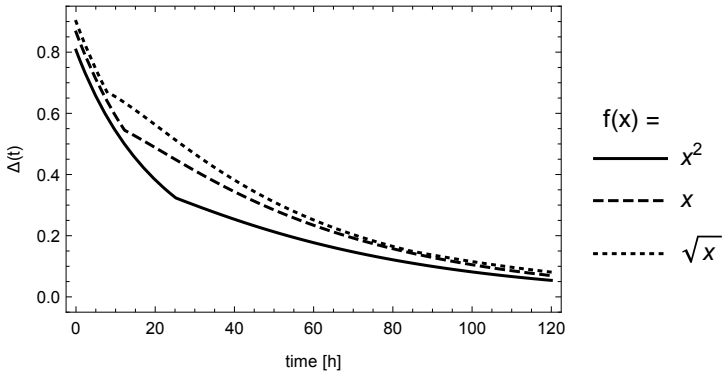


Figure 4.3: Time course of the total variation distance between the steady state and the solution of model (4.2.1) for the initial data $P_k(t) = 0$ and a division rate $p = \log(2)/(20\,\mathrm{h})$.

4.3 Division-Coupled Lengthening of the Cell Cycle

The previous section has shown that the assumptions made in Assumption 4.2.1 contradict the observed data of the time progression of BrdU labeled progenitors under steady state conditions. We thus refine (A6) of Assumption 4.2.1 by assuming that the division rate of progenitors depends on their capacity, i.e. the number of remaining divisions. This assumption is in line with the observation that neural progenitors lengthen their cell cycle during embryonic development [13].

The refined model takes the form

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}P_k(t) &= 2p\left(\frac{k+1}{N}\right)P_{k+1}(t) - p\left(\frac{k}{N}\right)P_k(t),\\ \frac{\mathrm{d}}{\mathrm{d}t}P_N(t) &= -p(1)P_N(t) + b\end{aligned} \tag{4.3.1}$$

where

$$p(x) = p_{\min} + x(p_{\max} - p_{\min}).$$

We now follow the modeling procedure of the previous section by computing the steady state of the above system to specify the initial data for the dynamics of BrdU labeled progenitors. The right-hand side of (4.3.1) is zero at

$$\bar{P}_N = \frac{b}{p(1)}$$

and

$$\bar{P}_k = \frac{2p\left(\frac{k+1}{N}\right)}{p\left(\frac{k}{N}\right)}\bar{P}_{k+1},$$

thus,

$$\begin{aligned}\bar{P}_k &= \frac{b}{p(1)}\prod_{i=k}^{N-1}\frac{2p\left(\frac{i+1}{N}\right)}{p\left(\frac{i}{N}\right)}\\ &= 2^{N-k}\frac{b}{p\left(\frac{k}{N}\right)}\end{aligned}$$

for $k = 0, \ldots, N-1$. The initial data of BrdU labeled progenitors is then given by

$$P_k(0) = \frac{t_s + \delta}{t_c}\bar{P}_k,$$

where t_c is the length of the cell cycle, t_s the length of the S-phase and δ the length of BrdU bioavailability. However, the length of the cell cycle corresponds to the proliferation rate p via $p = \log(2)/t_c$, which in turn decreases along with a decline of capacity. We assume that the length of the S-phase is preserved during cell cycle lengthening, since it has been shown that changes of the cell cycle length are caused

by a shrinking or stretching of the G_1- or G_2-phase [13, 14, 36]. Hence,

$$\begin{aligned} P_k(0) &= \frac{p\left(\frac{k}{N}\right)(t_s+\delta)}{\log(2)} 2^{N-k} \frac{b}{p\left(\frac{k}{N}\right)} \\ &= \frac{(t_s+\delta)b}{\log(2)} 2^{N-k}. \end{aligned}$$

The dynamics of BrdU labeled progenitors incorporating a division-coupled lengthening of the cell cycle is thus given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}P_k(t) &= 2p\left(\frac{k+1}{N}\right)P_{k+1}(t) - p\left(\frac{k}{N}\right)P_k(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}P_N(t) &= -p(1)P_N(t), \\ P_l(0) &= \frac{(t_s+\delta)b}{\log(2)} 2^{N-l}, \\ p(x) &= p_{\min} + x(p_{\max} - p_{\min}), \\ t_s &< \log(2)/p_{\max} \end{aligned} \tag{4.3.2}$$

for $k = 0, \ldots, N-1$ and $l = 0, \ldots, N$, where the last condition ensures that the S-phase is always shorter than the shortest cell cycle.

4.3.1 Parameter Estimation

Model (4.3.2) features six free parameters, b, δ, N, t_s, $p_{\min}$ and $p_{\max}$. The duration of BrdU bioavailability δ for the dosage used in Figure 4.1 was measured as $\delta = 15\,\mathrm{min}$ [39]. Moreover, measurements of cell cycle and S-phase length indicate that the S-phase makes up about half of the cell cycle, i.e. $t_s = \log(2)/(2p_{\max})$ [11, 21]. In addition, we assume that progenitors undergo $N = 2$ rounds of symmetric divisions, which was suggested by Encinas et al. [21]. The remaining three parameters b, $p_{\min}$ and $p_{\max}$ will be estimated.

As can be seen from Figure 4.4, model (4.3.2) can predict the increase of BrdU labeled progenitor numbers under steady state conditions. However, the best fit assumes a minimum cell cycle length of about 5 hours (Figure 4.4, left plot), which contradicts the experimentally measured length of about 20 hours [11, 21]. Conversely, we can restrict the fitting procedure to satisfy $t_c^{\min} = \log(2)/p_{\max} = 20\,\mathrm{h}$. In

that case however, the model cannot predict the increase of BrdU labeled progenitor numbers (Figure 4.4, right plot).

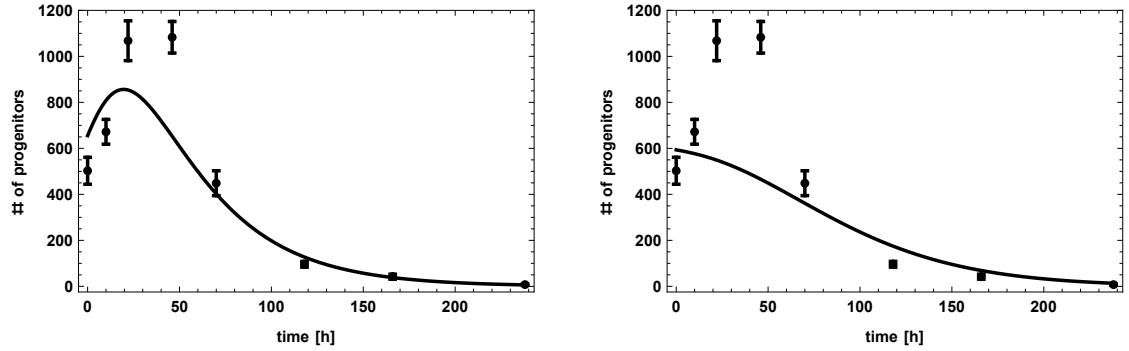


Figure 4.4: Fit of model (4.3.2) to the data of Figure 4.1. The fit of the left plot assumes no restrictions on the minimum length of the cell cycle, for the right plot, the assumption is $t_c^{\min} = \log(2)/p_{\max} = 20\,\mathrm{h}$. Estimated parameters are $t_c^{\min} = \log(2)/p_{\max} = 4.6\,\mathrm{h}$, $t_c^{\max} = \log(2)/p_{\min} = 27.5\,\mathrm{h}$ and $b = 11.9\,\mathrm{h}^{-1}$ (left plot, $R^2 = 0.89$) respectively $t_c^{\min} = 20\,\mathrm{h}$, $t_c^{\max} = 20.6\,\mathrm{h}$ and $b = 2.7\,\mathrm{h}^{-1}$ (right plot, $R^2 = 0.75$).

ε-Hybrid Initial Data

We have seen that model (4.3.2) can explain the experimentally observed increase of BrdU labeled progenitors, but not under the assumption of a biologically plausible cell cycle length. Moreover, even the best fit fails to explain that this increase of labeled progenitors is about 2-fold.

The dynamics of model (4.3.2) assume that the initial distribution of progenitors prior to BrdU labeling is in a steady state. As discussed earlier, it is possible that adult neurogenesis acted just a short amount of time and the system was not in a steady state but close to it. To incorporate this effect, we introduce a new parameter $\epsilon \in [0, 1]$ to measure the distance of the initial data to the respective steady state and assume that the BrdU experiment started in a so-called ε-hybrid state.

Definition 4.3.1 (ε-Hybrid state)**.** Let A be a matrix with all eigenvalues having negative real part, $b \neq 0$ and consider the system

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}x(t) &= Ax(t) + b, \\ x(0) &= 0 \end{aligned} \tag{4.3.3}$$

with asymptotically stable steady state

$$\bar{x} = -A^{-1}b$$

and solution

$$x(t) = \big(\mathbb{I} - \exp(tA)\big)\bar{x}.$$

Let

$$\delta(t) = \mathrm{d_{TV}}\left(\frac{x(t)}{\mathrm{e}^{\intercal}x(t)}, \frac{\bar{x}}{\mathrm{e}^{\intercal}\bar{x}}\right)$$

be the total variation distance between the steady state distribution and the distribution of the solution. For $\varepsilon \in [0,1]$, let

$$t_\varepsilon = \inf\{t > 0 \mid \delta(t) \leq \varepsilon\}$$

be the first time where this distance is below ε. We now define

$$\bar{x}_\varepsilon = x(t_\varepsilon)$$

as ε-hybrid state of (4.3.3).

Remark 4.3.2. Under the conditions of the preceding definition, we have $\bar{x}_1 = 0$ and

$$\lim_{\varepsilon \to 0} \bar{x}_\varepsilon = \bar{x}.$$

The first assertion follows from the distance of total variation being less or equal to 1, the second from

$$\lim_{t \to \infty} x(t) = \bar{x}.$$

We now refine the initial condition of model (4.3.2) and consider the model

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}P_k(t) &= 2p\left(\frac{k+1}{N}\right)P_{k+1}(t) - p\left(\frac{k}{N}\right)P_k(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}P_N(t) &= -p(1)P_N(t),\\
P_l(0) &= \bar{P}_\varepsilon^{(l)},\\
p(x) &= p_{\min} + x(p_{\max} - p_{\min}),\\
t_s &< \log(2)/p_{\max}
\end{aligned} \tag{4.3.4}$$

for $k = 0, \ldots, N-1$ and $l = 0, \ldots, N$, where $\bar{P}_\varepsilon$ is the ε-hybrid steady state of (4.3.1). This assumption is based on the idea that initially no progenitors are present at the start of adult neurogenesis, that the distribution of the capacity among progenitors converges to the steady state distribution during adult neurogenesis and that $\epsilon \in [0, 1]$ measures the discrepancy between the distribution of the capacity at the time point of BrdU labeling and the steady state distribution.

We now ask whether model (4.3.4) results in a better fit to the data and a biologically more plausible cell cycle length $t_c^{\min}$ than model (4.3.2). For this, model (4.3.4) was fitted to the data of Figure 4.1, assuming different values for $(\varepsilon, t_c^{\min})$ in the region $[0, 1] \times [8\,\mathrm{h}, 24\,\mathrm{h}]$ (Figure 4.5). As can be seen, the fit improves towards large values of ε (Figure 4.5, left plot). The explanation for this is that large values of ε correspond to a capacity distribution among progenitors shortly after the begin of neurogenesis. In that case, most of the progenitors have full capacity, leading to many consecutive divisions and an expansion of the progenitor pool. Interestingly, a very good fit to the data ($R^2 \geq 0.9$) can only be achieved with $t_c^{\min} < 16\,\mathrm{h}$, which is still shorter than the experimentally measured cell cycle length of about 20 h. Moreover, the fitting procedure shows that the ratio $\varphi = t_c^{\max}/t_c^{\min}$ increases towards small values of ε and $t_c^{\min}$ (Figure 4.5, right plot), indicating that model (4.3.4) predicts a division-coupled lengthening of the cell cycle under almost steady state conditions and a short minimum cell cycle length.

Another question is whether the maximum number N of progenitor divisions can be deduced from model (4.3.4). Figure 4.6 displays the goodness-of-fit measure R^2 dependent on $t_c^{\min}$ and N. The best fits are obtained for $N = 3$ or $N = 4$ with a small minimum cell cycle length, however only $N = 2$ admits a good fit ($R^2 \geq 0.9$) for a minimum cell cycle length which is close to the experimentally measured 20 h. This result is in line with the finding of Encinas et al. [21], stating that progenitors undergo on average $N = 2.3$ rounds of divisions.

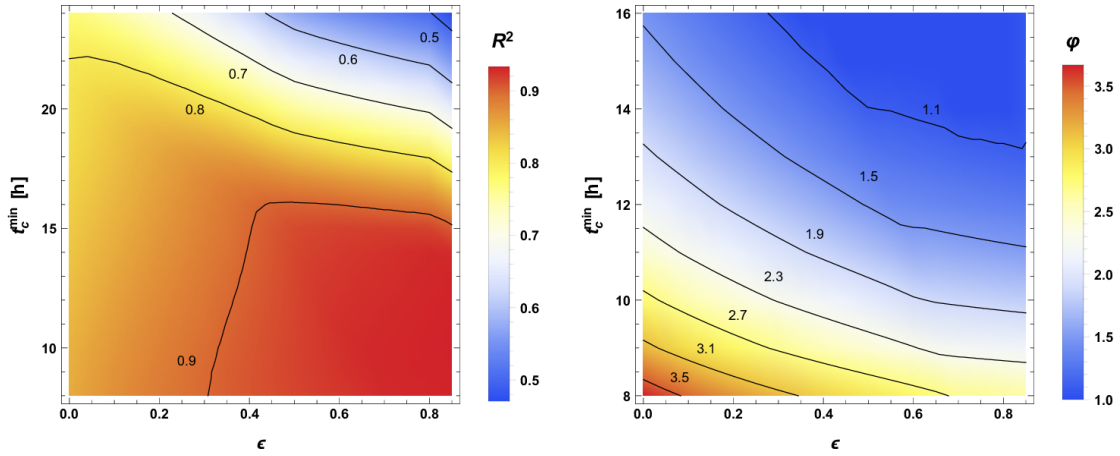


Figure 4.5: Characterization of the plausibility of model (4.3.4). Dependent on the closeness-to-steady-state measure ε and the minimum cell cycle length $t_c^{\min}$, model (4.3.4) was fitted to the data of Figure 4.1. Left plot shows the goodness-of-fit measure R^2, right plot the ratio $\varphi = t_c^{\max}/t_c^{\min}$. A very good fit to the data and a biologically plausible cell cycle length can be achieved with $t_c^{\min} = 16\,\text{h}$, in which case $t_c^{\max} = 16.2\,\text{h}$.

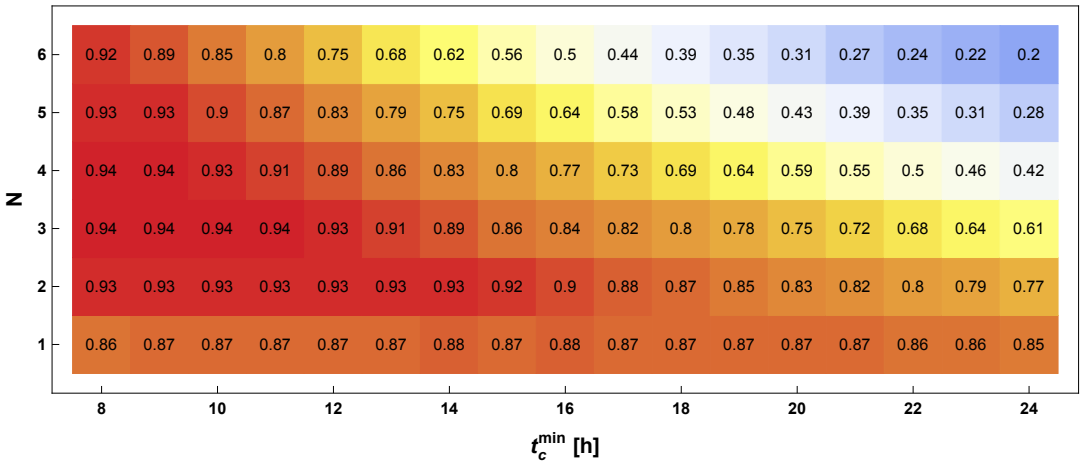


Figure 4.6: Dependency of the fit of model (4.3.4) on the shortest cell cycle length $t_c^{\min}$ and the number of divisions N as measured by the R^2 value of the fit. For each displayed combination of $t_c^{\min}$ and N, the largest R^2 value is shown, i.e. $R^2(t_c^{\min}, N) = \max_{\epsilon\in[0,1]} R^2(t_c^{\min}, N, \varepsilon)$.

4.4 Division-Coupled Death Rates

It has been reported that neural progenitors are subject to death via apoptosis [52]. To see whether this effect can explain the time course of BrdU labeled progenitors, we assume a model satisfying

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}P_k(t) &= 2pP_{k+1}(t) - \left(p + d\left(\frac{k}{N}\right)\right)P_k(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}P_N(t) &= -(p + d(1))P_N(t) + b, \\ d(x) &= d_{\max} + (d_{\min} - d_{\max})x^{\alpha} \end{aligned} \tag{4.4.1}$$

for $k = 0, \ldots, N-1$ and $d_{\min} < d_{\max}$. The reason for this form of death rate $d(x)$ is that progenitors are more likely to die in the transition phase to immature neurons [21, 52], i.e. after having completed their symmetric divisions. The steady state of (4.4.1) is given by

$$\bar{P}_k = b(2p)^{N-k}\prod_{j=k}^{N} 1/\big(p + d(j/N)\big). \tag{4.4.2}$$

We now follow the approach of the previous section by estimating parameters for model (4.4.1) under steady state initial conditions and without the source term ($b = 0$). As can be seen from Figure 4.7, the model cannot predict the increase of BrdU labeled progenitors.

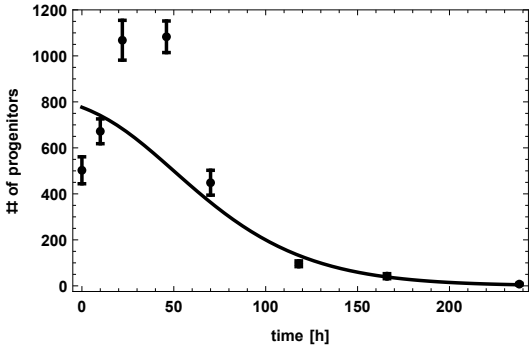


Figure 4.7: Fit of model (4.4.1) with steady state initial conditions to the data of Figure 4.1. Estimated parameters are $\alpha = 0.14$, $d_{\min} = 0.0029\,\mathrm{h}^{-1}$ and $d_{\max} = 0.022\,\mathrm{h}^{-1}$. $R^2 = 0.82$.

The reason for this discrepancy is that without death, the steady state number of progenitors with few divisions remaining is relatively high (Figure 4.8, left plot), thus causing a declining progression. On the other hand, the amount of apoptosis necessary to reverse the steady state distribution will then dominate the division rate (Figure 4.8, right plot), resulting also in a decline of labeled progenitors.

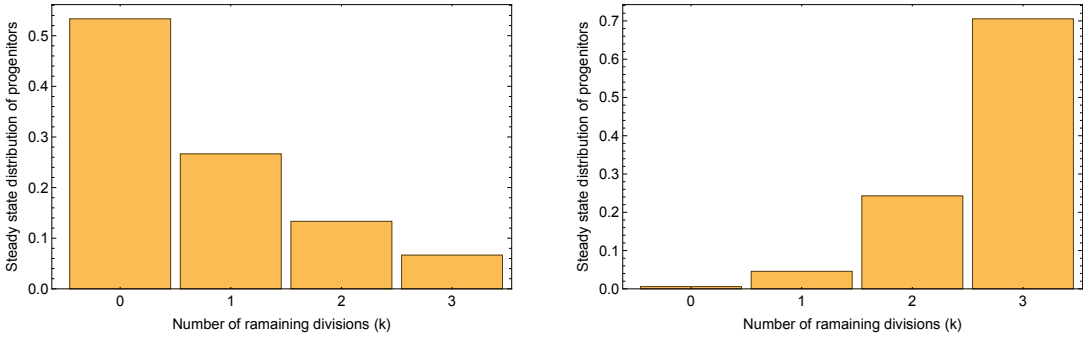


Figure 4.8: Steady state distribution of progenitors corresponding to equation (4.4.2). The parameters used are $p = \log(2)/(20\,\mathrm{h}) = 0.035\,\mathrm{h}^{-1}$ and $\alpha = 1$. Left plot shows the case $d_{\min} = d_{\max} = 0$, right plot $d_{\min} = 0$ and $d_{\max} = 0.5\,\mathrm{h}^{-1}$.

4.5 Asymmetric Progenitor Divisions

The previous chapters have shown that it is necessary to consider a progenitor dynamics admitting a steady state distribution with a high abundance of cells with many remaining divisions. Otherwise, the rapid increase of BrdU labeled progenitors cannot be explained. Another possibility to achieve the desired distribution is to consider asymmetric progenitor divisions. While usually an asymmetric division results in a cell with the same properties as the mother cell and a more differentiated cell, we consider asymmetric divisions where the mother cell gives rise to a cell with the same capacity (same number of remaining divisions) and one with lower capacity. We assume a model of the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}P_k(t) &= 2\left(1 - a\left(\frac{k+1}{N}\right)\right)pP_{k+1}(t) + \left(2a\left(\frac{k}{N}\right) - 1\right)pP_k(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}P_N(t) &= (2a(1) - 1)P_N(t) + b, \\ a(x) &= a_{\min} + (a_{\max} - a_{\min})x^{\alpha} \end{aligned} \tag{4.5.1}$$

for $k = 0, \ldots, N-1$ and $a_{\min} < a_{\max} < 1/2$, where $a(x)$ is the fraction of self-renewal. The reason for this choice of a is that the steady state abundance increases with increasing self-renewal and that we aim to increase the steady state number of cells with high capacity. The steady state of (4.5.1) is then given by

$$\bar{P}_k = \frac{b2^{N-k}}{p(1-2a(1))} \prod_{j=k+1}^{N} \frac{1-a\left(\frac{j}{N}\right)}{1-2a\left(\frac{j-1}{N}\right)}. \tag{4.5.2}$$

Fitting the model under steady state initial conditions to the data results in the plot of Figure 4.9. As can be seen, the model fails to reproduce the rapid increase of BrdU labeled progenitors. The reason for this discrepancy is again the form of

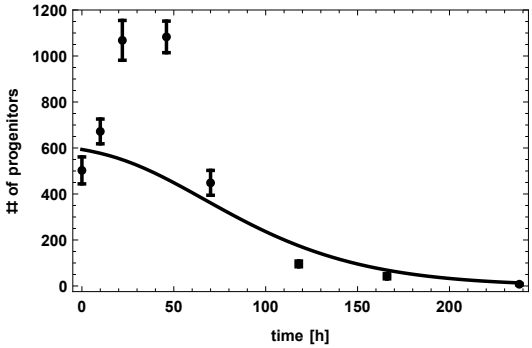


Figure 4.9: Fit of model (4.5.1) with steady state initial conditions to the data of Figure 4.1. Estimated parameters are $\alpha = 1.02$ and $a_{\min} = a_{\max} = 0\,\mathrm{h}^{-1}$. $R^2 = 0.75$.

the steady state distribution. If all cells have zero self-renewal, there exist only symmetric divisions leading to the distribution displayed in the left plot of Figure 4.10. On the other hand, even if progenitors with the maximum number of remaining divisions have high self-renewal and all other cells have zero self-renewal, the steady state distribution cannot be reversed (Figure 4.10, right plot).

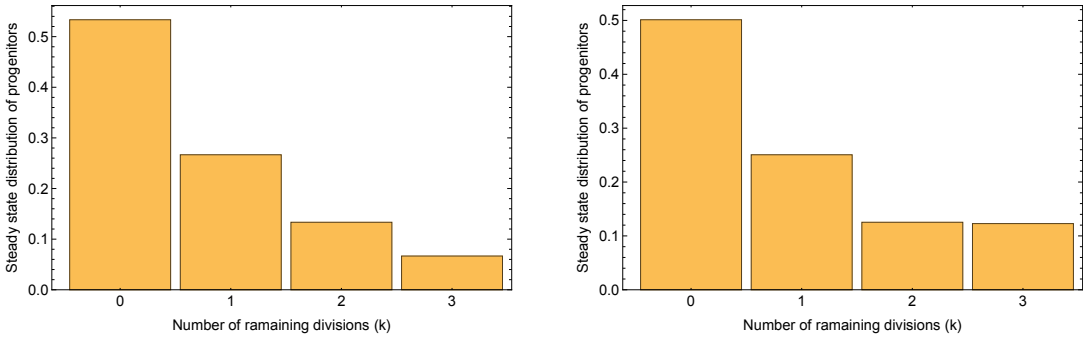


Figure 4.10: Steady state distribution of progenitors corresponding to equation (4.5.2), $p = \log(2)/(20\,\mathrm{h}) = 0.035\,\mathrm{h}^{-1}$. Left plot shows the case $a_{\min} = a_{\max} = 0$, right plot the case $a(1) = .49$ and $a(k/N) = 0$ for $k < N$.

4.6 Summary

We have investigated multiple scenarios to explain the experimental data on the time course of BrdU labeled progenitor numbers. It becomes apparent that at the time of the experiment, the majority of progenitor cells must have a high number of remaining divisions. Otherwise, the rapid increase of BrdU labeled progenitors cannot be explained. To model the data, we have used steady state initial conditions. This approach is based on the assumption that neurogenesis was active prior to the experiment and at the time point of labeling, progenitors attained the equilibrium distribution regarding the number of remaining divisions.

The dynamics suggested in ref. [21] lead to a steady state distribution where the relative proportion of cells with k remaining divisions is twice as high as the one of cells with $k+1$ remaining divisions. As a result, the rapid rise of BrdU labeled progenitor numbers cannot be explained. The scenarios 'Asymmetric Progenitor Divisions' and 'Division-Coupled Death Rates' are also not able to predict the observed data. The former fails to shift the equilibrium distribution towards a high abundance of cells with many remaining divisions. The latter can achieve that property, but the amount of death necessary then dominates the division rate, causing an always declining progression.

A lengthening of the cell cycle along with the number of already performed divisions can explain the rapid increase of BrdU labeled progenitors under steady state conditions. However, the predicted cell cycle length disagrees with the experimen-

tally measured one. This has led to the conclusion that system is not in a steady state at the start of the experiment but that it is close to it. In that case, the fit to the data improves as well as the agreement between the predicted and experimentally measured cell cycle length. However, it is possible that in this case a lengthening of the cell cycle is too small in order to be detected experimentally.

Chapter 5

Analysis of the Dkk1 Knockout

To study the role of Dickkopf-1 (Dkk1) in the regulation on neural stem cells (NSCs), Seib et al. [51] performed a knockout experiment in which the Dkk1 gene was deleted in adult animals. The authors suggest that a loss of Dkk1 causes NSCs to increase their self-renewal, counteracting the age-related decline of adult neurogenesis.

The aim of this chapter is to investigate how the dynamics of NSCs change upon deletion of Dkk1. To address this question, we use the repeated activation model of Chapter 3 and the progenitor model of Chapter 4 to reproduce the data of the wild type (WT) group as well as introducing additional parameters to account for the knockout (KO) dynamics. Our findings show that a shift of the balance between NSC activation and depletion towards a higher fraction activation events is the best explanation for the KO dynamics. In contrast, an increased self-renewal of NSCs results in a considerably worse fit to the KO data.

5.1 Experimental Procedure and Data

To compare the dynamics of Dkk1 deficient animals to the one of wild type individuals, Seib et al. [51] utilized the following experimental setup (Figure 5.1): 8-to 12-week-old mice were injected with Tamoxifen (TAM) for 5 days. Next, BrdU was administered 5 weeks after the Tamoxifen treatment and the number of BrdU labeled cells was examined at 24 hours and 4 weeks after the BrdU injection. The difference between the KO and WT group was that KO mice had both Dkk1 alleles floxed, such that Cre-mediated Dkk1 excision only occurred in those individuals.

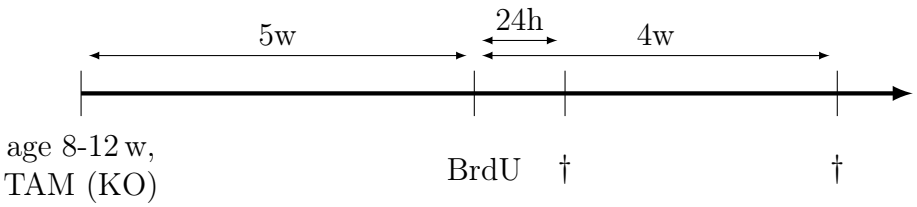


Figure 5.1: Graphical representation of the Dkk1 experiment of Seib et al. [51].

In order to study the effects of Dkk1 deletion during the neuron production process, cells at different stages of differentiation were quantified. At 24 hours after BrdU injection, the number of labeled stem cells, progenitors and neuroblasts was examined. Additionally, BrdU positive stem cells, neurons and astrocytes were quantified at 4 weeks after BrdU labeling. Moreover, apoptotic (Caspase positive) neuroblasts were analyzed at the 5 weeks after Tamoxifen time point. The data indicates a higher abundance of all neurogenesis related cell types except neurons. Moreover, the number of Caspase positive neuroblasts is increased in Dkk1 deficient animals (Figure 5.2).

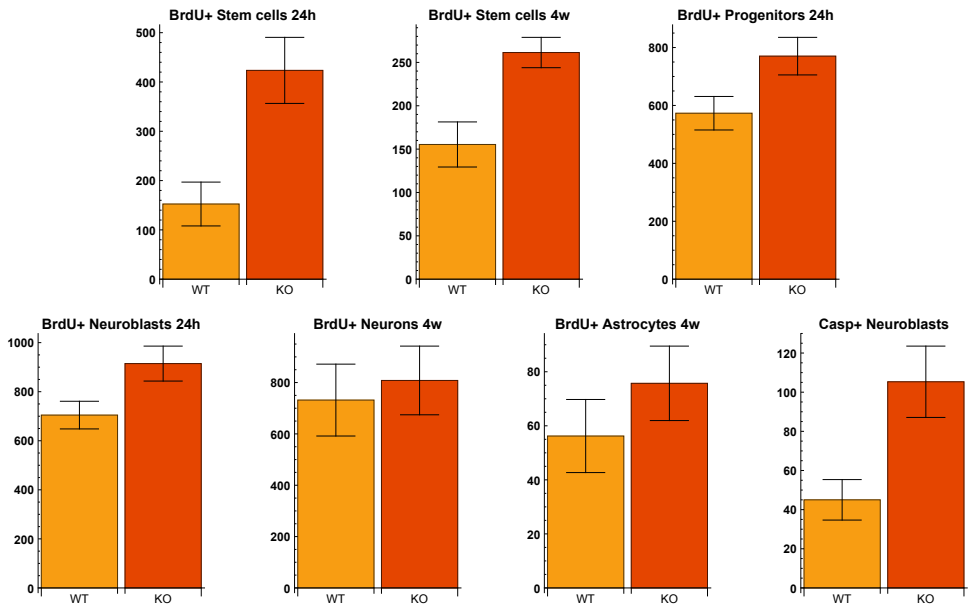


Figure 5.2: Reproduction of the Dkk1 data published by Seib et al. [51].

5.2 Modeling of the Knockout Experiment

To model the data of the Dkk1 experiment, we employ a coupled model consisting of the repeated activation model of Chapter 3 and the progenitor model of Chapter 4. In addition, since the Dkk1 data also contains neuroblasts and neurons, we extend the joint model to contain these cell types. We assume that progenitors of the lowest capacity directly differentiate into neuroblasts and that in turn neuroblasts either perform apoptosis or directly differentiate into mature neurons. Taken together, the full model of WT adult hippocampal neurogenesis is given by the set of equations

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\,\mathsf{stem}^{\mathrm{Q}}(t) &= -(r + q_{\max}e^{-\beta_q t})\,\mathsf{stem}^{\mathrm{Q}}(t) + 2ap^{\mathrm{stem}}\,\mathsf{stem}^{\mathrm{A}}(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}\,\mathsf{stem}^{\mathrm{A}}(t) &= r\,\mathsf{stem}^{\mathrm{Q}}(t) - p^{\mathrm{stem}}\,\mathsf{stem}^{\mathrm{A}}(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}\,\mathsf{prog}_N(t) &= -p^{\mathrm{prog}}(1)\,\mathsf{prog}_N(t) + \kappa 2(1-a)p^{\mathrm{stem}}\,\mathsf{stem}^{\mathrm{A}}(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}\,\mathsf{prog}_k(t) &= 2p^{\mathrm{prog}}\left(\frac{k+1}{N}\right)\mathsf{prog}_{k+1}(t) - p^{\mathrm{prog}}\left(\frac{k}{N}\right)\mathsf{prog}_k(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}\,\mathsf{nblast}(t) &= p^{\mathrm{prog}}(0)\,\mathsf{prog}_0(t) - (d+f)\,\mathsf{nblast}(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}\,\mathsf{neuron}(t) &= f\,\mathsf{nblast}(t),\\
\frac{\mathrm{d}}{\mathrm{d}t}\,\mathsf{astro}(t) &= \theta q_{\max}e^{-\beta_q t}\,\mathsf{stem}^{\mathrm{Q}}(t) + (1-\kappa)2(1-a)p^{\mathrm{stem}}\,\mathsf{stem}^{\mathrm{A}}(t),\\
p^{\mathrm{prog}}(x) &= p^{\mathrm{prog}}_{\min} + x\big(p^{\mathrm{prog}}_{\max} - p^{\mathrm{prog}}_{\min}\big),
\end{aligned} \tag{5.2.1}$$

where $\mathsf{stem}^{\mathrm{Q}}$ denotes quiescent stem cells, $\mathsf{stem}^{\mathrm{A}}$ active (cycling) stem cells, prog_i progenitors with i remaining divisions ($1 \le i \le N$), nblast neuroblasts, neuron mature neurons and astro astrocytes. The parameter values corresponding to the model are displayed in Table 5.3. As can be seen, the parameters κ, d and f are not known prior to the Dkk1 experiment and need to be inferred from fitting the above model to the WT part of the data.

In order to reconstruct the Dkk1 data, it is necessary to accurately simulate the experimental protocol of Figure 5.1 using our neurogenesis model: We make the simplifying assumption that all mice start at 10 weeks of age with a fixed number n_0 of stem cells, which needs to be estimated from the data. Next, we simulate the dynamics of these cells and their offspring according to (5.2.1) until the age

parameter	value	source
$q_{\max}$	$0.0431\,\mathrm{d}^{-1}$	Table 3.20
r	$0.0198\,\mathrm{d}^{-1}$	Table 3.20
β_q	$0.009\,21\,\mathrm{d}^{-1}$	Table 3.20
a	0.525	Section 3.4
θ	0.395	Section 3.6
p^{stem}	$\log(2)/(22.8\,\mathrm{h})$	Brandt et al. [11]
$p^{\mathrm{prog}}_{\min}$	$\log(2)/(16\,\mathrm{h})$	Figure 4.5
$p^{\mathrm{prog}}_{\max}$	$\log(2)/(16.2\,\mathrm{h})$	Figure 4.5
N	2	Encinas et al. [21]

Table 5.3: Parameter values of system (5.2.1), which is used to model the Dkk1 experiment of Seib et al. [51].

of 15 weeks, at which BrdU was administered. Since the dynamics of progenitors (cell cycle length of about 16 h) is relatively fast compared to this first time interval of 5 weeks, the initial number of progenitors is negligible because of their minor contribution to the number of existing cells at the BrdU time point. To calculate the fraction φ of dividing cells that acquire the BrdU label, we make use of the previously derived formula (4.2.3),

$$\varphi = \frac{t_s + \delta_{\mathrm{BrdU}}}{t_c},$$

where $\delta_{\mathrm{BrdU}} = 15\,\mathrm{min}$ has been measured by Mandyam et al. [39] for the BrdU dosage used during the Dkk1 experiment. Having computed the number of BrdU labeled cells at 15 weeks of age, we again use (5.2.1) to calculate the number of BrdU positive cells at 24 hours and 4 weeks after labeling. In addition, the number of apoptotic (Caspase positive) neuroblasts is given by

$$\mathsf{nblast}^{\dagger} = \frac{d}{d+f}\big(1 - e^{-(d+f)\delta_{\mathrm{phag}}}\big)\mathsf{nblast}.$$

Here, we use the fact that apoptotic cells are cleared via phagocytosis within $\delta_{\mathrm{phag}} = 1.35\,\mathrm{h}$ [52] and that leaving the neuroblast stage during that period is a joint decay process consisting of apoptosis (parameter d) and differentiation (parameter f) [38].

To model the effects of Dkk1 deletion, we introduce additional parameters for the WT model (5.2.1). Since the knockout of Dkk1 selectively occurred in NSCs,

we assume that knockout effects were the result of changes in the value of stem cell parameters. Thus, we consider changes of the parameters a, p^{stem}, $q_{\max}$ and r. In addition, the data indicates an increased death rate d of neuroblasts. For each parameter

$$\mathsf{p} \in \{a, p^{\text{stem}}, q_{\max}, r, d\},$$

we introduce a change Δ_{p} such that the corresponding WT and KO parameters are related via

$$\mathsf{p}^{\text{KO}} = (1 + \Delta_{\mathsf{p}})\mathsf{p}^{\text{WT}}.$$

Hence, Δ_{p} denotes the relative change of the WT parameter p.

5.3 Results

The first step in understanding Dkk1 deletion is to reconstruct the WT part of the Dkk1 data by estimating the non-quantified parameters of model (5.2.1). As can be seen from Figure 5.4, the model shows a very good fit to the data, indicating that it is a suitable description of WT neurogenesis. Moreover, the parameters estimated are biologically plausible: The value of $\kappa = 77\%$, reflecting the probability that an asymmetric NSCs division produces a progenitor rather than an astrocyte, is in line with experimental observations that most of the asymmetric NSC divisions generate progenitors [9]. Interestingly, $d/(d + f) = 61\%$, denoting the probability that a neuroblast dies rather than differentiating into a neuron, is consistent with experimental evidence stating that the major part of neuroblasts does not mature into neurons [21, 52].

Having achieved a very good fit to the WT part of the Dkk1 data, we next aim to understand how WT parameters change in order to explain the KO part. However, since the model fit to the WT part is not perfect, instead of the KO data we consider KO *effects*, defined as the ratio of the KO data to the corresponding WT data, i.e. (KO effect) = (KO data)/(WT data). To find the best explanation for the KO effects, we employ a nested approach by starting with simple explanations and later on transition to more complex scenarios (Table 5.5). At first, we assume that only one of the stem cell parameters a, p^{stem}, $q_{\max}$ or r and in addition the neuroblast death rate d changes (Table 5.5). The best fit to the data is achieved by a decrease of the depletion rate $q_{\max}$ ($R^2 = .923$) or an increase of the activation

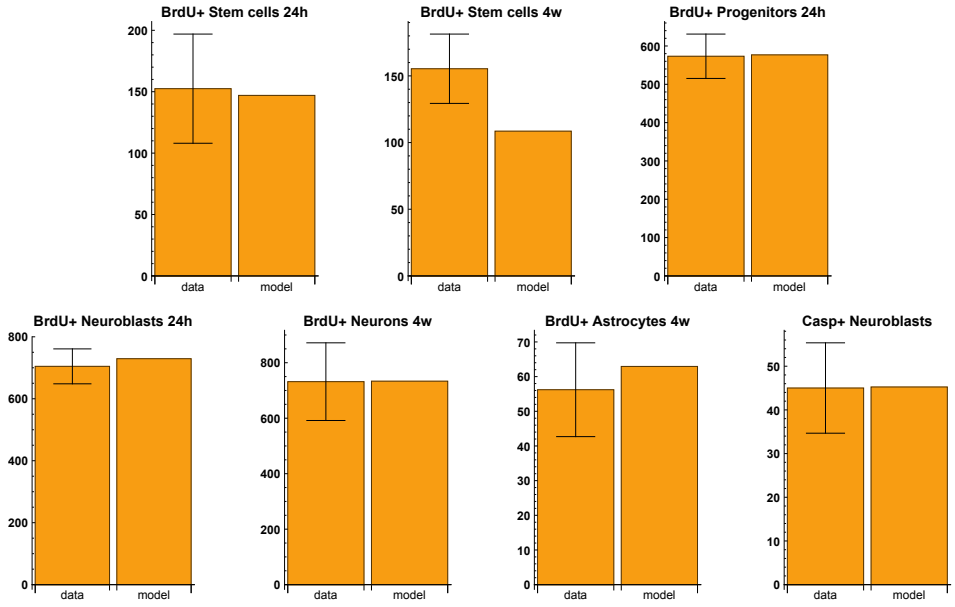


Figure 5.4: Plot of the WT part of the Dkk1 data (Figure 5.2) versus the best fit of model (5.2.1). Estimated parameters are $\kappa = 0.771$, $d = 0.0214\,\mathrm{d}^{-1}$, $f = 0.0137\,\mathrm{d}^{-1}$, $n_0 = 23300$. $R^2 = 0.89$.

rate r ($R^2 = .916$). In contrast, the by the authors suggested increased self-renewal scenario displays a considerably worse fit ($R^2 = 0.862$). The decreased depletion and increased activation scenario both lead to a shift of the balance between NSC activation and depletion towards a higher fraction activation events. We thus consider a scenario in which both parameters, $q_{\max}$ and r can change and find that it improves the fit ($R^2 = 0.923$). It is also possible that all NSC parameters change, resulting in a slightly better fit ($R^2 = 0.925$). Although the model fit improves with an increasing number of parameters, model selection theory suggests there to be a trade-off between the capability to predict the data and the complexity (number of free parameters) of a certain model [12]. This trade-off can be quantified using Akaike's information criterion (AIC) and the corresponding Akaike weight (ΔAIC). As previously discussed in Chapter 3, the recommendation for the level of empirical support of a certain model is substantial, if $0 \leq \Delta\mathrm{AIC} \leq 2$, considerably less, if $4 \leq \Delta\mathrm{AIC} \leq 7$ and essentially none, if $\Delta\mathrm{AIC} > 10$ [12]. Applying this recommendation to the cases displayed in Table 5.5, scenarios leading to the discussed shift of the balance between NSC activation and depletion towards a higher fraction activa-

tion events should be considered first. Moreover, the in ref. [51] suggested increased self-renewal of NSCs can be disregarded to explain the effects of Dkk1 deletion.

Δ_a	$\Delta_{p^{\text{stem}}}$	$\Delta_{q_{\max}}$	Δ_r	Δ_d	R^2	ΔAICc
.20	-	-	-	.94	.862	29
-	-.33	-	-	.97	.821	42
-	-	-.52	-	1.6	.923	0
-	-	-	.45	1.6	.916	4.2
-	-	-.42	.09	1.6	.923	2.0
.05	-.17	-.11	.29	1.5	.925	5.9

Table 5.5: Estimated changes of WT parameters to account for the Dkk1 KO effects. ΔAICc denotes the Akaike weight of the respective model corresponding to the small-sample-size corrected Akaike information criterion (AICc).

To further substantiate our findings, we plot the best explanation of the KO effects according to ΔAICc (Figure 5.6) as well as the in ref. [51] suggested increased activation scenario (Figure 5.7) versus the experimental data. Given our previous discussion, it is not surprising that Figure 5.6 displays a much better fit to the data than Figure 5.7. Interestingly, even our best explanation for the KO effects fails to accurately reproduce the 24h after BrdU time point of stem cells (Figure 5.6). Further studies using manual adjustments of NSC parameters show that both the 24h and 4w time point of stem cells cannot be matched. The reason is that reproducing the 24h time point would imply a much higher number of 4w BrdU positive stem cells. This discrepancy could possibly hint towards a negative feedback mechanism acting on stem cells.

As discussed in the previous chapters, scientific discovery is driven by deriving predictions that can be tested experimentally by future studies [46]. We thus quantify our best explanation for the KO effects ($\Delta_{q_{\max}} = -0.52$ and $\Delta_d = 1.6$) with respect to neurogenesis dynamics. Recall from equation (5.2.1) that leaving the quiescent stem cell stage is a joint decay process consisting of activation (parameter r) and age-dependent depletion (parameter $q(t) = q_{\max} e^{-\beta_q t}$). In the case of Dkk1 KO, the decreased depletion rate results in an increased fraction of NSC activation events, which is greater than 50% even at the start of adulthood (Figure 5.8a). In addition, the resulting increase of neurogenesis is counteracted by an increased fraction of apoptotic neuroblasts (Figure 5.8b).

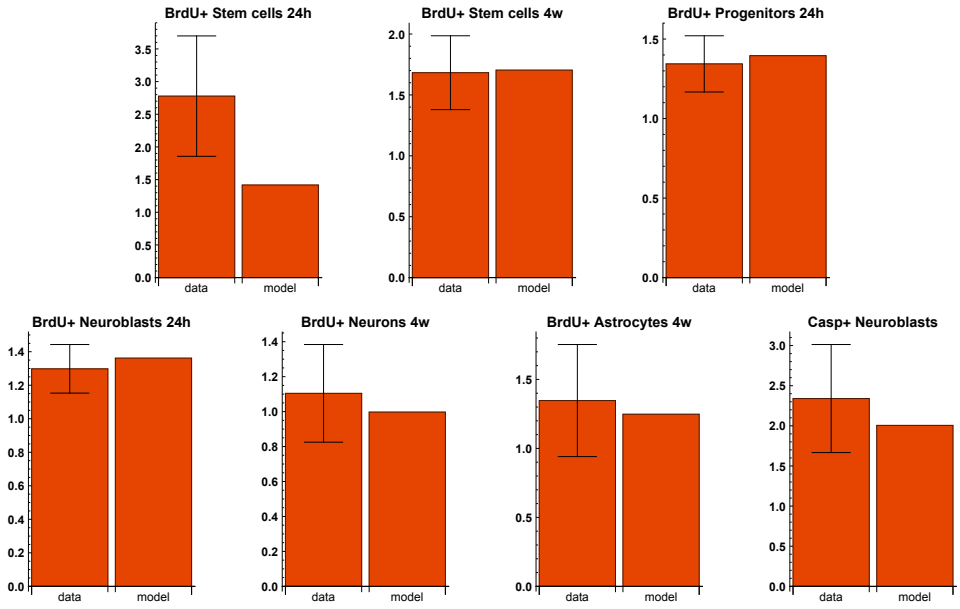


Figure 5.6: Plot of the KO effects of the Dkk1 data (Figure 5.2) versus the best fit of model (5.2.1), assuming that $q_{\max}$ and d change. Estimated parameters are $\Delta_{q_{\max}} = -0.52$ and $\Delta_d = 1.6$. $R^2 = 0.92$.

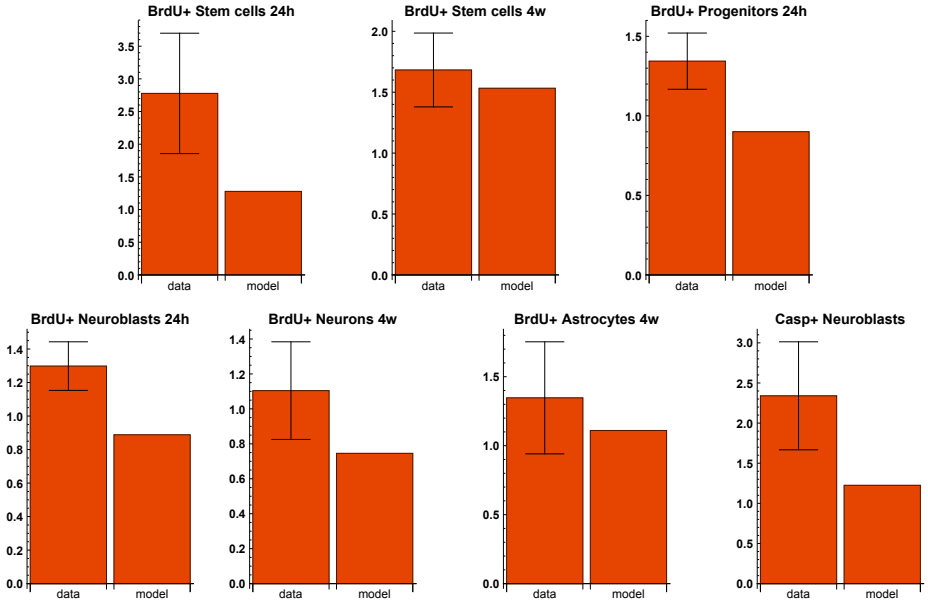
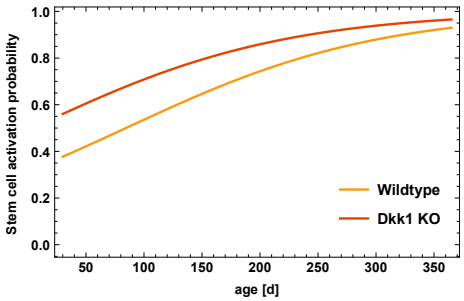
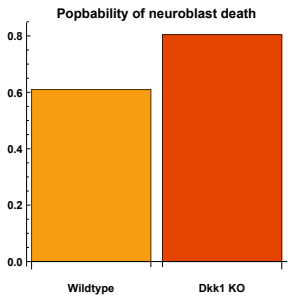


Figure 5.7: Plot of the KO effects of the Dkk1 data (Figure 5.2) versus the best fit of model (5.2.1), assuming that a and d change. Estimated parameters are $\Delta_a = 0.20$ and $\Delta_d = 0.94$. $R^2 = 0.86$.



(a) Age-dependent activation probability of stem cells. In the case of Dkk1 deletion, the probability is higher due to the decreased stem cell depletion rate $q_{\max}$.



(b) Probability of a neuroblast to die rather than maturing into a neuron. In the case of Dkk1 knockout, this probability is increased due to the increased death rate d.

Figure 5.8: Comparison of wild type and Dkk1 knockout neurogenesis dynamics.

5.4 Summary

Using the mathematical model, which has been established in the previous chapters, we were able to reconstruct the experimental data of the Dkk1 study of Seib et al. [51]. In a first step, we have used the model to reproduce the WT part of the data and the resulting very good fit confirms the validity of our model to describe WT neurogenesis. We then used a gradual approach to assess different scenarios regarding their ability to explain the KO effects. Our analysis indicates that the best explanation for the effects of Dkk1 deletion is a shift of the balance between NSC activation and depletion towards a higher fraction of activation events. Moreover, this shift is accompanied by an increased fraction of apoptotic neuroblasts. In contrast, we were able to rule out the by the authors suggested scenario of an increased self-renewal of NSCs using model selection theory. Additionally, we have quantified our best explanation for the KO data. This quantification can serve as a basis to test our new hypothesis by examining whether the fraction of active (non-quiescent) NSCs increases in the case of Dkk1 deletion.

Chapter 6

Summary and Outline of Future Research

The subgranular zone of the hippocampal dentate gyrus is one of the two best described regions in the brain where stem cells continuously give rise to new neurons during adulthood. Understanding the special micro environment of the dentate gyrus, which allows neuron production, has been subject of extensive research for the past decade. In order to study stem cell behavior, knockout experiments have become a widely used tool, because they allow to investigate the regulatory mechanisms of specific proteins. However, this approach comprises a black box system, since the knockout introduces an unknown change in the cellular dynamics, which can only be measured indirectly by quantifying cell counts. Mathematical modeling is needed in order to evaluate knockout experiments by reconstructing how experimental measurements are related to the unknown change of the biological system.

6.1 Summary of the Thesis

This thesis was motivated by the aim to understand the knockout of Dkk1, a study that has been published by the laboratory of our collaborator Prof. Martin-Villalba [51]. As a first step, we proposed a simplistic ODE model of hippocampal neurogenesis. The model was then used to investigate how a knockout-induced change of cellular dynamics, reflected by a changed parameter value, affects cell counts. Surprisingly, we have found that the effects of the knockout are time-dependent and that opposing effects can be observed at different time points after the knockout.

The analysis of the model was performed without an underlying quantification, using only minor assumptions on the parameters and computing the time-dependent signs of parameter-derivatives of the ODE solution. Accordingly, the obtained results were only qualitative and could not be applied to experimental data. This demonstrated that a quantification of the underlying model was needed to attribute observed knockout effects to changes in specific parameters.

We then proceeded with building a quantitative model of adult hippocampal neurogenesis, starting from the hypotheses of Bonaguidi et al. [9] and Encinas et al. [21]. The core of the model consists of the dynamics of neural stem cells, which are on top of the neurogenic differentiation cascade. Although both theories differ regarding the dynamics of neural stem cells, we were able to combine the two hypotheses into a unified model consisting of Bonaguidi's hypothesis of repeated stem cell activation and Encinas' suggestion that stem cells deplete by transforming into astrocytes. It became apparent that a return to quiescence is needed in order to maintain a fraction of proliferating stem cells, a feature that was observed in experiments. Also, we provided new data indicating that a part of the stem cell decline can be ascribed to astrocytic transformation. Interestingly, we could demonstrate that the two models of Bonaguidi et al. [9] and Encinas et al. [21] are indistinguishable when fitted to population level data. The reason is that in both models the ratio of the number of non-dividing to dividing stem cells converges to a globally asymptotically stable steady state. Moreover, if the ratio starts at the steady state, the total number of stem cells is described by a single exponential decay term. This finding demonstrates that neural stem cells exhibit a certain type of self-stabilizing dynamics, if instead of total cell counts, the composition of cell compartments relative to the whole population is considered. In addition, we have examined how stem cell dynamics change during aging and found that the best explanation for our data is an age-related decrease of the depletion rate of stem cells. This indicates that the survival chances of stem cells improve with increasing age.

The next step towards a quantitative model of hippocampal neurogenesis was to reconstruct the dynamics of progenitors, a cell population characterized by rapid symmetric divisions leading to an expansion of the pool of precursor cells. Using the BrdU data published by Encinas et al. [21], we examined several scenarios to explain the rapid increase of BrdU labeled progenitor numbers within the first 48 h of the experiment. Again it became apparent that the overall dynamics depend on the

initial distribution of the cell compartments within the whole population. To explain the data, the initial distribution must be such that at the time of the experiment, the number of progenitors with many remaining divisions is high compared to the one with few divisions remaining. However, the dynamics of consecutive symmetric divisions suggested in ref. [21] results in a distribution with many progenitors having few remaining divisions. This led to the conclusion that either at the time of the experiment, the system was close to the steady state or that progenitors exhibit a division-coupled lengthening of the cell cycle, a feature that can shift the steady state distribution towards a high abundance of cells with many remaining divisions.

Finally, to accomplish the initial objective of this thesis, we applied our quantitative neurogenesis model to the Dkk1 study [51] in order to investigate how neural stem cell dynamics change upon Dkk1 deletion. To evaluate the experiment, we first reconstructed the wild type part of the data by simulating the experimental protocol. The very good fit to the data independently confirmed the validity of our model to be a suitable description of hippocampal neurogenesis. We then analyzed different scenarios to account for the knockout data and found that the best explanation is an increased fraction of stem cell activation events in the case of Dkk1 deletion. Moreover, we were able to rule out the by the authors hypothesized increased self-renewal of stem cells.

The example of the Dkk1 knockout shows that multi-stage cell systems such as the neurogenic niche of the adult hippocampus harbor a deep complexity which cannot be fully understood using purely intuitive reasoning. In contrast, our study demonstrates that mathematical modeling is a powerful tool to handle this complexity and to provide novel insights that can be tested by future experiments.

6.2 Outline of Future Research

Our model of adult hippocampal neurogenesis consists of a system of linear ODEs or in the case of modeling single-cell level data a stochastic counterpart based on the Gillespie algorithm. Due to the linearity of the model, nonlinear feedbacks could not be taken into account. However, there has been a growing amount of knowledge about the signaling mechanisms regulating hippocampal neurogenesis [22]. One natural extension of our model is to introduce regulatory mechanisms, in particular

negative feedbacks acting on stem cells, and to study their influence on the stability of the system. Another possibility to extend the model is to introduce a spatial component. Since hippocampal neurogenesis takes place in a solid tissue, spatial effects need to be considered. Possible sources of such effects are for instance the cell-cell contact based Notch signaling pathway [1] or diffusible Wnt signaling [29, 51]. In addition, our model may be used to evaluate further knockout experiments in the same way as already demonstrated in the case of Dkk1 deletion.

Appendix

Parameter Estimation

In this section, we give a concise introduction into the theoretical foundation of estimating parameters for the developed ODE models. This question leads to the theory of nonlinear regression. By expanding the least-squares objective function into a Taylor series, the results of linear regression analysis can be applied in the nonlinear case via linear approximation. We start with some basic results from multivariate probability theory, which serve as a basis for the theoretical understanding of regression analysis.

The presented theory is a summary of the existing methods for linear and nonlinear regression analysis, tailored to our needs of parameter estimation for ODE models of neurogenesis dynamics [5, 6, 8, 43, 49, 50, 61].

A.1 Multivariate Distributions

Definition A.1.1 (Multivariate normal distribution)**.** Let $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n\times n}$ a symmetric and positive definite matrix. The multivariate normal distribution is given by the probability density function

$$p(x) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(x-\mu)^{\intercal}\Sigma^{-1}(x-\mu)\right]$$

Accordingly, a random vector X is said to be normally distributed with mean μ and covariance matrix Σ if it has probability density $p(x)$. In that case, we write

$$X \sim \mathbb{N}_n(\mu, \Sigma).$$

Proposition A.1.2. *Let $X \sim \mathbb{N}_n(\mu, \Sigma)$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $m \leq n$, then*

$$AX + b \sim \mathbb{N}_n(A\mu + b, A\Sigma A^\intercal).$$

Proof. See Moser [43, Theorem 2.1.2]. □

Definition A.1.3 (Chi-squared distribution)**.** Let $X_1, \ldots, X_n$ be independent standard normal distributed random variables. The chi-squared distribution with n degrees of freedom is defined as the distribution of the random variable

$$Z = \sum_{i=1}^{n} X_i^2$$

and we denote this by $Z \sim \chi_n^2$.

Definition A.1.4 (Noncentral chi-squared distribution)**.** Let $X \sim \mathbb{N}_n(\mu, \mathbb{I}_n)$ and $P \in \mathbb{R}^{n \times n}$ a rank p projection, i.e. $P^2 = P$. Then

$$Z = X^\intercal P X$$

has noncentral chi-squared distribution with p degrees of freedom and non-centrality parameter $\lambda = (\mu^\intercal P \mu)/2$. In that case, we write $Z \sim \chi_p^2(\lambda = (\mu^\intercal P \mu)/2)$.

Proposition A.1.5. *Let $X \sim \mathbb{N}_n(\mu, \Sigma)$, then*

$$(x - \mu)^\intercal \Sigma^{-1} (x - \mu) \sim \chi_n^2.$$

Proof. See Moser [43, Theorem 3.1.2]. □

Proposition A.1.6. *Let $Y \sim \mathbb{N}_n(\mu, \Sigma)$, $A \in \mathbb{R}^{m \times n}$ such that $A\Sigma = cP$ with $c > 0$ and P a rank p projection. Then*

$$Y^\intercal A Y \sim \chi_p^2(\lambda = \mu^\intercal A \mu/(2c)).$$

Proof. See Moser [43, Corollary 3.1.2a]. □

Proposition A.1.7. *Let $Y \sim \mathbb{N}_n(\mu, \Sigma)$, $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{l \times n}$. AY and BY are independent if and only if*

$$A\Sigma B^\intercal = 0.$$

Proof. See Seber and Lee [49, Theorem 2.5]. □

Definition A.1.8 (F-distribution)**.** Let $X_1 \sim \chi^2_{n_1}$ and $X_2 \sim \chi^2_{n_2}$ be independent. The F-distribution with n_1 and n_2 degrees of freedom is defined as the distribution of

$$Z = \frac{X_1/n_1}{X_2/n_2}$$

and we write $Z \sim F_{n_1,n_2}$.

A.2 Linear Regression

The simplest task in linear regression is fitting a straight line to a set of data. That is, assume we are given points $(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^2$ and want to find parameters $a, b \in \mathbb{R}$ such that

$$y_i = a + bx_i.$$

Usually, however, such a fit doesn't exist, since the y_i might be subject to random fluctuations. Assuming that these fluctuations are normally distributed with mean zero and *known* variance σ^2, the model becomes

$$y_i = a + bx_i + \varepsilon_i \tag{A.2.1}$$

with $\varepsilon_i \sim \mathbb{N}(0, \sigma^2)$. If we define

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \; X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \; \beta^\star = \begin{pmatrix} a \\ b \end{pmatrix}, \; \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \tag{A.2.2}$$

equation (A.2.1) can be written as

$$Y = X\beta + \varepsilon, \quad \varepsilon \sim \mathbb{N}_n(0, \sigma^2 \mathbb{I}_n), \tag{A.2.3}$$

which is the general form of a linear regression problem. In that context, $Y \in \mathbb{R}^{n\times 1}$ is the vector of observations (responses), $X \in \mathbb{R}^{n\times p}$ the design matrix of independent (explanatory) variables, $\beta^\star \in \mathbb{R}^{p\times 1}$ the parameter vector and ε the error vector.

The aim of linear regression is now to produce an estimator $\hat{\beta}$ of the true but unknown parameter $\beta^\star$. We will see that this estimator $\hat{\beta}$ minimizes a certain functional $S(\beta)$.

Considering equation (A.2.3), the probability of an observation Y, given a linear model with parameters β and variance σ^2 is given by the probability density

$$P\big(Y \mid \beta, \sigma^2\big) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\frac{(Y - X\beta)^\intercal (Y - X\beta)}{\sigma^2}\right],$$

which is equivalent to

$$Y \sim \mathbb{N}_n(X\beta^\star, \sigma^2 \mathbb{I}_n). \tag{A.2.4}$$

If we consider this probability density as a function of the parameters instead of a function of the data, we arrive at the likelihood of the parameters, given the data, i.e.

$$L(\beta, \sigma^2 \mid Y) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\frac{\|Y - X\beta\|^2}{\sigma^2}\right]. \tag{A.2.5}$$

The principle of *maximum likelihood* now states that the best estimate $\hat{\beta}$ for $\beta^\star$ is the one that maximizes the above likelihood function. Thus, we seek to minimize

$$S(\beta) := \|Y - X\beta\|^2 = \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2.$$

The following proposition gives the solution to the above problem.

Theorem A.2.1. *Let $Y \in \mathbb{R}^n$, $X \in \mathbb{R}^{n\times p}$ with $\operatorname{rank}(X) = p$. The unique solution of*

$$\min_{\beta\in\mathbb{R}^p} \quad S(\beta)$$

is given by

$$\hat{\beta} = (X^\intercal X)^{-1} X^\intercal Y.$$

Proof. Since the mapping $\beta \mapsto X\beta$ is linear, there is no solution with components of β tending to infinity. Thus, it suffices to consider only local solutions. It holds

$$S(\beta) = (Y - X\beta)^\intercal (Y - X\beta) = Y^\intercal Y - 2Y^\intercal X\beta + \beta^\intercal X^\intercal X\beta.$$

The gradient of S is zero if and only if

$$0 = \nabla S(\beta) = -2(Y^\intercal X)^\intercal + 2X^\intercal X\beta,$$

which holds only if

$$\hat{\beta} = (X^\intercal X)^{-1} X^\intercal Y.$$

Since $p = \operatorname{rank}(X) = \operatorname{rank}(X^\intercal X)$, the matrix $X^\intercal X$ has full rank and is invertible. □

Remark A.2.2.

(i) From now on, we will only consider regression problems satisfying $\operatorname{rank}(X) = p$. In the case of fitting a straight line (equation (A.2.2)), this condition merely means that at least two of the x_i are distinct, i.e. there is data for at least two different values of the independent variable.

(ii) There exist several approaches addressing the case $\operatorname{rank}(X) < p$ under the keyword non-full-rank models. Commonly used methods involve a generalized inverse of the matrix $X^\intercal X$ [61, Chapter 7] or introducing identifiability constraints [49, Section 3.9].

(iii) In the case (A.2.2) of fitting a straight line to data,

$$S(\beta) = \|Y - X\beta\|^2 = \sum_{i=1}^{n} (y_i - (a + bx_i))^2 =: \sum_{i=1}^{n} e_i^2.$$

Thus, $\hat{\beta} = (X^\intercal X)^{-1} X^\intercal Y$ is called the *least squares estimator* of β, since it minimizes the sum of the squared residuals e_i^2.

(iv) For later considerations, it is important to attain a geometrical view of the least squares estimation: The set

$$E = \{X\beta \mid \beta \in \mathbb{R}^n\}$$

contains all possible values of the true but unknown expectation $X\beta^\star$ of the random variable Y and is thus called *expectation surface*. It forms a

p-dimensional linear surface in $\mathbb{R}^n$ and the estimate $\hat{\beta}$ yields the point

$$\hat{Y} = X\hat{\beta}$$

on that surface, which is closest to Y in the euclidean norm.

The maximum likelihood estimator $\hat{\beta}$ is computed from the data vector Y, which itself is normally distributed via (A.2.4). Thus, $\hat{\beta}$ is normally distributed too.

Theorem A.2.3. *The maximum likelihood estimator $\hat{\beta}$ of* (A.2.3) *is an unbiased estimator for $\beta^\star$ and satisfies*

$$\hat{\beta} \sim \mathbb{N}_p(\beta^\star, \sigma^2(X^\intercal X)^{-1}).$$

Proof. Follows from Proposition A.1.2. □

In order to address the uncertainty of $\hat{\beta}$, one is interested in a confidence region of $\hat{\beta}$ that contains the true parameter $\beta^\star$ with a certain probability:

Theorem A.2.4. *Let $\hat{\beta}$ be the maximum likelihood estimator of* (A.2.3)*. The set*

$$\left\{ \beta \mid \left\| X\hat{\beta} - X\beta \right\|^2 \leq \sigma^2 \chi^2_{n,\alpha} \right\}$$

is a $100(1-\alpha)\%$-confidence region for $\beta^\star$, i.e. it contains the true parameter $\beta^\star$ with probability α. Here, $\chi^2_{n,\alpha}$ is the upper α-quantile of the χ^2_n-distribution.

Proof. Proposition A.1.5 and Theorem A.2.3 imply

$$\frac{\left\| X\hat{\beta} - X\beta^\star \right\|^2}{\sigma^2} = \frac{(\hat{\beta} - \beta^\star)^\intercal X^\intercal X(\hat{\beta} - \beta^\star)}{\sigma^2} \sim \chi^2_n.$$

Thus

$$P\left(\left\| X\hat{\beta} - X\beta^\star \right\|^2 \leq \sigma^2 \chi^2_{n,\alpha} \right) = 1 - \alpha.$$

□

Remark A.2.5. The assumption $\operatorname{rank}(X) = p$ that the columns of X are linearly independent implies that the corresponding Gram matrix $X^\intercal X$ is positive definite.

Thus, the confidence region defined in the preceding proposition forms an ellipsoid around $\hat{\beta}$, since it is defined by the quadratic form

$$(\hat{\beta} - \beta)X^\intercal X(\hat{\beta} - \beta) \leq \sigma^2 \chi^2_{n,\alpha}.$$

A.2.1 Unknown Variance

So far, we have assumed that the variance σ^2 of the error term ε was known additionally to the data Y. In our case of neurogenesis modelling as in most applications however, there is no a priori knowledge of the error variance and it needs to be estimated from the data as well.

Theorem A.2.6. *Consider the linear problem* (A.2.3) *and let $\hat{\beta}$ be the minimum of $S(\beta) = \|Y - X\beta\|^2$. The quantity*

$$s^2 = \frac{S(\hat{\beta})}{n - p}$$

is an unbiased estimator for σ^2 and satisfies

$$s^2 \sim \frac{\sigma^2}{n - p} \chi^2_{n-p}.$$

Proof. See Seber and Lee [49, Theorems 3.3 and 3.5]. □

In Theorem A.2.4, we have seen that the confidence region for the estimator $\hat{\beta}$ depends on the known error variance σ^2. In the case of unknown variance, s^2 takes the place of σ^2 and a similar result can be obtained.

Theorem A.2.7. *Let $\hat{\beta}$ be the maximum likelihood estimator of* (A.2.3). *The set*

$$\left\{ \beta \mid \left\| X\hat{\beta} - X\beta \right\|^2 \leq p s^2 F_{p,n-p,\alpha} \right\}$$

is a $100(1-\alpha)\%$-confidence region for $\beta^\star$. Here, $F_{p,n-p,\alpha}$ is the upper α-quantile of the $F_{p,n-p}$-distribution.

Proof. Analogously to Theorem A.2.4, it suffices to show that

$$\frac{(\hat{\beta} - \beta^\star)^\intercal X^\intercal X(\hat{\beta} - \beta^\star)}{p s^2} \sim F_{p,n-p}.$$

We have seen already that

$$\frac{(\hat{\beta}-\beta^\star)^\intercal X^\intercal X(\hat{\beta}-\beta^\star)}{\sigma^2} \sim \chi^2_n$$

and

$$\frac{s^2}{\sigma^2} \sim \frac{1}{n-p}\chi^2_{n-p}.$$

Thus,

$$\frac{(\hat{\beta}-\beta^\star)^\intercal X^\intercal X(\hat{\beta}-\beta^\star)}{ps^2} \sim \frac{\chi^2_n/p}{\chi^2_{n-p}/(n-p)}$$

and what remains to show is the independence of $(\hat{\beta}-\beta^\star)^\intercal X^\intercal X(\hat{\beta}-\beta^\star)$ and s^2. It holds

$$(X^\intercal X)^{-1}X^\intercal\sigma^2\mathbb{I}_n\left(\mathbb{I}_n - X(X^\intercal X)^{-1}X^\intercal\right)^\intercal = 0.$$

Using Proposition A.1.7,

$$\hat{\beta} = (X^\intercal X)^{-1}X^\intercal Y$$

and

$$Y - X\hat{\beta} = \left(\mathbb{I}_n - X(X^\intercal X)^{-1}X^\intercal\right)^\intercal Y$$

are independent of each other. Since $(\hat{\beta}-\beta^\star)^\intercal X^\intercal X(\hat{\beta}-\beta^\star)$ and s^2 continuously depend on $\hat{\beta}$ and $Y-X\hat{\beta}$ respectively, they are independent of each other as well. □

In linear regression analysis, one is sometimes interested in a confidence interval for a linear combination of the components of $\beta^\star$. The following theorem addresses this question.

Theorem A.2.8. *Let $\hat{\beta}$ be the maximum likelihood estimator of* (A.2.3) *and let $a \in \mathbb{R}^p$. A $100(1-\alpha)\%$-confidence interval for $a^\intercal\beta^\star$ is given by*

$$a^\intercal\hat{\beta} \pm \mathrm{se}(a^\intercal\hat{\beta})t^{\alpha/2}_{n-p},$$

where $t^{\alpha/2}_{n-p}$ is the upper $\alpha/2$ quantile of Student's t-distribution with $n-p$ degrees of freedom and

$$\mathrm{se}(a^\intercal\hat{\beta}) = s(a^\intercal(X^\intercal X)^{-1}a)^{1/2}$$

is the standard error of the estimate $a^\intercal\hat{\beta}$.

Proof. See Seber and Lee [49, Example 4.7]. □

An important corollary of the preceding theorem is the derivation of the standard error and confidence interval of a certain component $\beta_j^\star$ of the parameter vector $\beta^\star$, which follows from setting a equal to the j-the canonical eigenvector e_j.

Corollary A.2.9. *A $100(1-\alpha)\%$ confidence interval for the j-th parameter $\beta_j^\star$ of model* (A.2.3) *is given by*

$$\hat{\beta}_j \pm \mathrm{se}(\hat{\beta}_j) t_{n-p}^{\alpha/2},$$

where

$$\mathrm{se}(\hat{\beta}_j) = s\sqrt{(X^\intercal X)_{jj}^{-1}}$$

is the standard error of $\hat{\beta}_j$ and $(X^\intercal X)_{jj}$ the j-th diagonal element of the matrix $(X^\intercal X)^{-1}$.

A.2.2 Generalized Least Squares

In the previous two sections, it was assumed that the errors ε_i of all observations were i.i.d. normally distributed with zero mean and common variance σ^2. Dropping the latter assumption leads to the *generalized least squares* (GLS) estimation procedure. The modified linear model takes the form

$$Y = X\beta^\star + \varepsilon, \quad \varepsilon \sim \mathbb{N}_n(0, \sigma^2 V), \tag{A.2.6}$$

where $V \in \mathbb{R}^{n\times n}$ is a symmetrical positive definite matrix. The new model can be transformed to the form (A.2.3) via linear transformation: Since V is positive definite, there exists a Cholesky decomposition

$$V = TT^\intercal,$$

where T is an invertible matrix. If we define

$$Y_w = T^{-1}Y,\ X_w = T^{-1}X \text{ and } \varepsilon_w = T^{-1}\varepsilon,$$

Proposition A.1.2 implies

$$Y_w = X_w\beta^\star + \varepsilon_w, \quad \varepsilon_w \sim \mathbb{N}_n(0, \sigma^2\mathbb{I}_n). \tag{A.2.7}$$

From Theorem A.2.1, it follows that the least squares estimator of (A.2.7) satisfies

$$\begin{aligned}\hat{\beta}_w &= (X_w^\intercal X_w)^{-1} X_w^\intercal Y_w \\ &= ((T^{-1}X)^\intercal (T^{-1}X))^{-1} (T^{-1}X)^\intercal T^{-1} Y \\ &= (X^\intercal V^{-1} X)^{-1} X^\intercal V^{-1} Y.\end{aligned}$$

Analogously to the previous two sections, the following results can be shown by replacing $X^\intercal X$ with $X^\intercal V^{-1} X$:

Theorem A.2.10. *The least squares estimator $\hat{\beta}_w$ of* (A.2.7) *has the following properties:*

(i) $\hat{\beta}_w$ minimizes

$$S_w(\beta) := \|Y_w - X_w\beta\|^2 = (Y - X\beta)^\intercal V^{-1} (Y - X\beta)$$

and is given by

$$\hat{\beta}_w = (X^\intercal V^{-1} X)^{-1} X^\intercal V^{-1} Y.$$

(ii) $\hat{\beta}_w$ is an unbiased estimator for $\beta^\star$, i.e.

$$\hat{\beta}_w \sim \mathbb{N}_p(\beta^\star, \sigma^2 (X^\intercal V^{-1} X)^{-1}),$$

(iii)

$$s_w^2 := \frac{S(\beta_w)}{n-p} \sim \frac{\sigma^2}{n-p} \chi^2_{n-p},$$

(iv) The set

$$\left\{ \beta \mid \frac{(\hat{\beta}_w - \beta)^\intercal X^\intercal V^{-1} X (\hat{\beta}_w - \beta)}{p s_w^2} \leq F_{p,n-p,\alpha} \right\}$$

is a $100(1-\alpha)\%$-confidence region for $\beta^\star$.

A.2.3 Weighted Least Squares

A special case of (A.2.6) is the case of $\sigma^2 V$ being a diagonal matrix,

$$\sigma^2 V = \operatorname{diag}\big(\sigma_1^2, \ldots, \sigma_n^2\big),$$

and thus the assumption is that the i-th component of ε has variance σ_i^2. Considering Theorem A.2.10 (i), the least squares functional then takes the form

$$S(\beta) = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(y_i - (X\beta)_i\right)^2 . \tag{A.2.8}$$

with the variances σ_i^2 being weights associated to the i-th residual

$$e_i = y_i - (X\beta)_i.$$

The resulting minimizer of $S(\beta)$ is called *weighted least squares estimator*.

A.2.4 Linear Model with Replication

In our case of neurogenesis modeling, multiple measurements (usually five) of cell numbers originating from distinct animals have been taken at different time points. Such a situation leads to regression models based on replication [49, Section 10.4]. In the linear case, the corresponding model has the form

$$y_{ij} = x_i^\intercal \beta^\star + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathbb{N}(0, \sigma_i^2),$$

for $i = 1, \ldots, n_i$ and $j = 1, \ldots, m$. Thus, the assumption is that there are n_i independent responses for the vector of independent variables (observations) x_i and that the errors for the i-th observation are i.i.d. normally distributed with zero mean and common variance σ_i^2. This model is a special case of the weighted least squares model.

As previously discussed, the variances σ_i^2 might not be known and need to be estimated from the data. An obvious method would be to estimate σ_i^2 from the sample variance

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} \left(y_{ij} - \frac{1}{n_i} \sum_{k=1}^{n_i} y_{ik} \right)$$

and use the s_i^2 as weights instead of σ_i^2 in the weighted least squares functional. However, as discussed in [49, Section 10.4], this estimation is very imprecise in the case of $n_i < 6$ and leads to a much higher variance of the so obtained estimator $\hat{\beta}$.

A better approach is to estimate the variance σ_i^2 from

$$\tilde{s_i}^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} e_{ij}^2,$$

where e_{ij} is the residual of the ordinary least squares estimation and then use the $\tilde{s_i}^2$ as weights.

A.3 Nonlinear Regression

Up to now, the functional relationship between independent variables and observations was considered to be linear in the parameters. In the case of fitting differential equations to data however, this assumption will not hold in general.

We consider a non-linear model of the form

$$y_i = f(x_i, \theta^\star) + \varepsilon_i,$$

for $i = 1, \ldots, n$. Here, $x_i \in \mathbb{R}^k$ is the i-th experimental design, $\varepsilon_i \sim \mathbb{N}(0, \sigma^2)$ the error term and $\theta^\star$ is the true but unknown parameter vector, which is assumed to belong to a parameter space $\Theta \subset \mathbb{R}^p$. In vector notation, the nonlinear model can be written as

$$Y = \eta(\theta^\star) + \varepsilon, \tag{A.3.1}$$

with $\eta(\theta) = (f(x_1, \theta), \ldots, f(x_n, \theta))^\intercal$ and $\varepsilon \sim \mathbb{N}_n(0, \sigma^2 \mathbb{I}_n)$.

Analogously to the linear case, the least squares estimate $\hat{\theta}$ of $\theta^\star$ minimizes the functional

$$S(\theta) = \sum_{i=1}^{n} (y_i - f(x_i, \theta))^2 = \|Y - \eta(\theta)\|^2$$

and is also a maximum likelihood estimator of $\theta^\star$.

A.3.1 Linear Approximation

Following Seber and Wild [50], we motivate the use of a linear approximation of the function f in order to apply the linear theory of the previous section and make approximate statistical inferences about the least squares estimator $\hat{\theta}$.

Developing $\eta(\theta)$ into a Taylor series about $\theta^\star$ leads to

$$\eta(\theta) \approx \eta(\theta^\star) + J_\eta(\theta^\star)(\theta - \theta^\star),$$

where J_η is the Jacobian of η. It follows that

$$S(\theta) = \|Y - \eta(\theta)\|^2 \approx \|Y - \eta(\theta^\star) - J_\eta(\theta^\star)(\theta - \theta^\star)\|^2.$$

Assuming further that $\hat{\theta}$ is close to $\theta^\star$ so that $J_\eta(\theta^\star) \approx J_\eta(\hat{\theta})$, it follows

$$S(\theta) \approx \tilde{S}(\theta) = \left\| Y - \eta(\theta^\star) - J_\eta(\hat{\theta})(\theta - \theta^\star) \right\|^2$$

in the neighborhood of $\theta^\star$. We now apply Theorem A.2.1 to find that the minimum $\tilde{\theta}$ of $\tilde{S}(\theta)$, which is because of the above relation close to the minimum $\hat{\theta}$ of $S(\theta)$ satisfies

$$\hat{\theta} - \theta^\star \approx \tilde{\theta} - \theta^\star = (\hat{V}.^\intercal \hat{V}.)^{-1} \hat{V}.(Y - \eta(\theta^\star)),$$

where we have set $\hat{V}. = J_\eta(\hat{\theta})$. The matrix $\hat{V}.$ is called *velocity matrix* and has a physical interpretation which will be discussed in the next section. From (A.3.1), we have $Y - \eta(\theta^\star) \sim \mathbb{N}_n(0, \sigma^2 \mathbb{I})$ and Proposition A.1.2 implies

$$\hat{\theta} - \theta^\star \approx \tilde{\theta} - \theta^\star \sim \mathbb{N}_p(0, \sigma^2 (\hat{V}.^\intercal \hat{V}.)^{-1}).$$

Comparing the above relation with the one in Theorem A.2.3, it can be seen that in the linear approximation, the matrix $\hat{V}.$ takes the place of the design matrix X. Moreover, since $\hat{\theta}$ is approximately normally distributed, the preceding linear regression theory implies that an approximate $100(1-\alpha)\%$-confidence region for $\theta^\star$ is the ellipsoid

$$\left\{ \theta \mid \left\| \hat{V}.\hat{\theta} - \hat{V}.\theta \right\|^2 \leq p s^2 F_{p,n-p,\alpha} \right\}, \tag{A.3.2}$$

where $s^2 = S(\hat{\theta})/(n-p)$.

A.3.2 Geometrical Considerations

Recall from the linear regression section that the expectation surface

$$E = \{X\beta \mid \beta \in \mathbb{R}^p\}$$

is a p-dimensional linear surface in $\mathbb{R}^n$. Moreover, an α-confidence region for the true parameter $\beta^\star$ was given by

$$\left\{\beta \mid \left\|X\hat{\beta} - X\beta\right\|^2 \leq ps^2 F_{p,n-p,\alpha}\right\},$$

which is the set of all β in parameter space that are mapped by X into a sphere in the expectation surface with center $X\hat{\beta}$ and radius $\sqrt{ps^2 F_{p,n-p,\alpha}}$. In addition, this confidence region forms an ellipsoid centered at $\hat{\beta}$.

The reason that the definition of the confidence region is based on a sphere in E can be seen from (A.2.5): Points in $\mathbb{R}^n$ with equal distance to the data vector Y have equal likelihood of being the true but unknown expectation of Y. Since $X\hat{\beta}$ is the point closest to Y on E, points on E with same distance to $X\hat{\beta}$ have same distance to Y and have thus equal likelihood of being the true expectation of Y. We denote a region on the expectation surface that is used to define a confidence region as described above as the associated *reference region*.

As discussed by Bates and Watts [6], the nonlinear regression case differs from the linear one in two aspects. The first, which is called *intrinsic curvature*, is that the corresponding expectation surface

$$E_\eta = \{\eta(\theta) \mid \theta \in \Theta\}$$

is curved rather than flat. Thus, points on E_η with same distance to Y do in general not form a sphere around $\eta(\hat{\theta})$ and the goodness of a spherical approximation of the reference region as performed in (A.3.2) depends on the flatness of E_η in the neighborhood $\hat{\theta}$.

The other difference, called *parameter effects curvature*, is that the mapping $\theta \mapsto \eta(\theta)$ is nonlinear and hence the preimage of a sphere on E_η under η does in general

not form an ellipsoid. From the Taylor approximation

$$\eta(\hat{\theta}) - \eta(\theta) \approx \hat{V}.(\hat{\theta} - \theta),$$

it can be seen that the set of (A.3.2) is an ellipsoidal approximation of the confidence region

$$\left\{ \theta \mid \left\| \eta(\hat{\theta}) - \eta(\theta) \right\|^2 \leq p s^2 F_{p,n-p,\alpha} \right\}. \tag{A.3.3}$$

The name parameter effects curvature stems from the fact that the goodness of ellipsoidal approximation of (A.3.3) by (A.3.2) can change by a reparametrization of the nonlinear model, that is, by replacing η with $\eta \circ g$ where g is a bijection on Θ. In contrast, intrinsic curvature does not change by a reparametrization, since $E_\eta = E_{\eta \circ g}$.

In conclusion, intrinsic curvature means the degree of curvature of the expectation surface and is related to the deviation of the reference region from a sphere, while parameter effects curvature means the nonlinearity of the mapping $\theta \mapsto \eta(\theta)$ and is related to the ellipsoidal approximation of (A.3.3) by (A.3.2).

A.3.3 Measures of Curvature

In order to assess the validity of statistical inferences based on linear approximation, Bates and Watts [5] introduced measures for intrinsic and parameter effects curvature.

Expanding the model function η into a second order Taylor series about $\hat{\theta}$ leads to

$$\eta(\hat{\theta} + h) = \eta(\hat{\theta}) + \hat{V}.h + \frac{1}{2} h^\intercal \hat{V}..h,$$

where

$$\hat{V}. = \left(\frac{\partial \eta_i}{\partial \theta_j} \right) \Bigg|_{\theta = \hat{\theta}}$$

is the Jacobian of η and

$$\hat{V}.. = \left(\frac{\partial \eta_i}{\partial \theta_j \partial \theta_k} \right) \Bigg|_{\theta = \hat{\theta}}$$

is the $n \times p \times p$ array of the second derivative. In order to measure the nonlinearity of the function η, Bates and Watts [5] suggested to compare the size of the first-order term $\hat{V}.h$ with the one of the second order, $h^\intercal \hat{V}..h$.

To obtain a geometrical interpretation of the first and second order derivative, consider for a fixed direction $h \in \mathbb{R}^p$ the line in parameter space given by

$$\theta(t) = \hat{\theta} + th.$$

The model function η maps this line to

$$\eta_h(t) = \eta(\hat{\theta} + th).$$

The tangent to the curve η_h at $t = 0$ is then given by

$$\dot{\eta}_h = \left.\frac{\partial \eta_h}{\partial t}\right|_{t=0} = \sum_{i=1}^{p} \left.\frac{\partial \eta}{\partial \theta_i}\right|_{\hat{\theta}} \cdot \left.\frac{\partial \theta_i(t)}{\partial t}\right|_{t=0} = \hat{V}.h,$$

which can be interpreted as the velocity of a point at $t = 0$ moving along the line η_h. The second derivative at $t = 0$, given by

$$\ddot{\eta}_h = \left.\frac{\partial^2 \eta_h}{\partial t^2}\right|_{t=0} = h^\intercal \hat{V}..h$$

is the corresponding acceleration.

To measure intrinsic and parameter effects curvature, the acceleration vector $\ddot{\eta}_h$ needs to be decomposed into components $\ddot{\eta}_h^N$ normal to the tangent plane at $\eta(\hat{\theta})$ and $\ddot{\eta}_h^T$ tangential to the tangent plane, such that

$$\ddot{\eta}_h = \ddot{\eta}_h^N + \ddot{\eta}_h^T.$$

How to achieve such a decomposition is explained in [50]. Next, Bates and Watts [5] introduce the quantities

$$K_h^N = \left\|\ddot{\eta}_h^N\right\| / \|\dot{\eta}_h\|^2 \quad \text{and} \quad K_h^T = \left\|\ddot{\eta}_h^T\right\| / \|\dot{\eta}_h\|^2$$

and discuss that $1/K_h^N$ is the radius of the circle that best approximates E_η in the direction $\dot{\eta}_h$ at $\hat{\theta}$. Thus, a small value of K_h^N indicates a large radius and a *small* intrinsic curvature. Since K_h^N only depends on the curvature of the expectation surface, it measures intrinsic curvature and it can be seen that K_h^N is independent of

the parametrization of the nonlinear model [50]. In contrast, K_h^T is not independent of the parametrization and measures parameter effects curvature.

Both curvature quantities depend on the scale of the nonlinear model (A.3.1): Multiplying η and Y by a certain constant will divide K_h^N and K_h^T by that constant. Thus, Bates and Watts [5] use the so called *standard radius*

$$\rho = s\sqrt{p}$$

as a scaling factor such that

$$\gamma_h^N = \rho K_h^N \quad \text{and} \quad \gamma_h^T = \rho K_h^T$$

are dimensionless quantities defining the *relative intrinsic curvature* and *relative parameter effects curvature* respectively.

The advantage of scaling with ρ is that $\rho\sqrt{F_{p,n-p,\alpha}}$ is also the radius of the spherical reference region in the expectation surface used to define the confidence region (A.3.2). Analogously, the curvature of this sphere is $1/(\rho\sqrt{F_{p,n-p,\alpha}})$. To assess whether the intrinsic curvature of the expectation surface is small over the reference region, we are interested in the ratio

$$\frac{K_h^N}{1/(\rho\sqrt{F_{p,n-p,\alpha}})} = \frac{\gamma_h^N}{1/\sqrt{F_{p,n-p,\alpha}}}.$$

The quantities γ_h^N and γ_h^T both depend on the direction $h \in \mathbb{R}^p$. To overcome this dependency, Bates and Watts [5] define

$$\Gamma^N = \max_{h\in\mathbb{R}^p} \gamma_h^N \quad \text{and} \quad \Gamma^T = \max_{h\in\mathbb{R}^p} \gamma_h^T$$

as *maximum intrinsic curvature* and *maximum parameter effects curvature* respectively. From an analysis of 24 data sets, they found that the relation

$$\Gamma^N < \frac{1}{2\sqrt{F_{p,n-p,\alpha}}}$$

can be considered as a good rule to asses whether the linear approximation is acceptable, provided that the model is parametrized in a way such that

$$\Gamma^T < \Gamma^N.$$

Remark A.3.1. In the software package *Mathematica*, nonlinear regression analysis can be done with the `NonlinearModelFit` function, which returns a `FittedModel` object. If *nlm* is the variable storing that object, the function call

```
nlm["ParameterTable"]
```

lists the values of estimated parameters together with their corresponding (linearly approximated) confidence intervals. Moreover,

```
nlm["ParameterCurvatureTable"]
```

displays the values of Γ^N, Γ^T and $1/\sqrt{F_{p,n-p,\alpha}}$ respectively to assess the validity of the linear approximation. The default setting assumes a confidence level of $1-\alpha = 95\%$.

A.4 Error Propagation

An important aspect of mathematical modeling is to develop model-based predictions, which can be tested experimentally, in order to further validate a mathematical model. Since the model depends on parameters, θ, the prediction is a function $f(\theta)$ of the parameters as well. In order to assess the confidence one has in a certain prediction, it is necessary to carry over the uncertainty of the estimate $\hat{\theta}$ to an uncertainty of $f(\hat{\theta})$. In the case of f being linear, i.e. $f(\theta) = a^\intercal\theta$, Theorem A.2.8 shows how to compute a confidence interval and the standard error of $f(\hat{\theta})$. If this relationship is nonlinear however, approximate inferences can be made based on linear approximation of f.

Let θ be an n-dimensional real-valued random vector with expectation $\mathbb{E}(\theta) = \mu$, covariance matrix $\operatorname{cov}(\theta) = \Sigma$ and $f\colon \mathbb{R}^n \to \mathbb{R}$ a nonlinear function. Developing f about μ into a Taylor series leads to the first order approximation

$$f(x) \approx f(\mu) + \nabla_f(\mu)(x-\mu).$$

Thus, the random variable $f(\theta)$ has expectation $\mathbb{E}(f(\theta)) = f(\mu)$ and approximately variance

$$\operatorname{var}(f(\theta)) \approx \nabla_f(\mu)\Sigma\nabla_f(\mu)^\intercal. \tag{A.4.1}$$

In the case of uncorrelated components of θ, the covariance matrix has the form $\Sigma = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_n^2)$, with σ_i^2 being the variance of the i-th component of θ. Then, (A.4.1) implies that $f(\theta)$ has variance

$$\sigma_{f(\theta)}^2 = \operatorname{var}(f(\theta)) \approx \sum_{i=1}^{n} \left(\frac{\partial f(x)}{\partial x_i}\bigg|_\mu \right)^2 \sigma_i^2,$$

which is also known as the *error propagation formula* [8].

Bibliography

[1] J. L. Ables, N. A. Decarolis, M. A. Johnson, P. D. Rivera, Z. Gao, D. C. Cooper, F. Radtke, J. Hsieh, and A. J. Eisch. Notch1 is required for maintenance of the reservoir of adult hippocampal stem cells. *J. Neurosci.*, 30:10484–10492, Aug 2010.

[2] J. B. Aimone, J. Wiles, and F. H. Gage. Computational influence of adult neurogenesis on memory encoding. *Neuron*, 61:187–202, 2009.

[3] Joseph Altman and Gopal D Das. Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *Journal of Comparative Neurology*, 124(3):319–335, 1965.

[4] J. M. Ashbourn, J. J. Miller, V. Reumers, V. Baekelandt, and L. Geris. A mathematical model of adult subventricular neurogenesis. *J R Soc Interface*, 9:2414–2423, 2012. doi: 10.1098/rsif.2012.0193.

[5] Douglas M Bates and Donald G Watts. Relative curvature measures of non-linearity. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–25, 1980.

[6] Douglas M. Bates and Donald G. Watts. *Nonlinear regression analysis and its applications*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York, 1988. ISBN 0-471-81643-4.

[7] N. M. Ben Abdallah, L. Slomianka, A. L. Vyssotski, and H. P. Lipp. Early age-related changes in adult hippocampal neurogenesis in C57 mice. *Neurobiol. Aging*, 31:151–161, 2010. doi: 10.1016/j.neurobiolaging.2008.03.002.

[8] Philip R Bevington and D Keith Robinson. *Data reduction and error analysis for the physical sciences*, volume 336. McGraw-Hill New York, 1969.

[9] Michael A. Bonaguidi, Michael A. Wheeler, Jason S. Shapiro, Ryan P. Stadel, Gerald J. Sun, Guo li Ming, and Hongjun Song. In vivo clonal analysis reveals self-renewing and multipotent adult neural stem cell characteristics. *Cell*, 145: 1142–1155, 2011.

[10] Pascale Bouchard-Cannon, Lucia Mendoza-Viveros, Andrew Yuen, Mads Kærn, and Hai-Ying M Cheng. The circadian molecular clock regulates adult hippocampal neurogenesis by controlling the timing of cell-cycle entry and exit. *Cell reports*, 5(4):961–973, 2013.

[11] M. D. Brandt, M. Hubner, and A. Storch. Brief report: Adult hippocampal precursor cells shorten S-phase and total cell cycle length during neuronal differentiation. *Stem Cells*, 30:2843–2847, 2012.

[12] Kenneth P Burnham and David R Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, 2002.

[13] Federico Calegari, Wulf Haubensak, Christiane Haffner, and Wieland B Huttner. Selective lengthening of the cell cycle in the neurogenic subpopulation of neural progenitor cells during mouse brain development. *The Journal of neuroscience*, 25(28):6533–6538, 2005.

[14] VS Caviness, T Takahashi, and RS Nowakowski. Numbers, time and neocortical neuronogenesis: a general developmental and evolutionary model. *Trends in neurosciences*, 18(9):379–383, 1995.

[15] Kyung-Ok Cho, Zane R Lybrand, Naoki Ito, Rebecca Brulet, Farrah Tafacory, Ling Zhang, Levi Good, Kerstin Ure, Steven G Kernie, Shari G Birnbaum, et al. Aberrant hippocampal neurogenesis contributes to epilepsy and associated cognitive decline. *Nature communications*, 6, 2015.

[16] Elizabeth Clayton, David P Doupé, Allon M Klein, Douglas J Winton, Benjamin D Simons, and Philip H Jones. A single type of progenitor cell maintains normal epidermis. *Nature*, 446(7132):185–189, 2007.

[17] N. S. Corsini, I. Sancho-Martinez, S. Laudenklos, D. Glagow, S. Kumar, E. Letellier, P. Koch, M. Teodorczyk, S. Kleber, S. Klussmann, B. Wiestler, O. Brustle, W. Mueller, C. Gieffers, O. Hill, M. Thiemann, M. Seedorf, N. Gretz, R. Sprengel, T. Celikel, and A. Martin-Villalba. The death receptor CD95 activates adult neural stem cells for working memory formation and brain repair. *Cell Stem Cell*, 5:178–190, Aug 2009.

[18] Michael H Donovan, Umar Yazdani, Rebekah D Norris, Dora Games, Dwight C German, and Amelia J Eisch. Decreased adult hippocampal neurogenesis in the pdapp mouse model of alzheimer's disease. *Journal of Comparative Neurology*, 495(1):70–83, 2006.

[19] David P Doupé, Allon M Klein, Benjamin D Simons, and Philip H Jones. The ordered architecture of murine ear epidermis is maintained by progenitor cells with random fate. *Developmental cell*, 18(2):317–323, 2010.

[20] David P Doupé, Maria P Alcolea, Amit Roshan, Gen Zhang, Allon M Klein, Benjamin D Simons, and Philip H Jones. A single progenitor population switches behavior to maintain and repair esophageal epithelium. *Science*, 337 (6098):1091–1093, 2012.

[21] Juan M. Encinas, Tatyana V. Michurina, Natalia Peunova, June-Hee Park, Julie Tordo, Daniel A. Peterson, Gord Fishell, Alex Koulakov, and Grigori Enikolopov. Division-coupled astrocytic differentiation and age-related depletion of neural stem cells in the adult hippocampus. *Cell Stem Cell*, 8:566–579, 2011.

[22] Roland Faigle and Hongjun Song. Signaling mechanisms regulating adult neural stem cells and neurogenesis. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1830(2):2435–2448, 2013.

[23] Stefano Farioli-Vecchioli, Andrea Mattera, Laura Micheli, Manuela Ceccarelli, Luca Leonardi, Daniele Saraulli, Marco Costanzi, Vincenzo Cestari, Jean-Pierre Rouault, and Felice Tirone. Running rescues defective adult neurogenesis by shortening the length of the cell cycle of neural stem and progenitor cells. *Stem Cells*, 32(7):1968–1982, 2014.

[24] S. Furutachi, A. Matsumoto, K. I. Nakayama, and Y. Gotoh. p57 controls adult neural stem cell quiescence and modulates the pace of lifelong neurogenesis. *EMBO J.*, 32:970–981, Apr 2013.

[25] Fred H Gage. Mammalian neural stem cells. *Science*, 287(5457):1433–1438, 2000.

[26] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[27] Günter U Höglinger, Pamela Rizk, Marie P Muriel, Charles Duyckaerts, Wolfgang H Oertel, Isabelle Caille, and Etienne C Hirsch. Dopamine depletion impairs precursor cell proliferation in parkinson disease. *Nature neuroscience*, 7(7):726–735, 2004.

[28] Richard M. Huggins and Robert G. Staudte. Cell lineage analysis: Variance components models for dependent cell populations. *Journal of the American Statistical Association*, 89(425):19–29, 1994.

[29] M. H. Jang, M. A. Bonaguidi, Y. Kitabatake, J. Sun, J. Song, E. Kang, H. Jun, C. Zhong, Y. Su, J. U. Guo, M. X. Wang, K. A. Sailor, J. Y. Kim, Y. Gao, K. M. Christian, G. L. Ming, and H. Song. Secreted frizzled-related protein 3 regulates activity-dependent adult hippocampal neurogenesis. *Cell Stem Cell*, 12:215–223, Feb 2013.

[30] Sebastian Jessberger and Fred H Gage. Adult neurogenesis: bridging the gap between mice and humans. *Trends in cell biology*, 24(10):558–563, 2014.

[31] S. Jinno. Decline in adult neurogenesis during aging follows a topographic pattern in the mouse hippocampus. *J. Comp. Neurol.*, 519:451–466, 2011.

[32] Kieran M Jones, Nemanja Sarić, John P Russell, Cynthia L Andoniadou, Peter J Scambler, and M Albert Basson. Chd7 maintains neural stem cell quiescence and prevents premature stem cell depletion in the adult hippocampus. *Stem Cells*, 33(1):196–210, 2015.

[33] Gerd Kempermann, Daniela Gast, Golo Kronenberg, Masahiro Yamaguchi, and Fred H Gage. Early determination and long-term persistence of adult-generated new neurons in the hippocampus of mice. *Development*, 130(2):391–399, 2003.

[34] Marek Kimmel and David E. Axelrod. *Branching Processes in Biology.* Springer, 2002 edition, 6 2002. ISBN 9780387953403.

[35] Allon M Klein, Toshinori Nakagawa, Rie Ichikawa, Shosei Yoshida, and Benjamin D Simons. Mouse germ line stem cells undergo rapid and stochastic turnover. *Cell Stem Cell*, 7(2):214–224, 2010.

[36] Christian Lange, Wieland B Huttner, and Federico Calegari. Cdk4/cyclinD1 overexpression in neural stem cells shortens G1, delays neurogenesis, and promotes the generation and expansion of basal progenitors. *Cell stem cell*, 5(3): 320–331, 2009.

[37] S. W. Lee, G. D. Clemenson, and F. H. Gage. New neurons in an aged brain. *Behav. Brain Res.*, 227:497–507, 2012.

[38] Joseph Magill and Jean Galy. *Radioactivity, radionuclides, radiation.* Springer, 2005.

[39] Chitra D Mandyam, Gwyndolen C Harburg, and Amelia J Eisch. Determination of key aspects of precursor cell proliferation, cell cycle length and kinetics in the adult mouse subgranular zone. *Neuroscience*, 146(1):108–122, 2007.

[40] A. Marciniak-Czochra, T. Stiehl, A. D. Ho, W. Jager, and W. Wagner. Modeling of asymmetric cell division in hematopoietic stem cells–regulation of self-renewal is essential for efficient repopulation. *Stem Cells Dev.*, 18:377–385, 2009.

[41] Guo-Li Ming and Hongjun Song. Adult neurogenesis in the mammalian brain: significant answers and significant questions. *Neuron*, 70:687–702, 2011.

[42] Axel Montagne, Samuel R Barnes, Melanie D Sweeney, Matthew R Halliday, Abhay P Sagare, Zhen Zhao, Arthur W Toga, Russell E Jacobs, Collin Y Liu, Lilyana Amezcua, et al. Blood-brain barrier breakdown in the aging human hippocampus. *Neuron*, 85(2):296–302, 2015.

[43] William R Moser. *Linear models: A mean model approach.* Elsevier, 1996.

[44] Yangling Mu and Fred H Gage. Adult hippocampal neurogenesis and its role in alzheimer?s disease. *Mol Neurodegener*, 6(1):85, 2011.

[45] Nina Patzke, Muhammad A Spocter, Mads F Bertelsen, Mark Haagensen, Richard Chawana, Sonja Streicher, Consolate Kaswera, Emmanuel Gilissen, Abdulaziz N Alagaili, Osama B Mohammed, et al. In contrast to many other mammals, cetaceans have relatively small hippocampi that appear to lack adult neurogenesis. *Brain Structure and Function*, 220(1):361–383, 2015.

[46] Karl Popper. *The logic of scientific discovery*. Hutchinson, 1959.

[47] Kay L Richards, Nyoman D Kurniawan, Zhengyi Yang, Tae Hwan Kim, Marianne D Keller, Jun Low, Jeremy FP Ullmann, Stacey Cole, Samuel Foong, Graham J Galloway, et al. Hippocampal volume and cell density changes in a mouse model of human genetic epilepsy. *Neurology*, 80(13):1240–1246, 2013.

[48] A Roshan, PH Jones, and CD Greenman. Exact, time-independent estimation of clone size distributions in normal and mutated cells. *Journal of The Royal Society Interface*, 11(99):20140654, 2014.

[49] George A. F. Seber and Alan J. Lee. *Linear Regression Analysis*. Wiley, New York, 2nd edition, 2003. ISBN 978-0-471-41540-4.

[50] George A. F. Seber and C. J. Wild. *Nonlinear Regression*. John Wiley & Sons, New York, 1st edition, 2003. ISBN 978-0-471-47135-6.

[51] D. R. Seib, N. S. Corsini, K. Ellwanger, C. Plaas, A. Mateos, C. Pitzer, C. Niehrs, T. Celikel, and A. Martin-Villalba. Loss of dickkopf-1 restores neurogenesis in old age and counteracts cognitive decline. *Cell Stem Cell*, 12:204–214, 2013.

[52] A. Sierra, J. M. Encinas, J. J. Deudero, J. H. Chancey, G. Enikolopov, L. S. Overstreet-Wadiche, S. E. Tsirka, and M. Maletic-Savatic. Microglia shape adult hippocampal neurogenesis through apoptosis-coupled phagocytosis. *Cell Stem Cell*, 7:483–495, 2010.

[53] Jason S Snyder, Amélie Soumier, Michelle Brewer, James Pickel, and Heather A Cameron. Adult hippocampal neurogenesis buffers stress responses and depressive behaviour. *Nature*, 476(7361):458–461, 2011.

[54] Kirsty L Spalding, Olaf Bergmann, Kanar Alkass, Samuel Bernard, Mehran Salehpour, Hagen B Huttner, Emil Boström, Isabelle Westerlund, Céline Vial, Bruce A Buchholz, et al. Dynamics of hippocampal neurogenesis in adult humans. *Cell*, 153(6):1219–1227, 2013.

[55] Larry R Squire. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological review*, 99(2):195, 1992.

[56] T. Stiehl and A. Marciniak-Czochra. Characterization of stem cells using mathematical models of multistage cell lineages. *Mathematical and Computer Modelling*, 53:1505–1517, 2011. ISSN 0895-7177.

[57] Gerald Teschl. *Ordinary Differential Equations and Dynamical Systems (Graduate Studies in Mathematics)*. American Mathematical Society, 2012. ISBN 9780821883280.

[58] Josephine Walter, Silke Keiner, Otto W Witte, and Christoph Redecker. Age-related effects on hippocampal precursor cell subpopulations and neurogenesis. *Neurobiology of aging*, 32(10):1906–1914, 2011.

[59] Z. L. Whichard, C. A. Sarkar, M. Kimmel, and S. J. Corey. Hematopoiesis and its disorders: a systems biology approach. *Blood*, 115:2339–2347, 2010. doi: 10.1182/blood-2009-08-215798.

[60] Beate Winner, Martin Regensburger, Sebastian Schreglmann, Leah Boyer, Iryna Prots, Edward Rockenstein, Michael Mante, Chunmei Zhao, Jürgen Winkler, Eliezer Masliah, et al. Role of α-synuclein in adult neurogenesis and neuronal maturation in the dentate gyrus. *The Journal of Neuroscience*, 32(47): 16906–16916, 2012.

[61] Xin Yan. *Linear Regression Analysis - Theory and Computing*. World Scientific, Singapur, 2009. ISBN 978-9-812-83411-9.

[62] Frederik Ziebell, Ana Martin-Villalba, and Anna Marciniak-Czochra. Mathematical modelling of adult hippocampal neurogenesis: effects of altered stem cell dynamics on cell counts and bromodeoxyuridine-labelled cells. *Journal of The Royal Society Interface*, 11(94), 2014.