

# **Interaction of external forcing and noise in bio-inspired oscillator systems**

Zur Erlangung des akademischen Grades eines

**DOKTORS DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)**

von der KIT-Fakultät für Maschinenbau des  
Karlsruher Instituts für Technologie (KIT)

angenommene

**DISSERTATION**

von

**Vojtěch Kumpošt**

geboren am 2. Februar 1994 in Prag

Tag der mündlichen Prüfung: 1. Dezember 2023  
Hauptreferent: apl. Prof. Dr.-Ing. Ralf Mikut  
Korreferenten: Prof. Lennart Hilbert, PhD  
Prof. Dr. Veit Hagenmeyer



# Kurzfassung

Biologische und bioinspirierte Oszillatoren sind faszinierende Systeme, die aus technischer Sicht untersucht werden sollten. Biologische Oszillatoren sind von Natur aus verrauscht, erzeugen aber dennoch stabile Schwingungsrhythmen und passen ihre Periode an die Periode der Eingangssignale an, ein Prozess, der als Entrainment bekannt ist. Oszillatorsysteme kommen nicht nur in der Natur, sondern auch in vielen vom Menschen geschaffenen Systemen vor, zum Beispiel in Bewegungsgeneratoren für Roboter oder in elektrischen Schaltkreisen. Daher ist ein besseres Verständnis der Konstruktionsprinzipien biologischer Oszillatoren und ihrer Strategien, mit Rauschen umzugehen oder es sogar zu nutzen, um ihr Entrainment an die Eingangssignale zu unterstützen, für ein besseres Verständnis der oszillatorischen Systeme und ihrer technischen Anwendungen unerlässlich.

Der Hauptbeitrag der vorliegenden Arbeit ist die numerische Untersuchung und Analyse einer Population stochastischer Oszillatoren unter wechselnden Parametern des Eingangssignals und variierender Rauschintensität. Das theoretische Kapitel der Arbeit zeigt, dass Rauschen die Empfindlichkeit gegenüber schwachen externen Signalen erhöhen kann und somit die Anpassung an einen größeren Bereich von Eingangsamplituden und -perioden im Vergleich zu einem äquivalenten deterministischen System ermöglicht. Das Rauschen erhöht auch die Phasenreaktion auf einen stufenförmigen Eingangsimpuls und beschleunigt die Erholung von einer Jet-Lag-artigen Störung. Es wird ferner gezeigt, dass diese Effekte nicht nur auf biologische Oszillatoren beschränkt sind, sondern auch für eine größere Anzahl von generischen Oszillatorsystemen mit einem Grenzyklus zu gelten scheinen. Im letzten Teil des theoretischen Kapitels wird ein neuartiger schrittweiser Anpassungsalgorithmus vorgestellt, der eine Parameteranpassung von stochastischen Oszillatorpopulationen ermöglicht. Alle im theoretischen Kapitel entwickelten Methoden wurden als Open-Source-Softwarepaket zur Verfügung gestellt. Im letzten Teil der Arbeit wird eine praktische Anwendung für die entwickelten Methoden vorgestellt. Hier wird ein stochastisches Gleichungsmodell entwickelt, um die Tag-Nacht-Rhythmen in Zebrafisch-Zelllinien zu untersuchen. Das Modell wird anschließend verwendet, um festzustellen, wie verschiedene Medikamente die Synchronisation und Stochastizität der biologischen Uhr beeinflussen.



# Abstract

Biological and bio-inspired oscillators are fascinating systems to be studied from an engineering perspective. The biological oscillators are inherently noisy, but still, they generate stable oscillatory rhythms and adjust their period to the period of input signals, a process known as entrainment. Oscillator systems occur not only in nature but also in many human-made systems, such as movement generators for robots or electric circuits. Therefore, a better understanding of the design principles of biological oscillators and their strategies to cope with noise or even utilize it to support their entrainment to input signals is essential for a better understanding of the oscillatory systems and their engineering applications.

The major contribution of the presented thesis is the throughout numerical exploration and analysis of a population of stochastic oscillators under changing parameters of the input signal and varying noise intensity. The theoretical chapter of the thesis demonstrates that noise can increase the sensitivity to weak external signals and allows thus entrainment to a wider range of input amplitudes and periods in comparison to an equivalent deterministic system. The noise also increases phase response to a step-like input pulse and speeds up recovery from a jet-lag-like perturbation. It is further shown that those effects are not only limited to the biological oscillators but seem to apply to a wider range of generic oscillator systems with a limit cycle. In the last part of the theoretical chapter, a novel step-wise fitting algorithm is presented, which allows parameter fitting of stochastic oscillator populations to the traces from experimental data. All methods developed in the theoretical chapter were made available as an open-source software package. The last part of the thesis presents a practical application for the developed methods. Here, a stochastic equation model is developed to study the day-night rhythms in the zebrafish cell lines. The model is consequently used to determine how different drugs affect the synchronization and stochasticity of the biological clock.



# Acknowledgements

In the first place, I would like to thank apl. Prof. Dr.-Ing. Ralf Mikut for the unique opportunity to spend my time as a Ph.D. student at the Institute for Automation and Applied Informatics (IAI), for his supervision of my thesis, and for all the help and advice he gave me throughout the past years. Special thanks also to Prof. Dr. Lennart Hilbert for being my second supervisor, always being available to help, and making his lab at the Institute of Biological and Chemical Systems - Biological Information Processing (IBCS-BIP) the second home of my research journey. Many thanks also to Prof. Dr. Veit Hagenmeyer for reviewing the thesis and to Dr. Daniel Hübschmann who served along with my main supervisors as an independent member of my thesis advisory committee (TAC). I would also like to thank all the collaboration partners that helped to shape parts of this thesis. I thank Didier Gonze, Ph.D. for discussions on topics included in Section 2.6 and I thank Prof. Dr. Nicholas S. Foulkes and Dr. Daniela Vallone for the recorded data and biological expertise that made Chapter 4 possible. I would also like to thank the head office of the Helmholtz Information & Data Science School for Health (HIDSS4Health), for the opportunity to attend the exciting advanced training program and exchange with other doctoral researchers.





# Contents

<b>Kurzfassung</b> . . . . .	<b>i</b>
<b>Abstract</b> . . . . .	<b>iii</b>
<b>Acknowledgements</b> . . . . .	<b>v</b>
<b>Acronyms and symbols</b> . . . . .	<b>xi</b>
<b>1 Introduction</b> . . . . .	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Theoretical background and related work . . . . .	3
1.2.1 Limit cycle oscillator . . . . .	3
1.2.2 Design principles of biological oscillators . . . . .	5
1.2.3 Entrainment to input signal . . . . .	8
1.2.4 Stochastic effects on oscillation generation . . . . .	10
1.2.5 Population dynamics . . . . .	13
1.3 Open questions . . . . .	16
1.4 Objectives and thesis outline . . . . .	17
<b>2 Theory: Effect of noise on population-level entrainment</b> . . . . .	<b>19</b>
2.1 Bio-inspired stochastic oscillator model . . . . .	19
2.1.1 Model . . . . .	19
2.1.2 Stochastic simulations of biological oscillators . . . . .	21
2.1.3 Population-level output . . . . .	27
2.2 Metrics for entrainment quantification . . . . .	27
2.2.1 Autocorrelation . . . . .	28
2.2.2 Winding number . . . . .	29
2.2.3 Peak detection . . . . .	30
2.2.4 Circular cross-correlation . . . . .	31
2.2.5 Phase coherence . . . . .	31
2.2.6 Comparison of the entrainment metrics . . . . .	34
2.3 Noise widens range of entrainment . . . . .	35

2.4	Noise increases phase response . . . . .	38
2.4.1	Noise increases amplitude of phase response curve . . . . .	38
2.4.2	Noise increases slope of T-cycle phase response curve . . . . .	39
2.4.3	Noise allows faster recovery after jet lag . . . . .	40
2.5	Noise facilitates entrainment of generic limit cycle oscillators . . . . .	44
2.5.1	Generic oscillator models . . . . .	45
2.5.2	Noise facilitates entrainment of limit cycle but not noise-induced oscillators . . . . .	47
2.6	Population heterogeneity as a form of noise . . . . .	47
2.7	Parameter fitting algorithm for a stochastic oscillator population . . . . .	55
2.7.1	Benchmark model and data . . . . .	55
2.7.2	Step-wise fitting algorithm . . . . .	56
2.7.3	Evaluation step-wise optimization algorithm . . . . .	65
2.7.4	Goodness of fit metric . . . . .	66
2.8	Conclusion . . . . .	66
<b>3</b>	<b>Implementation: Software package for oscillator population analysis . . . . .</b>	<b>69</b>
3.1	Introduction . . . . .	69
3.2	Installation . . . . .	70
3.3	Usage examples . . . . .	71
3.3.1	Simulation of population of uncoupled stochastic oscillators . . . . .	71
3.3.2	Model and input parameter scans . . . . .	72
3.3.3	Optimization . . . . .	74
3.4	Package for peak detection . . . . .	77
3.5	Conclusion . . . . .	78
<b>4</b>	<b>Application: Entrainment of cellular clocks by light . . . . .</b>	<b>81</b>
4.1	Introduction . . . . .	81
4.2	Dataset description . . . . .	83
4.3	Dataset normalization . . . . .	85
4.4	Model derivation . . . . .	87
4.5	Model fitting and validation . . . . .	92
4.6	Model parameters quantify pharmacological treatments . . . . .	94
4.7	Using model to predict mechanisms of population-level amplitude loss . . . . .	100
4.8	Example of further application: synchronization of fungal hyphae . . . . .	105
4.9	Conclusion . . . . .	105
<b>5</b>	<b>Conclusion . . . . .</b>	<b>107</b>
<b>A</b>	<b>Supplementary material for Chapter 2 . . . . .</b>	<b>109</b>
<b>B</b>	<b>Supplementary material for Chapter 4 . . . . .</b>	<b>119</b>

<b>List of Figures</b> . . . . .	<b>125</b>
<b>List of Tables</b> . . . . .	<b>129</b>
<b>List of Publications</b> . . . . .	<b>131</b>
Journal articles . . . . .	131
Conference posters and talks . . . . .	131
<b>Bibliography</b> . . . . .	<b>133</b>



# Acronyms and symbols

## Acronyms

<b>CLE</b>	Chemical Langevin Equation
<b>DD</b>	Dark-dark, constant darkness, input signal at constant 0
<b>LD</b>	Light-dark, regular light-dark cycle, square signal
<b>LL</b>	Light-light, constant light, input signal at constant 1
<b>ODE</b>	Ordinary differential equations
<b>PRC</b>	Phase response curve
<b>PTC</b>	Phase transition curve
<b>PCA</b>	Principal component analysis
<b>SDE</b>	Stochastic differential equations
<b>SPE</b>	Stochastic population entrainment
<b>SSA</b>	Gillespie's stochastic simulation algorithm
<b>T-PRC</b>	T-cycle phase response curve

## Constants

$\pi$	Pi: 3.14159...
$e$	Euler's number: 2.71828...

## Latin symbols and variables

$A$	Free-running amplitude
-----	------------------------

$A_{kf}$	Activator concentration (Kim-Forger function)
$dt$	Integration step
$d$	Damping rate
$D$	Duty cycle of the input signal
$h$	Hours
$I$	Amplitude of the input signal
$I_{\square}$	Input square function
$K_{kf}$	Dissociation constant (Kim-Forger function)
$K_m$	Michaelis constant (Hill function)
$min$	Minutes
$ms$	Milliseconds
$n$	Population size
$n_h$	Hill coefficient
$n_e$	Number of equations in a model
$N$	Normal distribution, number of samples
$p$	General model parameter
$r$	Radial coordinate
$s$	Seconds
$t$	Time variable
$T$	Period of the input signal
$w$	Winding number
$W$	Gaussian process, Wiener process
$x, y, z$	General state variable
$Z$	Z-score

**Greek symbols and variables**

$\lambda$	Relaxation rate (amplitude-phase model)
$\sigma$	Noise intensity
$\tau$	Free-running period
$\varphi$	Angle, phase
$\omega$	Angular frequency
$\Omega$	System size parameter

**Operators and math symbols**

$x$	Vector
$\dot{x}$	Time derivative of $x$
$dx$	Infinitesimal change in $x$
$\Delta x$	Change in $x$
mod	Modulo
$[x]$	$\max(0, x)$
$ x $	Absolute value of $x$ , size of $x$
$\overline{AB}$	Line from point $A$ to point $B$
$x \sim N$	$x$ is drawn from distribution $N$
$x := y$	Assign the value of $y$ to $x$ , re-label $y$ as $x$
$\hat{x}$	Estimate of $x$
$x^*$	Scaled variable $x$
$\bar{x}$	Mean value of $x$
$S_x$	Standard deviation of $x$



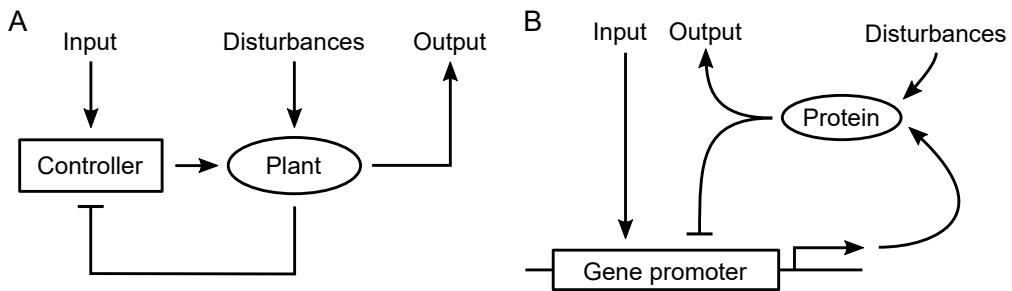


# 1 Introduction

## 1.1 Motivation

Oscillator systems are omnipresent ranging from simple mechanical systems, like mass on a spring and pendulum [7], over sophisticated man-made systems, like electrical circuits [8] and movement generators for robots [9], to complex biological systems, like day-night regulation [10] and heartbeat [11]. As oscillators are widely present in both engineering and biology, it comes as no surprise that they might be studied using the same tools [12]. For example, all biological oscillators consist of some form of negative feedback with a delay [13], mathematically not different from a negative feedback loop as we know it from control engineering (Figure 1.1). Control engineering theory can thus provide quantitative description and insights into the biological process [14]. In turn, the implementation of the control systems in living organisms provides a good opportunity to learn how to build robust and reliable man-made systems. Biological oscillator systems are inherently noisy [15] yet this noise does not appear detrimental to their function [16, 17]. To contrary, biological control systems exhibit a remarkable ability to adapt to noise environments, and sometimes even utilize it to improve their performance [18, 19]. From the perspective of control engineering, it remains, however, unclear how this robustness and utilization of noise is implemented in the biological control systems and how they could be potentially transferred to technical application on a long-term perspective.

The main aim of this thesis is to study the interaction of noise and external forcing in biological oscillator systems. The adaption of the oscillatory systems to external forcing is critical mainly to the correct function of many biological oscillators (Table 1.1) but finds its application also in the domain of the control of oscillatory mechanical systems [21]. A particular interest of my thesis is in the forcing of a population of noisy oscillators. Such a scenario is of especial importance as many biological oscillators occur in populations of thousands of oscillating cells whose effect on the organism is the result of their population output. This thesis presents a novel methodology to study this class of systems and provides a comprehensive and modular software tool to analyze various oscillator models under external forcing and noise. This tool is subsequently applied to infer new insights into the dynamics of a forced stochastic population showing that increasing noise intensity allows entrainment to a wider range of input signals and at the same time accelerate



**Figure 1.1:** The negative feedback loop is a central concept in the control engineering [7] as well as systems biology [20]. (A) The output of a plant (e.g. heater) is fed to the controller (e.g. thermostat) in a negative feedback loop (indicated with  $-$ ). The controller compares the input signal (e.g. room temperature) with the signal from the plant and regulates the plant through control signals. The plant is also affected by disturbances that might distort its function. (B) In gene regulation, a gene produces a protein that can bind to the gene promoter region and thus inhibit its production. The gene can be additionally regulated by other cellular elements acting as an external input signal. The protein levels depend also on disturbances such as noise in gene expression or changing environment.

Function	Period	Input signal	Reference
Circadian clock (day-night regulation)	24 h	Light	[22]
Cell cycle	24 h	Circadian clock	[23]
Cardiac pacemaker	200 ms	Electric current	[24]
Energy generation	50 s	Glucose	[25]
Signaling (NF- $\kappa$ B)	90 min	Signaling protein (TNF- $\alpha$ )	[26]

**Table 1.1:** Biological oscillators perform various functions and operate on various time scales. The common characteristics of those oscillators are that they are implemented as a negative feedback loop, they operate in a noisy environment, and they reliably adapt to an external input signal. This table is based on a recent review on the entrainment in biological oscillators [27].

the system response to those signals. The theoretical findings are supplemented by a practical application of the methods to the study of a population of zebrafish clock cells. Here, it is shown that the developed methods can be used to infer single-oscillators parameters from population-level recordings, which can be used to obtain insights into the effects of various drug treatments.

The following sections of this chapter provide an overview of essential theory, related work, and the context of the thesis. The main focus lies on the mathematical description of the oscillatory systems, the specifics of bio-inspired oscillators, and how those systems behave under noise-induced perturbation and periodic forcing. At the end of the chapter, it is shown how the population behavior differs from the single-oscillator behavior and what effects of noise on the population-level entrainment were described to this date. The chapter ends with an overview of open questions, aims, and major contributions of the thesis.

## 1.2 Theoretical background and related work

### 1.2.1 Limit cycle oscillator

Stable limit cycle oscillators represent a scientifically important class of oscillators that exhibit self-sustained oscillations even in free-running conditions without external forcing [28]. In practice, limit cycle oscillator models are used to describe many systems from various fields including neural activity [29], electrical circuits [8], earthquakes [30], and circadian clock [31]. Those models are represented, in general, by coupled differential equations of a general form

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{p}] \quad (1.1)$$

where  $\mathbf{x}$  is a vector of state variables,  $t$  is time,  $\mathbf{p}$  is a vector of model parameters, and  $\mathbf{f}$  is a nonlinear function of state variables and parameters. This model generates stable limit cycle oscillations if for its solution holds that

$$\lim_{t \rightarrow \infty} [\mathbf{x}(t) - \mathbf{x}(t - \tau)] = \mathbf{0} \quad (1.2)$$

where the smallest positive  $\tau$ , for which Equation 1.2 holds, is the period of the generated oscillations [32]. The free-running period of the oscillations can be conveniently adjusted without loss of dynamics by rescaling Equation 1.1 as [33]

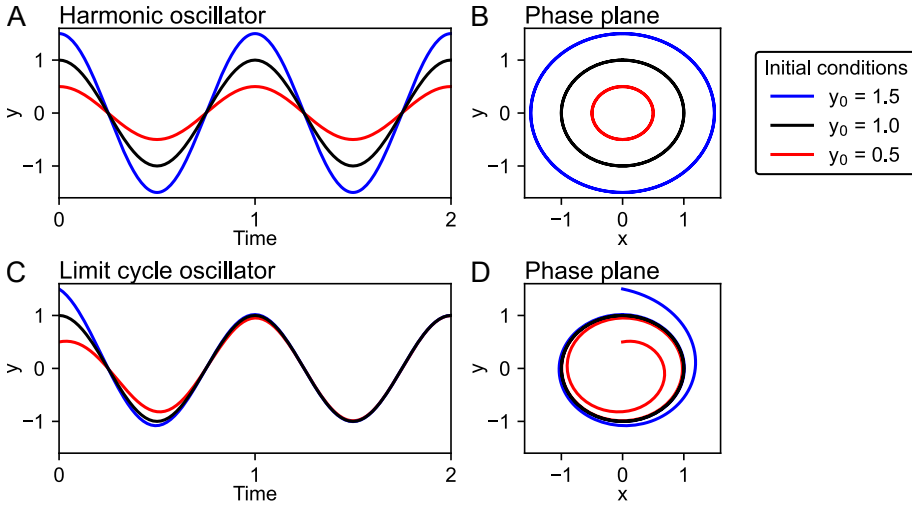
$$\dot{\mathbf{x}}(t) = \frac{\tau_{\text{new}}}{\tau} \{\mathbf{f}[\mathbf{x}(t), \mathbf{p}]\}, \quad (1.3)$$

where  $\tau_{\text{new}}$  is usually set to  $2\pi$  or 1.

A prototypical limit cycle oscillator can be characterized by its amplitude ( $A$ ), period ( $\tau$ ), and relaxation rate ( $\lambda$ ), which indicates how fast the oscillations return to the free-running amplitude after a perturbation. The amplitude-phase model (Poincaré oscillator) encodes those parameters directly in a set of ordinary differential equations (ODEs) in polar coordinates [34]

$$\dot{r} = \lambda r (A - r) \quad (1.4a)$$

$$\dot{\varphi} = \omega = \frac{2\pi}{\tau}, \quad (1.4b)$$



**Figure 1.2:** Amplitude of the harmonic oscillator depends on the initial conditions, whereas amplitude of the limit cycle oscillator converges to the stable orbit. (A) Linear system represented by the amplitude-phase model with parameters  $\lambda = 0, A = 1, \omega = -1$  exhibits oscillatory behavior that depends on the initial conditions. (B) In the phase plane corresponding to panel A, we can see three closed orbits, one for each simulated initial condition. (C) A limit cycle oscillator represented by the amplitude-phase model with parameters  $\lambda = 3, A = 1, \omega = -1$  has a stable limit cycle. Regardless of the initial conditions, the simulation converges to the same amplitude oscillations in finite time. (D) In the phase plane corresponding to panel C, we can see a single stable closed orbit that attracts all neighboring trajectories.

which can be for practical purposes also written in cartesian coordinates as

$$\dot{x} = \lambda x (A - r) - \omega y \quad (1.5a)$$

$$\dot{y} = \lambda y (A - r) + \omega x \quad (1.5b)$$

$$r = \sqrt{x^2 + y^2}. \quad (1.5c)$$

The important component of this model is the nonlinear term  $(A - r)$  that constantly evaluates the distance from the target amplitude  $A$  and adjusts the derivative so we always move in the direction of the stable orbit. This is in contrast with the simple harmonic oscillator ( $\lambda = 0$ ) whose amplitude depends on the initial conditions (Figure 1.2). The limit cycle is thus an inherently nonlinear phenomenon and can occur only in nonlinear systems [35]. Although this model is rather abstract and simple, it has a wide range of applications in studying the general properties of limit cycle oscillator systems. This model was used before as a conceptual model of biological oscillators [34, 36]. For its simplicity and generalization, this model is used in this introductory chapter to reproduce some of the fundamental properties of the oscillator systems.

An important class of limit cycle oscillators are relaxation oscillators that are phenomenologically described by the Van der Pol model [37, 38] as

$$\dot{x} = y \tag{1.6a}$$

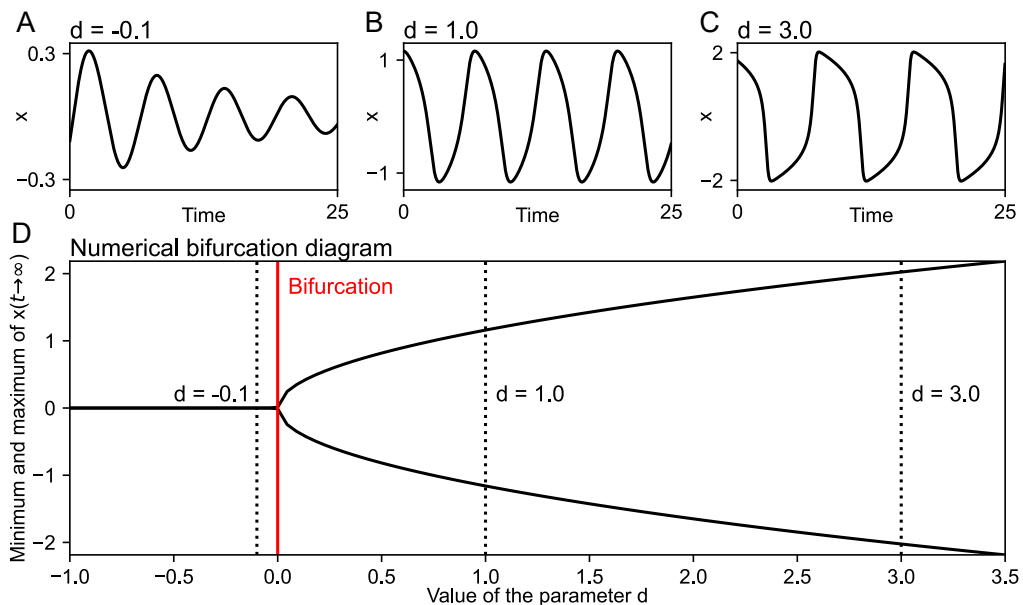
$$\dot{y} = -(Bx^2 - d)y - x, \tag{1.6b}$$

where  $B$  and  $d$  are free parameters. The special property of the relaxation oscillator is that it does not generate a sinusoidal rhythm as the amplitude-phase oscillator, but rather so-called relaxation oscillations. The term relaxation refers here to the time evolution within a cycle where we can observe a slow part (“stress” build-up) followed by a sudden jump (relaxation) [28]. Relaxation oscillations appear in various fields of science and engineering, including electrical circuits [8] as well as biological oscillators such as circadian clock [39] and cardiac pacemakers [40]. The shape of the generated oscillations also depends on the parameter values whose change can even lead to the disappearance of the limit cycle altogether (Figure 1.3). The parameter value, at which the limit cycle appears/disappears, is called Hopf bifurcation and mathematically represents a parameter value, for which a pair of complex conjugate eigenvalues of the linearized system crosses the imaginary axis [41]. The Hopf bifurcation is important in many disciplines as it indicates a threshold under which oscillations stop. In the real world, such parameter change can mean climate change in the ecosystem or wear out of components in mechanical systems [28]. The bifurcation diagram is also often used in the literature to provide better insight into the dynamics of the presented oscillator model [42].

## 1.2.2 Design principles of biological oscillators

The models presented in the last section provide a simple and valuable toolkit to study the general properties of limit cycle oscillators but do not consider any specific properties of biological systems. To make a step closer to the biological reality, we can consider that state variables  $\mathbf{x}$  represent concentrations of molecules and parameters  $\mathbf{p}$  are kinetic parameters that govern the interactions between those molecules [32]. Consequently, the ODEs represent the production and decay of those molecules as [41]

$$\dot{\mathbf{x}}(t) = f_{\text{production}}[\mathbf{x}(t), \mathbf{p}] - f_{\text{decay}}[\mathbf{x}(t), \mathbf{p}]. \tag{1.7}$$



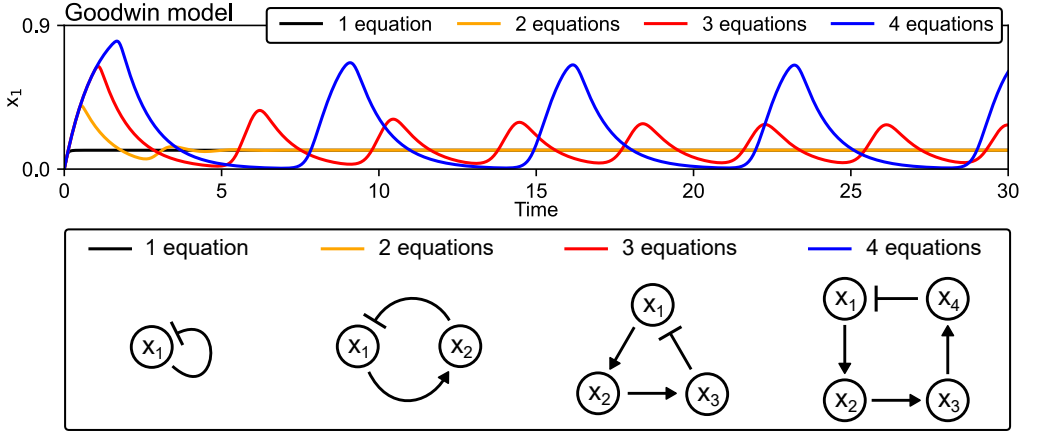
**Figure 1.3:** The limit cycle is born at the Hopf bifurcation. (A) The Van der Pol model with parameter values of  $B = 3.0$ ,  $d = -0.1$  generates only damped oscillations. (B) By increasing the value of  $d$  to  $d = 1.0$ , we cross the Hopf bifurcation, and the model becomes a limit cycle oscillator. (C) Increasing the value of  $d$  further to  $d = 3.0$  increases amplitude and period and makes the relaxation oscillations more prominent with clearly visible slow changes followed by sudden jumps. (D) Numerical estimation of the bifurcation diagram for varying parameter  $d$  shows the Hopf bifurcation at  $d = 0$ .

As  $x$  in Equation 1.7 represents molecular concentrations, it cannot reach negative values creating a constraint on the simulation of this class of models. A simple example of a mechanistic model is Goodwin model [43] that can be in the most general form written as [44]

$$\dot{x}_1 = g(x_{n_e}) - x_1 \quad (1.8a)$$

$$\dot{x}_i = x_{i-1} - x_i \quad i = 2, \dots, n_e, \quad (1.8b)$$

where  $n_e$  is the number of coupled equations and  $g$  is a nonlinear function with a strictly negative derivative [45] representing negative feedback. The number of equations in this model represents time delay that is essential to generate the oscillations (Figure 1.4) [13].



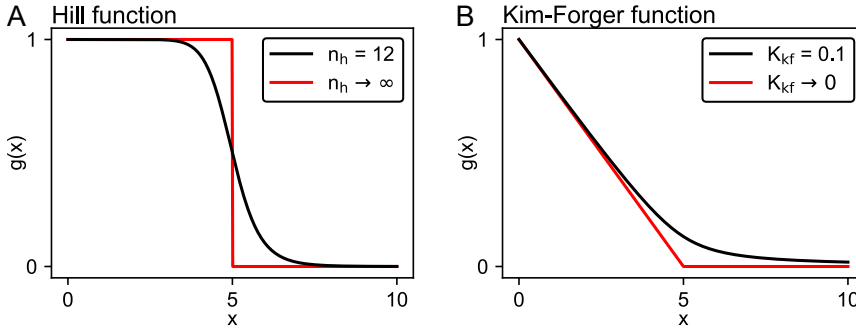
**Figure 1.4:** Oscillations can be generated only with a sufficient delay in the negative feedback loop. The Goodwin model was simulated with one, two, three, and four coupled equations. The more equations we consider the longer is the delay caused by the propagation of the signal through the individual state variables. To generate sustained oscillations, the Goodwin model must consist of at least three equations.

The function  $g$  represents inhibition, such as caused by a repressor molecule inhibiting gene expression, and can be chosen based on the inhibitory mechanism we want to model [46]. The most commonly used is the Hill function (Figure 1.5A) [41]

$$g(x) = \frac{1}{1 + \left(\frac{x}{K_h}\right)^{n_h}} \xrightarrow{n_h \rightarrow \infty} \begin{cases} 1 & x \leq K_h \\ 0 & x > K_h \end{cases}, \quad (1.9)$$

which has a shape of a decreasing sigmoid function. This represents cooperative binding that manifests as a switch-like process: with increasing concentration of  $x$  over the threshold  $K_h$  the function drops from 1 to 0.  $n_h$  indicates the steepness of the sigmoid [41] and must be sufficiently large for the oscillations to occur. For example, in the minimal Goodwin model with 3 equations we need  $n > 8$  [47]. The Hill function has been for decades de facto the standard function to use with the Goodwin model, especially in theoretical studies. Later, the Kim-Forger function (Figure 1.5B) has emerged as an alternative. The Kim-Forger function can be written as [48]

$$g(x) = \frac{A_{kf} - K_{kf} - x + \sqrt{(A_{kf} - K_{kf} - x)^2 + 4A_{kf}K_{kf}}}{2A_{kf}} \xrightarrow{K_{kf} \rightarrow 0} \begin{cases} 1 - \frac{x}{A_{kf}} & x \leq A_{kf} \\ 0 & x > A_{kf} \end{cases} \quad (1.10)$$



**Figure 1.5:** Common nonlinear functions that are used in the Goodwin model. (A) The Hill function has two free parameters  $n_h$  and  $K_h$ .  $n_h$  indicates the slope of the inhibition and  $K_h$  indicates the position of the half-maximal value (here  $K_h = 5$ ). (B) The Kim-Forger function has two free parameters  $K_{kf}$  and  $A_{kf}$ .  $K_{kf}$  indicates the "smoothness" of the elbow and  $A_{kf}$  indicates its position (here  $A_{kf} = 5$ ).

and represents 1:1 binding of activator ( $A_{kf}$ ) and repressor ( $x$ ) molecules: with increasing concentration of  $x$  more activator molecules  $A_{kf}$  are bound to  $x$  and the activation function decreases until it reaches 0 (no free activators remain) [49]. It has been suggested that transcriptional repression in multicellular organisms is better explained by the Kim-Forger function, rather than Hill function [50]. The Kim-Forger function also seems to be more reliable for stochastic simulations [51]. Similarly as with the Hill function, sufficient nonlinearity is necessary to produce oscillations ( $K_{kf} < 10^{-4}$ ) [48].

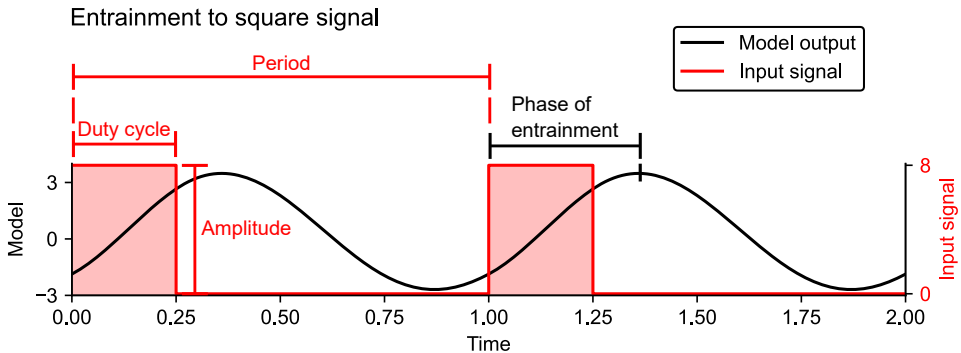
### 1.2.3 Entrainment to input signal

Entrainment is a phenomenon when an oscillator adjusts its period to the period of an input periodic signal. The fundamental property of entrainment, as compared to the general phenomenon of synchronization, is that the oscillator generates sustained oscillations also without external forcing [52]. Entrainment is common to many biological oscillators that must adjust their period to changes in the external environment, for example, entrainment of the auditory cortex oscillators by speech and music [53], circadian clock entrained by the day-night rhythm [10] or respiratory system entrained by a ventilator [54]. A common forcing signal used as an input to the oscillator in computational studies is a square-wave signal

$$I_{\square}(t) = \begin{cases} I & 0 \leq (t \bmod T) < DT \\ 0 & DT \leq (t \bmod T) < T \end{cases}, \quad I \geq 0, \quad T > 0, \quad D \in [0, 1], \quad (1.11)$$

where  $I$  is the input amplitude,  $T$  input period and  $D$  input duty cycle (Figure 1.6). Other shapes of the input signal can be also applied to force the oscillator, among which the sine wave is the



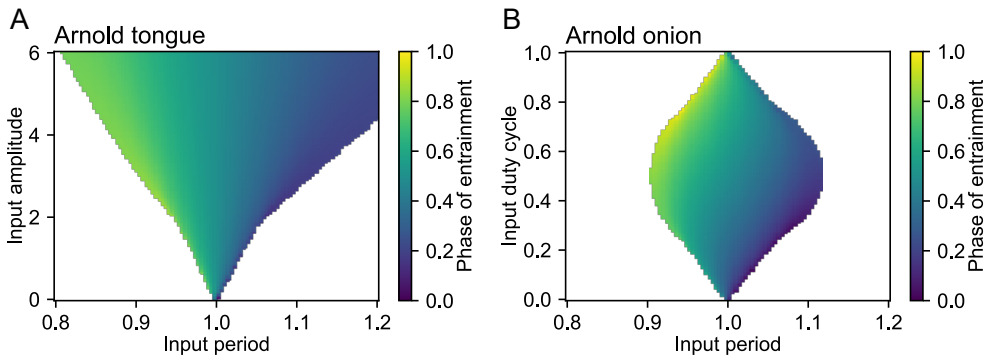


**Figure 1.6:** The square input signal (red) is characterized by three parameters: input amplitude ( $I$ ), period ( $T$ ), and duty cycle ( $D$ ). Parameters of the showed input signal are  $I = 8$ ,  $T = 1$ , and  $D = 0.25$ . The model output (black) oscillates with the same period as the input signal and is thus entrained. For an entrained oscillator, we can define the phase of entrainment as the distance from the input signal cycle onset to the nearest following output signal peak.

most popular. Although some differences between the response of the oscillator system to sine and square input signals were reported [55], the differences seem to be, for the context of this thesis, rather quantitative than qualitative. A square wave is also easier to realize in experimental settings. In simple conceptual models, input is usually implemented as an additive term to one of the state equations [56, 36, 57, 58]. In some special cases, such as light input to the mammalian circadian clock, the input signal might be also modeled multiplicatively by changing a value of one of the kinetic parameters already presented in the model [59].

The important property of the forced oscillator is the range of the input signal parameters that allow for entrainment. An Arnold tongue is a pictogram that visualizes the area of the entrainment in a 2D plane representing varying input amplitudes and periods of the input signal (Figure 1.7A). In general, if we increase the amplitude of the input signal the range of periods for which the system is entrained also increases. This results in the typical, tongue-shaped region of entrainment [60, 61]. Arnold tongues are a standard method to describe entrainment in oscillator systems including biological [55] as well as mechanical oscillators [62]. In the context of some oscillator systems, the input amplitude in the pictogram is switched for the input duty cycle creating a visualization known as Arnold onion (Figure 1.7B). Arnold onions stem from the circadian research where the duty cycle represents varying lengths of the day over the seasons [63, 36].

Another classical way to quantify the response of an oscillator to an input signal is the phase response curve (PRC). The PRC visualizes the phase shift caused by a pulse input applied at a specific time of a cycle. To measure the PRC the oscillator is simulated without the input signal oscillating with its free-running period. Then a pulse input is delivered at different times across the period (cycle) and the phase change is measured as the time shift of the oscillations from the reference output without the input signal [64]. The PRC is plotted as the pulse time versus the



**Figure 1.7:** Arnold tongues and onions quantify the range of entrainment for varying parameters of the input signal. (A) Arnold tongue plots the area of entrainment as a function of input amplitude and period. The duty cycle is fixed (here  $D = 0.5$ ). The phase of entrainment is color-coded. The white area indicates the values of the input amplitude and period for which the model is not entrained. (B) Arnold onion plots the area of entrainment as a function of input amplitude and duty cycle. Input amplitude is fixed (here  $I = 3$ ). Color-coded is the phase of entrainment. The white area indicates the values of the input amplitude and duty cycle for which the model is not entrained.

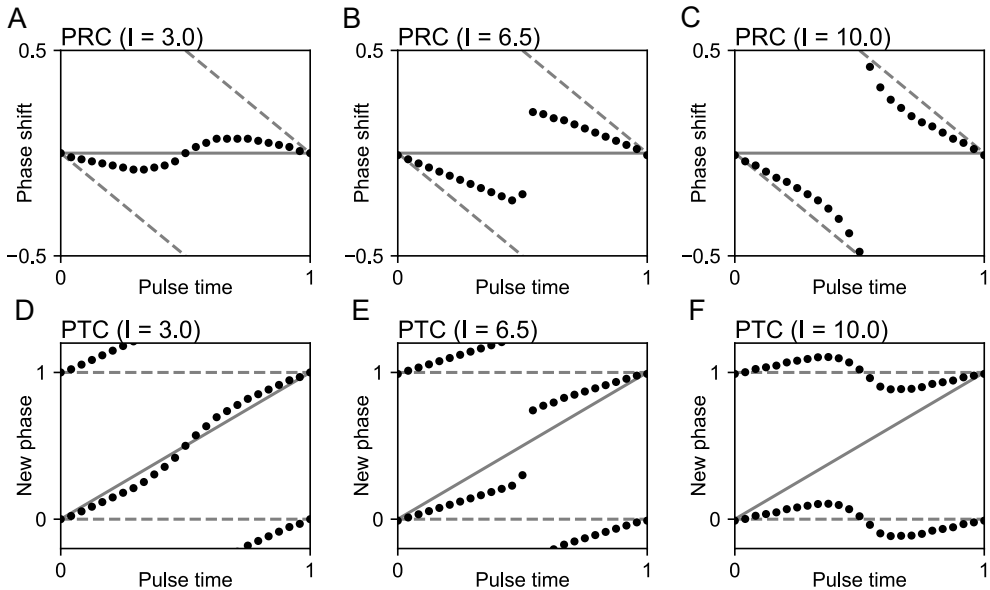
phase shift (Figure 1.8A, B, C) and its shape depends on the oscillator model structure as well as the applied pulse amplitude and length [65]. Based on the shape of the PRC we can call the phase response as type 1 or type 0 (Figure 1.8). The terms type 0 and type 1 refer to the average slope of the phase transition curve (PTC), an alternative to the PRC that plots the new phase versus the pulse time (Figure 1.8D, E, F) [64]. The new phase is calculated as new phase = pulse time + phase shift. The PRCs are extensively used to quantify pulse-response properties of biological and bio-inspired oscillators such as phase resetting of the circadian clocks [10], phase control of the beating heart cells [66], and adaptive control of the human walking rhythms [67, 68].

## 1.2.4 Stochastic effects on oscillation generation

Noise is an inherent component of biological control systems [15] and in many cases, it is necessary to consider, if we desire to capture the correct dynamics of the biological oscillators or fit precisely the experimental data [69, 70]. Noise in the computational models can be encoded using stochastic differential equations (SDEs) in the form of

$$d\mathbf{x}(t) = f[\mathbf{x}(t), \mathbf{p}]dt + \sigma h[\mathbf{x}(t), \mathbf{p}]d\mathbf{W}(t) \quad (1.12)$$

where  $f$  is called a drift function,  $\mathbf{x}$  are state variables,  $\mathbf{p}$  are model parameters,  $\sigma$  is noise intensity,  $h$  is called a diffusion or noise function, and  $\mathbf{W}$  are independent Gaussian processes. In biological systems, the noise stems from a low number of interacting molecules, which makes the molecular interactions scarce and occur at stochastically distributed times [71, 72]. We can

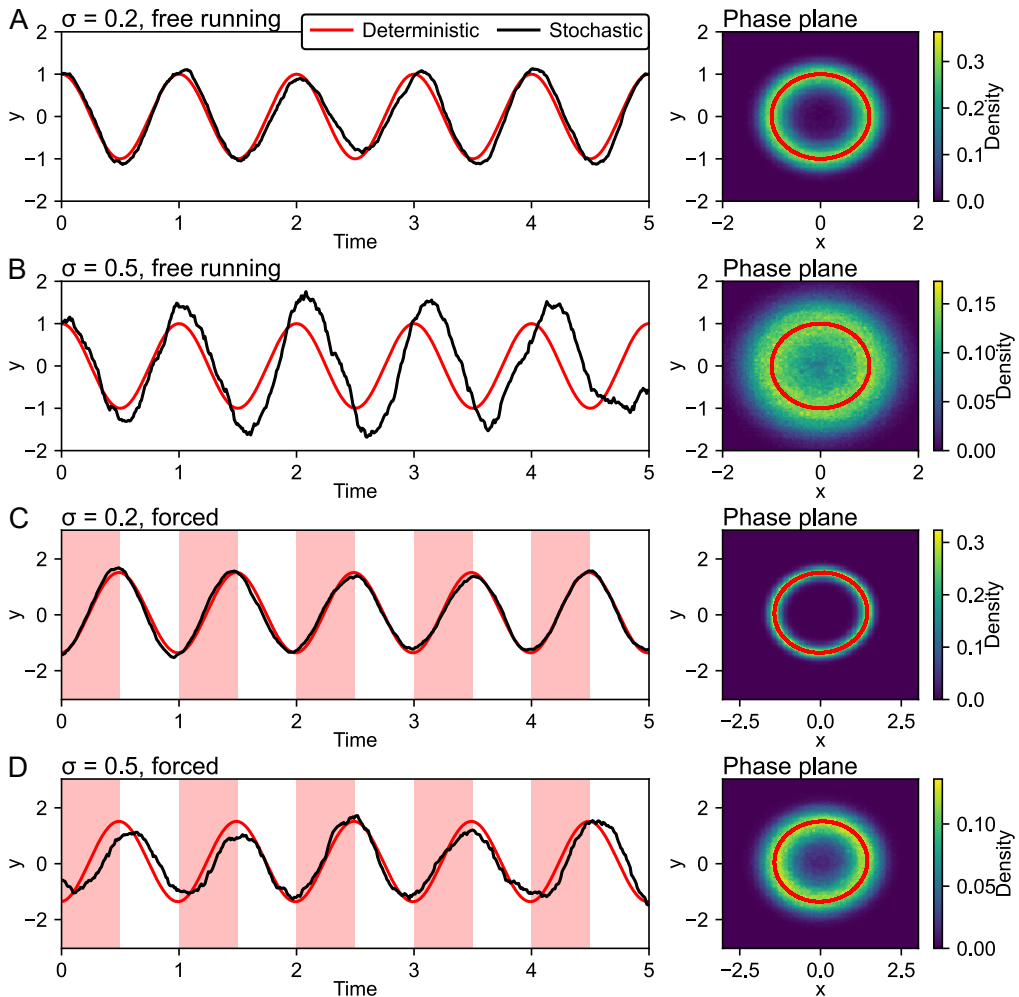


**Figure 1.8:** The phase response curve (PRC) changes its shape with the increasing amplitude  $I$  of the input signal. (A) PRC for low input amplitude shows only a minimal change in phase shift (type 1). (B) An increasing input amplitude increases also the phase shift. (C) For high input amplitude the PRC "wrap-around" the cycle (type 0). (D) Phase transition curve (PTC) to the PRC from panel A. The new phase is calculated as "new phase = pulse time + phase shift" and is technically just the PRC rotated by 45 degrees (see the guiding lines). The type 1 resetting manifests itself in the PTC plot as a curve with an average slope of 1. (E) The PTC equivalent to the PRC from panel B. (F) The PTC to the PRC from panel C. The type 0 resetting manifests itself in the PTC plot as a curve with an average slope of 0.

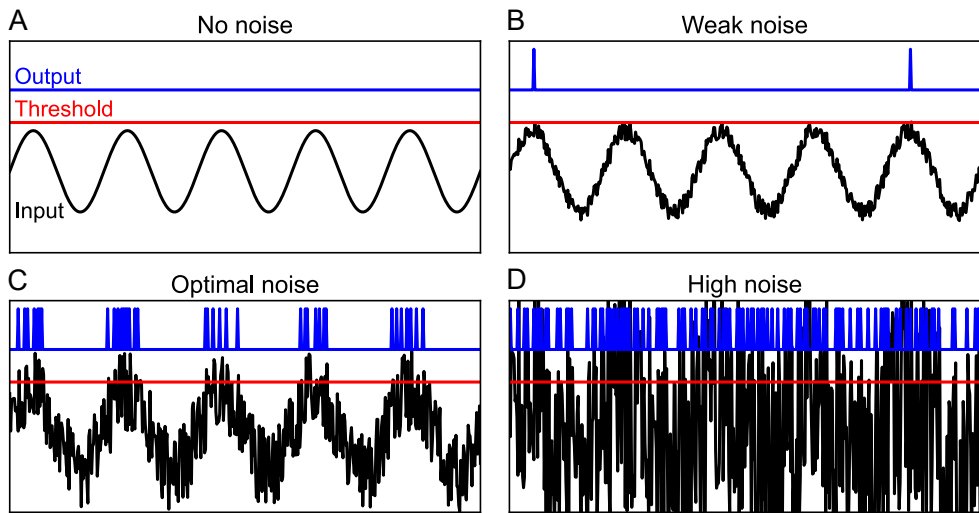
therefore write the noise intensity as a function of system size ( $\Omega$ ), which represents the number of molecules, as [73]

$$\sigma = \frac{1}{\sqrt{\Omega}}. \quad (1.13)$$

Another important notion is that the stochastic processes  $\mathbf{W}$  in Equation 1.12 are not differentiable by time and thus classical ODE solvers (such as ode45 in Matlab) cannot be used and suitable SDE solvers must be applied to obtain a precise approximation of the solution [74]. This also means that the stochastic models are considerably more computationally demanding than deterministic models and more challenging to fit to experimental data. Despite this challenge, a great body of work focused on the interference of noise with oscillator generation showing that oscillations can occur also in systems with high noise intensities [75, 76, 77, 78] and that also noisy systems can be entrained by external forcing [79, 80]. The main consequence of increasing noise intensity in the limit cycle oscillator model is an increase in random perturbation of the system from its limit cycle (Figure 1.9).



**Figure 1.9:** Noise causes perturbations from the limit cycle. (A) For a relatively small noise intensity, the perturbations are minimal. In the phase plane, we can conveniently see the range of values around the deterministic limit cycle that the stochastic model visited during a simulation over 10,000 periods. The red ring indicates the limit cycle and the color-coding indicates the fraction of time that the stochastic model spent in the given state during a simulation of 10,000 periods. (B) Increasing noise intensity increases also the range of deviations from the limit cycle. Notice that the stochastic time series also shows a visibly larger amplitude and longer period than the deterministic model. (C) Periodic forcing decreases the deviations of the stochastic simulation from the deterministic limit cycle (compare to panel A). (D) Increased noise intensity increases also the range of deviations from the limit cycle. But the deviations are visibly smaller than for the free-running model (compare to panel B). As the input signal amplitude in panels C and D is sufficiently large, the oscillator is entrained to the input period, and thus the oscillator period is not changed by higher noise intensity.

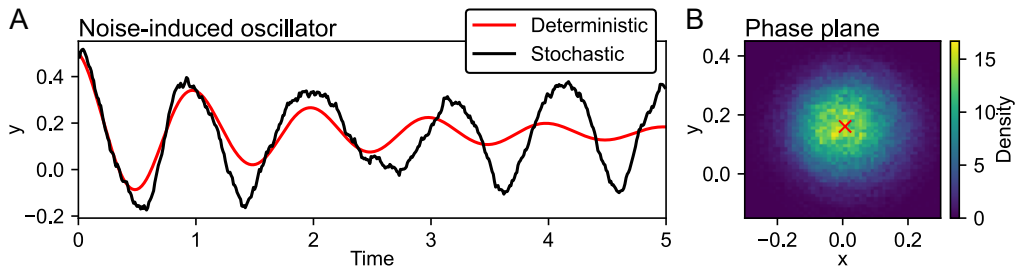


**Figure 1.10:** Noise improves the detection of weak signals in threshold-like systems. (A) The input signal without noise does not reach a threshold and thus does not generate any input signal. (B) The addition of noise makes the input signal reach the threshold and generates output spikes. Each spike corresponds to the time when the input signal crossed the threshold. (C) Optimal noise intensity leads to the generation of spike trains in periodic intervals that match the period of the input signal. (D) If the noise intensity is too high, the spikes are generated perpetually and the input pattern is difficult to recognize. This figure is inspired by a previous description of the threshold stochastic resonance in the context of neural synchrony [85].

In the state-of-the-art man-made control systems, noise is usually considered undesirable, but in biological control systems, noise is often utilized to facilitate information transfer [81, 82]. The effect of increasing the performance of the system by increasing noise intensity is called stochastic resonance [83] and was observed in various disciplines including cell biology, ecology, and physics [84]. In the context of the oscillator systems, the term stochastic resonance is traditionally used in neuroscience to describe improved detection of weak signals in threshold-like systems (Figure 1.10) [85, 86]. In practice, this means enhanced response of sensory neurons to an input signal [87, 88]. Noise can also play a key role in the generation of oscillations. Specifically, in the proximity to the Hopf bifurcation, the noise can induce oscillations [89, 90, 91] even if the deterministic model has no limit cycle and exhibits only damped oscillations without noise (Figure 1.11).

## 1.2.5 Population dynamics

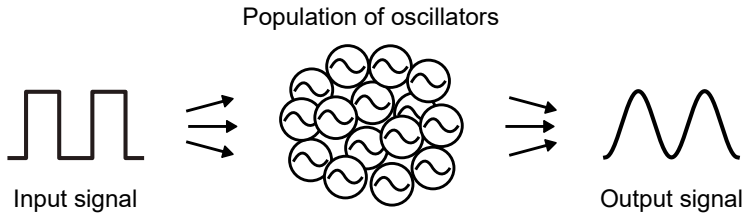
Let us consider a population of  $n$  uncoupled stochastic oscillators forced by a common input signal (Figure 1.12). A population of uncoupled forced stochastic oscillations can be used to model many biological systems, including oscillating signaling proteins [26, 92], populations of



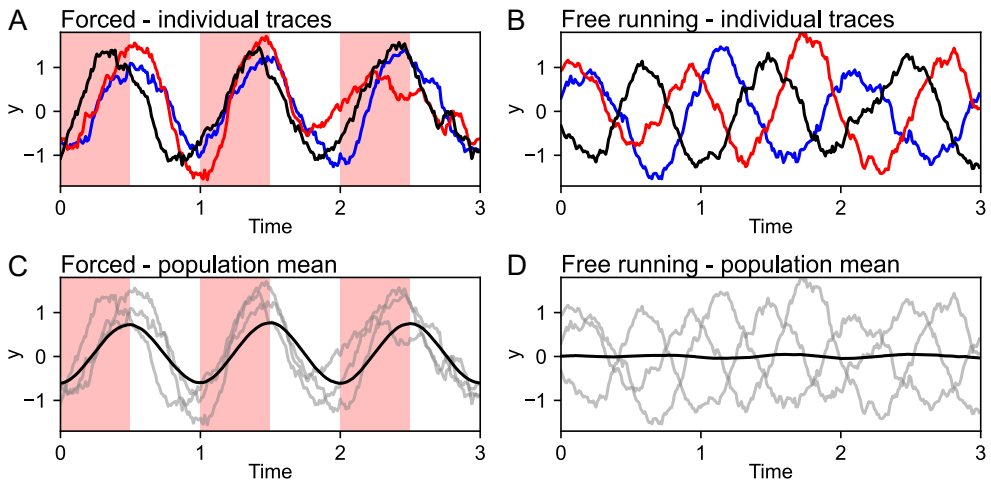
**Figure 1.11:** Noise can induce oscillations in a system that does not have a limit cycle. (A) The output of the oscillator without noise produces only damped oscillations that after a transient period reach a steady state. However, if noise is added to the system, the model generates sustained oscillations. (B) In the phase plane, noise leads to a generation of a circular pattern around the deterministic steady state (red x). The phase plane values correspond to the stochastic system from panel A (black curve). The red x corresponds to the steady state of the deterministic system from panel A (red curve).

synthetic oscillators [93], and many examples from the field of circadian clocks, such as circadian oscillators in plants [94], mammalian peripheral tissues [95], and laboratory cell cultures [96]. The importance of studying the population dynamics as opposed to studying individual oscillators can be illustrated with two examples when the population-level output and dynamics of individual cells differ substantially. The first example is the desynchronization of the individual oscillators in the free-running conditions (Figure 1.13). Without external forcing, the individual oscillators desynchronize, and thus their population-level average will display no oscillations even though the individual oscillators keep oscillating [96]. The second example is the phenomenon of the stochastic population entrainment (SPE, Figure 1.14). The SPE refers to the phenomenon when the population-level average is entrained by the input signal even though the individual stochastic oscillators are not [92]. Those two examples of the discrepancy between the population-level average and single-oscillator traces demonstrate the importance of studying the population-level output and how it deviates from the dynamics of the individual cells. A better understanding of the relationship between the population-level output and the single-cell dynamics would be beneficial, for example, for the correct analysis of population-level recording methods that are often easier, cheaper, and faster than performing single-cell imaging [97].

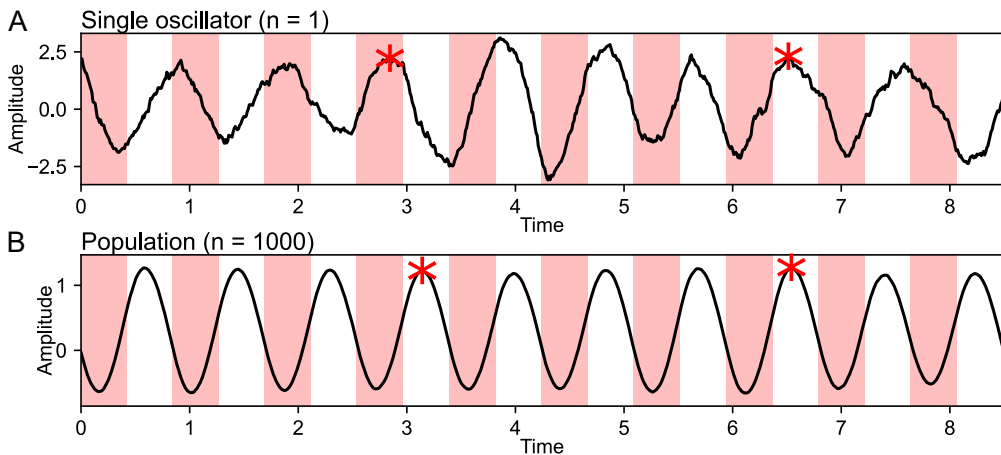
The system of uncoupled stochastic oscillators presented in the paragraph above is in contrast to a more studied case of a population of coupled oscillators. Much previous work has been devoted to the study of the effects of coupling among the individual oscillator [98] and the effect of external forcing in such systems [99, 100]. The coupling has been shown to provide the network with an increased level of resistance to noise and external perturbation [34]. Weakening coupling between the individual oscillators allows entrainment to a wider range of input periods and weaker external input signal [101]. The noise plays also an important role in the dynamics of a coupled population. Noise has been shown to enhance synchronization between networks of oscillators [102] and can



**Figure 1.12:** The main interest of this thesis is a population of uncoupled stochastic oscillators forced by a square signal. The same input signal acts on all members of the population that in turn produce an oscillatory output. The output signal is obtained as average of the outputs of the individual oscillators. If the input and output signals oscillate with the same period, the population is entrained.



**Figure 1.13:** Noise desynchronize individual oscillators leading to the loss of oscillations at the population level. (A) Under a periodic forcing, the individual oscillators are synchronized and oscillate with similar phase. (B) Without external forcing, the individual oscillators desynchronize and peaks occur at any time. (C) Under periodic forcing, population-level oscillations can be observed. In gray, the example traces from panel A. The population mean is constructed from 1000 independent traces. (D) The lack of periodic forcing leads to the loss of oscillatory behavior at the population level even though the individual oscillators keep oscillating. In gray, the example traces from panel B. The population mean is constructed from 1000 independent traces.



**Figure 1.14:** At the population level, a population of oscillators can be entrained even if individual oscillators in the population are not. (A) A single stochastic oscillator is not entrained to the input signal. Notice that the relative peak position (indicated with stars) changes in the course of several periods. (B) An average of 1000 independent stochastic oscillators show robust entrainment with peaks (indicated with stars) occurring at the same relative position in the cycle.

shorten the time needed for the coupled population to recover from a jet-lag-like phase reversal of the input signal [103]. Although those findings are intriguing for further exploration, the main focus of this thesis is the case without coupling. Uncoupled oscillators provide reduced mathematical complexity and faster computational simulations. This simplification allows one to computationally explore a wide range of noise intensities under various scenarios of the input forcing and for large populations of thousands of cells including fitting those large populations to experimental data. Given the abundance of oscillators in biology and other fields, this thesis will contribute to the general knowledge of oscillatory systems as well as to the analysis of a specific class of oscillatory systems that consists of uncoupled stochastic oscillators [33].

### 1.3 Open questions

Although previous research on oscillator systems offers a tremendous amount of theoretical, computational, and experimental insights, there are still several open questions in the field of forced stochastic oscillators that remain to be addressed:

- Entrainment, period, and other parameters of forced oscillators are challenging to reliably quantify in stochastic systems. Many common methods used to quantify entrainment (Arnold tongues, PRCs) work well for deterministic systems, but their employment for the



analysis of stochastic systems with significant level of noise remains unambiguous in the face of stochastic uncertainties.

- The implications of noise on population-level entrainment are still poorly understood. Although noise and entrainment are individually recognized to influence oscillator dynamics, it remains unclear whether noise is generally detrimental to entrainment, or might be co-opted in the entrainment of oscillator populations.
- High computational costs of stochastic simulations keep fitting stochastic oscillator models to real-world data challenging. Here, it is necessary to explore methods that would keep the necessary reevaluations of the equations to a minimum but could still provide robust and reliable estimation for the model parameters.
- Biological as well as man-made systems adjust their function as a response to variations in environmental conditions. A suitable parameter identification strategy that could map those environmental changes to the changes in parameter values of a computational model would be beneficial. In the biological context, such a framework could be used to infer the effects of pharmacological treatments or genetic mutants. In the context of man-made systems, the changes in environmental conditions could be quantified.
- In the field of the circadian clock, experimental measurement of the degree of single-cell clock synchronization requires long-term imaging of individual cells, which is technically challenging and not suited to large-scale screening experiments. Here, a mathematical modeling approach that could predict the degree of individual cell clock synchronization from population-level recordings would be valuable.

## 1.4 Objectives and thesis outline

Concerning the open question presented in the previous section, the main objectives of the present thesis are:

1. Provide methodology than can be used to reliably quantify the entrainment of stochastic oscillators to the input signal.
2. Explore and quantify the behavior of a population of uncoupled stochastic oscillators under varying parameters of the input signal (period, amplitude, duty cycle) and different noise intensities.
3. Build an efficient data-fitting pipeline that can be used to infer stochastic oscillator parameters from real-world data.

4. Implement all developed methods in a comprehensive and modular open-source software package that enables computational exploration of oscillator models and their fitting to the real-world data.
5. Use the implemented tool to fit, analyze and interpret real-world data from the field of the circadian clock.

Those objectives are addressed in the following chapters of the thesis. Chapter 2 presents some theoretical findings and methods for the quantification of entrainment in stochastic oscillator populations and a pipeline for fitting experimental data. This answers the Objectives 1, 2, and 3. Chapter 3 describes an open-source implementation of the methods developed in Chapter 2 and thus directly relates to Objective 4. In Chapter 4, the developed methods are applied to the analysis of experimental data, which answers the last objective, Objective 5, of the thesis. Chapter 5 summarizes the work and provides an outlook for further research.

## 2 Theory: Effect of noise on population-level entrainment

This chapter presents new theoretical findings in the domain of uncoupled stochastic oscillator and their entrainment. The chapter begins with the evaluation of existing numerical methods for the simulation of stochastic models of biological oscillators and gives a recommendation on the simulation approach that is accurate and computationally effective (Section 2.1). With the established simulation approach, the chapter continues with the development of the methodology to reliably quantify the entrainment of stochastic oscillators under periodic forcing (Section 2.2). This addresses Objective 1 of the thesis. Once the optimal approach to the numerical simulation and entrainment quantification is established, the chapter continues with an extensive theoretical exploration of the entrainment dynamics for a population of stochastic oscillators (Sections 2.3, 2.4, 2.5, 2.6). This addresses Objective 2 of the thesis. The chapter concludes with a presentation of a novel parameter-fitting algorithm (Section 2.7). This addresses Objective 3 of the thesis and builds a bridge to the practical application described later in Chapter 4. Parts of this chapter have been adapted from a previously published journal article [1].

### 2.1 Bio-inspired stochastic oscillator model

#### 2.1.1 Model

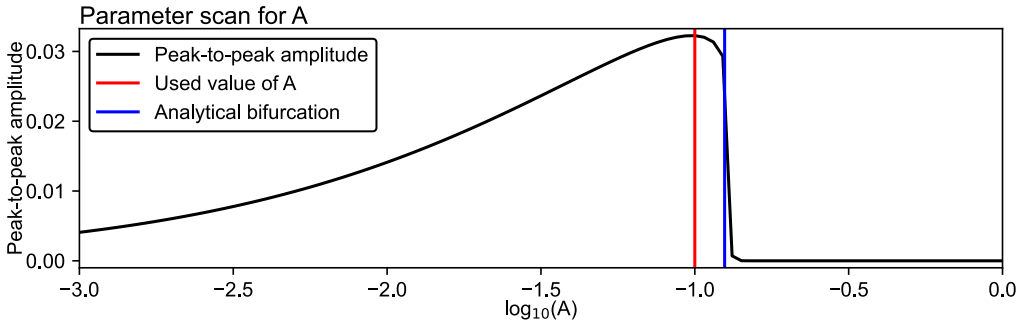
In this chapter, the simple Kim-Forger model [48] is used as a prototype of a generic biological oscillator. In its minimal form, the Kim-Forger model consists of three equations as

$$\dot{x} = f(z, A) - x + I_{\Gamma} \tag{2.1a}$$

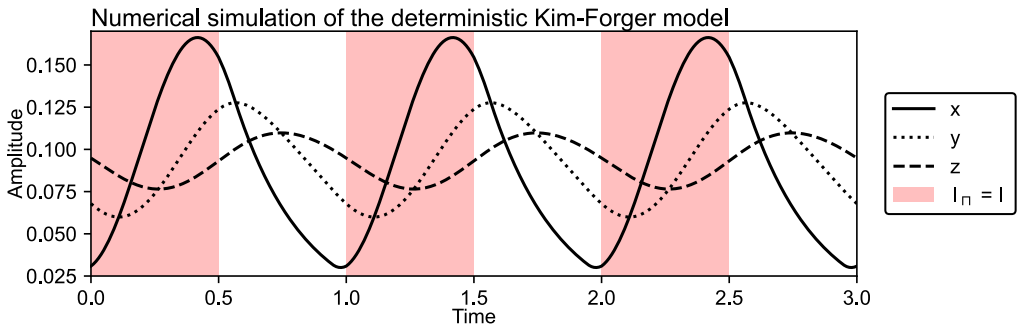
$$\dot{y} = x - y \tag{2.1b}$$

$$\dot{z} = y - z \tag{2.1c}$$

$$f(z, A) = \begin{cases} 1 - \frac{z}{A} & \frac{z}{A} \leq 1 \\ 0 & \frac{z}{A} > 1 \end{cases} \tag{2.1d}$$



**Figure 2.1:** Bifurcation diagram for parameter  $A$  of the Kim-Forger model (Equation 2.1). This figure is adapted from the previous publication [1].

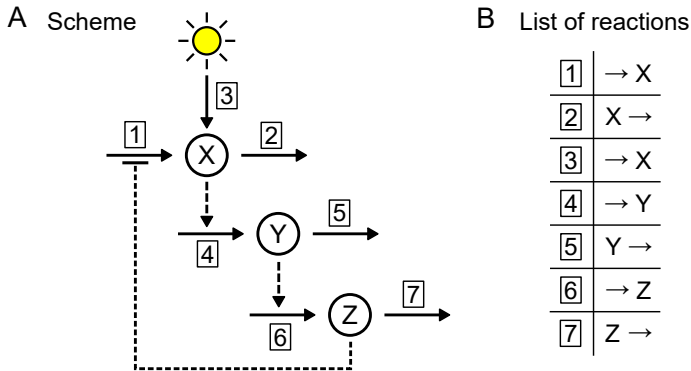


**Figure 2.2:** Numerical simulation of the deterministic Kim-Forger model (Equation 2.1) for parameters  $A = 0.1$  and  $I = 0.01$  and period of the input signal  $T = 1$  (Equation 2.2).

where  $x$ ,  $y$ ,  $z$  represent concentrations of molecular entities that form a negative feedback loop (see also Section 1.2.2).  $A$  is the only free parameter that is set for this chapter constant as  $A = 0.1$ , which corresponds to a limit cycle oscillator that can generate high-amplitude sustained oscillations (Figure 2.1).  $I_\Gamma$  is the input square signal, for example, day-night rhythm in the example of the circadian clock, which can be written as

$$I_\Gamma(t) = \begin{cases} I & 0 \leq (t \bmod T) < \frac{T}{2} \\ 0 & \frac{T}{2} \leq (t \bmod T) < T \end{cases}, \quad (2.2)$$

where  $T$  is the input period and  $I$  is the input amplitude (see also Section 1.2.3). An example numerical simulation of the Kim-Forger model is shown in Figure 2.2.



**Figure 2.3:** Chemical reaction network based on the Kim-Forger model (Equation 2.1). (A) Full-line arrows indicate synthesis (when pointing in) or degradation (when pointing out) of the respective molecular entity. The dashed lines indicate that the given entity facilitates the given reaction but is not itself affected by it. Reactions 1, 2, and 3 represent terms in Equation 2.1a, reactions 4 and 5 represent terms in Equation 2.1b, and reactions 6 and 7 represent terms in Equation 2.1c. (B) List of reactions that represents the scheme in panel A and provides the basis for the discrete computational model (Table 2.1).

## 2.1.2 Stochastic simulations of biological oscillators

### Biological oscillator as a network of chemical reactions

The molecular noise affecting biological oscillators stems from the low absolute number of reacting molecules [15]. As the molecules are subject to thermal noise, they move erratically around and collide at random times [104]. In consequence, the molecular interactions occur at stochastically distributed times, resulting in stochastic variations in the oscillatory dynamics. To model this kind of stochasticity, the deterministic model in Equation 2.1 needs to be decomposed into the individual chemical reactions (Figure 2.3). Particularly, each state variable  $x$ ,  $y$ , and  $z$  represents the concentration of one molecular entity and each equation describes synthesis and degradation reactions by positive and negative terms, respectively. Following this logic, Equation 2.1 can be decomposed in 7 individual reactions, where each reaction represents synthesis or removal of a specific molecule (Figure 2.3). The substantial difference of the chemical reaction model from the ODE model is that the chemical reaction models represents exact number of molecules ( $X$ ,  $Y$ ,  $Z$ ), which relate to the concentrations ( $x$ ,  $y$ ,  $z$ ) as  $X = \Omega x$ ,  $Y = \Omega y$ ,  $Z = \Omega z$ , where  $\Omega$  is the system size parameter that determines the overall number of molecules in the system [75, 76, 77]. To finalize the discrete model, we use the parameter  $\Omega$  to also scale the model parameters and prepare a list of transitions that describe probabilities with which the individual reactions occur (Table 2.1).

Several algorithms to perform the numerical simulation of the stochastic models exist. However, it is not clear, which algorithm provides the best accuracy and reasonable computation times,

Transition number	Transition	Transition rate
$k = 1$	$X \rightarrow X + 1$	$tr_1 = \Omega \text{kfr}(Z, \Omega A)$
$k = 2$	$X \rightarrow X - 1$	$tr_2 = X$
$k = 3$	$X \rightarrow X + 1$	$tr_3 = \Omega I_{\Gamma}$
$k = 4$	$Y \rightarrow Y + 1$	$tr_4 = X$
$k = 5$	$Z \rightarrow Z - 1$	$tr_5 = Y$
$k = 6$	$Z \rightarrow Z + 1$	$tr_6 = Y$
$k = 7$	$Z \rightarrow Z - 1$	$tr_7 = Z$

**Table 2.1:** Table of transitions based on the Kim-Forger model (Equation 2.1) and its reaction model (Figure 2.3).

especially in the context of this thesis, where a larger range of noise intensities should be explored. This section compares two common methods for stochastic simulations of biological systems, specifically Gillespie's Stochastic Simulation Algorithm (SSA) and Chemical Langevin Equation (CLE). In the end, a combined approach is proposed, which uses SSA for high and CLE for low noise intensities. This approach leads to optimal results in terms of accuracy and computational efficiency of the stochastic simulations.

### Gillespie's Stochastic Simulation Algorithm (SSA)

Gillespie's Stochastic Simulation Algorithm (SSA) is a discrete Monte Carlo method that models molecular interactions as discontinuous jumps among individual states [72]. Meaning, SSA estimates discrete times at which an integer is added or subtracted from a state variable to model synthesis or degradation of the given molecular entity, respectively (Figure 2.4A). Given the system of reactions in Table 2.1, the SSA runs as [104]:

1. Initialize the state vector  $\mathbf{X}(t = 0) = [X_0, Y_0, Z_0]$ .
2. Evaluate the individual transition rates  $tr_k(\mathbf{X})$  according to Table 2.1.
3. Calculate time for next reaction as

$$\Delta t = -\frac{1}{\sum_{k=1}^K tr_k(\mathbf{X})} \ln r_1, \quad (2.3)$$

where  $r_1$  is a random number drawn from a uniform random interval  $[0, 1]$ , and  $K = 7$  is the number of reactions on the model.

4. Select which reaction will occur next by rolling a biased dice. That is, find reaction  $j$  for which holds that

$$\frac{\sum_{k=1}^{j-1} tr_k(\mathbf{X})}{\sum_{k=1}^K tr_k(\mathbf{X})} \leq r_2 < \frac{\sum_{k=1}^j tr_k(\mathbf{X})}{\sum_{k=1}^K tr_k(\mathbf{X})}, \quad (2.4)$$

where  $r_2$  is a random number drawn from a uniform random interval  $[0, 1]$

5. Update the state vector  $\mathbf{X}$  according to the transition for transition number  $k = j$  from Table 2.1. For the Kim-Forger model that means to add or subtract 1 from one of the state variables  $X, Y, Z$ .
6. Update time by  $t \rightarrow t + \Delta t$ .
7. Repeat steps 2  $\rightarrow$  6 until desired time  $t$  is reached.

The main advantage of SSA is that it produces a statistically accurate realization of a reaction model. The disadvantage is high computational costs that depend strongly on the system size  $\Omega$  (Figure 2.5A). This drawback is also obvious from the algorithm described above as higher  $\Omega$  leads to shorter time steps ( $\Delta t$ ) in step 3.

### Chemical Langevin Equation (CLE)

The SSA produces a statistically correct realization of the given reaction model but becomes computationally expensive for increasing system size  $\Omega$ . However, for high  $\Omega$ , the absolute numbers of molecules become larger and relative changes between the individual states smaller. The discrete states can be thus approximated by a continuous stochastic differential model known as Chemical Langevin Equation (CLE) (Figure 2.4) [73]. The CLE approximation of the reaction model from Table 2.1 is

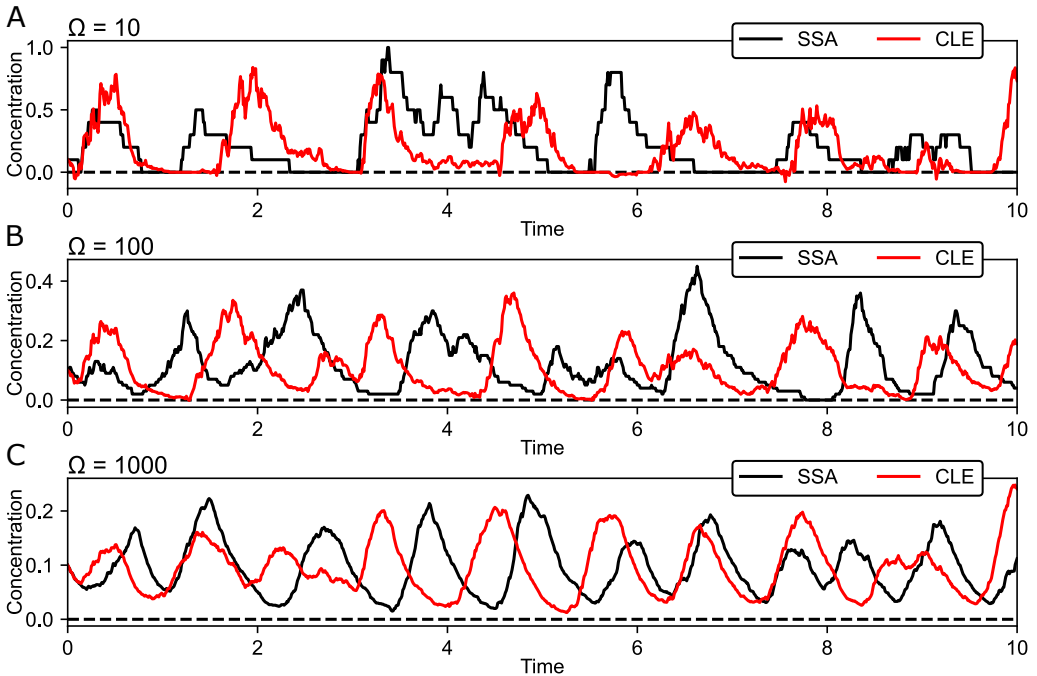
$$\dot{x} = \text{kfr}(z, A) - x + I_{\Gamma} + \sigma \left( \sqrt{[\text{kfr}(z, A)]} W_1 + \sqrt{[x]} W_2 + \sqrt{[I_{\Gamma}]} W_3 \right) \quad (2.5a)$$

$$\dot{y} = x - y + \sigma \left( \sqrt{[x]} W_4 + \sqrt{[y]} W_5 \right) \quad (2.5b)$$

$$\dot{z} = y - z + \sigma \left( \sqrt{[y]} W_6 + \sqrt{[z]} W_7 \right) \quad (2.5c)$$

where  $[a] = \max(0, a)$  prevents the terms under square roots to become negative (discussed below),  $\sigma$  is noise intensity that depends on the system size  $\Omega$  as

$$\sigma = \frac{1}{\sqrt{\Omega}}, \quad (2.6)$$



**Figure 2.4:** Comparison of simulated time series with Gillespie's Stochastic Simulation Algorithm (SSA, black) and Chemical Langevin Equation (CLE, red). The dashed line indicates 0. (A) For a very low system size ( $\Omega = 10$ ) the exact numbers of estimated molecules are important. This is visible by a step-like time series of the SSA. CLE gives an inaccurate approximation, especially around 0, where it often estimates negative concentrations. (B) Increased system size ( $\Omega = 100$ ) increases also the number of molecules in the system. The step-like behavior is thus becoming less prominent and the approximation by CLE becoming more accurate. (C) For a higher system size ( $\Omega = 1000$ ) give both methods comparable solutions.

and  $W_i$  are independent Wiener processes for which apply that

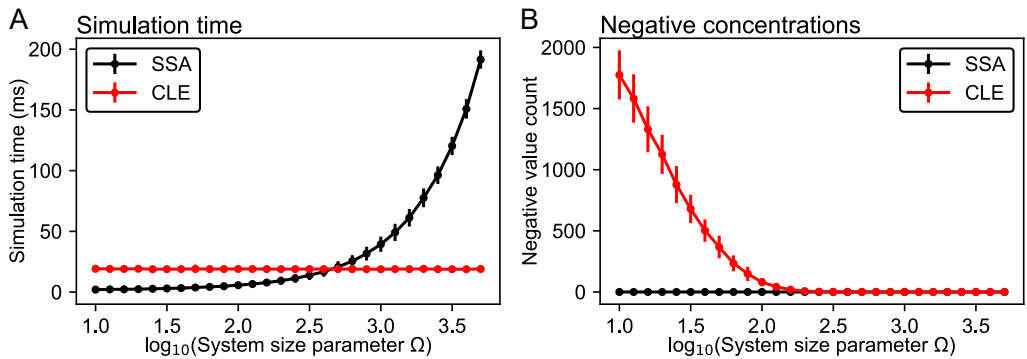
$$W(t = 0) = 0 \quad (2.7)$$

$$W(t + \Delta t) = W(t) + \sqrt{\Delta t}N(0, 1), \quad (2.8)$$

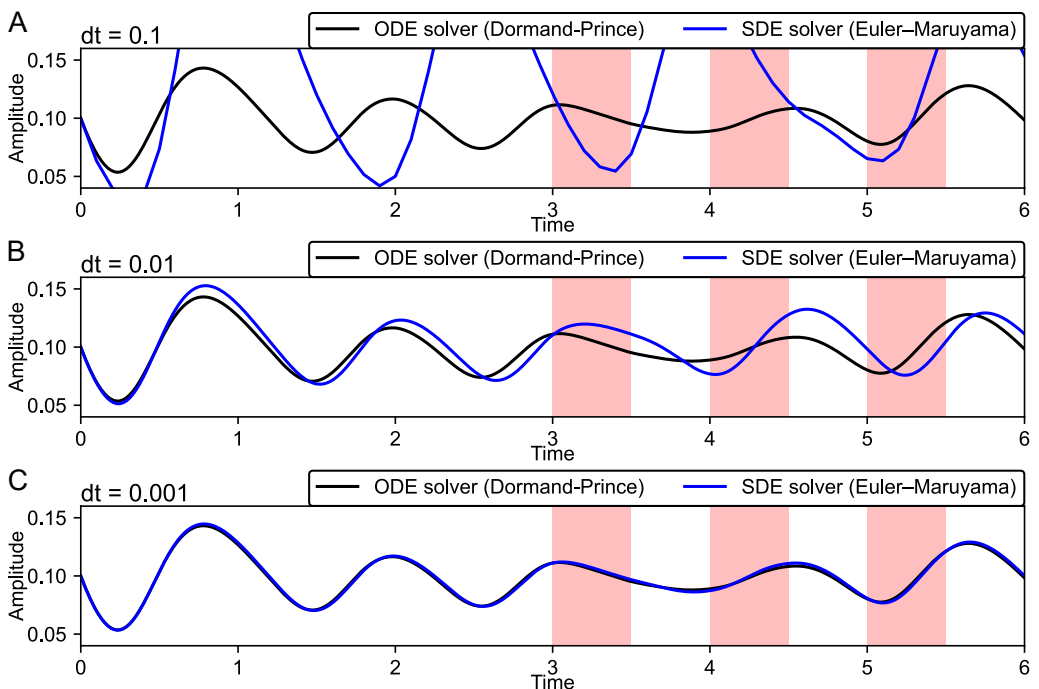
where  $N(0, 1)$  is a normally distributed random variable with zero mean and unit variance and  $\Delta t$  is a time increment [74]. The CLE model was simulated using the Euler-Maruyama method with an integration step  $\Delta t = 0.001$ . The integration step is sufficiently low to give comparable results with an accurate deterministic adaptive method for  $\sigma = 0$  (Figure 2.6) as well as with the accurate SSA method for higher values of  $\sigma$  (Figure 2.7).

One common problem associated with the CLE is that it breaks down if the terms under the square roots reach negative values [105]. The negative concentrations are not biologically possible and thus this should mathematically never happen. However, in numerical practice, the differential

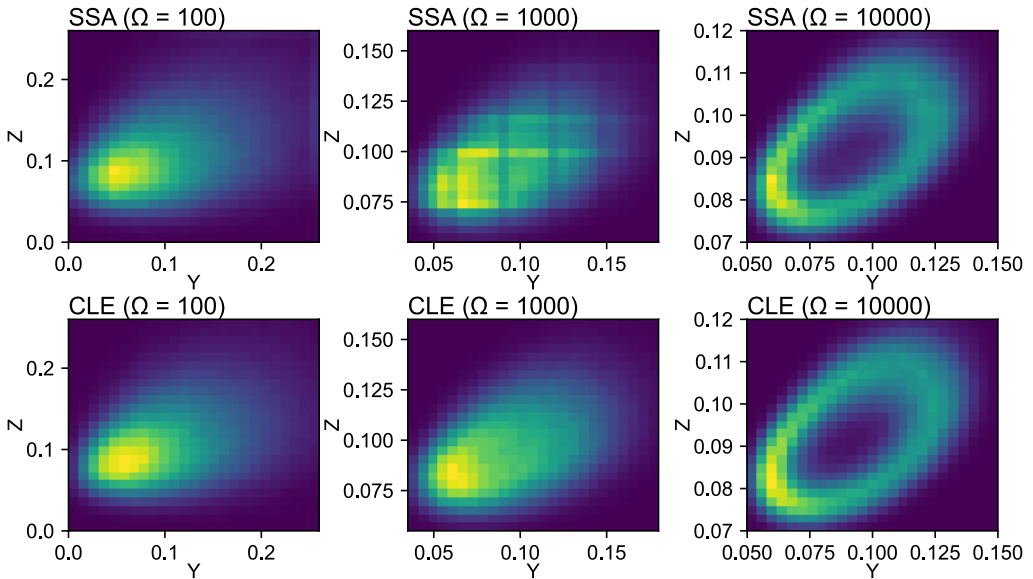




**Figure 2.5:** Simulation times and accuracy for Gillespie's Stochastic Simulation Algorithm (SSA) and Chemical Langevin Equation (CLE). The dots and error bars represent the mean and standard deviation over 10 repeated simulations. The methods were benchmarked on a standard office computer with Intel Core i7-9700 CPU. (A) The simulation time of SSA increases rapidly with the system size parameter  $\Omega$ . In contrast, the simulation time of the CLE is constant and depends on the chosen integration step of the Euler–Maruyama method (here  $dt = 0.001$ ). (B) The SSA provides a precise realization for the reaction model. In contrast, the CLE loses its accuracy with decreasing  $\Omega$ . Here, the losing accuracy is illustrated by the number of estimated negative concentrations, which is biologically impossible and is the result of the inaccurate approximation.

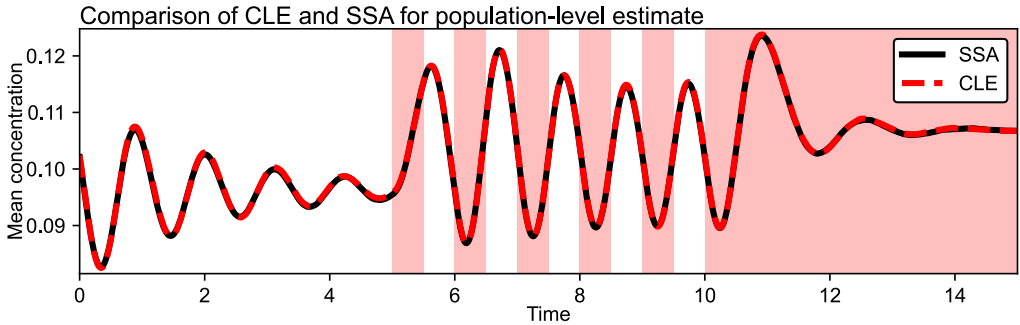


**Figure 2.6:** The chosen integration step for the SDE solver gives for  $\sigma = 0$  a comparable estimation as an adaptive ODE solver. (A) The integration step  $dt = 0.1$  is too large to give a usable approximation. (B) Decreasing the integration step to  $dt = 0.01$  improves the estimation markedly. (C) The integration step  $dt = 0.001$  gives almost the same result as an adaptive ODE solver.



**Figure 2.7:** Phase-plane histograms for different system sizes ( $\Omega$ ) estimated with SSA and CLE. Bright yellow and dark blue indicate that the states were visited with high and low frequency, respectively.

equations solver might step in the negative domain, if the solution lives close to zero due to numerical inaccuracies and extrapolations [106]. Additionally, the CLE becomes increasingly less accurate for a decreasing system size  $\Omega$ , which leads to an increased probability of stepping in the negative plane. This problem has no rigorous solution and thus CLE should not be used for a small system size  $\Omega$  and its accuracy must be always proved. For our application, this problem is circumvented by using CLE only for high values of the system size  $\Omega$  and for lower values switching to SSA. In particular, for the here-presented Kim-Forgner model, this breaking down of the CLE occurs around  $\Omega = 1000$ . Accidentally, SSA also becomes numerically manageable around this point (Figure 2.5). Thus, SSA is used as a fallback for  $\Omega < 1000$ , and CLE is used otherwise. Taking together that for  $\Omega \approx 1000$  both methods give the same results (Figure 2.7), this approach leads to accurate as well as fast numerical simulations for a wide range of  $\Omega$ .



**Figure 2.8:** SSA and CLE give an equal estimate of the population-level mean under a complicated input signal. Here, the noise intensity was set to  $\sigma = 0.03$  ( $\Omega \approx 1111$ ) and population size to  $n = 30000$ .

### 2.1.3 Population-level output

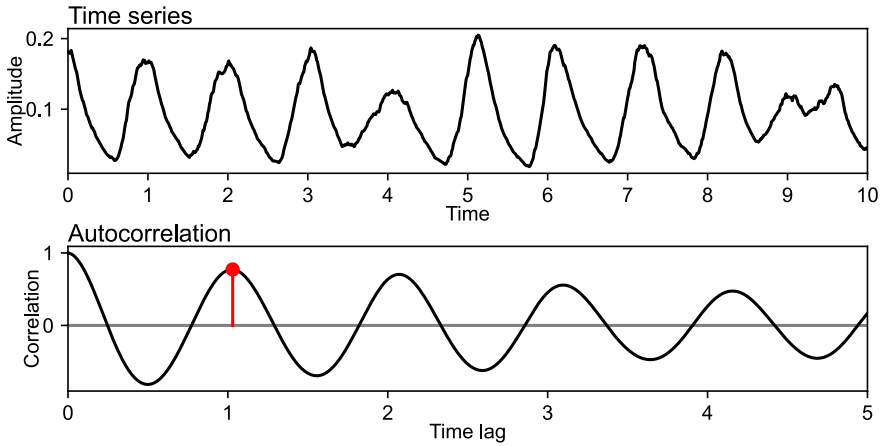
A population of uncoupled oscillators is simulated by calculating a mean of repeated independent numerical simulations as

$$x = \frac{1}{n} \sum_{i=1}^n x_i, \quad (2.9)$$

where  $n$  is the number of oscillators in the population. Also for a population of oscillators with a non-zero input signal and noise intensity close to the breaking point of the CLE method ( $\sigma = 0.03$ ,  $\Omega \approx 1111$ ), the population-level mean estimated by CLE and SSA is the same (Figure 2.8), further validating the modeling approach outlined above.

## 2.2 Metrics for entrainment quantification

Entrainment is a phenomenon when an oscillator adjusts its period to match the period of an input periodic signal (Section 1.2.3). Although this definition seems quite straightforward, in the context of stochastic oscillators it might be not always obvious what is a "matching" period. The stochasticity blurs the border between entrained and not entrained and makes it less obvious whether the observed trajectory is entrained or not. Thus, a continuous metric that would quantify the quality of entrainment rather than giving a hard threshold would be of particular interest. Here, four such continuous metrics are proposed, two based on the period estimation (autocorrelation, winding number) and two based on the phase estimation (peak detection, circular cross-correlation). Even though this list of metrics is not necessarily exhaustive, it provides a



**Figure 2.9:** Period estimation with autocorrelation function. The marked peak is the first peak of the autocorrelation function, and its location ( $\approx 1$ ) defines the period of the signal.

good overview of possible quantification approaches and their limitations. After the presentation of the methods, those are compared for the construction of Arnold tongues.

## 2.2.1 Autocorrelation

The period of oscillations can be identified as the dominant lag of its autocorrelation function [107]. The sample autocorrelation of a signal  $x$  is defined as

$$R_{xx}(k) = \sum_{i=1}^{N-k} z_x[i+k]z_x[i], \quad (2.10)$$

where  $z_x$  is Z-score defined as

$$z_x = \frac{x - \bar{x}}{S_x}, \quad \bar{x} = \frac{1}{N} \sum_{i=1}^N x[i], \quad S_x = \sqrt{\frac{1}{1-N} \sum_{i=1}^N (x[i] - \bar{x})^2}, \quad (2.11)$$

where  $\bar{x}$  is mean and  $S_x$  is standard deviation. Autocorrelation defined according to Equation 2.10 is a function that starts in  $R_{xx}(k=0) = 1$ , and is bounded to interval  $(-1, 1)$  for all  $k > 0$ . The period of the signal is estimated as the lag  $k$  of the first peak (Figure 2.9), which is the argument of the first local maximum after the time lag 0. The estimated period can be compared with the known period of the input signal to determine the degree of entrainment.

## 2.2.2 Winding number

In geometry, the winding number  $w$  is the number of revolutions made around a point  $P$  while traveling along a closed curve  $C$ . If we assume  $P = [0, 0]$  and a closed discrete trajectory  $C = [C_1, C_2, \dots, C_n, C_{n+1} = C_1]$  then the winding number is defined as [108]

$$w = \frac{1}{2\pi} \sum_{i=1}^n \varphi_i \quad (2.12)$$

where  $\varphi_i$  is the signed angle between the edges  $\overline{PC_i}$  and  $\overline{PC_{i+1}}$ . The angle  $\varphi_i$  can be calculated for two neighboring points  $C_i = [x_i, y_i]$  and  $C_{i+1} = [x_{i+1}, y_{i+1}]$  using the two-parameter arctangent function as

$$\varphi_i = \text{atan2}(x_i * y_{i+1} - y_i * x_{i+1}, x_i * x_{i+1} + y_i * y_{i+1}), \quad (2.13)$$

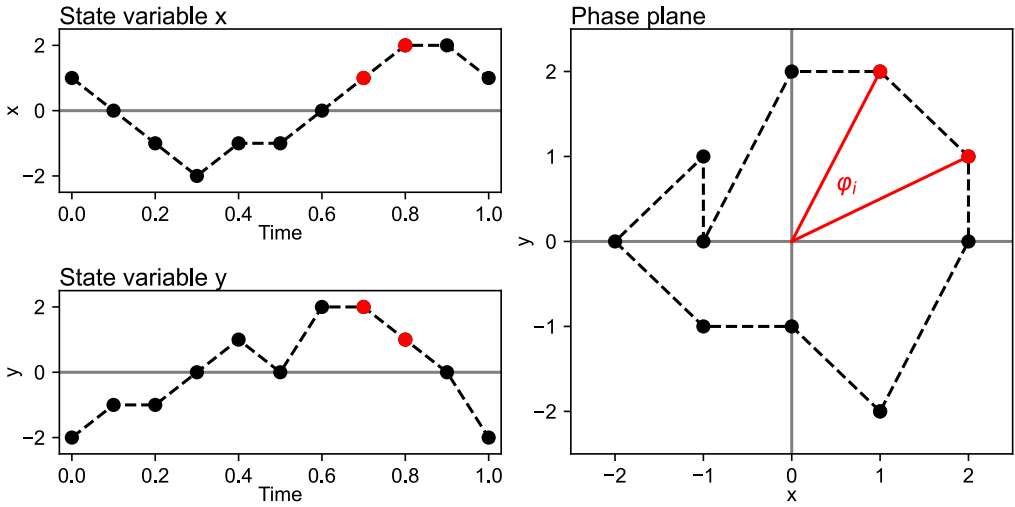
where  $\text{atan2}$  is defined as

$$\text{atan2}(x, y) = \begin{cases} \text{atan}\left(\frac{y}{x}\right) & x > 0 \\ \text{atan}\left(\frac{y}{x}\right) + \pi & x < 0, y \geq 0 \\ \text{atan}\left(\frac{y}{x}\right) - \pi & x < 0, y < 0 \\ \frac{\pi}{2} & x = 0, y > 0 \\ -\frac{\pi}{2} & x = 0, y < 0 \\ \pi & x = 0, y = 0 \end{cases}. \quad (2.14)$$

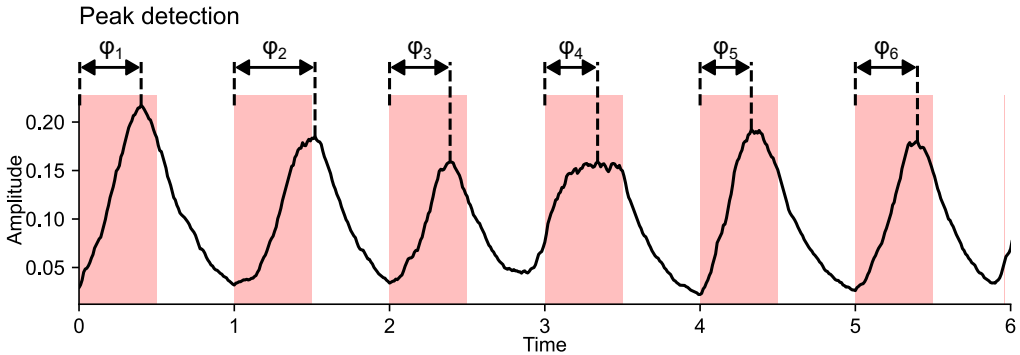
For a numerical simulation of a differential equation model, we can estimate the winding number from the phase plane of two state variables  $x = [x_1, x_2, \dots, x_n]$  and  $y = [y_1, y_2, \dots, y_n]$  (Figure 2.10). As  $w$  in Equation 2.12 is calculated as the number of revolutions around the point  $[0, 0]$ , the mean is subtracted from the  $x$  and  $y$  to ensure that the phase plane revolves around the origin. It is also important to mention that the two assumptions made above about the geometric winding number will not apply to the numerical estimation of the period. First,  $w$  estimated from data does not necessarily need to be an integer as the phase plane might not form a closed trajectory. Second, the period does not have a sign, therefore, the absolute value  $|w|$  is used. The period is then calculated as

$$\tau = \frac{t_{\text{end}}}{|w|}, \quad (2.15)$$

where  $t_{\text{end}}$  is the duration of the simulation.



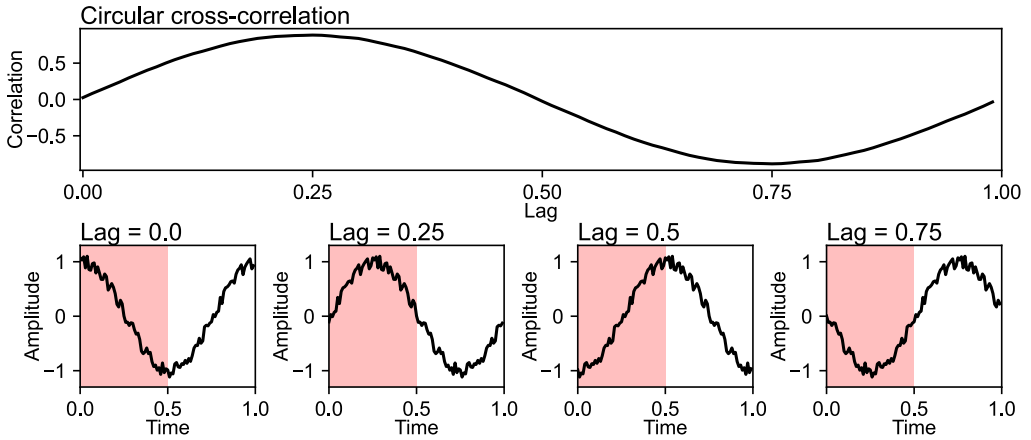
**Figure 2.10:** Two state variables ( $x, y$ ) can be used to draw a phase plane. The phase plane is used to calculate the winding number as the sum of the angles of the individual data points ( $\varphi_i$ ). In this specific case, the winding number would be exactly  $w = 1$  as the phase plane shows exactly one revolution around the origin. Considering that the maximal recorded time in the figure is  $t_{\text{end}} = 1$ , we get an estimated period  $\tau = 1$ , according to Equation 2.15.



**Figure 2.11:** The entrainment phase ( $\varphi_i$ ) can be quantified as the distance from the beginning of a cycle to the highest peak within the cycle. For a stochastic oscillator, the phase can vary substantially between the cycles.

### 2.2.3 Peak detection

The phase of entrainment can be defined as the relative timing of an event within the period of the input signal [109]. This "event" can have various definitions and depend on the studied system, but one of the most straightforward definitions is to consider the peak of the output signal. The phase is then defined as the distance of the highest peak within an input cycle from the beginning of the input cycle (Figure 2.11). In the context of this thesis, the input signal switch from 0 to 1 is considered the beginning of a cycle.



**Figure 2.12:** Circular cross-correlation for a noisy cosine function and square input signal indicated by the red boxes. The circular cross-correlation has a maximum at time 0.25, which would be the estimated entrainment phase.

## 2.2.4 Circular cross-correlation

As an alternative to peak detection, the phase of entrainment can be estimated using circular cross-correlation. The circular cross-correlation is defined for two periodic signals  $x = [x_0, x_1, x_2, \dots, x_N = x_0]$  and  $y = [y_0, y_1, y_2, \dots, y_N = y_0]$  as

$$R_{xy}^c(k) = \frac{1}{N} \sum_{i=0}^{N-1} z_x[(i+k) \bmod N] z_y[i], \quad (2.16)$$

where  $N$  is the number of samples in a cycle, and  $z_x$  and  $z_y$  are z-scores of  $x$  and  $y$  according to Equation 2.11, respectively. The phase of entrainment is defined as the lag  $k$ , for which is  $R_{xy}^c(k)$  maximal (Figure 2.12).

## 2.2.5 Phase coherence

The output of the period-based methods, here autocorrelation and winding number, return the period of the analyzed signal that can be directly compared with the period of the input signal. The phase-based methods, here peak detection and circular cross-correlation, however, estimate the phase for each cycle of the input signal separately (Figure 2.11). A summarizing technique is thus needed that quantifies the stability of the phase estimated from several consequent cycles. In

the literature, the problem of phase variability is often studied in the context of a population of oscillators using the complex-valued order parameter, which is defined as [110, 111]

$$r \cdot e^{i\psi} = \frac{1}{n} \sum_{j=1}^n e^{i\phi_j} \quad (2.17)$$

where  $r$  is the degree of coherence,  $i$  is the imaginary number,  $\psi$  is the collective phase,  $\phi_j$  is the phase of the individual oscillators in the population, and  $n$  is the number of oscillators in the population. The degree of coherence  $r$  is a number between 0 and 1, where  $r = 1$  indicates that all oscillators are in the same phase and  $r = 0$  indicates complete incoherence when the individual oscillators average out at the population-level mean. Based on the order parameter, two metrics of entrainment are derived to specifically target the problem of phase stability: phase coherence, quantifying period-period phase variability, and population phase coherence, quantifying phase variability among the individual oscillators in the population.

Phase coherence (PC) quantifies the period-period variability of the estimated phases as

$$PC = \left| \frac{1}{N} \sum_{k=1}^N e^{i\varphi_k} \right| \quad (2.18)$$

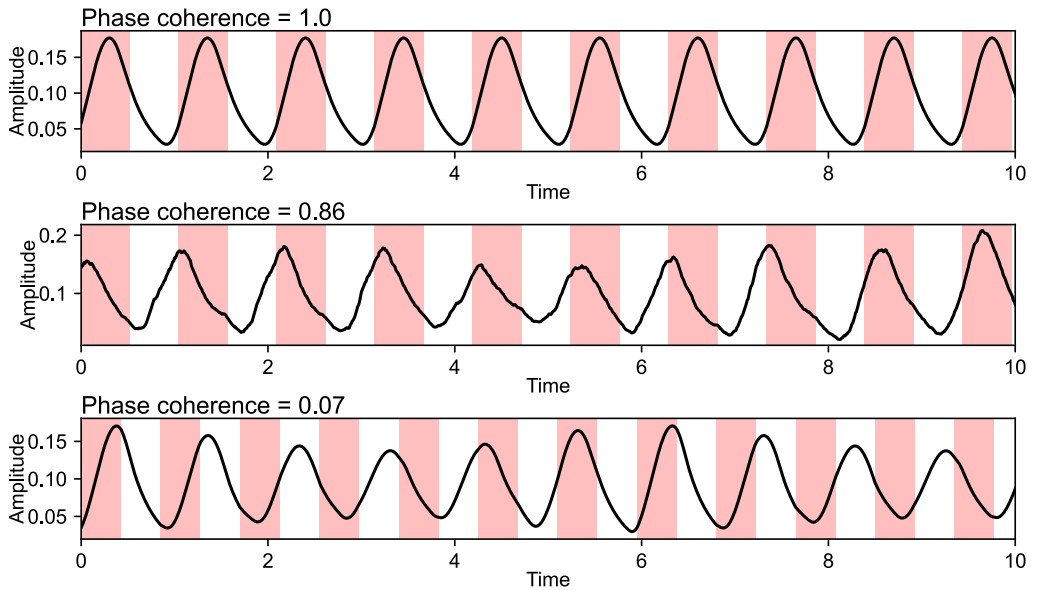
where  $N$  is the number of periods of the input signal, and  $\varphi_k$  is the estimated phase in  $k$ -th cycle of the input signal. The phase coherence ranges from 0 to 1. A value of 1 means that the phase of the population-averaged signal occurs exactly at the same time in every cycle. Values close to 0 indicate an unentrained signal with peaks occurring randomly within the cycle (Figure 2.13).

Population phase coherence (PPC) quantifies the desynchronization of the individual oscillators in the population and is equivalent to  $r$  from Equation 2.17 as

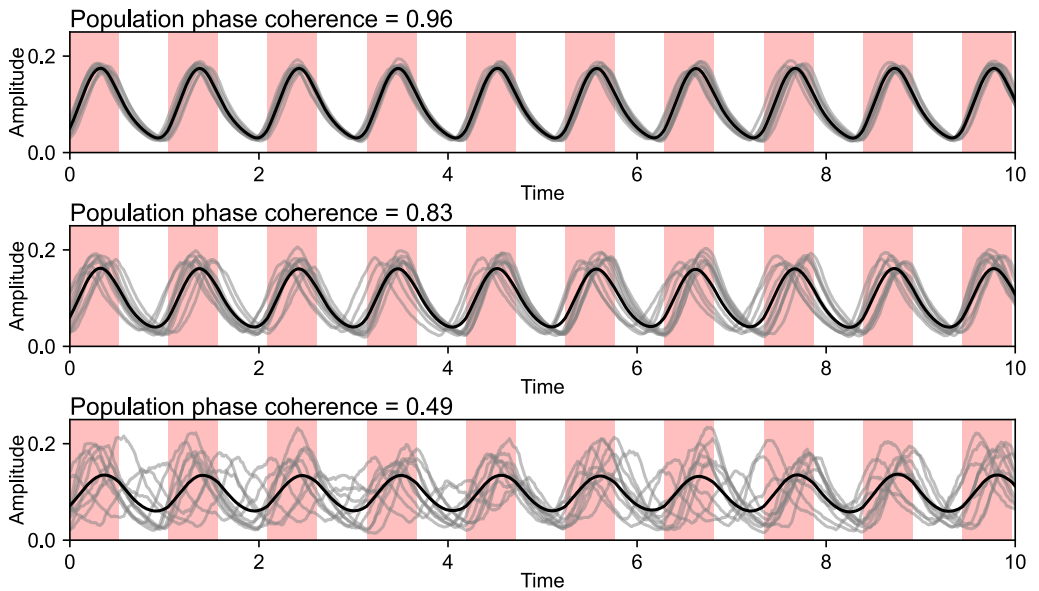
$$PPC = \left| \frac{1}{n} \sum_{j=1}^n e^{i\phi_j} \right| \quad (2.19)$$

where  $n$  is the number of oscillators in the population, and  $\phi_j$  is phase for  $j$ -th oscillator in the population. The precision of this metric can be increased by calculating the circular mean of the population phase coherence over several periods of the input signal. The final value ranges from 0 to 1. A value of 1 indicates that phases of all individual oscillators within an input cycle are at the same time. Values close to 0 indicate a highly disperse population where each oscillator has a different phase (Figure 2.14).

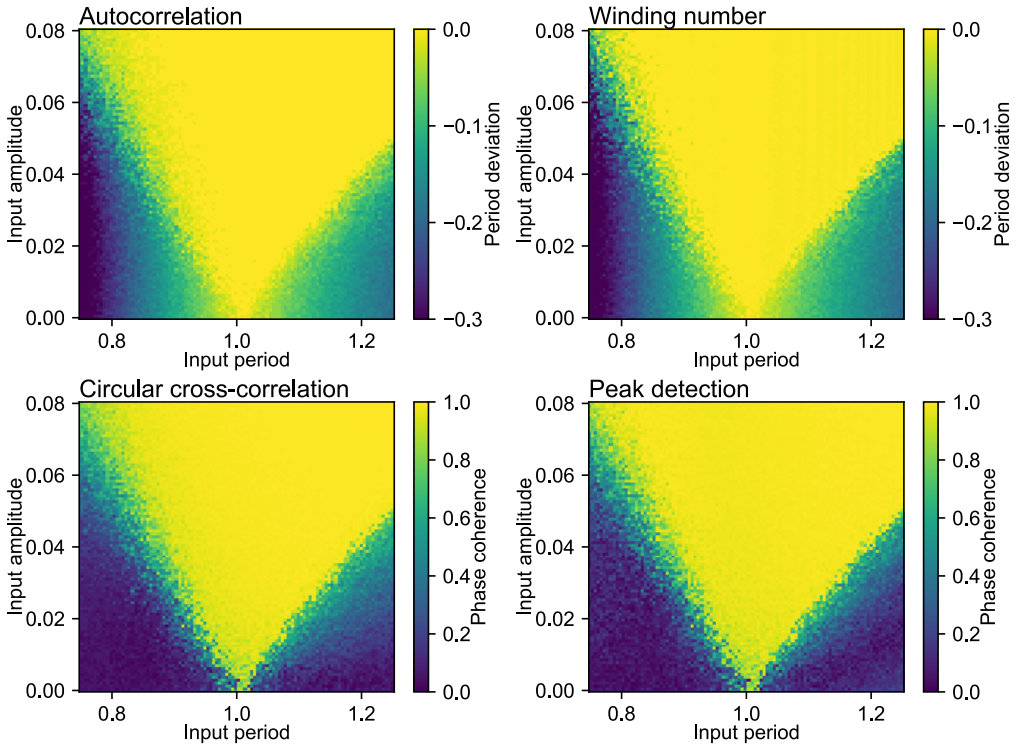




**Figure 2.13:** Phase coherence is decreasing from the top to the bottom panel. The decreasing phase coherence corresponds to increasing cycle-to-cycle variability of the peak locations within the individual cycles.



**Figure 2.14:** Population phase coherence is decreasing from the top to the bottom panel. The decreasing population phase coherence corresponds to increasing variability of the individual oscillators (gray) that form the population-level mean (black). The phase coherence of the mean trajectory is 1 in all panels, as the mean exhibit no cycle-to-cycle variations in any of the panels.



**Figure 2.15:** Arnold tongues estimated with various entrainment metrics. For the period-based metrics (autocorrelation and winding number) the color shows the deviation of the oscillator period from the period of the input signal. For the phase-based method (peak detection and circular cross-correlation) the color shows phase coherence.

## 2.2.6 Comparison of the entrainment metrics

To compare the metrics presented above, Arnold tongues for input periods between 0.75 and 1.25 and input amplitudes between 0.0 and 0.08 were constructed using the Kim-Forger model (Equation 2.5) with noise intensity  $\sigma = 0.005$ . All metrics gave comparable results with no obvious differences (Figure 2.15). Further evaluation of the entrainment metrics, this time in the context of heterogeneous populations, is provided in Section 2.6.

Another important aspect of the developed methods is their computational expense. The presented methods were benchmarked on a standard office computer with Intel Core i7-9700 CPU and a time trace with 10000 samples and 100 cycles of the input signal. The algorithms were run 10000 times and the resulting mean and standard deviation of the execution time were noted. In this test, autocorrelation was the slowest method ( $6.95 \pm 1.51$  ms) and the winding number the fastest ( $0.17 \pm 0.34$  ms). The phase estimation methods were slower than the winding number but still considerably faster than autocorrelation (circular cross-correlation,  $1.63 \pm 1.06$  ms, peak

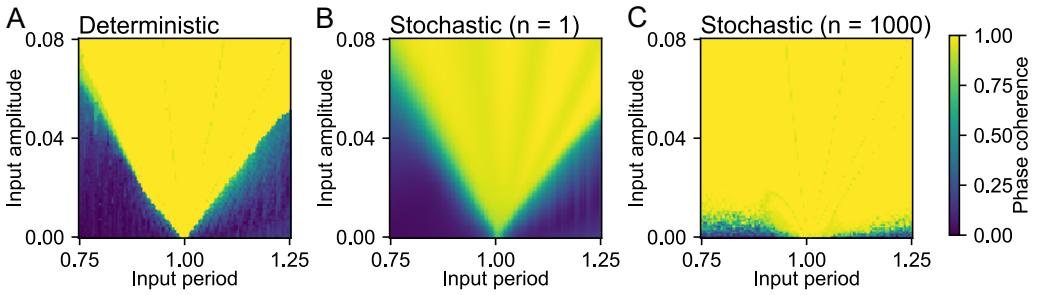
detection,  $1.05 \pm 1.46$  ms). At first sight, the computation differences might not seem significant, as all numbers are in milliseconds, but consider that for an Arnold tongue as shown in Figure 2.15 one already needs  $100 \times 100$  evaluations. Adding that a population of up to 1000 oscillators should be explored in the next section, this already adds up to  $10^7$  evaluations of the metric for the estimation of a single Arnold tongue.

For the next sections, the phase coherence estimated with peak detection will be used as the metric of choice. The advantages of peak detection are its reasonable computational time and straightforward interpretation (distance of peak from the beginning of the cycle). The resulting phase coherence also allows for easy quantification and comparison of multiple Arnold tongues as we can calculate the average phase coherence of the tongue, which will always be a number between 0 and 1 indicating the entrainment range of the population.

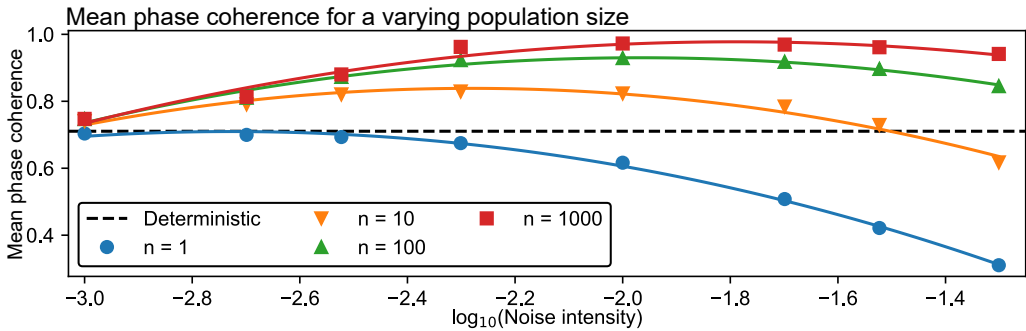
## 2.3 Noise widens range of entrainment

Arnold tongues visualize how entrainment depends on the amplitude and period of the input signal [60, 61]. In general, a greater amplitude of the input signal increases the range of periods for which the system is entrained, resulting in the typical, tongue-shaped regions of entrainment (Figure 2.16). It has been previously shown that if a population mean is considered, the Arnold tongue is wider in comparison to the Arnold tongue estimated with a single stochastic or deterministic model [92]. This section explores how this widening of the entrainment range depends on the noise intensity ( $\sigma$ ) and the number of oscillators ( $n$ ) used to construct the population mean. The Arnold tongues were quantified by calculating the mean phase coherence of the tongue (Figure 2.16). Applying this metric to the output of a deterministic model ( $\sigma = 0$ ) and a stochastic model ( $\sigma = 0.005$ ) showed that the range of entrainment for the deterministic system is lower than for a stochastic population of 1000 oscillators (Figure 2.16A, C). This is in line with previously published findings [92]. However, if a single stochastic oscillator is considered, the entrainment area drops to a lower value close to the deterministic model (Figure 2.16B). This suggests that noise must be compensated for with a sufficiently large population size to allow the widening of the Arnold tongue by noise.

To extend this observation, the entrainment area for four different population sizes and seven noise intensities was measured (Figure 2.17, Figure A.1). Considering the mean phase coherence for a deterministic model as a reference, one can see that for a single oscillator ( $n = 1$ ) the range of entrainment decreases with increasing noise intensity, showing the detrimental effect of noise for a single-oscillator model described also by a previous study [80]. However, with increasing population size ( $n > 1$ ), an optimal value of noise intensity exists, for which the range of entrainment is maximal. With increasing population size, this optimal noise intensity moves



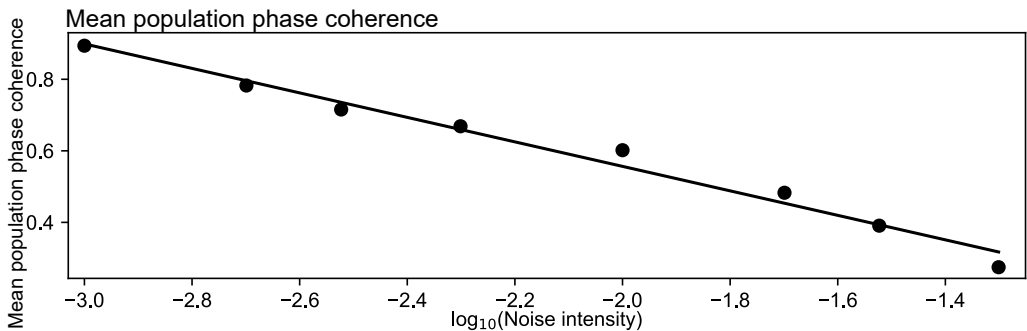
**Figure 2.16:** Arnold tongues and corresponding mean phase coherence (PC) for a varying population size  $n$ . (A) Deterministic model ( $PC = 0.72$ ). (B) Stochastic model representing one oscillator with noise intensity  $\sigma = 0.005$  ( $PC = 0.67$ ). (C) Stochastic model representing a population of stochastic oscillators with the same noise intensity as in panel B ( $PC = 0.96$ ).



**Figure 2.17:** Mean phase coherence as a function of population size ( $n$ ) and noise intensity ( $\sigma$ ). The individual Arnold tongues used to generate this image are shown in Figure A.1.

to higher values, and also the maximal range of entrainment at this optimum noise intensity is increased. The results thus suggest an optimal noise intensity for the entrainment of a population of a given size. When this optimum noise intensity is exceeded, the entrainment capacity is again compromised.

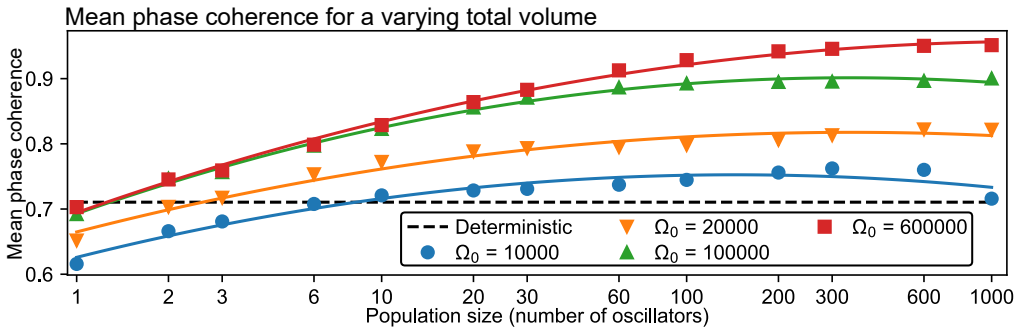
To better understand how the observed population-level effect relates to the dynamics of the individual oscillators in the population, population phase coherence was used as a metric of desynchronization among the individual oscillators within the population (Equation 2.19, Figure 2.14). Utilizing population phase coherence showed that with increasing noise intensity the population phase coherence steadily decreases, indicating progressively more desynchronization among the individual oscillators (Figure 2.18, Figure A.2). A moderate desynchronization of the individual oscillators thus supports the population-level entrainment, however, the amplitude of the population mean is progressively less prominent with increasing noise in comparison to the amplitude of the individual oscillators in the population.



**Figure 2.18:** Mean population phase coherence as a function of noise intensity ( $\sigma$ ) estimated for a population of  $n = 1000$  oscillators. The individual Arnold tongues used to generate this image are shown in Figure A.2.

To understand how a given population of cellular oscillators (cells) might be tuned towards optimal entrainment, a hypothetical case was investigated where a total volume  $\Omega_0$  of cell material is available, but the number of cells ( $n$ ) and their individual system size ( $\Omega$ ) can change. In other words, the experiment starts with one big cell ( $\Omega = \Omega_0$ ) that is subsequently divided into many small cells ( $\Omega = \Omega_0/n$ ) while the total volume  $\Omega_0$  remains constant. Here, the noise intensity ( $\sigma$ ) depends on the system size of the individual cells ( $\Omega$ ) as  $\sigma = 1/\sqrt{\Omega}$ . The simulation results suggest an optimal population size for each total volume (Figure 2.19, Figure A.3). Specifically, with increasing total volume, the maximal range of entrainment occurs at a higher population size and the average phase coherence is also larger. In other words, a higher total volume can support a higher number of noisy oscillators, which in turn allows the population to take advantage of high noise intensities. Considering those results more generally, if the goal is maximal sensitivity in entrainment to input signals, it is more advantageous to distribute the resources to several noisy units rather than maintain a single unit with minimal noise. However, when the population size is increased beyond the optimum, the individual units become too noisy and the ability of entrainment is again compromised.

The existence of an optimal noise intensity is a typical characteristic of a phenomenon known as stochastic resonance [83]. The term stochastic resonance is traditionally used in neuroscience to describe improved detection of weak signals in threshold-like systems [86]. The term stochastic resonance is also used more generally to describe improvement in output performance of a noisy system in various disciplines including cell biology, ecology, and physics [84]. In the system of uncoupled stochastic oscillators under period forcing, noise improves the population-level entrainment but only to the point where the population size is sufficiently large to compensate for the noise-induced fluctuations at the population-level read-out. Thus, in comparison to the previous studies, the section shows not only that noise widens the range of entrainment, but also



**Figure 2.19:** Mean phase coherence as a function of the total volume ( $\Omega_0$ ) and population size ( $n$ ). The individual Arnold tongues used to generate this image are shown in Figure A.3.

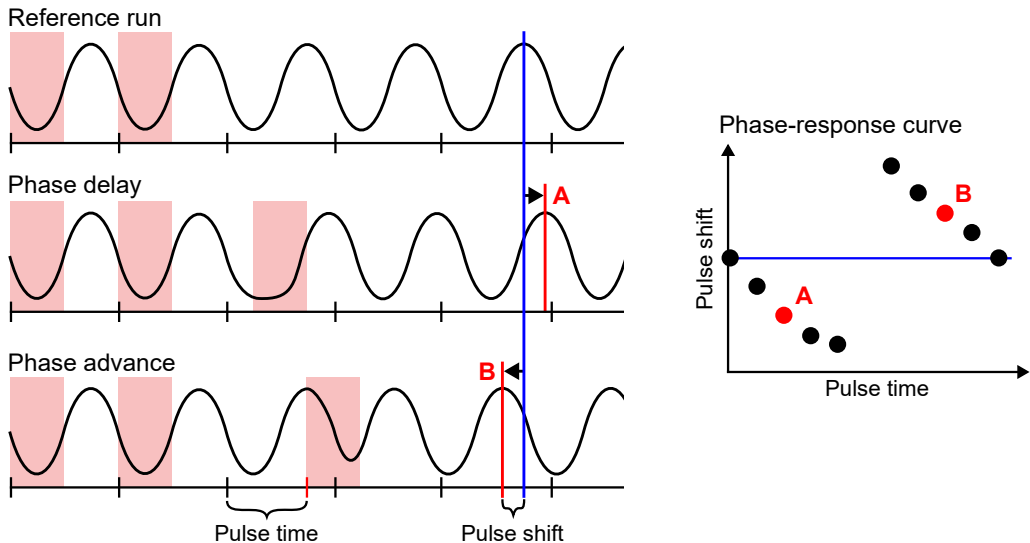
that there exists an optimal value of the noise intensity, for which the range of entrainment is the widest.

## 2.4 Noise increases phase response

### 2.4.1 Noise increases amplitude of phase response curve

The phase response curve (PRC) is a plot that displays the change in the oscillator phase caused by a step pulse as a function of the time during the oscillation cycle at which the pulse was applied (Section 1.2.3). In practice, the PRC is measured by a repeated run of the oscillator and application of a step pulse at different phases at each run [112, 64]. The algorithm starts with a reference simulation that consists of a square input signal followed by a constant zero. The initial input signal is necessary to synchronize the individual oscillators in the population and also to provide time reference. In the following simulations, the same square signal is applied at the beginning of the simulation followed by a step signal at various times of the oscillation cycle. After the pulse, the input signal returns to constant zero and the phase shift compared to the reference simulation is calculated (Figure 2.20).

The PRCs are often characterized based on their amplitude, which represents the extent of the pulse-induced phase shift, as type 1 or type 0. Type 1 PRCs exhibit relatively small phase shifts and appear continuous in the PRC plot, whereas type 0 PRCs show large phase shifts and appear visually discontinuous [112, 64]. Increasing the PRC amplitude in the consequence of increasing input amplitude is well-documented [65], however, this section shows that the PRC amplitude surprisingly increases also as a consequence of increasing noise intensity (Figure 2.21). Specifically, for low input amplitude and low noise intensity, relatively small phase shifts can

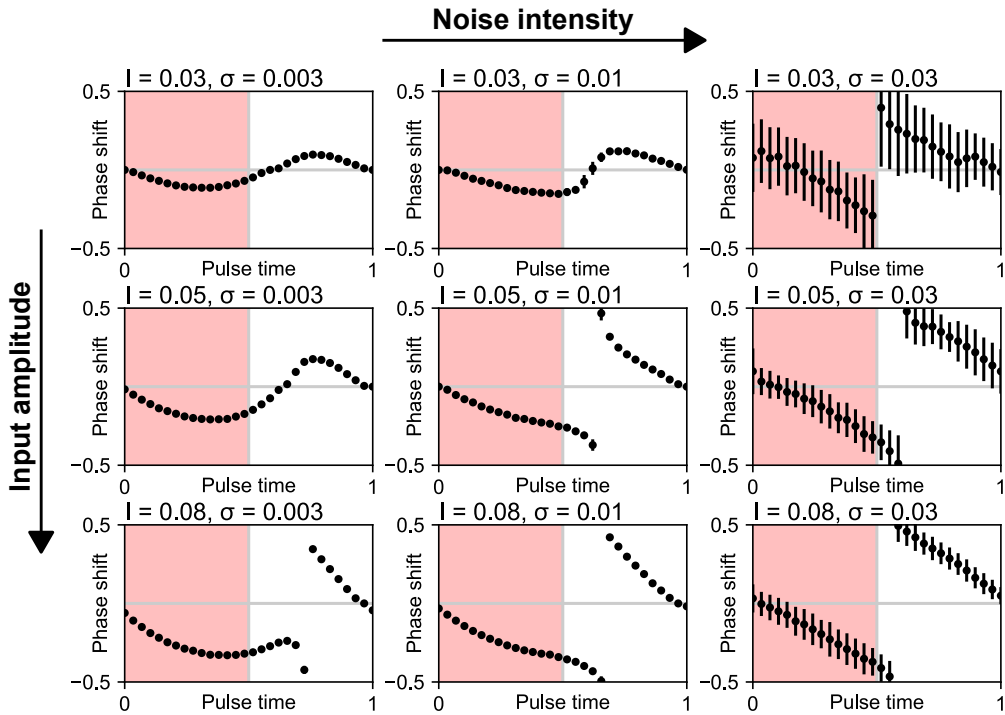


**Figure 2.20:** Estimation of the phase response curve (PRC) by comparing pulse shifts caused by varying pulse times. This example corresponds to type 0 PRC.

be observed (type 1 PRC). As expected, when the input amplitude is increased, the phase shifts also increase. Interestingly, the same effect also applies to increasing the noise intensity. For high noise intensities, the resulting PRCs have large phase shifts (type 0 PRC) regardless of the amplitude of the input signal. Accordingly, increasing noise intensity allows the transition from low-amplitude (type 1) to high-amplitude (type 0) PRC, even if the input amplitude remains low.

## 2.4.2 Noise increases slope of T-cycle phase response curve

A disadvantage of the PRC related to the populations of stochastic oscillators is that it cannot be estimated if the individual oscillators are subject to high noise intensity and thus desynchronize too fast. In such cases, the population-level mean loses its amplitude rapidly and the analysis based on the signal peaks cannot be performed. This section proposes a T-cycle phase response curve (T-PRC) as an alternative. The T-PRC is a plot that shows the phase of entrainment as a function of the input period  $T$ , thus T-PRC (Figure 2.22). The simulation from the previous section was repeated to estimate the relationship between the PRC and T-PRC (Figure 2.23). Whereas the PRC showed increasing amplitude following increasing input amplitude and noise intensity, the T-PRC exhibited decreasing slope with increasing input amplitude and noise intensity. The decreasing slope of the T-PRC potentially relates to the previous observation that noise increases the range of entrainment. The increased range of entrainment would have a smaller slope of the T-PRC as the phase is changing over a larger range of input periods before entrainment is lost.



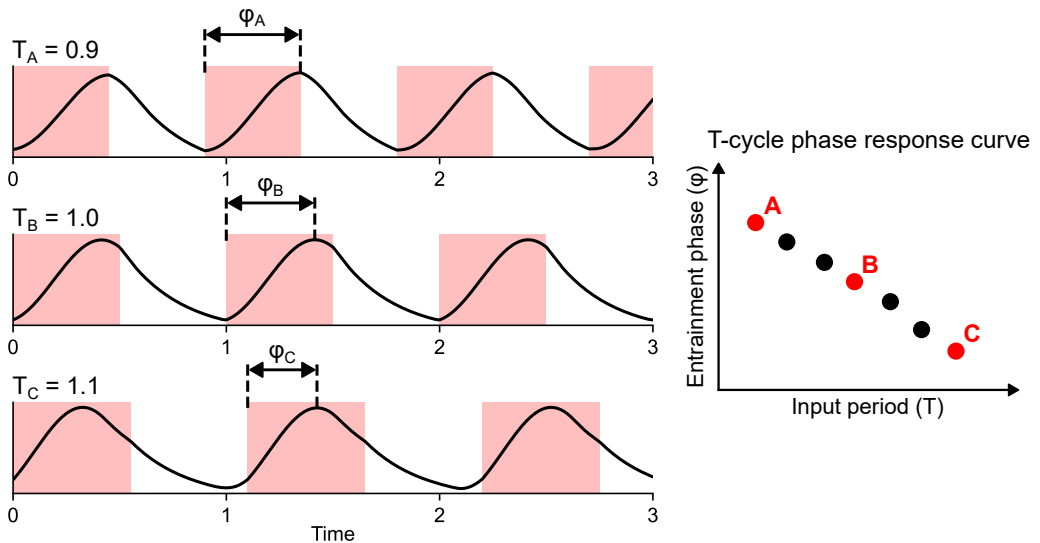
**Figure 2.21:** Phase response curves (PRCs) for a varying noise intensity ( $\sigma$ ) and input amplitude ( $I$ ) for a population-level mean of 1000 oscillators. Phase shift and pulse time are cyclic quantities normalized to the unit circle. The pulse length is 0.5, which is half of the free-running period for the deterministic model ( $\sigma = 0$ ). The dots represent the mean and error bars the standard deviation over 10 simulations.

Taken together, the T-PRC can be used as an alternative to the PRC for rapidly desynchronizing populations. Section 4.7 also shows a practical demonstration of the T-PRC in the context of cellular entrainment.

### 2.4.3 Noise allows faster recovery after jet lag

Jet lag is a famous phenomenon that occurs when traveling between time zones. This section explores the effect of a varying noise intensity on the reentrainment to a persistent phase-shift in the input signal, or, “recovery from jet lag”. In these simulations, the population is first entrained by a regular input cycle representing a day-night cycle. After the output of the population ( $n = 1000$ ) is phase-locked to the input signal, an abrupt shift in the phase of the input signal is introduced and the time, until the population output is locked to the new cycle, is measured (Figure 2.24). The results show that noise shortens this time, allowing faster recovery from jet

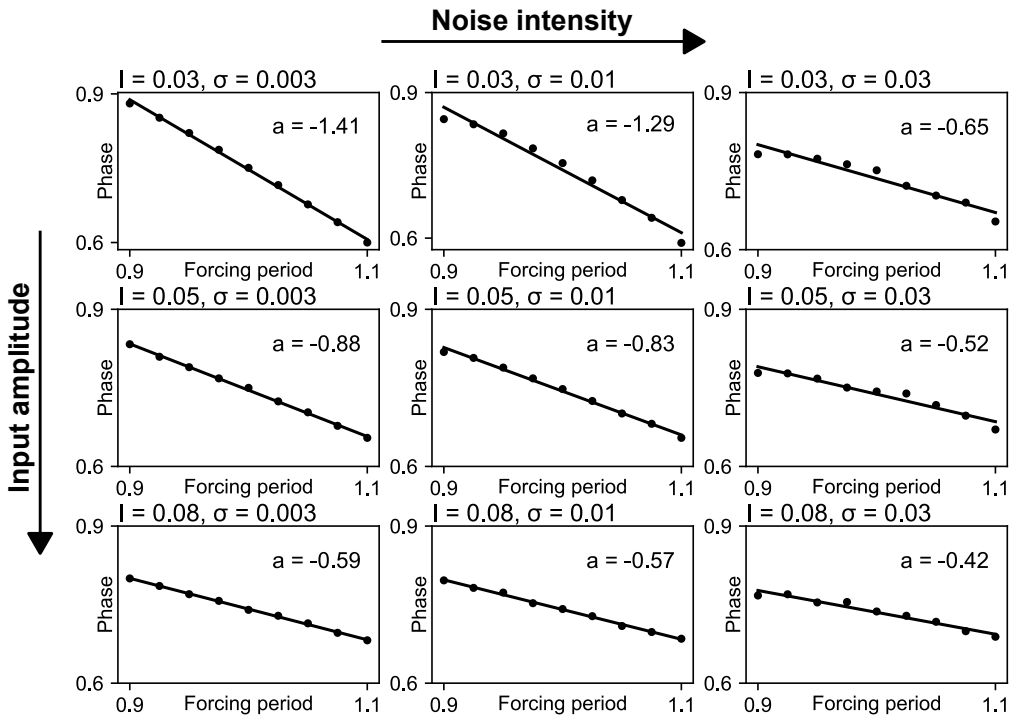




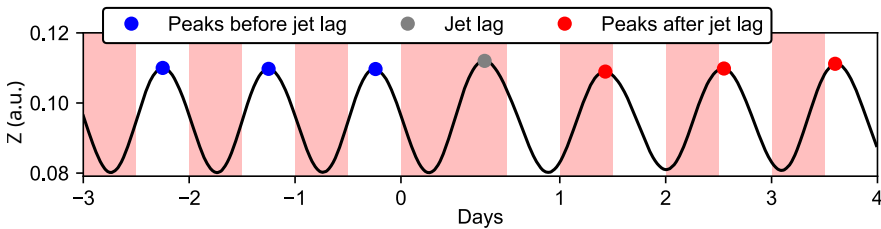
**Figure 2.22:** Estimation of the T-cycle phase response curve (T-PRC) by comparing entrainment phase ( $\varphi$ ) for a varying input period ( $T$ ).

lag. This is well visible especially when the input amplitude is low (Figure 2.25). With increasing input amplitude, reentrainment to the new cycle is fast for all noise intensities, so the effect of the increasing noise is not as apparent. Accordingly, noise allows faster recovery from jet lag, especially if the input amplitude is low.

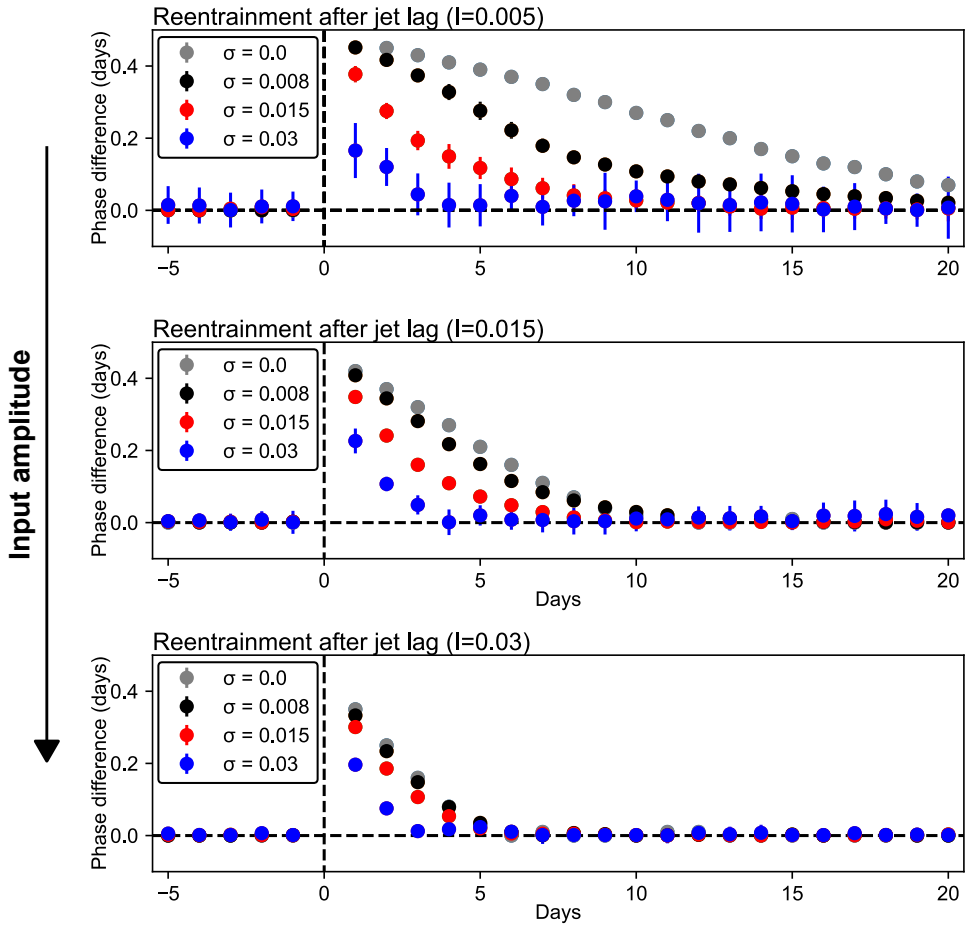
These results might be particularly interesting in the context of the entrainment dynamics of the circadian clock. In mammals, the clock is thought to be organized hierarchically, with a central pacemaker of highly coupled oscillators in the suprachiasmatic nucleus (SCN) and less coupled peripheral oscillators that are entrained by signals from the SCN [95]. Pharmacologically increased noise in the SCN cells shortens jet lag even without explicitly weakening coupling among the cells [103]. This hints at the possible relevance of the results also to the domain of coupled oscillators. The coupling also provides the SCN with a certain level of resistance to noise and external perturbation [34, 101]. The results show that for increasing noise intensities the population becomes very sensitive to the input signals. This property would not seem useful to the SCN in maintaining a steady rhythm but might be advantageous for peripheral clocks that need to adjust to a manifold of external cues.



**Figure 2.23:** T-cycle phase response curves (T-PRCs) for a varying noise intensity ( $\sigma$ ) and input amplitude ( $I$ ) for a population-level mean of 1000 oscillators. Increasing input amplitude ( $I$ ) and increasing noise intensity ( $\sigma$ ) flatten the slope ( $a$ ) of the T-cycle phase response curve (T-PRC).



**Figure 2.24:** Jet lag is simulated by keeping the input square signal in its high state for doubled time and then proceeding again with a regular rhythm.



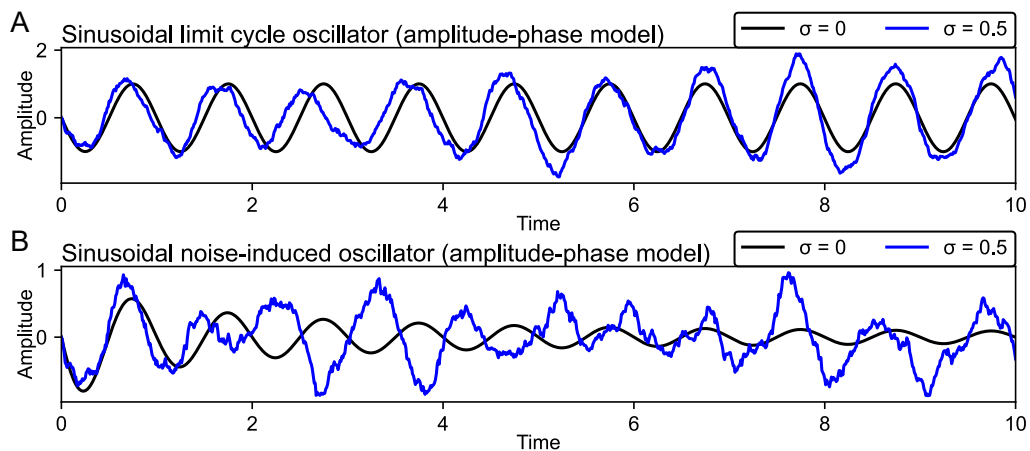
**Figure 2.25:** Increasing noise intensity ( $\sigma$ ) allows faster reentrainment after jet lag, especially for low input amplitudes ( $I$ ). The phase difference indicates the difference between the phase on a specific day and the average phase of the days before the jet lag (Figure 2.24). Dots represent the mean and error bars the standard deviation over 10 simulations for all panels.

## 2.5 Noise facilitates entrainment of generic limit cycle oscillators

This section demonstrates that the positive effects of noise are not limited to bio-inspired oscillator systems but emerge also in a broad class of limit cycle oscillators. Limit cycle oscillators can be classified as either sinusoidal oscillators, which have a harmonic limit cycle, or as relaxation oscillators, which generate non-sinusoidal oscillations (Section 1.2.1). This distinction is important in the study of entrainment as different classes of oscillators react differently to external forcing. For example, relaxation oscillators allow for faster entrainment [113]. In this section, the population-level entrainment under noise is studied for two generic limit cycle oscillator models representing both classes: the Amplitude-phase model as a model of sinusoidal oscillations and the Van der Pol model as a generic model of relaxation oscillations.

Another way to classify oscillators based on their dynamics is to differentiate between limit cycle oscillators, which generate sustained oscillations, and noise-induced oscillators, whose parameters are set below the Hopf bifurcation, and thus their deterministic dynamics correspond only to damped oscillations. In the second case, sustained oscillations might occur under noise that provides constant disruptions driving the system away from the stable state [89, 114]. This class of noise-induced oscillators is important for biological oscillators, where it is not often clear if the system is noise-induced or limit cycle [115], as well as in engineering applications, for example in the control of semiconductor superlattices [116]. Regarding entrainment, it has been shown that a noise-induced oscillator can be entrained to a wider range of input amplitudes than a limit cycle oscillator [117].

To explore the all above-mentioned oscillator types and regimes, this section presents numerical simulations with two generic models: the amplitude-phase model representing a sinusoidal oscillator and the Van der Pol model representing a relaxation oscillator. Furthermore, both models are considered in two regimes, first as a limit cycle (sustained) oscillator and second as noise-induced oscillators that produce sustained oscillations only in the presence of noise. The novelty of the here-presented numerical experiments lies mainly in considering the individual classes of oscillators in the context of population-level dynamics and under external input signals and a range of noise intensities. This should provide more insights into the mathematical properties of the oscillator systems that are necessary to reproduce the results achieved with the bio-inspired oscillator model in the previous sections of this chapter.



**Figure 2.26:** Comparison of the deterministic and stochastic simulations for the amplitude-phase model in two regimes: limit cycle oscillator (panel A) and noise-induced oscillator (panel B).

## 2.5.1 Generic oscillator models

### Amplitude-phase model

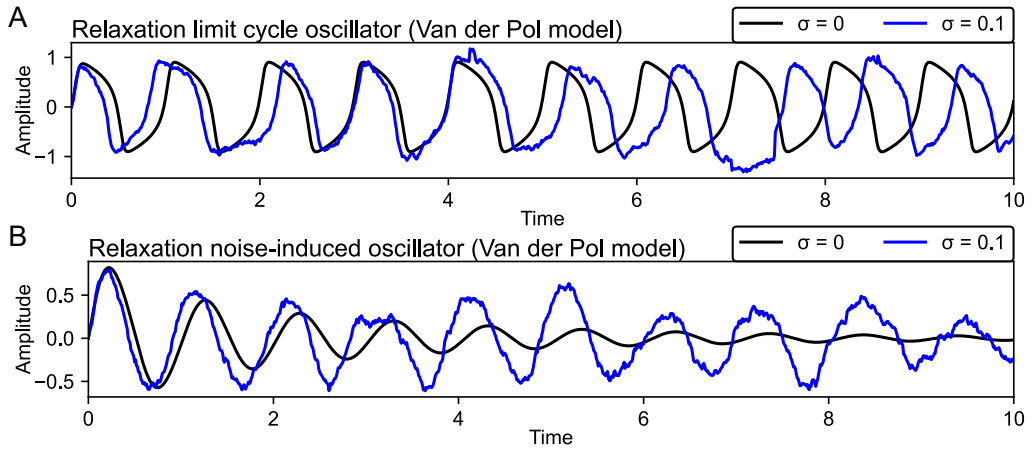
The amplitude-phase model is a generic model of an oscillator with sinusoidal oscillations and has equations [36]

$$\dot{X} = \lambda X (A - R) - \omega Y + I + \sigma W_1, \quad (2.20a)$$

$$\dot{Y} = \lambda Y (A - R) + \omega X + \sigma W_2, \quad (2.20b)$$

$$R = \sqrt{X^2 + Y^2}, \quad (2.20c)$$

where  $\sigma$  is noise intensity,  $W_i$  are independent Wiener processes (Equation 2.7),  $I$  is input,  $\lambda$  is the time rate of the return to the limit cycle,  $A$  is amplitude, and  $\omega$  is the angular frequency (see also Section 1.2.1 for detailed description). To represent a limit cycle oscillator, the free parameters were set to  $\lambda = A = \omega = 1$  according to the previous study (Figure 2.26A) [34]. To explore the model also in the regime of a noise-induced oscillator, the parameter  $A$ , representing the amplitude of the oscillations, can be set to  $A = 0$  (Figure 2.26B). The amplitude-phase model is a generic model whose equations do not have a direct biological interpretation. Therefore, the noise terms are purely additive and represent a general stochastic disturbance rather than specifically molecular noise as in the Kim-Forger model. The SDE model was simulated using the Euler-Maruyama method with an integration step  $dt = 0.001$ .



**Figure 2.27:** Comparison of the deterministic and stochastic simulations for the Van der Pol model in two regimes: limit cycle oscillator (panel A) and noise-induced oscillator (panel B).

## Van der Pol model

The Van der Pol model is a generic model of an oscillator with relaxation oscillations and has equations [38]

$$\frac{dX}{dt} = Y + \sigma W_1 \quad (2.21a)$$

$$\frac{dY}{dt} = -(BX^2 - d)Y - X + I + \sigma W_2 \quad (2.21b)$$

where  $\sigma$  is noise intensity,  $W_i$  are independent Wiener processes (Equation 2.7),  $I$  is the input, and  $d$  and  $B$  are free parameters that were set, according to the previous study [38], to represent a relaxation limit cycle oscillator ( $d = 2$ ,  $B = 10$ , Figure 2.27A) and a noise-driven oscillator ( $d = -0.1$ ,  $B = 1$ , Figure 2.27B). This model, including its parameters, was adapted from a previous study on the entrainment of stochastic oscillators (see also Section 1.2.1 for detailed description) [38]. Similar to the amplitude-phase model, the Van der Pol model is a generic model whose equations do not have a direct biological interpretation. Therefore, the noise terms are purely additive and represent a general stochastic disturbance rather than specifically molecular noise as in the Kim-Forgner model. The SDE model was simulated using the Euler-Maruyama method with an integration step  $dt = 0.001$ .

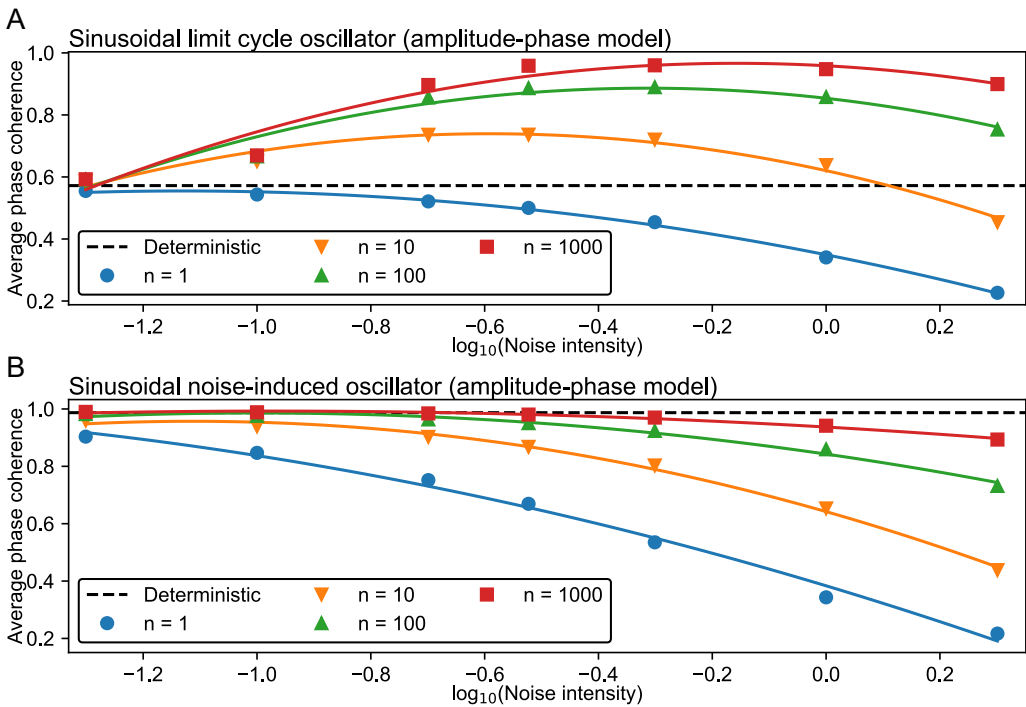
## 2.5.2 Noise facilitates entrainment of limit cycle but not noise-induced oscillators

The repetition of the experiments from Sections 2.3 and 2.4 with the generic oscillator models showed that the previously described results work also for generic sinusoidal and relaxation oscillator models with the parameters corresponding to the limit cycle regime. Distinctive behavior, however, appears for noise-induced types of oscillators. Specifically, the sinusoidal and relaxation limit cycle oscillators showed the same results as the bio-inspired oscillator in the previous sections: increasing noise intensity increases the range of entrainment (Figures 2.28A and Figures 2.29A), the amplitude of the phase response curves becomes larger (Figures 2.30A and Figures 2.31A), and recovery from jet lag faster (Figures 2.32A and 2.33A). On the other hand, repeating the experiments with parameters corresponding to noise-induced regimes, showed distinctively different behavior. Here, already the deterministic model with zero noise intensity showed a large range of entrainment (Figures 2.28B and Figures 2.29B), high-amplitude PRCs (Figures 2.30B and Figures 2.31B), and short jet lags (Figures 2.32B and 2.33B). Those properties were consequently not further improved by increasing the noise intensity. We thus conclude that the observations made in this chapter are general properties of limit cycle oscillators, both sinusoidal and relaxation, but do not apply to noise-induced oscillators without a limit cycle.

## 2.6 Population heterogeneity as a form of noise

The previous sections of this chapter investigated a population of identical stochastic oscillators whose noisy behavior stemmed from the uncertainty in the oscillation generation. In real oscillator systems, however, the oscillators might not be identical and a certain level of heterogeneity can be expected [118, 119]. This heterogeneity is effectively also a form of noise, which raises an intriguing question of whether this so-called extrinsic noise would have the same impact on the entrainment dynamics as intrinsic noise considered so far. A previous study has already shown that heterogeneity facilitates a wider range of population-level entrainment similar to the intrinsic noise [26]. This section aims to investigate the effects of heterogeneity more in detail. Particularly, the focus is on how the dynamics of the individual oscillators map to its population-level mean and how the entrainment metrics developed in Section 2.2 can be interpreted in the context of more complicated and ambiguous waveforms that stem from averaging non-identical oscillators.

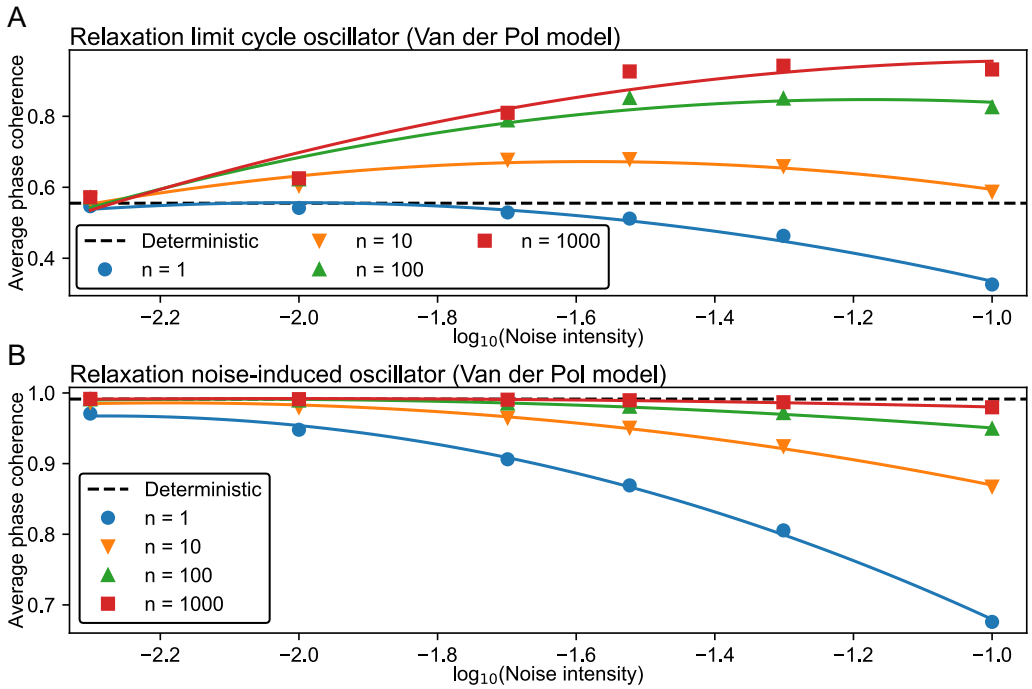
To get better insight into the complexity that stems from population heterogeneity, first, a case of a small population of only two oscillators is considered. The oscillators are modeled using the deterministic Kim-Forger model (Equation 2.1) with the population heterogeneity implemented by a time-scaling parameter, which scales the free-running period of the oscillator (Equation 1.3).



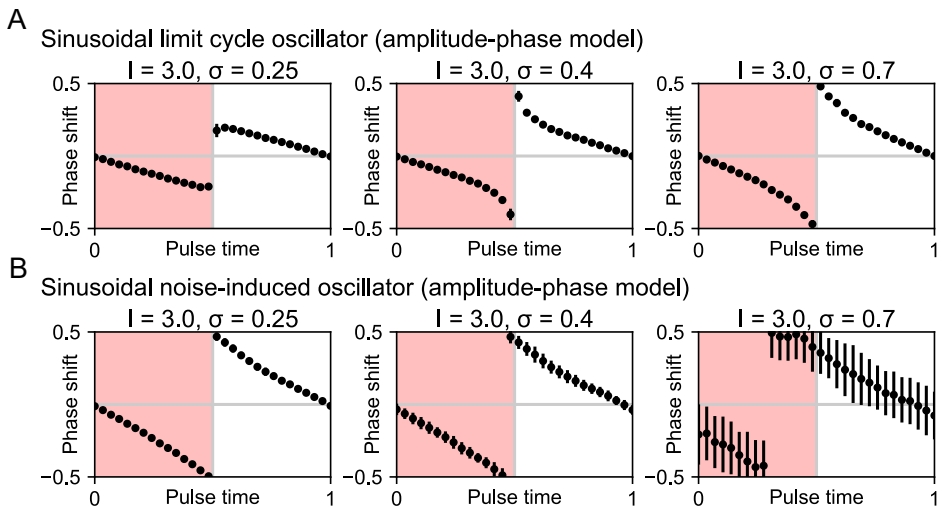
**Figure 2.28:** Mean phase coherence for the limit cycle (panel A) and noise-induced (panel B) regimes of the amplitude-phase model. The mean phase coherence is shown as a function of the noise intensity ( $\sigma$ ) and population size ( $n$ ). Solid lines represent second-order polynomial fits. The individual Arnold tongues used to generate those images are shown in Figures A.4 and A.5 for panels A and B, respectively.

The free-running period is denoted  $\tau$  and, in this example, is chosen as  $\tau = 0.9$  and  $\tau = 1.1$ , leading to two distinctive single-cell Arnold tongues that superimpose at the population level (Figure 2.34). The superposition of two non-identical oscillators leads to the emergence of rather complicated non-sinusoidal waveforms (Figure 2.35). To estimate the entrainment, the Arnold tongues were estimated using all four metrics presented in Section 2.2: winding number, autocorrelation, peak detection, and circular cross-correlation. As expected from the previous results, all four metrics gave the same estimate for the single-cell Arnold tongues with a clear distinction between entrained and unentrained areas. However, at the population level, the situation is considerably more complex as the population-level Arnold tongue depends on the selected metric. The population-level Arnold tongues also show generally three regimes of entrainment: full entrainment, when both oscillators are entrained (point A in Figure 2.34), quasi entrainment, when only one oscillator is entrained (points B and C in Figure 2.34), and no entrainment when neither of the oscillators is entrained (point D in Figure 2.34). The difference among the individual metrics occurs mainly in the area of the Arnold tongue, where only one of the oscillators is entrained. Thus in the context of non-identical oscillators, the definition of

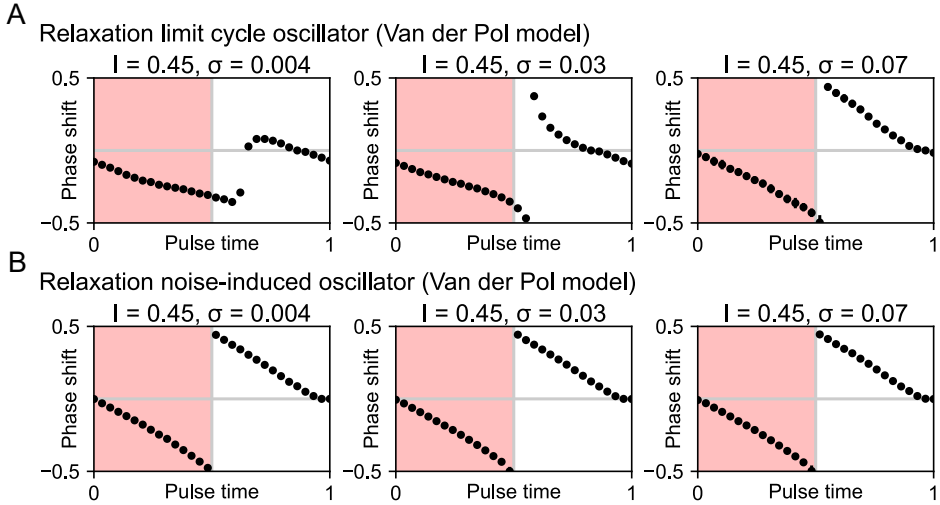




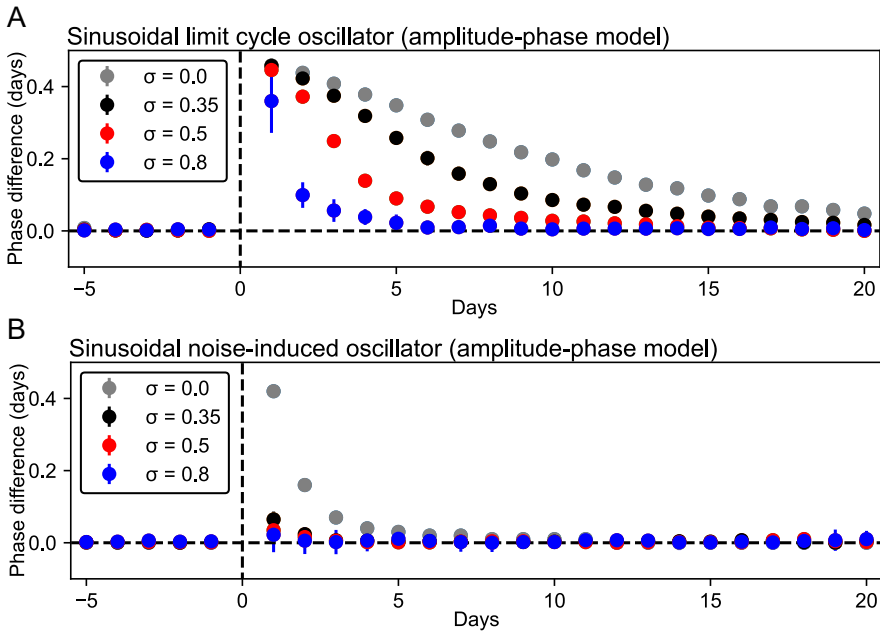
**Figure 2.29:** Mean phase coherence for the limit cycle (panel A) and noise-induced (panel B) regimes of the Van der Pol model. The mean phase coherence is shown as a function of the noise intensity ( $\sigma$ ) and population size ( $n$ ). Solid lines represent second-order polynomial fits. The individual Arnold tongues used to generate those images are shown in Figures A.6 and A.7 for panels A and B, respectively.



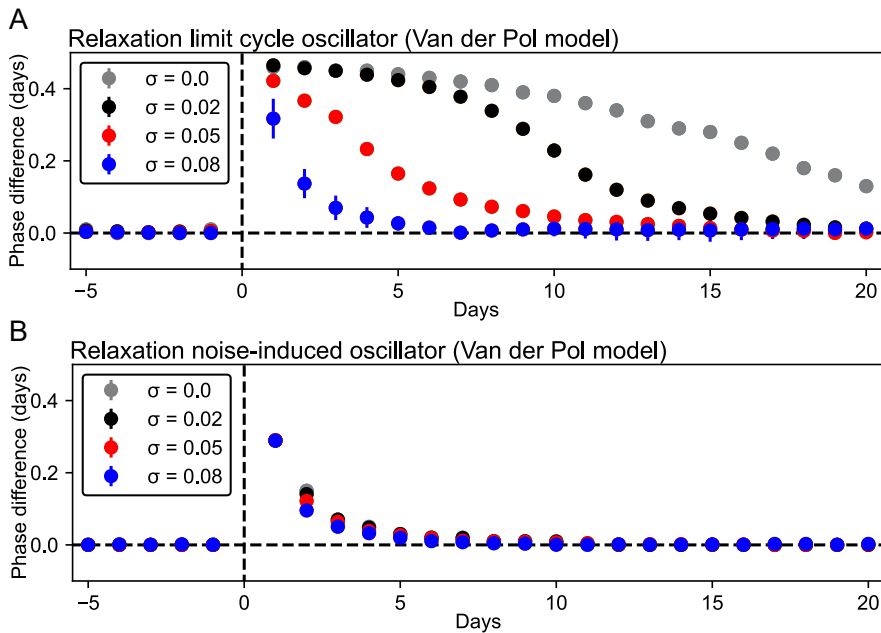
**Figure 2.30:** Phase response curves for the limit cycle (panel A) and noise-induced (panel B) regimes of the amplitude-phase model.



**Figure 2.31:** Phase response curves for the limit cycle (panel A) and noise-induced (panel B) regimes of the Van der Pol model. The applied values of the input amplitude and noise intensity parameters differ from the parameters used for the amplitude-phase model (Figure 2.30) as the models have different structures and thus the observed effects appear at different parameter values.



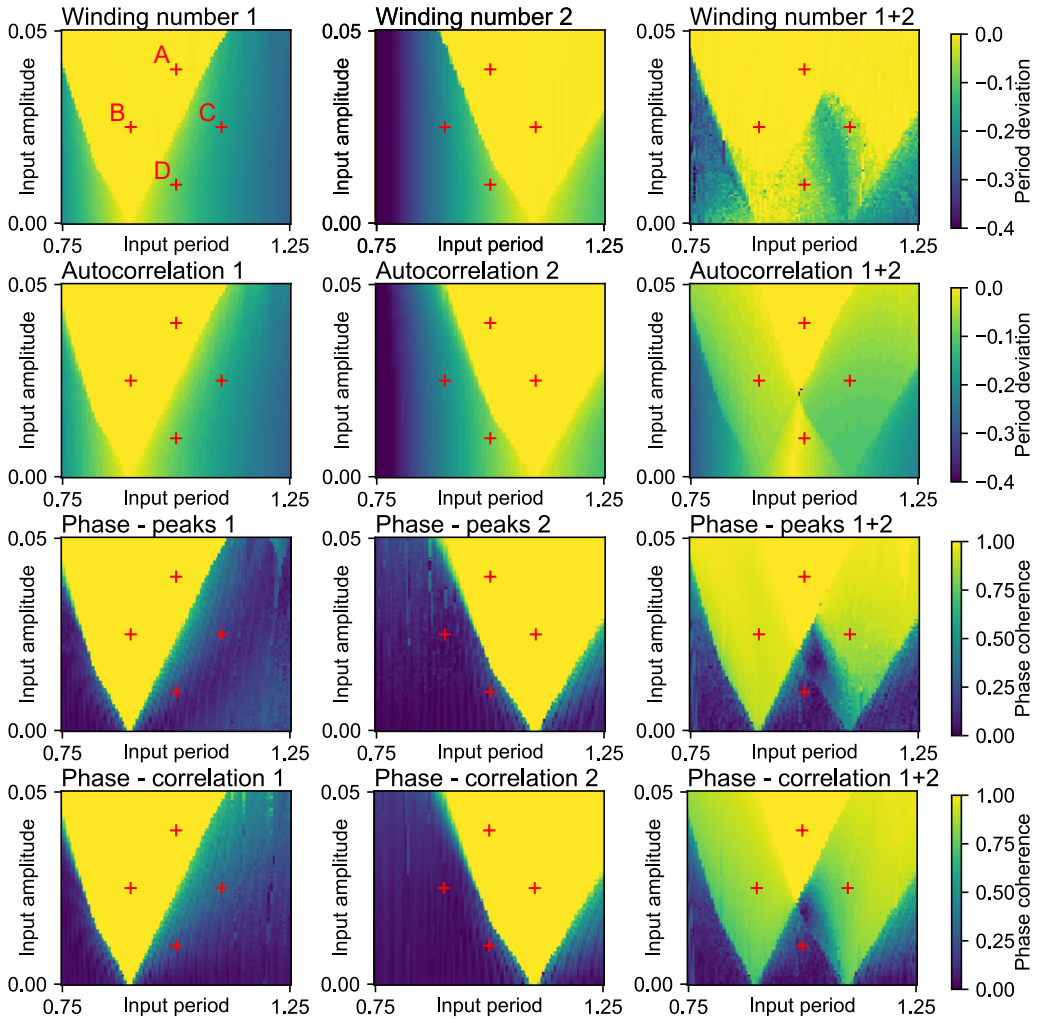
**Figure 2.32:** Recovery from jet lag for the limit cycle (panel A) and noise-induced (panel B) regimes of the amplitude-phase model.



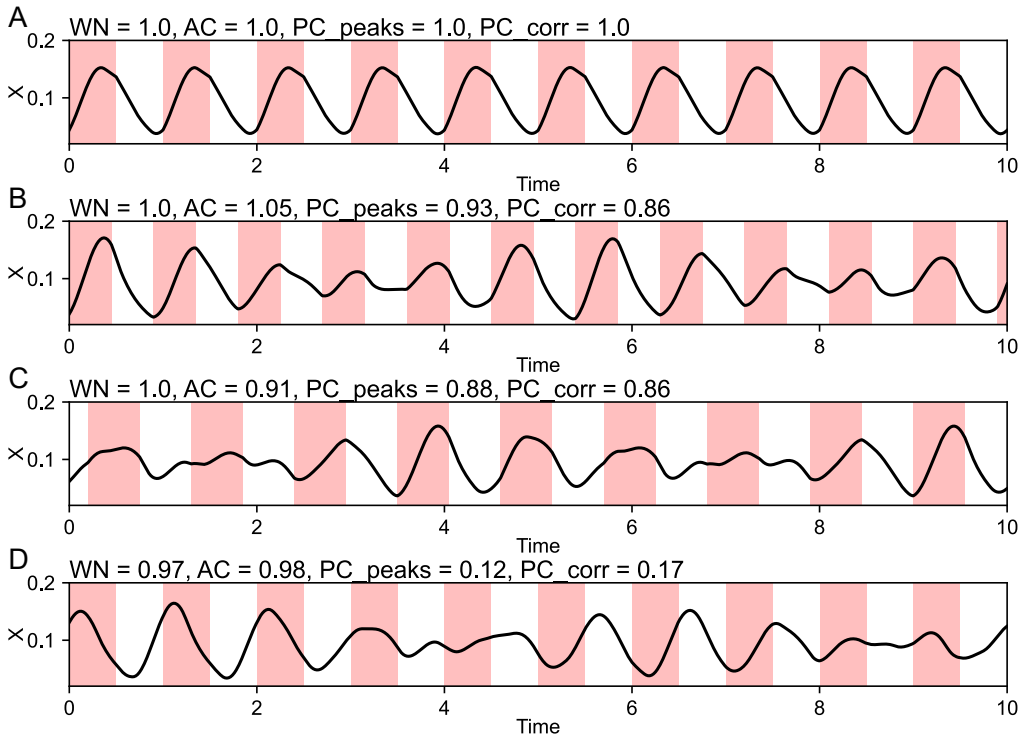
**Figure 2.33:** Recovery from jet lag for the limit cycle (panel A) and noise-induced (panel B) regimes of the Van der Pol model. The applied values of the input amplitude and noise intensity parameters differ from the parameters used for the amplitude-phase model (Figure 2.32) as the models have different structures and thus the observed effects appear at different parameter values.

entrainment is more ambiguous as the estimated degree of entrainment depends strongly on the selected metric.

In the next step, the results from the previous paragraph were extended for a population of three, five, and seven oscillators (Figure 2.36). In this experiment, two metrics of entrainment were used: winding number as a period estimator and peak detection as a phase estimator. The results show two qualitative differences, depending on if the period or phase estimator is used. Specifically, when winding number (period estimator) is used, two regions of the population-level Arnold tongue can be recognized: entrained and unentrained. The border between those regions approximately follows the border where at least half of the oscillators are entrained. In contrast, quantifying entrainment with peak detection leads to the emergence of multiple entrainment zones with various levels of entrainment that overlay the number of overlapping single-cell Arnold tongues. Taken together, if at least half of the oscillators are entrained then the output oscillates with the period of the input signal, however, the phase coherence improves markedly with each oscillator that becomes entrained.



**Figure 2.34:** Superposition of Arnold tongues for two oscillators (right column) with free-running periods  $\tau = 0.9$  (left column) and  $\tau = 1.1$  (middle column). The Arnold tongues are estimated using the four metrics presented in Section 2.2 (individual rows in the figure). The labels A, B, C, D refer to plotted trajectories shown in Figure 2.35.

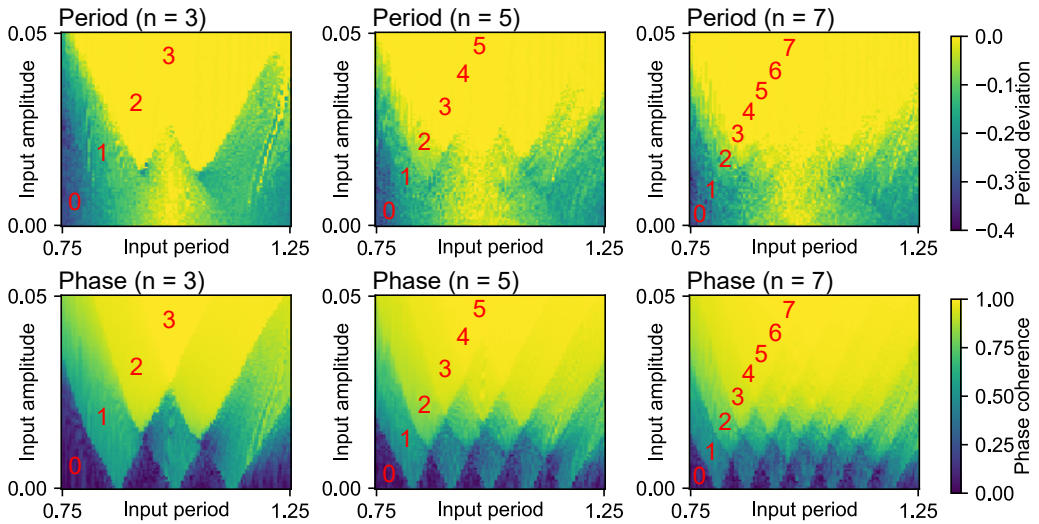


**Figure 2.35:** Time trajectories illustrating the range of dynamical behavior achieved with a system consisting of two heterogeneous oscillators. The panels A, B, C, D correspond to points A, B, C, D from Figure 2.34. Titles of the individual panels show values of entrainment metrics from Section 2.2: winding number (WN), autocorrelation (AC), phase coherence using peak detection (PC\_peaks), and phase coherence using circular cross-correlation (PC\_corr). A value of 1 for WN, AC indicates that the input and oscillator period are the same. A value of 1 for AC and PC\_peaks indicates stable phase-locked entrainment.

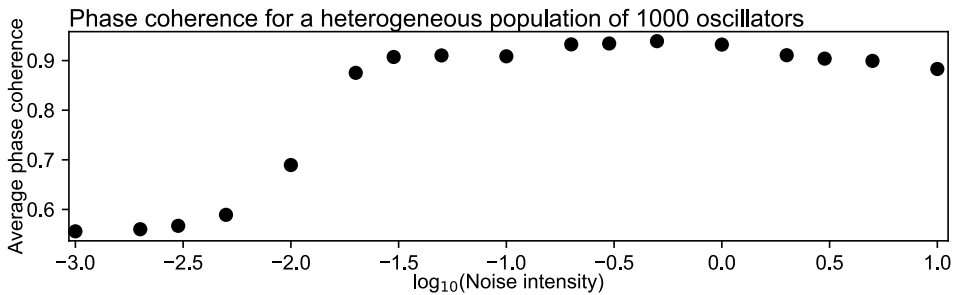
As the last step presented in this section, a population of 1000 heterogeneous oscillators was considered (Figure 2.37). The period of each oscillator was set by drawing a random number from a normal distribution  $N$  as

$$\tau \sim N(\tau_0, \sigma), \quad (2.22)$$

where  $\tau_0 = 1$  is mean and  $\sigma$  is standard deviation representing the noise intensity. Plotting the mean phase coherence as a function of the noise intensity  $\sigma$  showed that the mean phase coherence is increasing with increasing  $\sigma$ . The effect thus seems to be similar to the intrinsic noise (Section 2.3). However, given the complexity of the population-level waveforms, as described in the previous paragraphs, a more throughout investigation of this phenomenon is required before drawing any conclusions. For example, here only 1:1 Arnold tongues are considered, but considering a system that is driven by a periodic input signal with period  $T_{IN}$ , the system can be also entrained with arbitrary m:n ratio if the output is periodic with a period  $T_{OUT}$



**Figure 2.36:** Superposition of three ( $n = 3$ ), five ( $n = 5$ ) and seven ( $n = 7$ ) Arnold tongues. The first row shows Arnold tongues estimated by winding number, an estimator of the oscillator period. The second row shows Arnold tongues estimated by phase coherence with phase estimated by the circular cross-correlation. The red numbers show, visually approximated, how many Arnold tongues are overlapping in the giving region.



**Figure 2.37:** Mean phase coherence as a function of noise intensity for a heterogeneous population of 1000 oscillators. Noise intensity refers to the standard deviation of the normal distribution from which the parameter values were drawn. The individual Arnold tongues used to generate this figure are shown in Figure A.8.

and  $T_{\text{OUT}}/T_{\text{IN}} = m/n$  [120]. The phenomenon of  $m:n$  entrainment is especially important in the context of a heterogeneous population as the Arnold tongues can move far away from its "default" position and thus enter the range of input amplitudes and periods captured by the experiment [55]. Another shortcoming is, that here only the time-scaling parameter  $\tau$  is varied, but varying other parameters will lead to widely different waveforms and some might even cross the Hopf bifurcation or exhibit chaotic behavior. Therefore, the results presented here should be considered rather preliminary and should be developed in detail in future work.

## 2.7 Parameter fitting algorithm for a stochastic oscillator population

The optimization of stochastic models is a notoriously difficult problem because of the uncertainties caused by stochastic terms as well as computational demands of the numerical methods for evaluation of the stochastic models [121]. This section presents an optimizing approach that aims to minimize the computational effort by allowing pre-estimation of the parameter values using a deterministic model and subsequently using stochastic simulations only to fine-tune the model parameters to the data.

### 2.7.1 Benchmark model and data

The model used as a benchmark for the presented fitting algorithm represents a population of 1000 synthetic gene oscillators each consisting of a network of two genes. The model equations for a single oscillator read [122]

$$\dot{x} = F - \gamma_x x + I_{\square} + \sigma W_x \quad (2.23a)$$

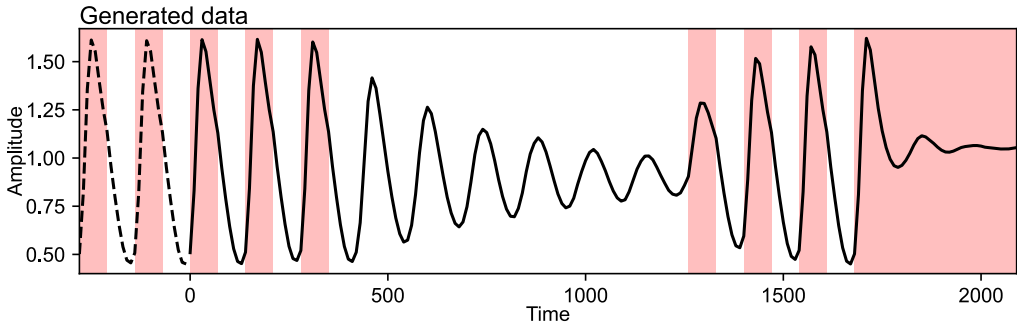
$$\dot{y} = \frac{1}{\tau_y} (F - \gamma_y y) + \sigma W_y \quad (2.23b)$$

$$F = \frac{1 + x^2 + \alpha\beta x^4}{(1 + x^2 + \beta x^4)(1 + y^4)}, \quad (2.23c)$$

where  $x$  and  $y$  present two gene products whose genes contain the same promoter represented by function  $F$ .  $y$  acts as a negative element inhibiting gene production and  $x$  is a positive element enhancing the production of both genes. The free parameters  $\alpha$  and  $\beta$  control the gene expression.  $\gamma_x$  and  $\gamma_y$  are degradation rates for  $x$  and  $y$  respectively.  $\tau_y$  is time scale for  $y$ .  $\sigma$  is noise intensity,  $I_{\square}$  square input signal. In comparison to the original paper [122], the benchmark model in Equation 2.23 was extended for the purpose of this section by independent noise sources  $W_x$  and  $W_y$ . The extension by the noise terms allows us to use the model in the context of stochastic population-level fitting. The parameters of the model were set to

$$\alpha = 11, \beta = 2, \gamma_x = 0.105, \gamma_y = 0.036, \tau_y = 5, I = 0.01, \sigma = 0.01. \quad (2.24)$$

The model was simulated with those parameters under a varying input signal consisting of periods of a regular square signal, and periods when the signal was kept at one or zero for a longer time span (Figure 2.38). For the following optimization, it is assumed that only the value of  $\tau_y = 5$  is known. The remaining parameters will be considered unknown and searched in the interval



**Figure 2.38:** Generated benchmark data for the optimization algorithm. The system of 1000 uncoupled stochastic oscillators is synchronized to the input cycle before the recording starts. The dashed line is not part of the recorded data.

[0, 100]. The Hasty model was chosen for this section as a generic example of an oscillator. In Section 4.5 the same fitting algorithm is used also for fitting of the Kim-Forger model to the real experimental data.

## 2.7.2 Step-wise fitting algorithm

Differential evolution is used as an optimization algorithm for the parameter estimation [123]. Differential evolution is a stochastic population-based algorithm that starts with a randomly initialized population of vectors. The values of the individual vectors evolve through several generations using random mutations of the vector values and crossovers with other population members until the whole population converges to an optimum. The differential evolution is an efficient black-box optimization method that does not require derivatives of the cost function to be known [124]. Computation experiments also showed that the differential evolution optimizer works also well for stochastic functions with undefined regions, which is important because the numerical simulation of the stochastic differential equations can fail or the equations might be simply not defined for some parameter values (Figure 2.39).

The advantages of the differential evolution algorithm can be illustrated by an example of the Gomez-Levy function [125]

$$f(x, y) = 4x^2 - 2.1x^4 + \frac{1}{3}x^6 + xy - 4y^2 - 1 + y^4 + \sigma N(0, 1), \quad (2.25)$$

where  $\sigma = 0.1$  is noise intensity and  $N(0, 1)$  is a random number with normal distribution with zero mean and unit standard deviation. The extension of the Gomez-Levy function by the noise terms was done purposefully for this section, to study the behavior of the optimizer under



uncertainty. The optimizer also needs to be able to deal with discontinuities and undefined regions that occur for the Gomez-Levy function after applying condition [125]

$$g = -\sin(4\pi x) + 2\sin^2(2\pi y) \leq a, \quad (2.26)$$

where  $a = 1.5$  (Figure 2.39A).

Running the differential evolution algorithm to find the optimum of Equation 2.25 subjected to the condition in Equation 2.26, scenarios for various initial populations can be studied. When the initial population covers the whole search range (Figure 2.39B), all members of the population converge to the global optimum, with uncertainty caused by the noise terms in Equation 2.25. When the initial population is limited to a sub-interval of the total search range in the vicinity of a local optimum (Figure 2.39C), the population might get stuck in this local minimum. However, even a restricted initial population can converge to the global optimum, even if this global optimum is outside of the initial range (Figure 2.39D). Those observations are obvious, but also important for the below-introduced optimization algorithm, which takes a pre-optimized population from a previous step and applies it as an initial population for the following step.

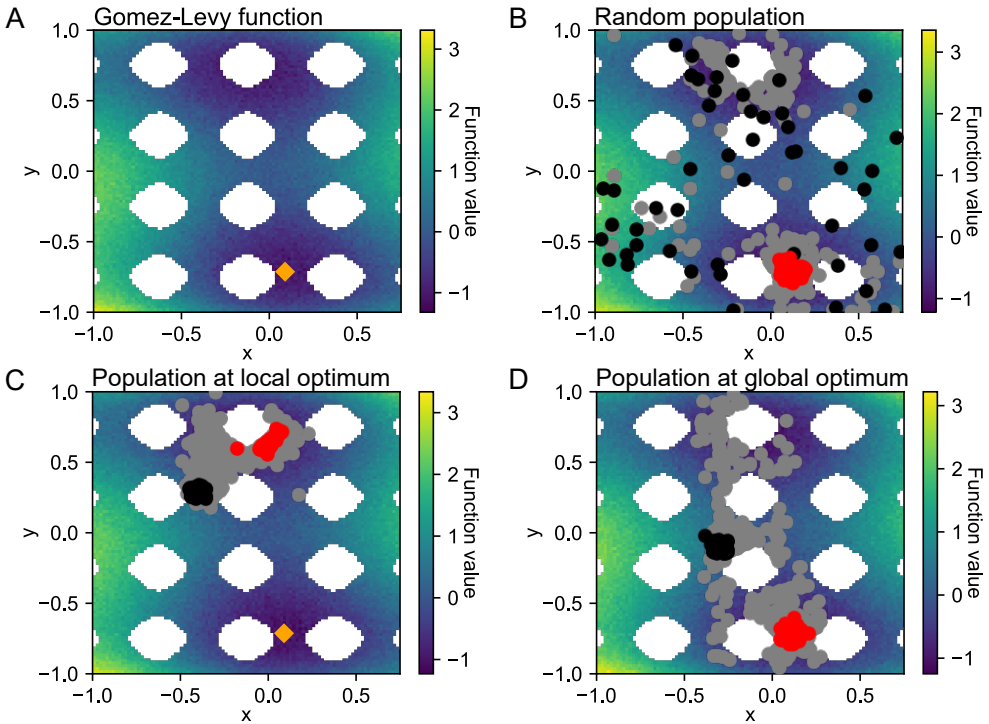
Taken together, differential evolution is a powerful and robust algorithm that can efficiently navigate very complicated landscapes. This section builds on the differential evolution optimizer by presenting a step-wise fitting algorithm that consists of three steps where each step starts with an initial population that was estimated by a previous step. The steps are:

1. Pre-estimation of parameter values for a deterministic model without noise.
2. Pre-estimation of noise intensity based on the population-level loss of synchrony.
3. Fine-tuning the parameter values to the experimental data.

The following subsections describe the outlined steps in detail and Section 2.7.3 compares the step-wise algorithm with direct fitting using the differential evolution without the pre-estimation steps 1 and 2.

### Step 1: Pre-estimation of parameter values for deterministic model

In the first step, all noise intensities in the model equations are set to zero and only the parameters of the resulting deterministic model are estimated to find an oscillating dynamic that can be entrained by the input signal. A lot of heavy lifting is thus made using the deterministic model that can be evaluated using fast adaptive ODE algorithms and can be simulated only once as opposed to the population of 1000 stochastic oscillators. The cost function consists of terms that



**Figure 2.39:** Pre-estimating the starting population of a population-based algorithm can reduce the time needed to find the optimum. (A) As an example, a Gomez-Levy function is considered [125]. The orange diamond marker indicates the global minimum. The white spaces indicate undefined values of the function. Random noise was also added to the function resulting in the grainy look of the plot. (B) With an initial population covering the whole search space, the differential evolution algorithm can easily find the global optimum even if noise and undefined intervals are present. Black, gray, and red dots indicate the initial population, points visited by the search algorithm, and final population, respectively. (C) An initial population restricted to a small interval at a local optimum can get stuck in this local optimum. (D) In a different scenario, the optimum can be found even if not covered by the initial population. In the context of the optimization of the SDE-model parameters, a pre-estimated initial population, as compared to a very widespread population, leads to faster convergence and thus lower number of time-expensive SDE evaluations.

represent experimental features that can be approximated also without stochasticity [126]. The cost function has form

$$C_{\text{ODE}} = C_T + C_\varphi + C_{\min} + C_{\max} \quad (2.27)$$

and denotes various oscillator properties observable from the target data or known from the literature.  $C_T$  indicates the error of the simulation from the target oscillation period with the constant input signal at 0 (Figure 2.40A)

$$C_T = \frac{|T_{\text{model}} - T_{\text{data}}|}{T_{\text{data}}}, \quad (2.28)$$

where  $T_{\text{model}}$  and  $T_{\text{data}}$  are periods estimated from the model output and target data, respectively. The second term  $C_\varphi$  concerns the ability of the model to be entrained by an input signal as

$$C_\varphi = |\varphi_1 - \varphi_2|, \quad (2.29)$$

where  $\varphi_1$  and  $\varphi_2$  are the phases of the oscillator under two cycles shifted by half of the period (Figure 2.40B, C). The phase shift by half of the period means that cycle 1 starts at time  $t = 0$ , whereas cycle 2 starts at the  $t = T/2$ , where  $T$  is the period of both cycles. Evaluation the phases for the two cycles, and making sure there are equal, ensure that the model is really entrained by the input signal and does have a constant entrainment phase simply due to the fact that the free-running period of the oscillator matches the period of the input signal [126]. The third term and fourth terms decode the minimum and maximum of the entrained signal as

$$C_{\min} = \frac{|\min(x_{\text{model}}) - \min(x_{\text{data}})|}{\min(x_{\text{data}})}, \quad (2.30)$$

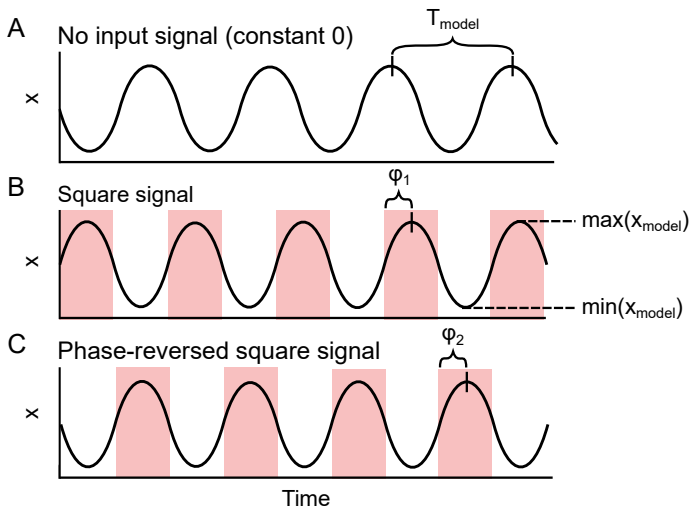
$$C_{\max} = \frac{|\max(x_{\text{model}}) - \max(x_{\text{data}})|}{\max(x_{\text{data}})}, \quad (2.31)$$

where  $\min(x_{\text{model}})$  and  $\min(x_{\text{data}})$  are minimum estimated from model output and target data, respectively. One assumes the considered oscillators generate non-zero values by definition. To allow some variance in the final population, the optimum is considered with a tolerance of 10 % for  $C_T$ ,  $C_{\min}$ , and  $C_{\max}$  as

$$g = \frac{|\text{estimated value} - \text{reference value}|}{\text{reference value}} < 0.1. \quad (2.32)$$

The individual terms of the cost function in Equation 2.27 can be adjusted based on the system studied and the information we have about the system.

The optimization was performed in the range of  $[0, 100]$  for all parameters. This parameter range was chosen as a toy example based on the fact that the largest parameter has a two-digit value.

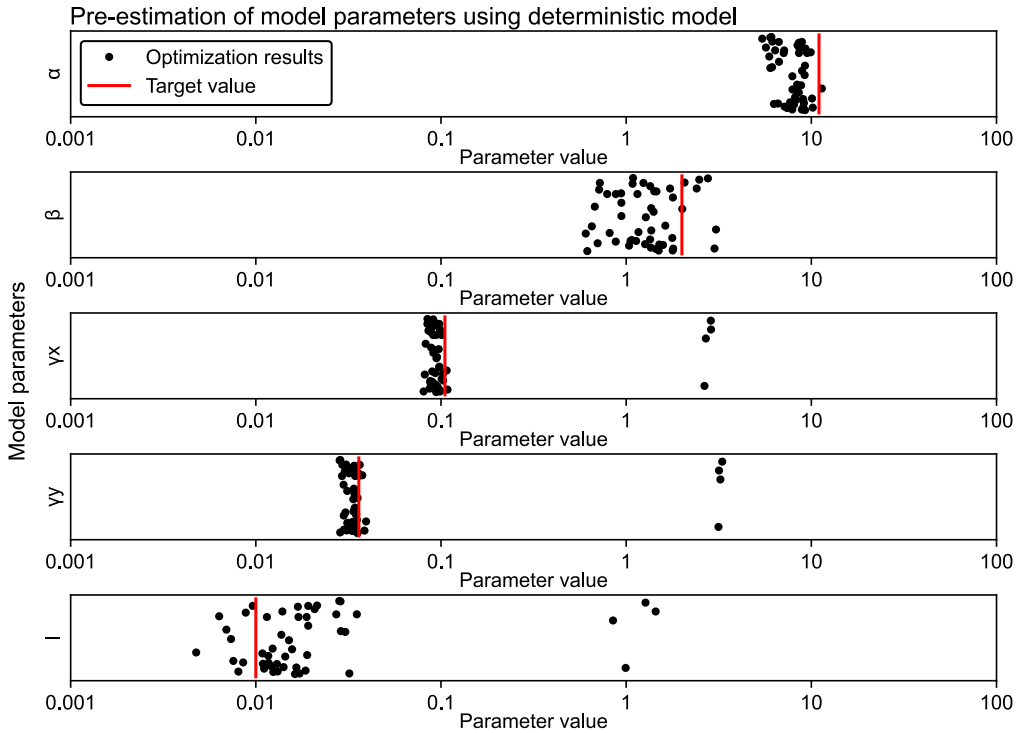


**Figure 2.40:** Visualization of metrics used for the optimization of the parameters of the deterministic model. Each panel represents an independent simulation of the model. (A) Simulation with the input signal at 0 to estimate free-running period  $T_{\text{model}}$ . (B) Simulation with square signal to estimate maximum and minimum of oscillations under the entrainment and the phase of entrainment ( $\varphi_1$ ). (C) Simulation with square signal shifted by half of the period to ensure that the new phase of entrainment ( $\varphi_2$ ) will be the same as in panel B.

In the real-world application, the parameter range must be estimated from literature or one must try multiple parameter ranges. The initial population was let to evolve for 100,000 generations. After 100,000 steps, the best candidate was saved. Those are consequently used as an initial population for the next fitting step. Comparison of the 50 pre-estimated values, each estimated with an independent run of the differential evolution algorithm, with the target parameter values (Figure 2.41) showed a significantly narrowed down search space from the original range of  $[0, 100]$ . Using the pre-estimation with a deterministic model can thus significantly reduce the search space that needs to be covered by the stochastic simulations later.

## Step 2: Pre-estimation of noise intensity

In this step, noise intensity  $\sigma$  is estimated for each parameter set from the previous step separately. The noise intensity can be estimated from the population-level recording as increasing noise intensity causes faster desynchronization of individual oscillators. If the input signal is set to 0, the faster desynchronization of individual oscillators leads to faster damping of oscillation



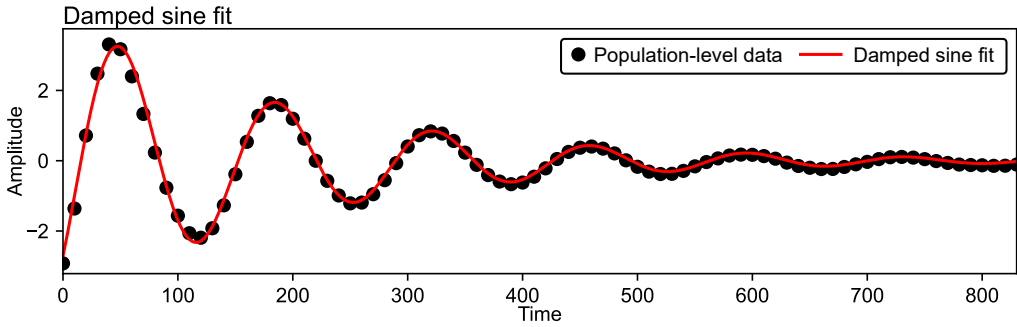
**Figure 2.41:** Results of the first optimization step for the four free parameters  $\alpha$ ,  $\gamma_x$ ,  $\gamma_y$ ,  $I$  of the Hasty model (Equation 2.23).

amplitude at the population level [127]. The damping rate can be estimated from data by fitting a damped sine (Figure 2.42) of form

$$f(t) = Ae^{-dt} \sin\left(\frac{2\pi t}{T} + \varphi\right), \quad (2.33)$$

where  $A$  is amplitude,  $d$  is damping rate,  $T$  is period, and  $\varphi$  is phase.

To estimate the damping rate  $d$  for a computational model, one can use the following approach. First, the parameter of  $\sigma$  is set to an arbitrary value. Second, the model is run under ten cycles of a regular square signal to achieve synchronization of the individual oscillators in the population. Third, the input signal is turned to 0. Consequently, the individual oscillators of the population start to desynchronize and a damped-sine-like population average emerges. The resulting damped-sine-like model output is Z-score normalized and damped sine is fitted. From the damped sine fit, one can estimate the value of the damping rate  $d$  for the given noise intensity  $\sigma$ . Repeating the algorithm for different values of  $\sigma$ , one can estimate a function  $d(\sigma)$ . There is, however, a



**Figure 2.42:** Damped sine fit to the population-level data provides estimation of noise intensity at the level of individual oscillators.

pitfall related to high values of  $\sigma$ . If  $\sigma$  is larger than can be supported by the population size, the output will not lead to a damped sine wave but rather to a random stochastic trajectory and the damped sine cannot be fitted well. Therefore, the estimated  $d$  is accepted only if the damped-sine fit reaches an efficiency coefficient of at least 0.8 according to Equation 2.38.

The damping rate  $d$  is a monotonically increasing function of noise intensity  $\sigma$  (Figure 2.43). One can take advantage of this property and find the target noise intensity using a fast converging binary search. For this algorithm first, the search range is defined as

$$[\sigma_{\min}, \sigma_{\max}], d(\sigma_{\min}) < d_{\text{target}} < d(\sigma_{\max}), \quad (2.34)$$

where  $\sigma_{\min}$  and  $\sigma_{\max}$  are search boundaries with the corresponding damping rates  $d(\sigma_{\min})$  and  $d(\sigma_{\max})$ , and  $d_{\text{target}}$  is the damping rate estimated from the data. A new candidate  $\sigma$  is then assigned as

$$\sigma_{\text{new}} = \frac{\sigma_{\min} + \sigma_{\max}}{2}. \quad (2.35)$$

In the last step, the original search range is collapsed toward the target value by assigning

$$\sigma_{\min} := \sigma_{\text{new}} \quad \text{if } d(\sigma_{\text{new}}) < d_{\text{target}} \quad \text{OR} \quad \sigma_{\max} := \sigma_{\text{new}} \quad \text{if } d(\sigma_{\text{new}}) > d_{\text{target}}. \quad (2.36)$$

Repeating Equations 2.35 and 2.36 for several iterations will give a value of noise intensity that corresponds to the damping rate estimated from the data.

In the previous step, a deterministic model was used to obtain 50 parameter sets, which should be now extended by the values for the noise intensities. The target damping rate can be estimated from the benchmark data (Figure 2.38) from the section where the input signal was kept 0 for a long time. The search range was set to  $[0, 100]$  and the binary search ran for each of the 50

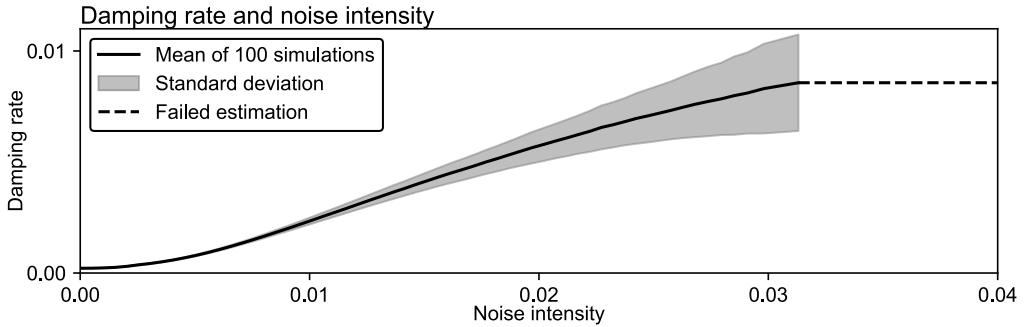


Figure 2.43: Damping rate ( $d$ ) is monotonically increasing function of noise intensity ( $\sigma$ ).

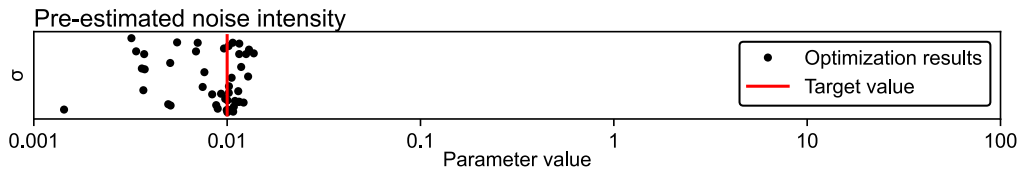


Figure 2.44: Results of the second optimization step for the noise intensity  $\sigma$  of the Hasty model (Equation 2.23).

parameter sets for 25 iterations estimating  $\sigma$  for all parameter sets independently. So estimated values of  $\sigma$  lie mostly in the close vicinity of the target value (Figure 2.44).

### Step 3: Fine-tuning the parameter values to experimental data

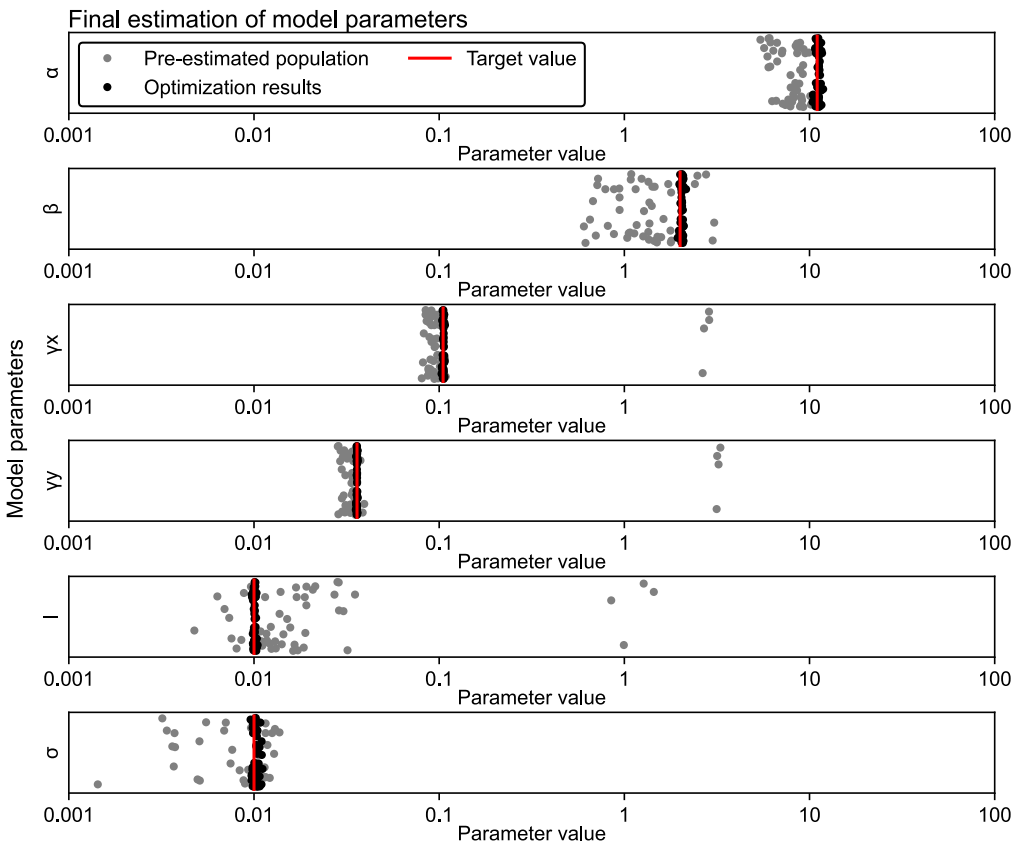
In the last step of the fitting algorithm, the pre-estimated parameter sets from the previous steps are used as the initial population for the differential evolution algorithm to fit the data. The cost function for the optimization is the squared error of the solution from the data defined as

$$C = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2, \quad (2.37)$$

where  $x_i$  denotes data points and  $\hat{x}_i$  denotes the output of the model. The initial conditions for the numerical simulation were estimated by letting the model run under 10 cycles of a square signal, which corresponds to the assumption of synchronized oscillators at the beginning of data (Figure 2.38). Performing this last optimization step on the benchmark data from the Hasty model (Equation 2.23), the original parameter set was found accurately (Table 2.2, Figures 2.45 and 2.46).

Parameter name	Target value	Search interval	Estimation
$\alpha$	11.000	[0, 100]	11.154
$\beta$	2.000	[0, 100]	2.049
$\gamma_x$	0.105	[0, 100]	0.105
$\gamma_y$	0.036	[0, 100]	0.036
$I$	0.010	[0, 100]	0.010
$\sigma$	0.010	[0, 100]	0.010

**Table 2.2:** Final fitting results for the generated data. The column "Estimation" is rounded to 3 decimal places. The parameters refer to the Hasty model showed in Equation 2.23.



**Figure 2.45:** The final optimized population of parameter sets after the last step of the optimization algorithm. The gray values show the pre-estimated values from the previous steps. The parameters refer to the parameters of the Hasty model showed in Equation 2.23.



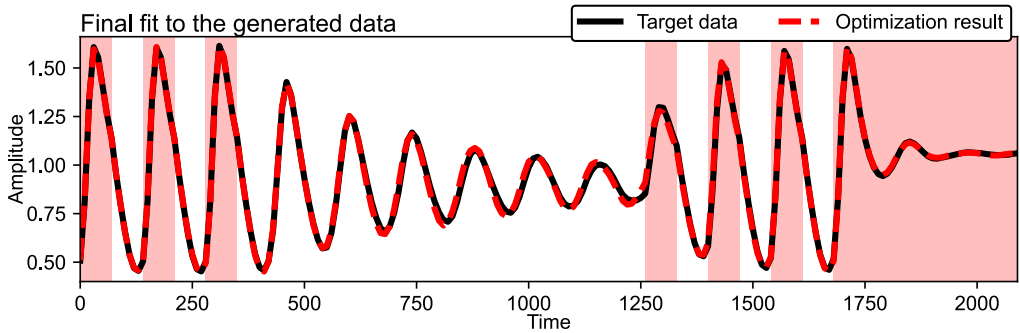


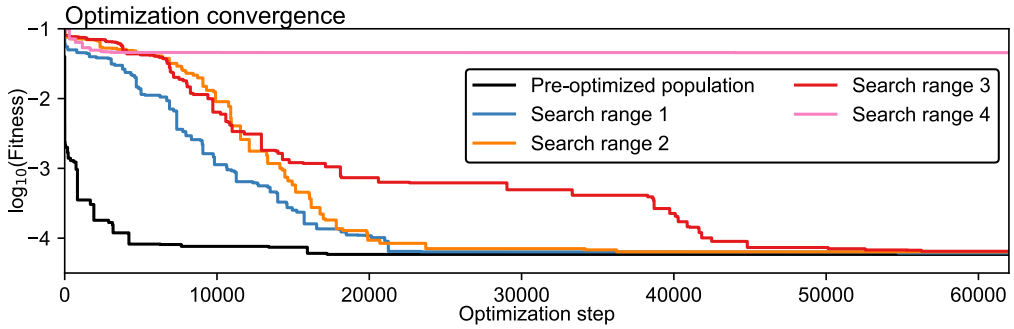
Figure 2.46: Final fit to the generated benchmark data after the step-wise optimization.

Search range	$\alpha$	$\beta$	$\gamma_x$	$\gamma_y$	$I$	$\sigma$
Step-wise algorithm	[0, 100]	[0, 100]	[0, 100]	[0, 100]	[0, 100]	[0, 100]
Search range 1	[0, 10]	[0, 10]	[0, 1]	[0, 1]	[0, 0.1]	[0, 0.1]
Search range 2	[0, 100]	[0, 100]	[0, 1]	[0, 1]	[0, 0.1]	[0, 0.1]
Search range 3	[0, 100]	[0, 100]	[0, 10]	[0, 10]	[0, 0.1]	[0, 0.1]
Search range 4	[0, 100]	[0, 100]	[0, 100]	[0, 100]	[0, 0.1]	[0, 0.1]
Search range 5	[0, 100]	[0, 100]	[0, 100]	[0, 100]	[0, 1]	[0, 0.1]

Table 2.3: Search ranges for the comparison of optimization convergence. The convergence can be seen in Figure 2.47. The parameters refer to the parameters of the Hasty model showed in Equation 2.23.

### 2.7.3 Evaluation step-wise optimization algorithm

As shown in the previous section, the presented step-wise algorithm can find the correct parameter values, but does it present a significant improvement from using the differential evolution algorithm directly? To answer this question the model parameters were fitted also directly using varying starting search ranges. Direct fitting would be here to use the third step of the optimization algorithm with a random initial population instead of the one pre-estimated with steps one and two. Specifically, 5 different search ranges were tested (Table 2.3) and corresponding optimization convergence was recorded (Figure 2.47). In comparison to the proposed algorithm, direct fitting stopped converging before the search range of [0, 100] could have been reached for all the parameters. For narrower search ranges, the direct fitting converged but not as fast as fitting with the pre-estimated population. The presented algorithm thus provides not only improvement in evaluation time but also allows for the exploration of parameter ranges that would not be possible with a conventional algorithm. For wider applicability of the proposed algorithm, the reader can also see Section 4.5 where the same algorithm was applied to fit real-world data with a different computational model.



**Figure 2.47:** Comparison of the optimization convergence with the pre-optimized population using the here-presented step-wise algorithm and with direct fitting using random populations in ranges specified in Table 2.3. For a rough reference, the evaluation of 10,000 optimization steps takes an order of hours of processor time. Due to the long computational time and because the effect captured in the figure is already obvious, multiple runs of the optimization algorithm were not performed.

## 2.7.4 Goodness of fit metric

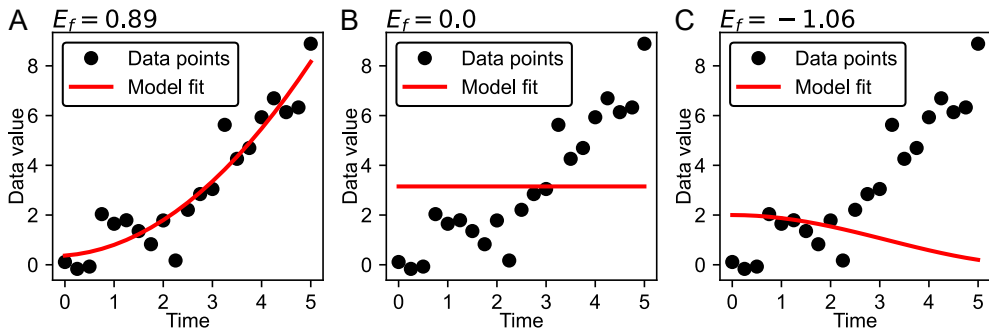
In the real-world applications it is not only important to fit the model parameters to data, but also be able to quantify the goodness of fit. This can be useful, when the user wishes to compare various parameter sets or model structures or simply to evaluate, whether the optimization algorithm gave an acceptable result. The goodness of fit can be assessed using the model efficiency coefficient defined as [128]

$$E_f = 1 - \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2.38)$$

where  $n$  is the number of data points,  $x_i$  are the individual data points,  $\hat{x}_i$  are model-estimated values and  $\bar{x}$  is mean of data points. The values of the model efficiency coefficient can range from  $-\infty$  to 1. Values near 1 indicate a high predictive value of the model while negative values indicate that data mean  $\bar{x}$  is a better predictor than model output  $\hat{x}$  (Figure 2.48).

## 2.8 Conclusion

This chapter explored the entrainment of a population of uncoupled stochastic oscillators represented by a minimal model of the circadian clock. It was found that noise allows for population-level entrainment to a wider range of input signal periods and amplitudes. Noise also facilitates a larger response to external stimuli and faster recovery from a jet-lag-like phase reversal in the input signal. These effects emerge specifically at the population level, and cannot be observed in single oscillators. In the next steps, the canonical Amplitude-phase and Van der Pol models



**Figure 2.48:** A toy example demonstrating the quantification of the goodness of fit by the model efficiency coefficient ( $E_f$ ). (A)  $E_f$  values close to 1 indicate a high predictive value of the model and thus a good fit to data. (B) A model that simply predicts the mean of data has  $E_f = 0$ . (C) Negative values of  $E_f$  indicate that the model fit is worse than simply taking the mean of the data and has thus no predictive power.

were used to show that this behavior emerges also for generic limit-cycle oscillators, but not for noise-induced oscillators without a deterministic limit cycle. Similar positive effects of noise can be observed also in heterogeneous populations with increasing heterogeneity having an analogous effect as increasing intensity of the intrinsic noise. In the last section, a step-wise fitting algorithm was introduced that enables the fitting of stochastic populations to the experimental data and thus provides a necessary link between this theoretical chapter and practical application in Chapter 4.



# 3 Implementation: Software package for oscillator population analysis

## 3.1 Introduction

There are many software solutions available for the numerical simulation of differential equations and also specifically for the simulation of chemical reaction networks [129]. Those include, among others, COPASI, XPPAUTO, Matlab, Python, and Julia. COPASI<sup>1</sup> is a user-friendly biochemical simulator, which provides numerical methods for deterministic and stochastic simulations of biochemical reaction networks. COPASI provides functions for fitting the models to experimental data, parameter scans, and visualization routines. For more flexibility, COPASI can be used as a command line tool or as an application programming interface (API) to integrate COPASI routines in various programming languages [130]. XPPAUTO<sup>2</sup> is a powerful tool for the analysis of dynamical systems. XPPAUTO can integrate differential equations, perform stability and phase plane analyses, and solve bifurcation problems for deterministic systems [129]. MATLAB<sup>3</sup> is a propriety programming language that is designed for mathematical and technical computing. MATLAB contains many toolboxes that make simulation, analysis, and fitting of mathematical models easy and convenient and includes also SimBiology toolbox designed for modeling, simulating, and analyzing chemical network systems. Python<sup>4</sup> is a general-purpose programming language with a great number of available libraries designed for scientific computing including numerical simulations of differential equations and optimization.

The software for this thesis is implemented in Julia<sup>5</sup>. Julia is a language designed specifically for scientific computing with a special focus on performance. Julia is a compiled language but implements also REPL (read-eval-print loop) for a quick and easy evaluation of statements similar to Python or MATLAB console [131]. Also, the popular Jupyter notebooks

---

<sup>1</sup> <https://copasi.org>  
<sup>2</sup> <http://math.pitt.edu/~bard/xpp/xpp.html>  
<sup>3</sup> <https://mathworks.com>  
<sup>4</sup> <https://python.org>  
<sup>5</sup> <https://julialang.org>

Functionality	Implementation	Theory
Mathematical models and simulation	3.3.1	2.1, 2.5, 2.7.1
Entrainment metrics	3.3.2	2.2
Arnold tongues	3.3.2	2.3
Phase response curves	3.3.2	2.4
Data fitting algorithm	3.3.3	2.7.2

**Table 3.1:** Overview of key methods from Chapter 2 with their implementation described in this chapter.

are natively equipped with the Julia kernel. Another factor for choosing Julia was the native `DifferentialEquations.jl` package that implements modern numerical methods for solving deterministic as well as stochastic differential equation models and also chemical reaction networks model [132]. To make the results of this thesis as reproducible as possible and also to allow the use of the methods in other projects, a Julia package<sup>6</sup> was developed. The developed package contains all methods described in the previous chapter (Table 3.1) and provides thus a framework to simulate and analyze deterministic and stochastic oscillators and their populations. This software was developed as open source under MIT License and can be thus used by anybody without any limitations.

## 3.2 Installation

The package can be installed directly from the Julia REPL by referring to its GitHub location (Listing 3.1). Once the package is installed, the user can import all package functions by calling using `OscillatorPopulation`. The list of all implemented functions with their brief description can be accessed by typing `?OscillatorPopulation`. Each function also includes its own detailed documentation that can be accessed simply by typing `?function_name`. Part of the package are also unit tests that can be used to ensure that the implemented code work as expected.

---

<sup>6</sup> <https://github.com/vkumpost/OscillatorPopulation>, DOI: 10.5281/zenodo.7620823

```

1  ## Package installation
2  import Pkg
3  Pkg.add(url="https://github.com/vkumpost/OscillatorPopulation")
4  Pkg.test("OscillatorPopulation") # optional: run package tests

```

**Listing 3.1:** The package can be installed directly from Julia REPL by calling its GitHub URL.

## 3.3 Usage examples

### 3.3.1 Simulation of population of uncoupled stochastic oscillators

The first example illustrates setting up and simulating a model of a population of uncoupled stochastic oscillators (Figure 3.1, Listing 3.3). To do this, first `load_model` is used to load a model as an instance of `Model` (Listing 3.3 line 6, Figure 3.1A). The currently implemented models are all models used in this thesis, which includes amplitude-phase model (Equation 2.20), Goodwin model (Equation 1.8), Van der Pol model (Equation 2.21), and Kim-Forger model (Equation 2.5). All models include their deterministic and stochastic implementations as described in the previous chapter. To implement new models the file `src/model_library.jl` should be edited following the example of already implemented models. An instance of `Model`, which `load_model` returns, represents a specified model together with its parameters, initial conditions, and simulation settings. Those can be displayed by `print_info` function (Listing 3.2). All implemented models contain, by convention, an input parameter called  $I$ , a noise intensity parameter called  $\sigma$  and a scaling parameter called  $\tau$ .  $\tau$  is used to scale time without changing the dynamics of the equations (Equation 1.3).

The parameters of the loaded model, such as parameter values, numerical solver, and numerical solver parameters can be changed by a set of functions (Listing 3.3 lines 7, 13, 16, Figure 3.1B). This is in detail explained in the documentation of the individual functions. For details, it is also advised to read the documentation of `DifferentialEquations.jl`<sup>7</sup>, which is used to perform the numerical simulations. Two particularities that are worth extra attention are functions `set_output!` and `set_input!`. `set_output!` is used to set the output of the model, which can be any transformation on the state variables. `set_input!` adds an input signal to the model based on the `events` matrix. `events` is a matrix with two columns, the first column indicating times when the signal turns on and the second column times when the signal turns off. For convenience,

<sup>7</sup> <https://diffeq.sciml.ai/stable>

```
1 SDE Model
2   tspan = (0.0, 100.0)
3 Solver
4   StochasticDiffEq.SOSRI()
5 Solver parameters
6   saveat = 0.01
7 Initial conditions
8   x = 0.1
9   y = 0.1
10 Parameter values
11   B = 10.0
12   d = 2.0
13   I = 0.0
14    $\tau$  = 7.63
15    $\sigma$  = 0.1
```

**Listing 3.2:** Example of calling `print_info(model)` showing properties of a `Model` object.

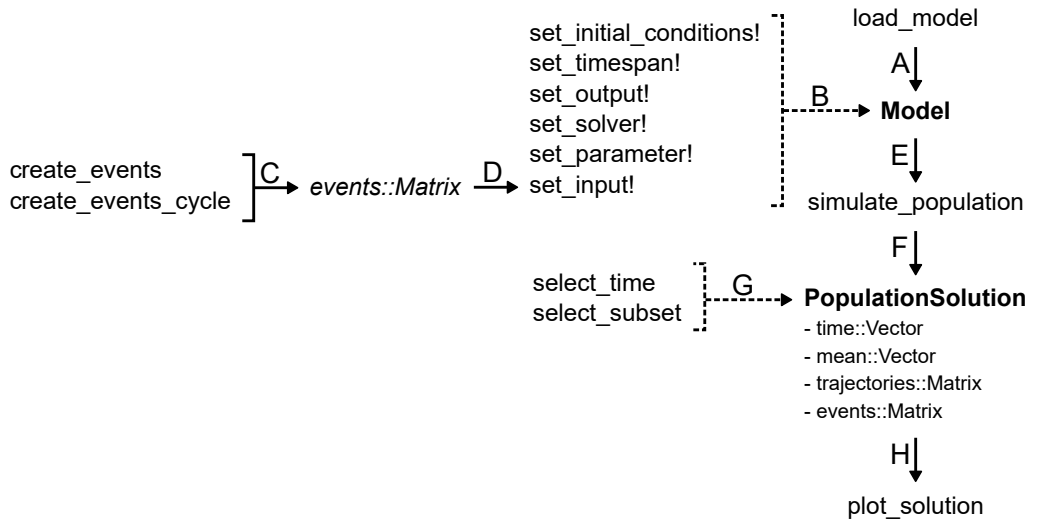
this matrix can be generated automatically using corresponding functions (Listing 3.3 lines 10-12, Figure 3.1C, D).

Once the model is ready, it can be passed to `simulate_population` to perform a simulation of the model and in turn to return an instance of `PopulationSolution`, which represents the numerical solution (Listing 3.3 lines 19, Figure 3.1E, F). `simulate_population` can also perform repeated parallel simulations to simulate a population of identical oscillators. A heterogeneous population can be simulated as well by passing additionally a vector with varying parameter values. The result of the simulation is an instance of `PopulationSolution` that contains information on individual trajectories, population-level mean as well as time vector and input signal in the matrix form described above. The instance of `PopulationSolution` can be further edited by corresponding functions, for example, to select a specific time frame or a subset of oscillators (Figure 3.1G). Finally, the results can be plotted using `plot_solution` (Listing 3.3 line 22, Figure 3.1H).

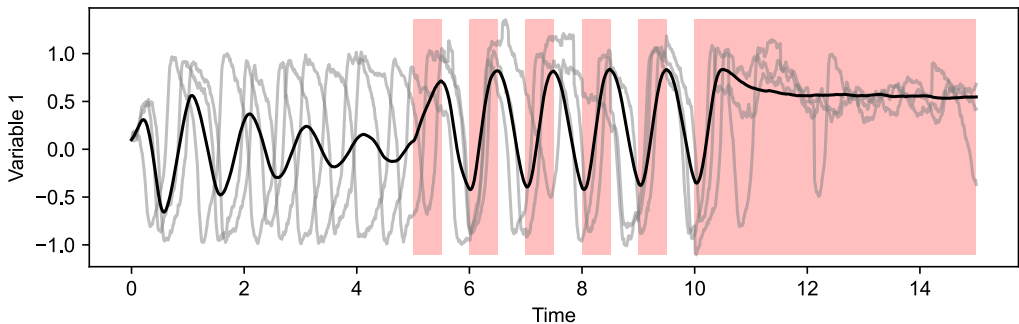
### 3.3.2 Model and input parameter scans

To perform a parameter scan, first, a model needs to be loaded and set up (Listing 3.4 lines 5-10, Figure 3.3A), which was described in detail in the previous section. As a second step a simulation function is created (Listing 3.4 lines 12-16, Figure 3.3B). The simulation function takes a model as input, performs simulation, and returns values for metrics described in Section 2.2. The simulation





**Figure 3.1:** Workflow for a simple population simulation. A code example following this workflow is in Listing 3.3.



**Figure 3.2:** The output of the code showed in Listing 3.3. Black depicts the population mean, gray lines depict a random selection of 5 oscillators from the population, and red boxes indicate when the input signal is active.

function is together with the model passed to a scan and returns a dataframe that provides a list of iterated parameters and corresponding metrics values (Listing 3.4 lines 18-23, Figure 3.3C, D, E). Depending on the chosen scanning function, the final results can represent a parameter scan (`scan`, Figure 2.1), Arnold tongue (`scan_arnold`, Figure 2.16), phase response curve (`estimate_prc`, Figure 2.21) or T-cycle phase response curve (`estimate_T_prc`, Figure 2.23). For Arnold tongues also a dedicated plotting function is provided to transform the dataframe into a color-coded matrix plot (Listing 3.4 lines 25-28, Figure 3.3F). The function `scan_arnold` can be also used to estimate and plot Arnold onions (Figure 1.7) as it allows to iterate over all parameters of the input signal, that is the input amplitude, input period, and input duty cycle. On

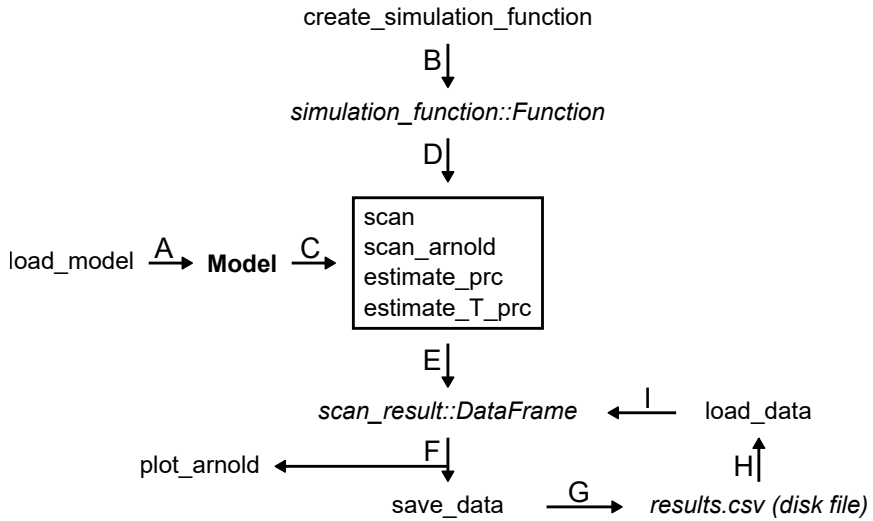
```
1  ## Simulate a population of uncoupled stochastic oscillators
2  # Load the package
3  using OscillatorPopulation
4
5  # Load model and change some of its parameters
6  model = load_model("van-der-pol", "sde")
7  set_parameter!(model, ["I", " $\sigma$ "], [0.5, 0.08])
8
9  # Generate input signal (events) and add it to the model
10 events = create_events(
11     [(:DD, 5), (:LD, 5, 0.5, 0.5), (:LL, 5)]
12 )
13 set_input!(model, events)
14
15 # Set maximum integration time to the end of the input signal
16 set_timespan!(model, events[end])
17
18 # Simulate population of 1000 oscillators
19 solution = simulate_population(model, 1000, seed=3)
20
21 # Plot solution
22 plot_solution(solution)
```

**Listing 3.3:** This example demonstrates how to set up and simulate a population of oscillators under external forcing. (Line 11): The array of tuples passed to `create_events` specifies the input signal. Here, `(:DD, 5)` means signal at 0 for 5 time units, `(:LD, 5, 0.5, 0.5)` means 5 cycles of a square signal consisting of cycles with 0.5 time units at 1 and 0.5 time units at 0, and `(:LL, 5)` means signal at 1 for 5 time units. The symbols are adapted from circadian research (DD - dark-dark, LD - light-dark, LL - light-light). (Line 19): The keyword `seed` specifies the seed for the random number generator, allowing thus for reproducible stochastic calculations. (Line 22): The output of this line can be seen in Figure 3.2.

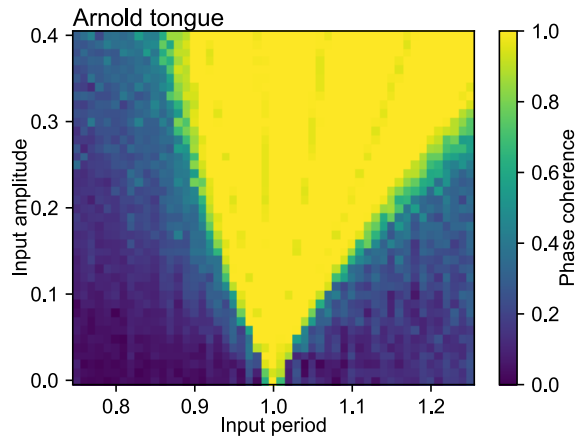
multi-core systems, the individual points of a scan are estimated in parallel. The scan result can be saved to a CSV file for later reuse (Figure 3.3F, G, H, I).

### 3.3.3 Optimization

As a first step, one needs to load and set up a model as shown in Section 3.3.1 (Listing 3.5 lines 5-7, Figure 3.5A). The prepared model, possibly together with fitting data, can be used to generate a cost function for the optimization (Listing 3.5 lines 9-20, Figure 3.5B, C, D). There are three optimization functions implemented, each corresponding to a single step of the step-wise fitting algorithm outlined in Section 2.7.2. The created cost function is next passed to `optimize`



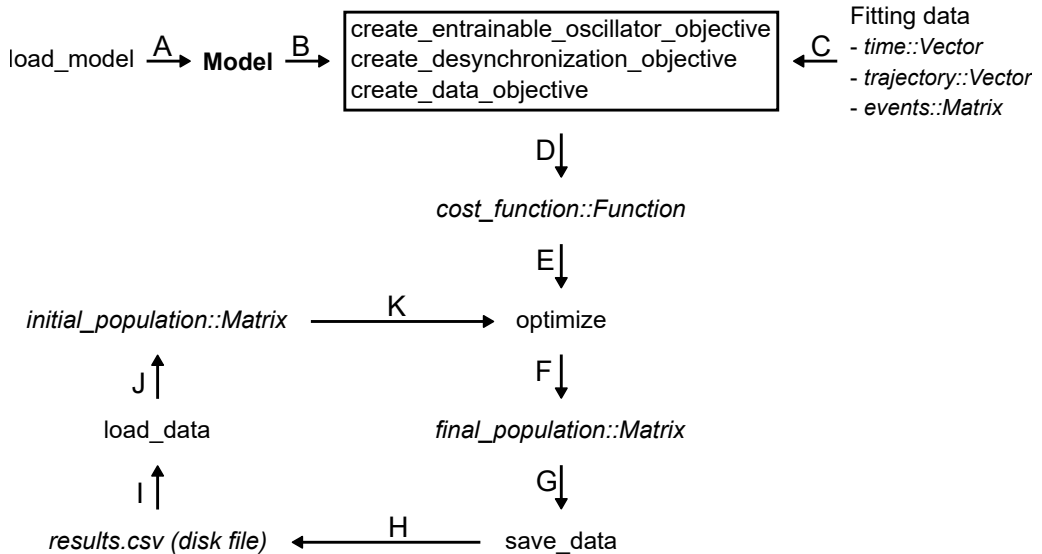
**Figure 3.3:** Workflow for scans over model and input signal parameters. A code example following this workflow is in Listing 3.4.



**Figure 3.4:** Output of the code showed in Listing 3.4.

```
1  ## Estimate and plot Arnold tongue
2  # Load the package
3  using OscillatorPopulation
4
5  # Load model
6  model = load_model("van-der-pol", "sde")
7
8  # Set noise intensity and end simulation time
9  set_parameter!(model, " $\sigma$ ", 0.005)
10 set_timespan!(model, 100.0)
11
12 # Simulation function
13 simulation_function = create_simulation_function(
14     ["phase_coherence"],
15     trajectories=10
16 )
17
18 # Scan a range of input amplitudes and periods
19 arnold_tongue = scan_arnold(model, simulation_function,
20     input_amplitudes=0:0.01:0.4,
21     input_periods=0.75:0.01:1.25,
22     show_progress=true
23 )
24
25 # Plot the resulting Arnold tongue
26 plot_arnold(arnold_tongue, "tongue",
27     property_name="phase_coherence"
28 )
```

**Listing 3.4:** This script illustrates how to estimate and plot the Arnold tongue for a population of 10 stochastic Van der Pol oscillators. (Line 14): The array of strings specifies, which metrics to estimate. (Line 15): Keyword `trajectories` specifies the number of oscillators in the population. (Line 22): Keyword `show_progress` will display a progress bar in the terminal, which is convenient, if the simulation time is expected to be long. (Lines 25-28): The string "tongue" specifies that the user wish to plot an Arnold tongue. Another option can be "onion" to plot an Arnold onion. Keyword `property_name` indicates, which estimated metric should be plotted. The resulting plot can be seen in Figure 3.4.



**Figure 3.5:** Workflow for parameter optimization. A code example following this workflow is in Listing 3.5.

function (Listing 3.5 lines 22-26, Figure 3.5E), which provides essentially a wrapper around `BlackBoxOptim.jl` package<sup>8</sup> implementing the differential evolution algorithm. The output of the optimization function is a matrix with all members of the final population and their fitness as determined by the cost function (Figure 3.5F). This final population can be saved as a CSV file and used later as a starting population for another optimization (Figure 3.5G, H, I, J, K).

## 3.4 Package for peak detection

Many algorithms presented in Chapter 2 rely on peak detection in either the model output directly to estimate phase based on peak detection or in a correlation function to find its dominant lag. As peak detection can be important also for other related projects, beyond this thesis, a Julia package<sup>9</sup> inspired by the Matlab's `findpeaks` function<sup>10</sup> was written. The package implements `findpeaks` function that takes a time series as an input and returns its local maxima together with their locations, prominences, and widths (Figure 3.6). The found peaks can be sorted and filtered based on their properties, which provides a powerful tool for peak detection and consequent analysis based on the oscillation peaks.

<sup>8</sup> <https://github.com/robertfeldt/BlackBoxOptim.jl>

<sup>9</sup> <https://github.com/vkumpost/FindPeaks>

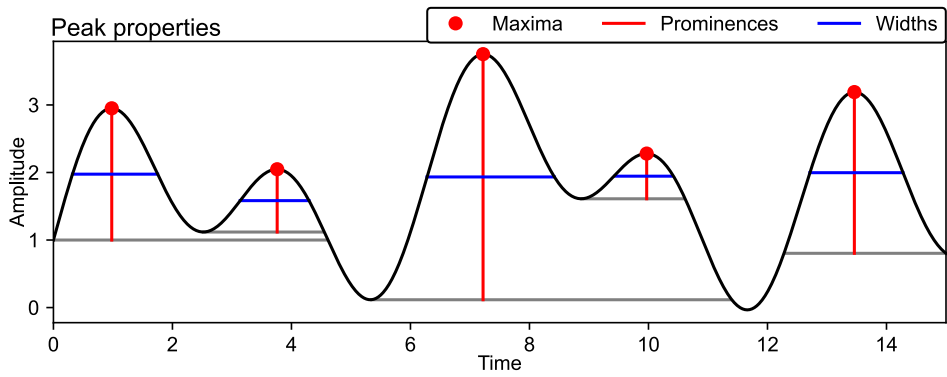
<sup>10</sup> <https://www.mathworks.com/help/signal/ref/findpeaks.html>

```
1  ## Find parameter values that give 24-hour oscillations
2  # Load the package
3  using OscillatorPopulation
4
5  # Load model
6  model = load_model("kim-forger-full", "ode")
7  set_output!(model, 2)
8
9  # Specify parameters and their search range
10 parameter_names = ["A", "dM=dP=dR", "I"]
11 search_range = [(0.0, 100.0), (0.0, 1.0), (0.0, 0.1)]
12
13 # Target period estimated from data
14 target_period = 24.0
15
16 # Construct cost function
17 cost_function = create_entrainable_oscillator_objective(model,
18     target_period,
19     parameter_names=parameter_names
20 )
21
22 # Perform optimization
23 best_candidate, final_population = optimize(cost_function,
24     search_range=search_range,
25     max_steps=5_000
26 )
```

**Listing 3.5:** Find parameter values that give 24-hour oscillations. (Line 7): This line specifies that we are interested only in the second state variable. (Line 10): Placing = between the parameter names as in "dM=dP=dR" says that the parameters  $dM$ ,  $dP$ ,  $dR$  should be equal and thus estimated as a single parameter.

## 3.5 Conclusion

A Julia package for the analysis of the stochastic oscillator populations was implemented. This package allows easy and fast simulation and manipulation of stochastic oscillator models and provides functions for analysis and quantification of the simulation results. The package also implements routines for parameter scans and estimations of PRCs and Arnold tongues. Finally, optimization routines are provided to fit the models to the experimental data. This package should thus comprehend all necessary routines to reproduce all the results in this thesis.



**Figure 3.6:** Visualization of peak properties that can be estimated using the `findpeaks` function from the `FindPeaks` package. The width is measured at the half of prominence.





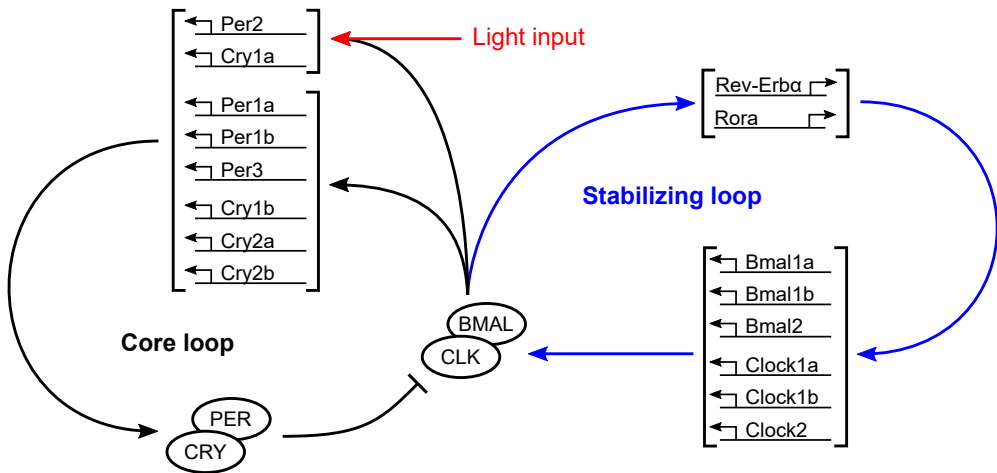
## 4 Application: Entrainment of cellular clocks by light

After the new theoretical findings, methods, and associated implementations have been presented in the previous chapters of the thesis, the following chapter presents an application of the developed methods in the analysis of experimental data from the field of the circadian (24-hour) clock. This addresses the last objective, Objective 5, of the thesis. Parts of this chapter have been adapted from a previously published journal article [2].

### 4.1 Introduction

The circadian clock is an inner timekeeping mechanism that coordinates various biological processes with respect to the time of the day [10]. The circadian timekeeping is maintained and regulated at the level of individual cells by a set of clock genes and proteins that functionally connect in a negative feedback loop (Figure 4.1) [133]. To ensure the correct timing of the circadian events, the inner clock can adjust its phase to environmental signals (zeitgebers) in the process known as entrainment [109]. Although various environmental cues (food, temperature, social interactions) can act as a zeitgeber, light is the most prominent [134]. Light entrainment is an important research area as misalignment of the circadian clock with the environmental light-dark cycle can lead to mood and cognitive disorders [135] as well as metabolic disorders and cancer [136].

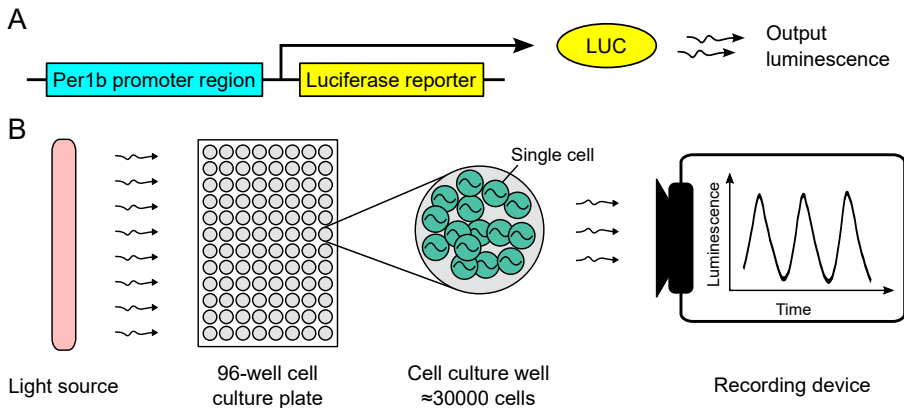
The circadian clock in zebrafish cells is directly light-sensitive, thus providing an attractive laboratory model for the analysis of light entrainment [138]. This is in contrast to mammalian cells that need to be synchronized by pharmacological treatments [139, 140] and cannot be regulated noninvasively simply by changing lighting conditions. The zebrafish light-responsive cell lines are suitable for high-throughput screening as well as studies of the transcriptional control mechanisms mediating light entrainment [141]. In this regard, bioluminescent reporter assays (Figure 4.2) have been established in zebrafish cell lines and enable the non-invasive assessment of dynamic changes in clock gene transcription at a high temporal resolution over the course of light exposure protocols [142]. The circadian regulation in those assays can be easily



**Figure 4.1:** The circadian clock is generated at the level of individual cells by a complex network of genes. The core motif consists of a negative feedback loop of multiple Period (Per) and Cryptochrome (Cry) genes activated by the active CLK-BMAL complexes. In turn, Per and Cry genes produce PER and CRY proteins that act as negative elements that inhibit the function of CLK-BMAL complexes thus closing the negative feedback loop. In zebrafish, a subset of Per and Cry genes is light-driven conveying information about the environmental light levels to the clock. The core loop is supplemented with a stabilizing loop that regulates the levels of the active elements BMAL and CLOCK (CLK) [137].

perturbed by mutations or pharmacological treatments, and the effect can be seen as changes in the luminescence data. However, the time traces obtained from the bioluminescent reporter assays represent a population-level mean from thousands of noisy, uncoupled oscillators, which obscures the actual effect of mutations or pharmacological treatments on the changes in the underlying genetic regulatory network [33]. Here, a mathematical modeling approach that could map the population-level changes to the changes in the single-cell parameters would be beneficial in the correct analysis and quantification of the bioluminescent reporter assays.

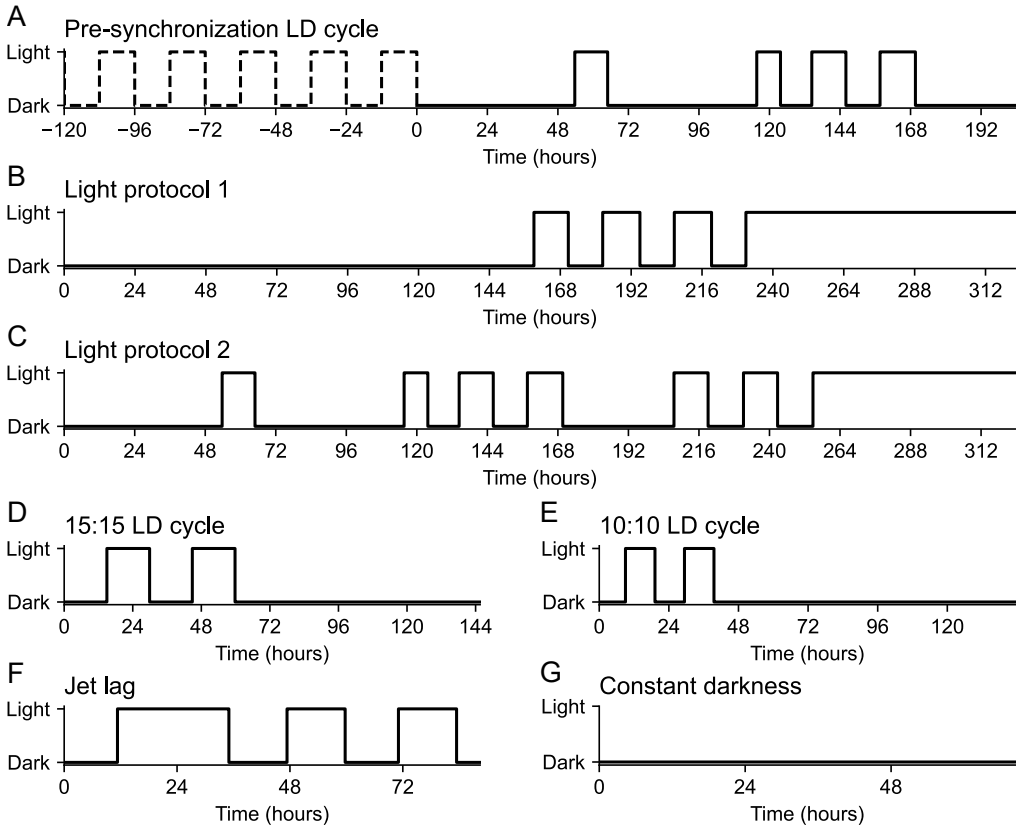
In this chapter, a minimal mathematical model of the zebrafish circadian regulation is used to quantify the time traces obtained from the bioluminescent reporter assays. The model is used to map changes in population-level luminescence data to changes in the model parameters that represent single cells. In particular, the model parameters are adjusted to infer the impact of various pharmacological treatments on core clock dynamics at the cellular level. This work thus provides a tool to characterize core clock dynamics, with the ability to provide hypotheses on how population-level luminescence signals emerge from single-cell behavior. This paves the way toward model-based, large-scale screens for genetically or pharmacologically-induced modifications affecting the degree of synchronization of single-cell circadian oscillators.



**Figure 4.2:** Continuous changes in the circadian gene expression can be recorded using bioluminescent reporter assays. (A) The zebrafish cells are transfected with the luciferase reporter gene attached to the Per1b promoter region. As the luciferase reporter is regulated by the same promoter region as Per1b, the output luminescence corresponds directly to the activation of the Per1b gene. (B) In bioluminescent reporter assay experiments, a controllable light source illuminates a 96-well cell culture plate. Each of the 96 wells of this plate can contain a cell culture of approximately 30000 cells transfected with a luciferase reporter (panel A). In regular intervals, the cell culture plate is automatically moved into a recording chamber and output luminescence is recorded from the individual wells. This leads to a final recording of up to 96 time traces, where each represents the population-level dynamics of 30000 independent cells. Panel B is adapted from the previous publication [2].

## 4.2 Dataset description

The luminescence data used in this chapter come from zebrafish cell lines stably transfected with a luciferase construct of a major circadian gene Per1b [141]. These zebrafish cell lines can be entrained by direct light exposure and are widely used to study the transcriptional mechanisms of light entrainment [143, 144]. The resulting time traces represent periodical activation of the Per1b gene. However, one must keep in mind that the recorded gene activation represents a population average of 30000 cells. The data were recorded under various light-forcing protocols (Figure 4.3) and contain multiple experimental repeats and pharmacological treatments (Table 4.1). The luminescence time traces were recorded with a resolution of approximately 40 minutes, which provides a fine time resolution compared to the circadian periods of 24 hours. Except for untreated (wild-type) cell lines also the effect of 6 pharmacological compounds has been tested: Forskolin (FOR), dibutyryl cAMP (DBC), epidermal growth factor (EGF), U0126, phorbol-12-myristate-13-acetate (PMA), and ro-318220 (RO) [145]. Those pharmacological compounds have been chosen to target major signaling pathways that are involved in the regulation of light entrainment [146, 147, 148, 149]. In summary, the experiments provide a sufficient number of experimental data to train and test a mathematical model.



**Figure 4.3:** The recorded data include a variety of forcing light protocols to explore the dynamics of light entrainment. (A) Before the beginning of each protocol, the cells were entrained using a regular light-dark (LD) cycle (dashed line). These first LD repetitions are not part of the final recording. The beginning of each protocol starts at the end of the light period (lights went off at time 0). (B) Light protocol 1 starts with 6 days of constant darkness followed by three 12:12 LD cycles and ends with 4 days of constant light. (C) Light protocol 2 consists of several irregular LD cycles and ends with constant light. (D) 15:15 LD cycle protocol consists of two 30-hour days followed by constant darkness. (E) 10:10 LD cycle consists of two 20-hour days followed by constant darkness. (F) The jet lag protocol consists of a jet-lag-like phase reversal (long day) followed by two days of the 12:12 LD cycle. (G) In the constant darkness protocol, the cells were kept in constant darkness for the whole time of the recording. The figure is partly adapted from the previous publication [2].

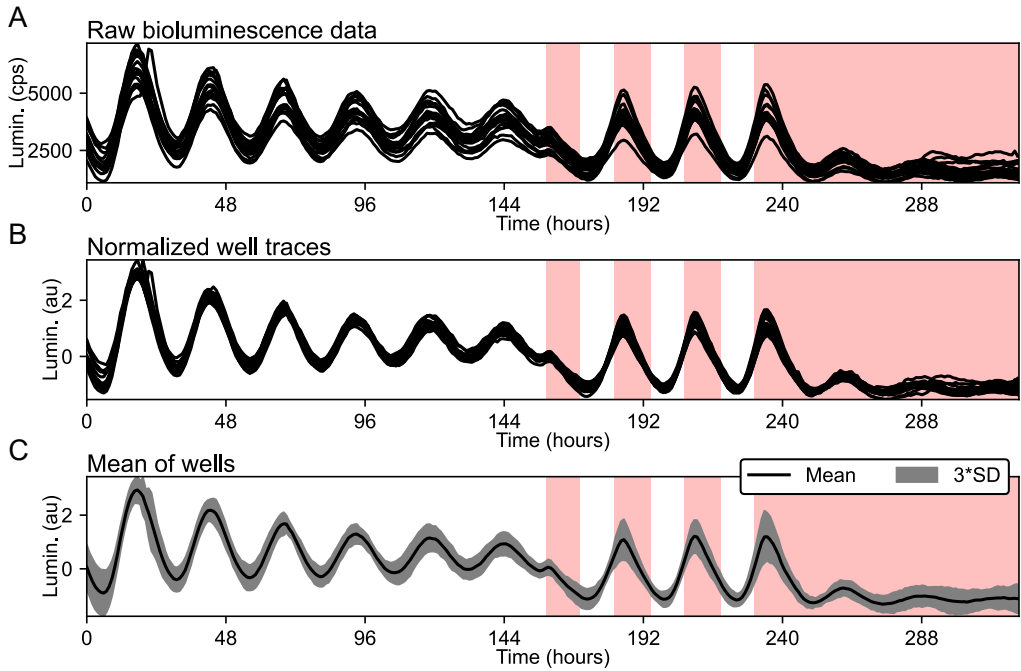
Plate	Well repeats
Light protocol 1	16x untreated cells
Light protocol 2	4x untreated cells
15:15 LD cycle	32x untreated cells
10:10 LD cycle	16x untreated cells
Jet lag (set A)	4x DMSO control treatment 12x FOR (4x 5 $\mu$ M, 4x 10 $\mu$ M, 4x 15 $\mu$ M) 12x DBC (4x 0.5 mM, 4x 1 mM, 4x 3 mM) 12x U0126 (4x 10 $\mu$ M, 4x 20 $\mu$ M, 4x 40 $\mu$ M)
Constant darkness (set A)	4x DMSO control treatment 12x FOR (4x 5 $\mu$ M, 4x 10 $\mu$ M, 4x 15 $\mu$ M) 12x DBC (4x 0.5 mM, 4x 1 mM, 4x 3 mM) 12x U0126 (4x 10 $\mu$ M, 4x 20 $\mu$ M, 4x 40 $\mu$ M)
Jet lag (set B)	4x DMSO control treatment 12x EGF (4x 30 ng/ml, 4x 50 ng/ml, 4x 80 ng/ml) 12x PMA (4x 0.5 $\mu$ M, 4x 1 $\mu$ M, 4x 3 $\mu$ M) 12x RO (4x 2 $\mu$ M, 4x 5 $\mu$ M, 4x 8 $\mu$ M)
Constant darkness (set B)	4x DMSO control treatment 12x EGF (4x 30 ng/ml, 4x 50 ng/ml, 4x 80 ng/ml) 12x PMA (4x 0.5 $\mu$ M, 4x 1 $\mu$ M, 4x 3 $\mu$ M) 12x RO (4x 2 $\mu$ M, 4x 5 $\mu$ M, 4x 8 $\mu$ M)

**Table 4.1:** Description of experimental repeats in the recorded luminescence data. Each cell culture plate consists of 96 wells (Figure 4.2) that can be used to record up to 96 cell cultures in parallel under the same light protocol. The first four light protocols (light protocol 1, light protocol 2, 15:15 LD cycle, and 10:10 LD cycle) were used to explore the light entrainment of untreated cells. The last two light protocols (jet lag and constant darkness) were used to explore the light entrainment under various pharmacological treatments. As the number of tested pharmacological treatments was high, those had to be split between two plates (set A and B). This table is adapted from the previous publication [2].

## 4.3 Dataset normalization

We want normalization so the pharmacological treatments are comparable among the plates.

In the luminescence recordings, the shape of the time traces seems to be well conserved for individual repeats but the absolute amplitude shows great variability (Figure 4.4A). Furthermore, inherent experimental errors cause variations in mean and amplitude also between individual plates [145]. In order to be able rigorously to compare the individual traces and use them for model fitting and validation, a reliable normalization strategy is required. Z-score removes the



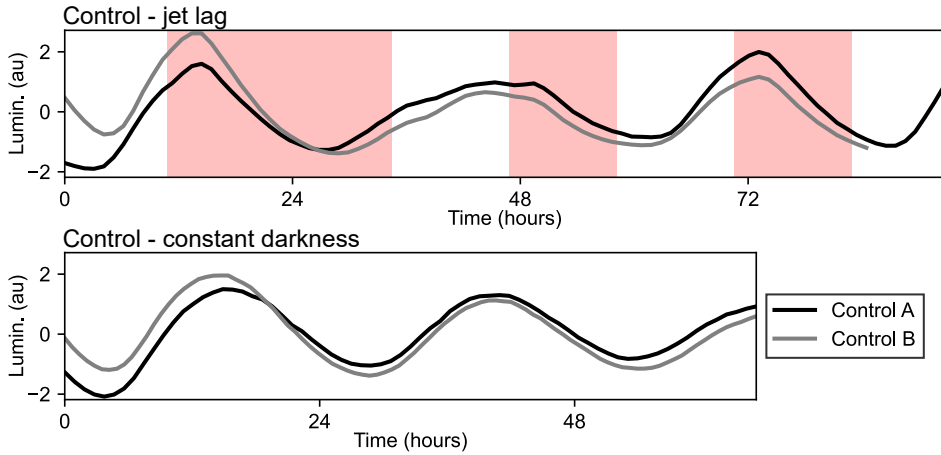
**Figure 4.4:** The luminescence data were normalized by Z-score to reduce the variability among the individual traces. This example corresponds to Light protocol 1 (Figure 4.3). (A) Data from 16 recording wells in the raw units as obtained from the recording device (counts per second). (B) Applying Z-score to the individual traces decreases the variability among the measurements from different wells. (C) The mean of normalized well traces was used for further analysis and data fitting (SD = standard deviation). This figure is adapted from the previous publication [2].

mean and normalized the standard deviation of the time trace to 1 while maintaining the shape of the traces unchanged. Z-score is defined as

$$z(x) = \frac{x - \bar{x}}{s_x}, \quad \bar{x} = \frac{1}{N} \sum_{i=1}^N x[i], \quad S_x = \sqrt{\frac{1}{1-N} \sum_{i=1}^N (x[i] - \bar{x})^2}, \quad (4.1)$$

where  $\bar{x}$  is sample mean and  $S_x$  is sample standard deviation of trace  $x$ . Normalizing the traces using the Z-score reduced the variability among the traces to the minimum (Figure 4.4).

The Z-score normalization works well for the recordings of untreated cell lines but cannot be directly used for pharmacological treatments whose effect can manifest in changes in the mean and amplitude of the recorded gene expression (Figures B.1, B.2). Thus a normalization strategy that would minimize the variations among the repeats of the same treatment but would still



**Figure 4.5:** Control treatments from different plates remain the similar waveform after the normalization.

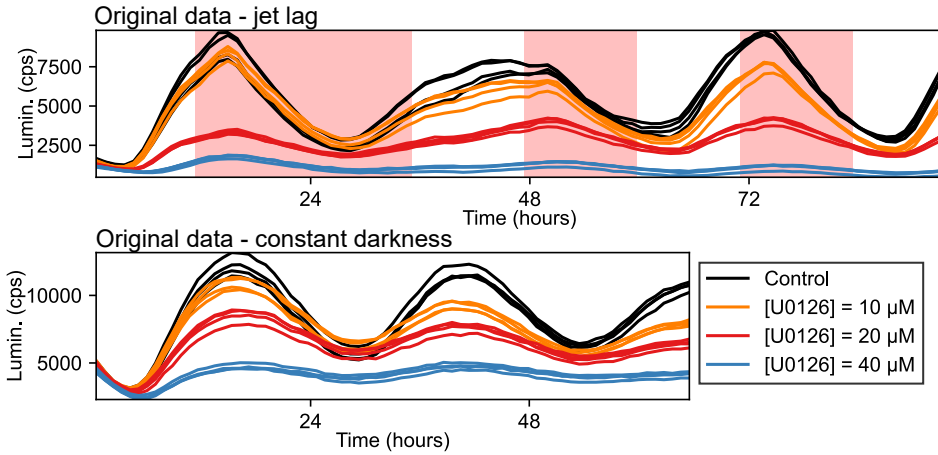
preserve the differences among the individual treatments is required. To achieve this, one can use an adjusted Z-score defined as

$$z_d(x) = \frac{x - \bar{x}}{S_x} \frac{S_d}{S_c} + \frac{\bar{x}_d - \bar{x}_c}{S_c} \quad (4.2)$$

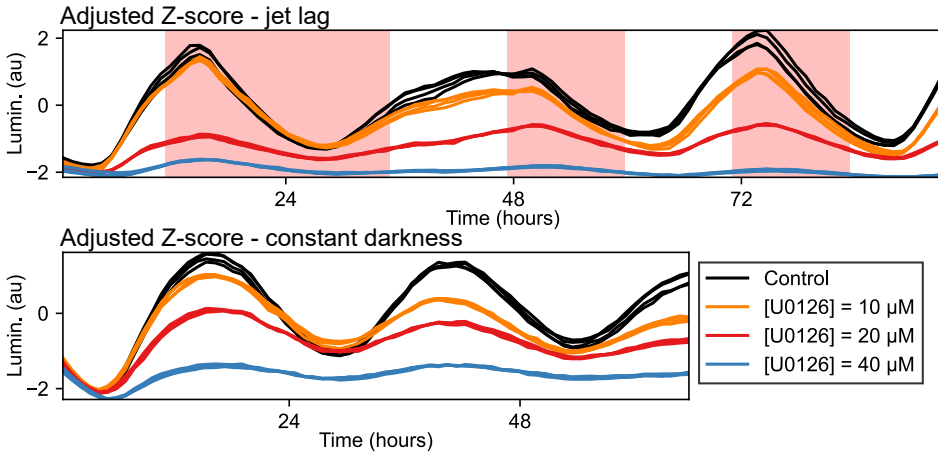
where  $\bar{x}$  and  $S_x$  are the sample mean and sample standard deviation of trace  $x$  as showed in Equation 4.1,  $\bar{x}_d$  and  $S_d$  are the mean of means and mean of standard deviations for all traces of the corresponding compound and dose on the plate, and  $\bar{x}_c$  and  $S_c$  are mean of means and mean of standard deviations for all control traces (without treatment) on the plate. An important property of this normalization strategy is that it reduces to standard Z-score if  $S_d = S_c$  and  $\bar{x}_d = \bar{x}_c$ . The values for control can be thus directly compared to the values of untreated cell lines from other plates normalized by the standard Z-score (Figure 4.5). Visual comparison of raw and normalized data showed low variance in traces of the same compound while correctly preserving the relative changes of the treated traces to control (Figure 4.6 and 4.7). If we used a standard Z-score, the difference in amplitude diminish (Figure 4.8). As a consequence, we have a normalization method that allows us to compare control traces across the different plates and pharmacological treatments relative to those control traces.

## 4.4 Model derivation

The model of the zebrafish circadian regulation is based on the simple Kim-Forger model [48] that was already introduced and applied in Sections 1.2.2 and 2.1. In the application, the Kim-Forger



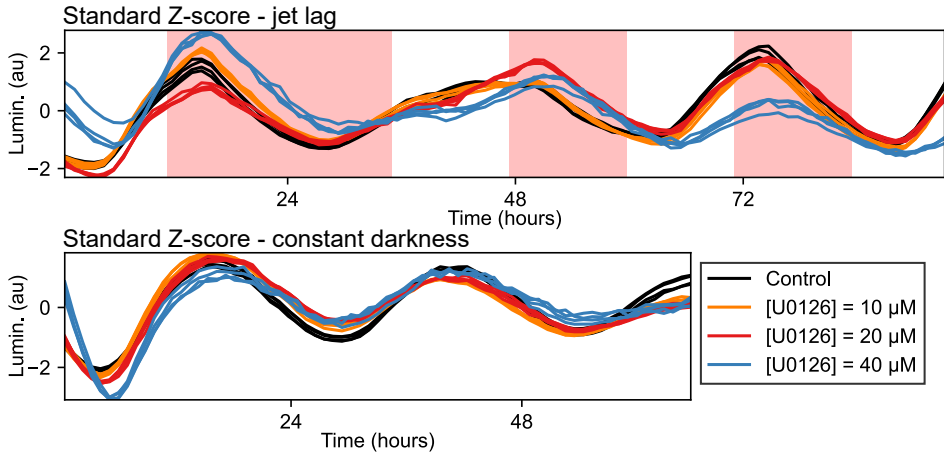
**Figure 4.6:** Example of unprocessed data for a pharmacological treatment with U0126.



**Figure 4.7:** Example of normalized data (adjusted Z-score) for a pharmacological treatment with U0126.

model can be thought as representing the oscillatory behavior of three clock elements connected in a negative feedback loop. The variables of the model are arbitrarily denoted as mRNA  $M$ , protein  $P$ , and repressor  $R$  (Figure 4.9). The three state variables implicitly encode the delay that





**Figure 4.8:** Standard Z-score is not suitable for the pharmacological treatments as it removes the amplitude difference among the individual treatments (compare with the adjusted Z-score normalization in Figure 4.7).

is necessary to produce the oscillatory behavior [13] and thereby provide a minimal model of the circadian clock [43, 46]. The simple Kim-Forger model reads

$$\frac{dM}{dt} = v_M f(R, A) - d_M M \quad (4.3a)$$

$$\frac{dP}{dt} = v_P M - d_P P \quad (4.3b)$$

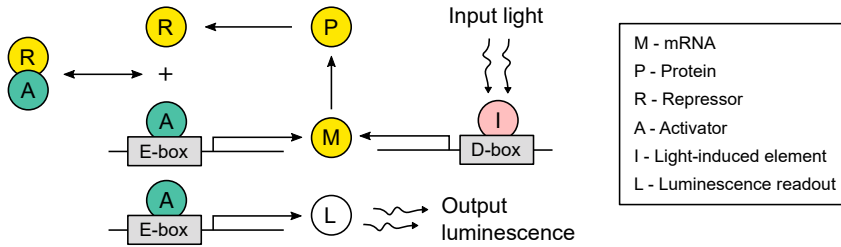
$$\frac{dR}{dt} = v_R P - d_R R \quad (4.3c)$$

where  $A$  denotes activator concentration,  $v_*$  are production rates, and  $d_*$  are degradation rates.

In the zebrafish circadian clock, the repressors physically interact with the activators and thereby inhibit the transcriptional induction of their own genes [137]. This corresponds to the physical meaning of the Kim-Forger model, where the repressor  $R$  binds to the activator  $A$ , forming an inactive  $RA$  complex and thus preventing the binding of the activator to the E-box enhancers, resulting in inhibition of gene expression. The according protein sequestration function that describes the transcriptional activation is defined as [48, 49]

$$f(R, A) = \begin{cases} 1 - \frac{R}{A} & R \leq A \\ 0 & R > A \end{cases} \quad (4.4)$$

where  $R$  and  $A$  are repressor and activator concentrations, respectively.



**Figure 4.9:** A mathematical model represents the transcription-translation feedback loop that generates the circadian rhythm in individual zebrafish cells. The core loop consists of positive (green) and negative (yellow) elements. The activator (A) binds to the promoter region (E-box) of a circadian gene, which starts the gene transcription and so production of the mRNA (M). M translates into a protein (P), which is transported back to the cell nucleus where it acts as a repressor (R). In the nucleus, R binds A forming an inactive RA complex. This prevents A from binding to the promoter of the gene leading to a decrease in the production of negative elements and the cycle repeats. Zebrafish possess additional light-driven genes that are regulated by D-box elements. Increased light levels in the environment lead to the activation of the D-box elements and thus increases the concentration of the negative elements. In the bioluminescence reporter assays, the cells are transfected with a reporter gene that has the same promoter as the circadian gene. The luminescence output is assumed to be proportional to the gene activation. This figure is adapted from the previous publication [2].

As the luminescence data are normalized and thus the amplitude has arbitrary units, the variables of the model can be scaled as

$$M = v_M M^*, P = v_M v_P P^*, R = v_M v_P v_R R^*, \quad (4.5)$$

where  $M, P, R$  are the original model variables and  $M^*, P^*, R^*$  are the scaled model variables. The rescaling of the model variables leads to the removal of production rates from the equations. This reduces the number of parameters by 3 effectively avoiding over-parametrization and connected problems of non-unique results of the fitting. For convenience, the scaled variables were relabeled back as  $M := M^*, P := P^*, R := R^*$ , which leads to the scaled model equations

$$\frac{dM}{dt} = f(R, A) - d_M M \quad (4.6a)$$

$$\frac{dP}{dt} = M - d_P P \quad (4.6b)$$

$$\frac{dR}{dt} = P - d_R R. \quad (4.6c)$$

The light input is implemented as a light-induced element  $I$ , whose activation increases the production of mRNA  $M$ . In the equations, this is encoded as an additive term to the first equation as

$$\frac{dM}{dt} = f(R, A) - d_M M + iI \quad (4.7a)$$

$$\frac{dP}{dt} = M - d_P P \quad (4.7b)$$

$$\frac{dR}{dt} = P - d_R R, \quad (4.7c)$$

where parameter  $i$  is light sensitivity and  $I$  is a time-dependent function that changes its value between 0 and 1 and represents turning on and off of the light source. Similar extensions to the minimal models of the circadian clock are commonly used to represent light stimuli [56, 150, 151]. In our model, the additive light term  $I$  represents light-driven activation of an additional set of light-driven Per and Cry clock genes that is characteristic of the zebrafish circadian clock and is mediated by D-box enhancers [143, 144]. These light-driven genes are specifically activated by light with minimal activation in darkness, and thus most accurately represented by an additive term.

Cellular noise is an intrinsic part of the clock mechanism [76] and is essential to consider when interpreting observations made at the cell population level [33]. We have therefore implemented noise at the single-cell level by additive noise terms [115] in the model as

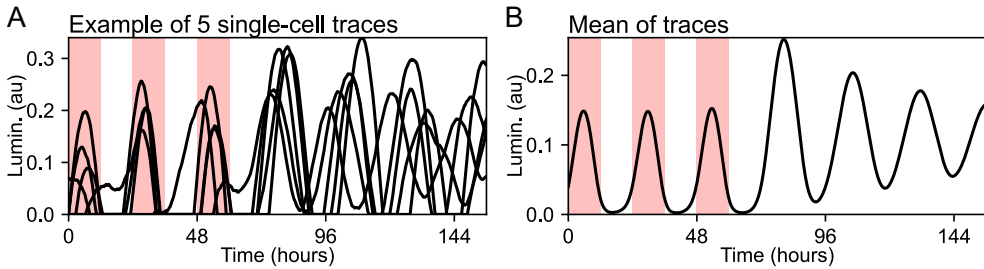
$$\frac{dM}{dt} = f(R, A) - d_M M + iI + \sigma W_M \quad (4.8a)$$

$$\frac{dP}{dt} = M - d_P P + \sigma W_P \quad (4.8b)$$

$$\frac{dR}{dt} = P - d_R R + \sigma W_R, \quad (4.8c)$$

where  $\sigma$  is noise intensity and  $W_i$  are independent Wiener processes (Equation 2.7).

In the experimental setup, cells presumably function as independent oscillators and are implemented accordingly in our model. Each well of a 96-well culture plate contains approximately 30000 independently oscillating cells that contribute to the luminescence signal measured from that well. To mimic the averaging over a population of single-cell oscillators in the luminescence assays, we simulate 30000 instances of the model and obtain the mean value as the final output [127, 58]. The function  $f$  is the quantity in the model that most closely represents the induction of the Per1b-luciferase reporter by E-box enhancers. We thus use the value of the function  $f$  to



**Figure 4.10:** The model represents a population average of 30000 independent cells. (A) Example of 5 single-cell traces. (B) Average of 30000 cells. Notice the loss of oscillation amplitude in the constant darkness that is caused by the desynchronization of the individual oscillators. This figure is adapted from the previous publication [2].

represent the luminescence read-out obtained in our experiments. The luminescence produced by a single cell is then equal to

$$L_j = f(R, A). \quad (4.9)$$

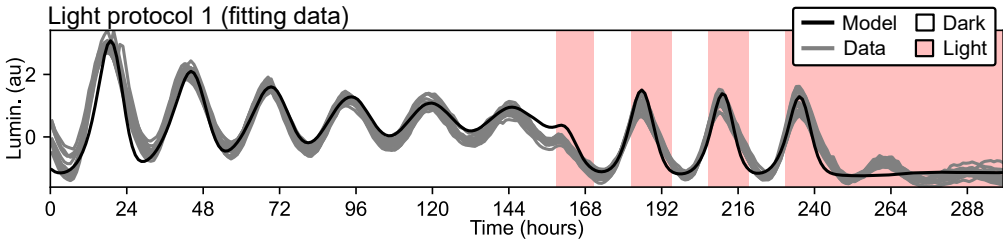
The final model output that corresponds to the luminescence output of a cell culture well is

$$L = \frac{1}{n} \sum_{j=1}^n L_j, \quad (4.10)$$

where  $j$  are independent evaluations of the stochastic model and  $n$  is the number of cells in the population. The necessity to address single-cell dynamics for the population-level recordings was illustrated previously by experiments that show that desynchronization at the level of individual cells leads to the loss of the population-level oscillations although the single cells exhibit sustained oscillations [96]. These experimental results can be reproduced using the presented model (Figure 4.10).

## 4.5 Model fitting and validation

The proposed mathematical model contains 6 free parameters: activator concentration ( $A$ ), degradation rates ( $d_M$ ,  $d_P$ ,  $d_R$ ), light sensitivity ( $i$ ) and noise intensity ( $\sigma$ ). For validation, the model parameters were fitted to the data from untreated cell lines experiments using a custom fitting algorithm based on an evolutionary optimizer (Section 2.7.2). Previous experiments have shown that, in the case of untreated cell lines, the individual cells behave as sustained oscillators [96].



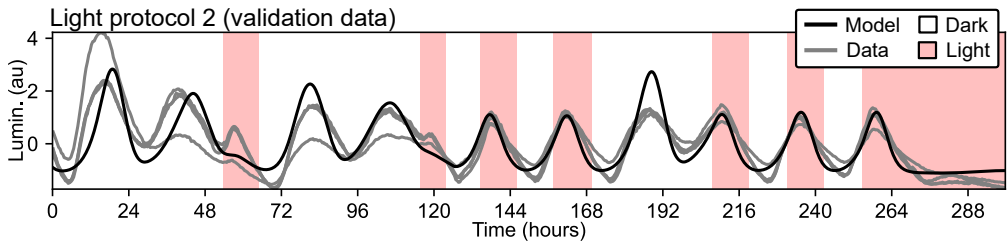
**Figure 4.11:** Data from light protocol 1 were used as fitting data for the model. Visual inspection of the results shows that the model can well reproduce all characteristics of the original data: damping of the oscillation amplitude in constant darkness, synchronization of the phase of the rhythm upon transfer from constant darkness to an LD cycle, and oscillation repression under constant light. This figure is adapted from the previous publication [2].

As the probability of the oscillations is the highest when the degradation terms are equal [152], a constraint on equal degradation terms

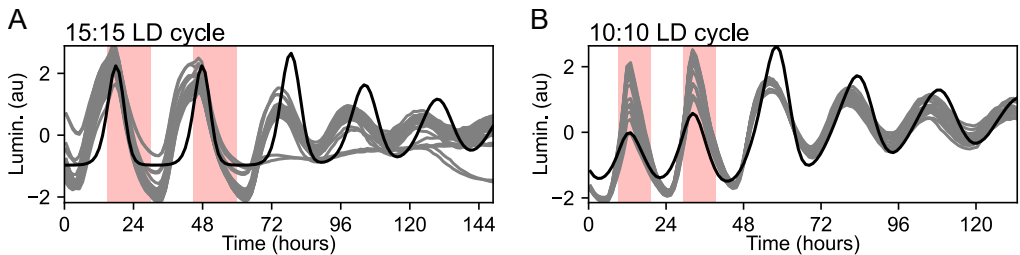
$$d = d_M = d_P = d_R \quad (4.11)$$

was used to simplify the fitting process by reducing the number of free parameters from 6 to 4. The constraint in Equation 4.11 is based on preliminary experiments that showed that unequal degradation terms do not lead to a better fit to the presented data. Of course, if other data are available and the fit is not satisfactory, this assumption can be relaxed. The model parameters were first fitted to the untreated cell lines under light protocol 1 (Figure 4.11) showing three characteristic phenomena: attenuation of the oscillation amplitude during free-running conditions in constant darkness, synchronization of the phase of the rhythm upon transfer from constant darkness to an LD cycle, and oscillation repression under constant light. The goodness of fit was evaluated by the model efficiency coefficient ( $E_f$ , Section 2.7.4). The resulting model efficiency on this fitting data was high (16 experimental repeats, mean  $\pm$  standard deviation,  $E_f = 0.89 \pm 0.03$ ).

For the validation of the estimated parameters, the model was simulated without further readjustment of model parameters for experiments under light protocol 2 (Figure 4.12), resulting in fair model efficiency (4 experimental repeats,  $E_f = 0.63 \pm 0.21$ ). Upon visual inspection, the model captured all key behaviors observed in the experiments. The performance of the model was further tested under LD cycles with period lengths significantly longer and shorter than 24 hours (30 hours, 15:15 LD cycle; and 20 hours, 10:10 LD cycle). In this case, the goodness of fit was lower than under the 12:12 LD cycle experiments (15:15 LD cycle, 32 experimental repeats,  $E_f = 0.24 \pm 0.20$ , Figure 4.13A; 10:10 LD cycle, 16 experimental repeats,  $E_f = 0.60 \pm 0.11$ , Figure 4.13B), mainly due to mismatched profiles of each expression cycle. Nevertheless, the model was still entrained and oscillated with the corresponding period of these long and short LD cycles, which corresponds to the previous experiments [141]. As the last validation step, the phase response curve (PRC, Section 2.4.1) was estimated for the fitted model and 12-hour



**Figure 4.12:** Data from light protocol 2 were used as validation data to verify the parameter fit for the model. Visual inspection of the results shows that the model can well reproduce data that were not used for the parameter fitting. This figure is adapted from the previous publication [2].

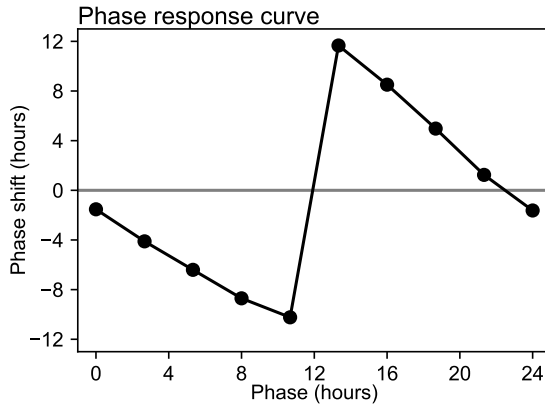


**Figure 4.13:** Simulating the model under LD cycles that differ significantly from the 12:12 LD cycle leads to a mismatch in the data and predicted curve shapes. (A) For the 15:15 LD cycle, the model curve differs significantly from the data, however, the main characteristics are maintained. The model is still entrained by this cycle and the peak seems to occur at the same time as in the data. (B) For the 10:10 LD cycle, the entrainment amplitude is significantly smaller than in the data. However, the model is still entrained by this LD cycle. This figure is adapted from the previous publication [2].

long light pulses (Figure 4.14). The resulting high-amplitude type 0 PRC [64] is also consistent with previously reported experimental PRC using those zebrafish cell lines [141]. Therefore, the model can reproduce and predict characteristics of the entrainment of the zebrafish cellular clock under complex lighting conditions and loses its accuracy only for entrainment period lengths that differ significantly from the natural 12:12 LD cycle.

## 4.6 Model parameters quantify pharmacological treatments

The parameters of the model were next adjusted to infer the effect of different pharmacological compounds on the core clock dynamics. The luminescence data for the pharmacologically-treated



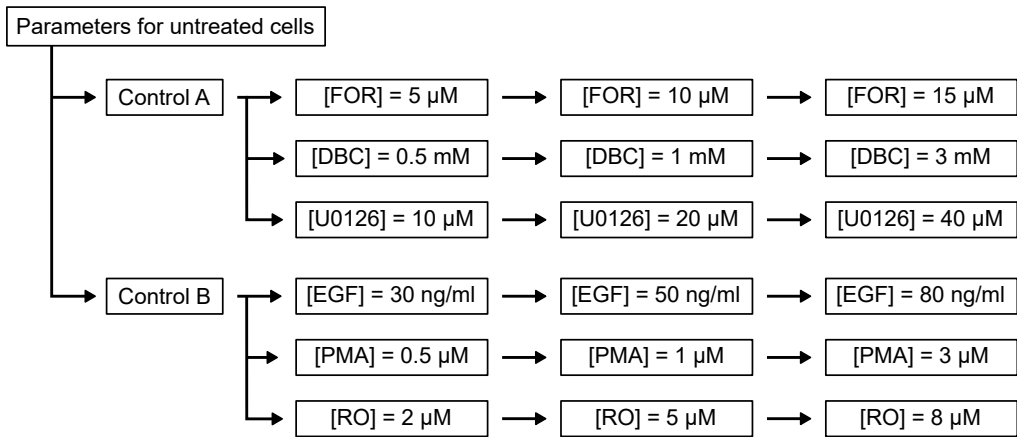
**Figure 4.14:** The estimated phase response curve with large phase shifts (type 0 PRC) corresponds to the previously published experimental PRC using the same zebrafish cell lines [141]. This figure is adapted from the previous publication [2].

cells were recorded under two light conditions: jet lag and constant darkness (Figure 4.7). The cost function considers both experiments in a single function as

$$C_{\text{treatment}} = \frac{1}{N_{JL}} \sum_{i=1}^{N_{JL}} (y_{JL}[i] - \hat{y}_{JL}[i])^2 + \frac{1}{N_{CD}} \sum_{i=1}^{N_{CD}} (y_{CD}[i] - \hat{y}_{CD}[i])^2, \quad (4.12)$$

where  $y_{JL}, y_{CD}$  denote luminescence recordings recorded under the jet lag protocol and in constant darkness respectively and  $\hat{y}_{JL}, \hat{y}_{CD}$  denote model outputs for the jet lag and constant darkness respectively. As the pharmacological treatment can substantially distort the regulatory mechanisms of the circadian clock and thus the individual cells might stop exhibiting sustain oscillations, the constraint on equal degradation rates (Equation 4.11) was relaxed so all three degradation rates can have different values.

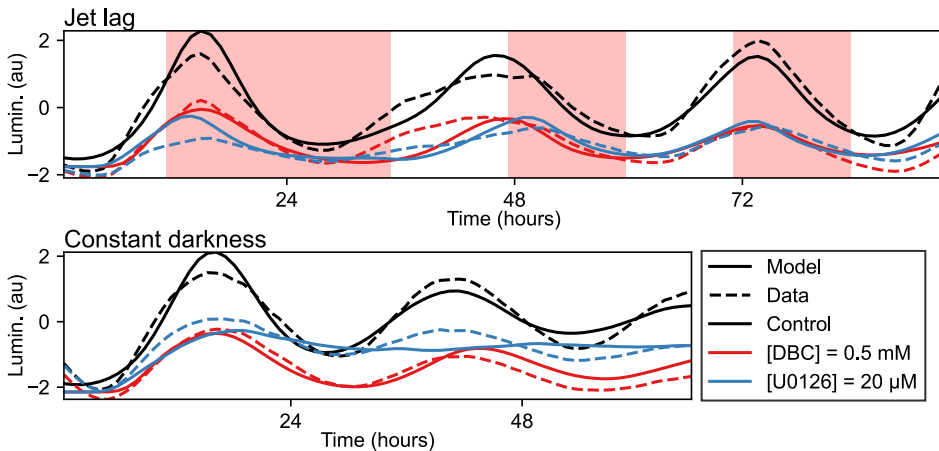
The compounds were administrated right before the beginning of the recording which caused transient behavior observable in the data. To incorporate this into our simulations, we precalculated initial conditions by running the model of the untreated cells for 10 days and then taking the last values as initial conditions for the simulations of pharmacological treatments. The resulted initial conditions correspond to the state of the oscillator at the transition from light to dark. This corresponds to the experimental data used in our work that starts at the beginning of the dark phase. The initial conditions were fixed for the whole process of optimization. Individual compound doses were fitted iteratively from control to the highest concentration of the specific compound (Figure 4.15). At each step, the final optimized population from the previous step was used as the initial population for the next step (Section 2.7.2). Representative fits to the selected



**Figure 4.15:** The data for the pharmacological treatments were fitted step-wise. First, the final population of candidates fitted to the untreated data was used as starting population for the control treatments. The final candidate population for the control treatments was used as the initial population for the lowest compound dose, continuing iteratively for the higher compound doses. This figure is adapted from the previous publication [2].

pharmacological treatments are in Figures 4.16 and 4.17. Table 4.2 shows the goodness of fit for all pharmacological treatments.

The principal component analysis (PCA) of the parameter sets representing the individual compounds revealed that the two main principal components together account for 86 % of the explained

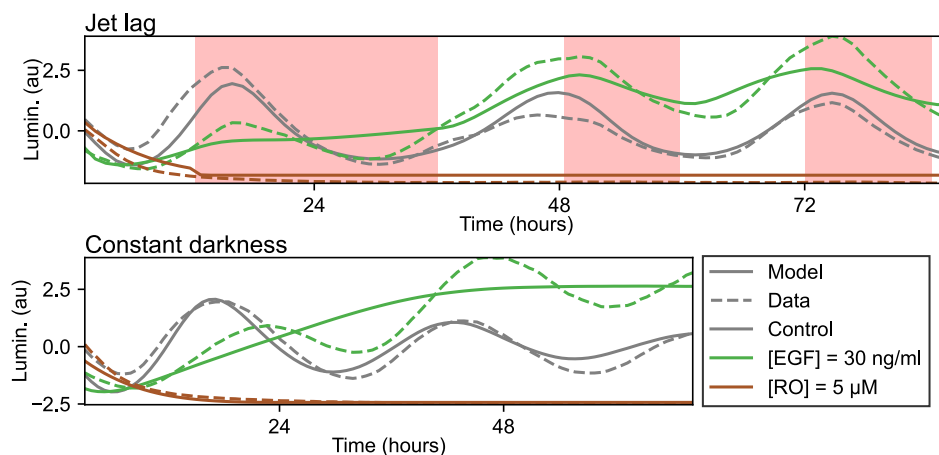


**Figure 4.16:** Pharmacological compounds DBC and U0126 have a similar effect of attenuating the oscillation amplitude under both LD cycle (panel A) and constant darkness (panel B). The model parameters can be adjusted to this. This figure is adapted from the previous publication [2].



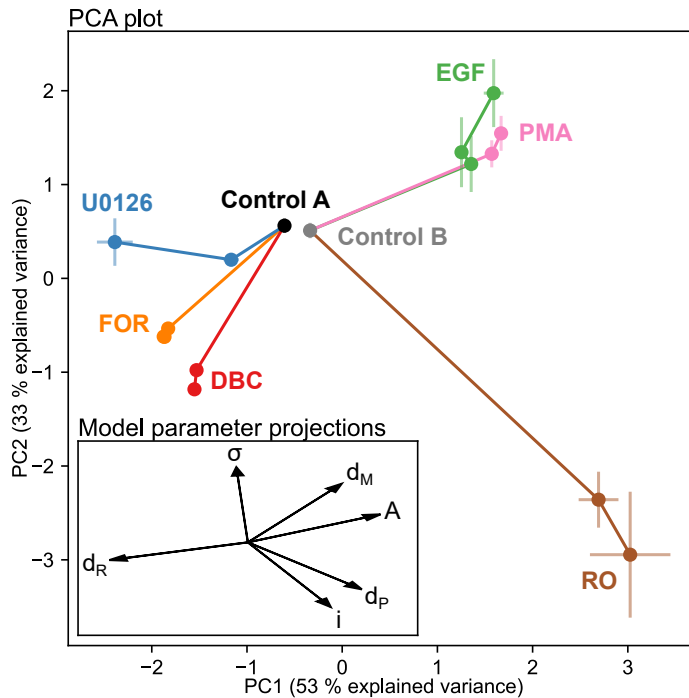
Compound	$E_f$ jet lag	$E_f$ constant darkness
Control A	0.87	0.89
Control B	0.72	0.81
[FOR] = 5 $\mu$ M	0.74	0.86
[FOR] = 10 $\mu$ M	0.74	0.87
[FOR] = 15 $\mu$ M	0.70	0.83
[DBC] = 0.5 mM	0.62	0.73
[DBC] = 1 mM	0.11	0.20
[DBC] = 3 mM	-1.56	-0.20
[U0126] = 10 $\mu$ M	0.86	0.86
[U0126] = 20 $\mu$ M	0.68	0.50
[U0126] = 40 $\mu$ M	-18.19	0.60
[EGF] = 30 ng/ml	0.73	0.73
[EGF] = 50 ng/ml	0.50	0.67
[EGF] = 80 ng/ml	0.69	0.77
[PMA] = 0.5 $\mu$ M	-1.59	-3.98
[PMA] = 1 $\mu$ M	0.83	0.89
[PMA] = 3 $\mu$ M	0.89	0.86
[RO] = 2 $\mu$ M	0.12	-0.04
[RO] = 5 $\mu$ M	0.46	0.84
[RO] = 8 $\mu$ M	0.27	0.84

**Table 4.2:** The goodness of fit for all pharmacological treatments. The goodness of fit was quantified by the model efficiency coefficient defined ( $E_f$ , Equation 2.38) that lives on the interval  $(-\infty, 1]$ . Values near 1 indicate high predictive value of the model while negative values indicate that data mean is a better predictor than model output (Figure 2.48). Most of the fits show a relatively high model efficiency coefficient. Only a few compounds at the highest concentrations could not be fitted ( $E_f < 0$ ) using the simple model. This table is adapted from the previous publication [2].



**Figure 4.17:** Application of the pharmacological compound EGF (green line) leads to a transient behavior with increased amplitude of oscillations. Administration of RO (brown line) leads to the complete removal of oscillations. The model can reproduce the effect of RO and partly also of EGF but struggles with a combination of low-frequency transient and high oscillation amplitude as shown in panel B. This figure is adapted from the previous publication [2].

variance. The relative changes of the six free parameters can be, therefore, depicted in the two-dimensional space defined by the two main principal components (Figure 4.18). In the PCA plot, every compound exhibits a change away from the control conditions. Compound-specific changes were increasingly pronounced at higher concentrations, while the overall direction of parameter change was conserved for each of the compounds. Based on the direction of displacement in the principal component plot, different compounds can be grouped. FOR, DBC, and U0126 are all associated with a decrease in principal component 1, reflected in a loss of the oscillation amplitude that is similar for all three compounds (Figure 4.16, Figure B.3A, B, C). PMA and EGF are associated with an increase in principal component 1, reflected in a decreased amplitude of the oscillations during the first day of the recording, followed by a sudden transition to a higher amplitude in the subsequent days of the assay (Figure 4.17, Figure B.3D, E). RO showed a distinct decrease in principal component 2 and lies considerably further away from all other compounds. This reflects the complete absence of oscillatory behavior associated with this specific treatment (Figure 4.17, Figure B.3F). Taken together, our refitting approach allows the categorization of these compounds based on their inferred effects on the model parameters, as indicated by the proximity of compounds with similar effects in the PCA plot. This also provides a rapid visual validation for our refitting approach, as it shows that parameter values that represent similar treatment effects are in fact placed close to each other in the PCA plot. In the case more compounds should be explored, the PCA has to be recomputed for new experiments using all pharmacological treatments available and the explained variance of the first two PCA components evaluated.



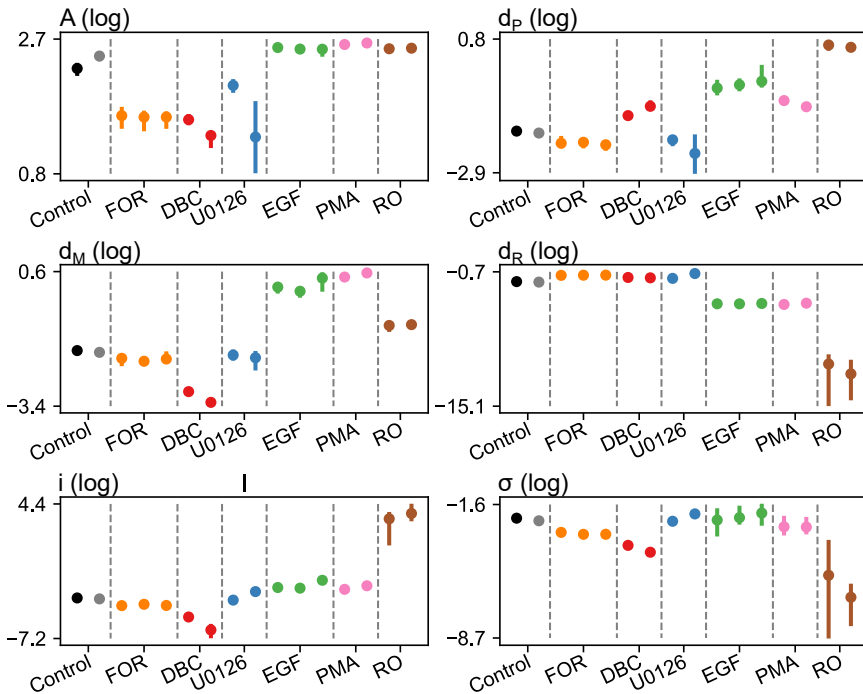
**Figure 4.18:** Principal component analysis can be used to visualize model parameter changes related to all pharmacological treatments in a single overview figure. Control A refers to untreated cells recorded at the same plate as FOR, DBC, and U0128. Control B refers to untreated cells recorded at the same plate as EGF, PMA, and RO. The lines indicate the increasing concentration of the pharmacological treatments from control  $\rightarrow$  low dose  $\rightarrow$  middle dose  $\rightarrow$  high dose. Dots indicate the mean and error bars standard deviation of the final population obtained from the evolutionary algorithm. Individual values used to calculate the median points are shown in Figure B.4. Only pharmacological treatments with positive model efficiency coefficient  $E_f$  (Table 4.2) were included in the PCA analysis. The explained variance for the first three PCs was 53 %, 33 %, and 8 %. This figure is adapted from the previous publication [2].

The results summarized by the PCA plot also correctly reflect the known molecular mechanisms of working of those pharmacological treatment. Inspecting the resulting parameter space, the parameter sets representing the treatments with FOR and DBC were closer to each other than to other compounds. This is expected, as both FOR and DBC act as activators of the cAMP pathway and should thus similarly affect the core clock mechanism [153]. The parameter sets representing EGF and PMA were also close to each other. This is consistent with previous work showing that the PKC pathway (activated by PMA) is an upstream activator of the MAPK pathway (activated by EGF) [146, 154]. Those results indicate that the presented model as well as the fitting procedure can accurately classify changes in core clock parameters. Thus the PCA plot provides a useful tool to easily assess an arbitrary number of experimental conditions in a single image. This kind of assessment would be beneficial when large numbers of compounds are tested, for example, as part of a high throughput screen.

Beyond the overview provided by the PCA plot, the changes in individual parameter values can be also related to changes in the luminescence traces upon pharmacological treatments (Figure 4.19). For example, all compounds leading to a decrease in oscillatory amplitude (FOR, DBC, U0126) exhibit a decrease in parameter  $A$ , representing activator concentration. In contrast, compounds resulting in higher amplitude (EGF, PMA) exhibit an increase in  $A$ . In the case of RO, a compound that abolishes oscillatory behavior, the parameter  $d_R$  is decreased.  $d_R$  represents the degradation rate of the repressor, so the decrease of  $d_R$  indicates that a stabilization of the repressor drives constant inhibition of the transcription activation. Leaving aside the fully repressed case of RO, the parameter  $i$  (light sensitivity) seems positively correlated with the parameter  $\sigma$  (noise intensity). This relationship suggests that entrainment due to higher light sensitivity can counteract the faster desynchronization of individual oscillators that results from higher noise intensity. This balance between light sensitivity and noise strength allows the possibility that FOR, DBC, and U0126, even though they all similarly decrease the amplitude of oscillation, have different noise intensities. Note that the analysis implies that the compounds typically result in a change of several model parameters, making it ultimately problematic to use our model to identify the exact regulatory element within the clock mechanism targeted by a given compound.

## 4.7 Using model to predict mechanisms of population-level amplitude loss

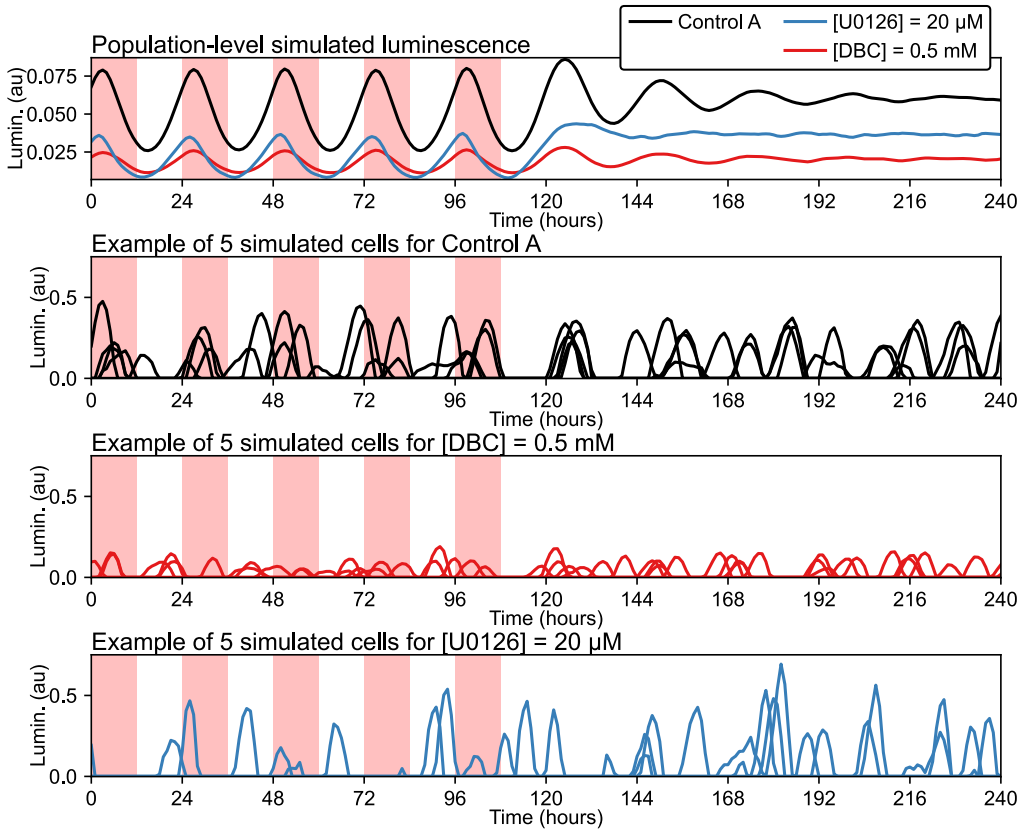
The impact of the FOR, DBC and U0126 compounds on the rhythmic parameters of the luminescence recordings is similar. However, close inspection of the PCA plot reveals that U0126 lies further away from FOR and DBC (Figure 4.18). This relative displacement appears to be mainly in the direction of parameter  $\sigma$ , suggesting that U0126 differs from FOR and DBC by higher noise intensity, as is also supported by the assessment of changes in individual parameters (Figure 4.19). As the model consists of an ensemble of oscillators, the observed difference can be mapped to the altered behavior of the simulated individual cells. As an illustrative example, one can compare the treatments with 0.5 mM DBC and 20  $\mu$ M U0126. When synchronized cell populations were transferred to constant darkness, both treatments resulted in a more rapid loss of population amplitude than in control-treated cultures. At the level of single cells, however, the model suggests that DBC treatment resulted in reduced single-cell oscillator amplitude, while incubation with U0126 caused a more pronounced desynchronization of individual oscillators due to higher noise (Figure 4.20). This effect of differences in the noise intensity was further investigated by visualizing the histograms of phases of individual cells (Figure 4.21). During the regular light-dark cycle, histograms of phases in all three cases (Control, DBC, U0126) markedly differ from the uniform distribution, indicating synchronized oscillations. After transfer into



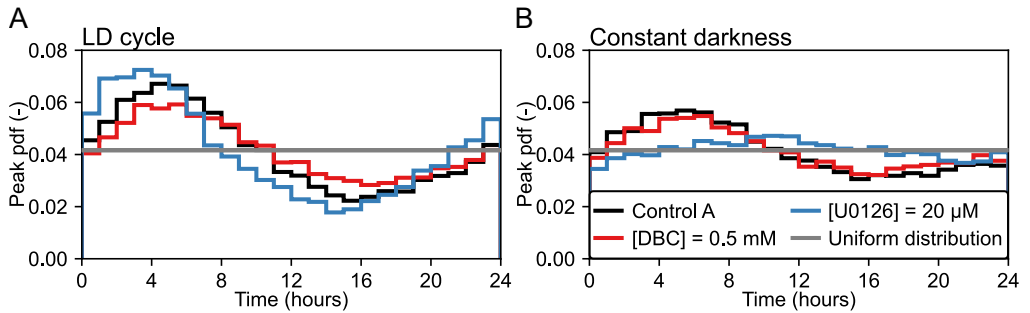
**Figure 4.19:** Parameter values for all pharmacological treatments. Each dot indicates the median value and lines maximal and minimal values of the parameter sets obtained by a differential evolution algorithm. Multiple dots per compound indicate increasing concentrations of the same compound from left to right. Only pharmacological treatments with positive model efficiency coefficient  $E_f$  (Table 4.2) are included. The numerical values of the parameters are in Table B.1. This figure is adapted from the previous publication [2].

the constant darkness, the histogram of phases for U0126 is more uniformly distributed than the one of control and DBC treatment, indicating a more pronounced loss of synchrony. This hints that DBC causes the loss of population-level amplitude by inhibition of gene expression at the single-cell level, whereas U0126 causes a more rapid desynchronization of the single-cell oscillators, which was further confirmed by comparing the average transcription activation and noise intensity of the two compounds (Figure 4.22). Taken together, the presented simulations imply two distinct sets of single-cell dynamics that both result in the accelerated loss of rhythm amplitude at the population level.

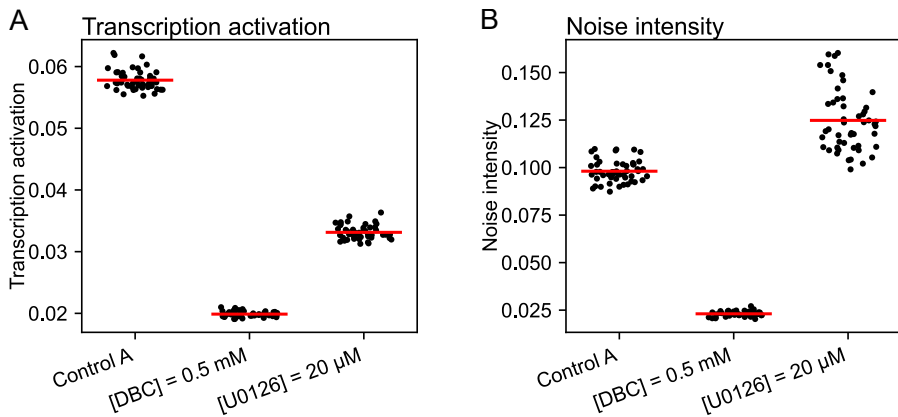
The effect of pharmacological treatments strongly depends on environmental light conditions [155, 156, 157]. To assess how the capacity for entrainment by light differs upon DBC and U0126 treatment, the phase of entrainment was extracted from simulations forced by light-dark cycles with different periods ( $T$ ). In this analysis, the model was simulated for various  $T$  and the phase angle as the time interval between the light onset and the oscillation peak was calculated (Figure 4.23) [64]. The conventional PRCs could not be constructed because, for some treatments,



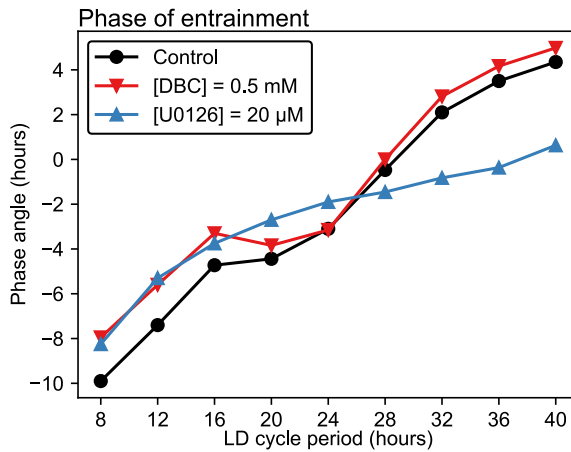
**Figure 4.20:** Model simulation with parameters corresponding to control treatment and DBC and U0126 treatments. (A) Population mean of 30000 oscillators show decreased amplitude for DBC and U0126 compounds. DBC shows also a later peak under the light-dark cycle and lower amplitude in constant darkness than U0126. (B) 10 individual cells that form the population mean for control-treated cells in panel A. (C) 10 individual cells that form the population mean for DBC in panel A. Notice that the individual cells have lower amplitude than the control in panel B but otherwise the waveforms look the same. (D) 10 individual cells that form the population mean for U0126 in panel A. U0126 treated cells have higher amplitude than DBC treated cells but the peaks seem to occur with lower probability. This figure is adapted from the previous publication [2].



**Figure 4.21:** Probability density function (PDF) of the occurrence of peaks during 24 hours. (A) In the light-dark cycle, control and DBC show similar PDFs. The PDF for U0126 differs slightly, showing an earlier peak. Those data were collected from times between 72 and 96 in Figure 4.20. (B) In constant darkness, control and DBC show similar PDFs. The PDF for U0126 differs distinctly from the control and DBC as it is uniformly distributed. Those data were collected from times between 122 and 146 in Figure 4.20. This figure is adapted from the previous publication [2].



**Figure 4.22:** Treatments with DBC and U0126 lead to lower average transcription activation (average value of  $f(R, A)$  in Equation 4.8) but DBC lead to decrease while U0126 to increase of the noise intensity (value of  $\sigma$  in Equation 4.8). (A) Transcription activation was calculated as mean of the output luminescence in constant darkness. The values for control ( $L = 0.0576 \pm 0.0017$ ) are higher than the values for DBC ( $L = 0.0198 \pm 0.0004$ ) and U0126 ( $L = 0.0331 \pm 0.0012$ ). (B) The noise intensity value for control ( $\sigma = 0.098 \pm 0.006$ ) was lower than for U0126 ( $\sigma = 0.124 \pm 0.016$ ) but higher than for DBC ( $\sigma = 0.023 \pm 0.001$ ). Mean and standard deviation were calculated across the final population of parameter sets obtained by the evolutionary optimizer. This figure is adapted from the previous publication [2].



**Figure 4.23:** Simulated effect of the light-dark cycle period ( $T$ ) on the phase of entrainment. The phase angle is calculated as the time interval between the light onset and the oscillation peak normalized by  $T/24$ . Negative values indicate that the oscillation peak is delayed with respect to the light onset. The phase of entrainment curves for all pharmacological treatments can be seen in Figure B.5. This figure is adapted from the previous publication [2].

oscillations in constant darkness diminished too rapidly, making an analysis based on luminescence peaks impossible (Section 2.4.1). The analysis revealed that, for  $T$  smaller than 24 hours both compounds, DBC, and U0126, reduce the phase angle relative to the control treatment. For  $T$  higher than 24 hours, DBC had no apparent effect, while U0126 markedly reduced the phase angle. U0126 leads to increased noise intensity in comparison to the treatment with DBC and thus also T-PRC related to U0126 has a smaller slope than T-PRC related to DBC. This corresponds to the theoretical observations from Section 2.4.2 that increasing noise intensity decreases the slope of the T-PRC. Considering that these simulated phase angle curves provide highly specific predictions for future experiments, the same type of curves were generated also for the other compounds (Figure B.5). Interestingly, these simulations predict that the entrainment dynamics upon U0126 treatment are similar to those upon EGF and PMA treatment, even though the loss of amplitude upon U0126 treatment under a 24-hour light-dark cycle is closer to the results obtained upon DBC and FOR treatment. Interestingly, U0126 is an inhibitor of the MAPK pathway, the same pathway that is positively regulated by EGF and PMA whereas FOR and DBC are inhibitors of the cAMP pathway.

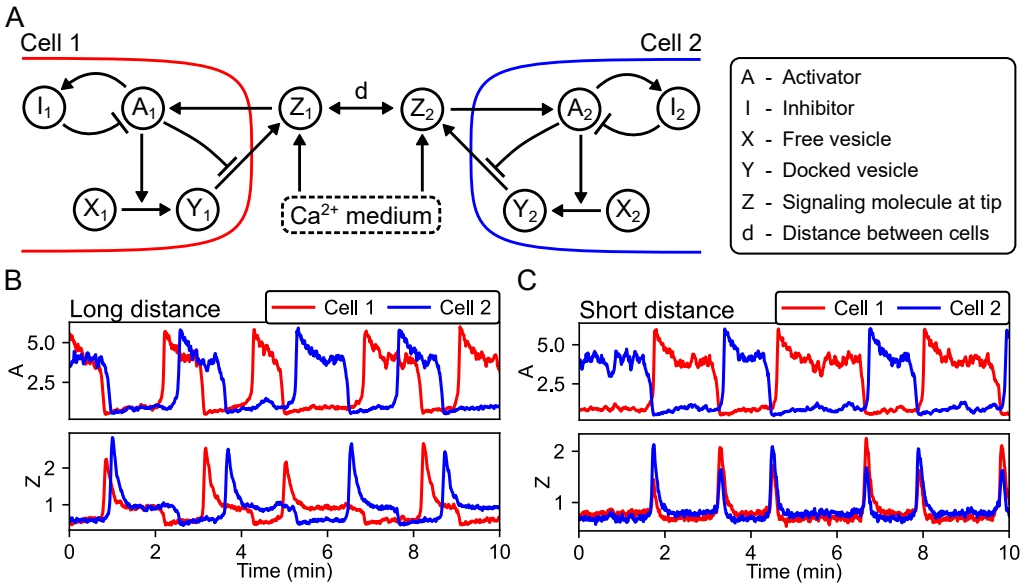


## 4.8 Example of further application: synchronization of fungal hyphae

This chapter focused on the application of the developed methods in the domain of circadian timekeeping, however, in many other systems stochastic oscillators and their synchronization play an important role. As an example and a sneak peek of a possible other application, a project on the modeling of cellular communications in fungal hyphae can be presented [3]. In this project, not a population of oscillators was explored but rather two identical stochastic oscillators that have the ability to synchronize in anti-phase when they are spatially close to each other. This kind of cell-to-cell communication occurs in growing fungal hyphae that can then navigate themselves toward each other and consequently fuse into a single cell. The cells rely on a single chemical signaling channel to exchange its location, which raises the question of how communication is initiated and how a cell knows it just does not speak with itself but there is actually a partner in the area. Implementing a stochastic equation model revealed that the system can work as two coupled excitatory oscillators whose cycle includes a refractory period that makes the cell temporarily numb to the input signal (Figure 4.24). When one cell is active, it sucks in all available signaling molecules in the shared medium not allowing the other cell to be activated. Only at the end of the cycle, the signaling molecules are released, which provides the basis for the activation of the partner cell. Here the refractory period prevents the immediate auto-excitation giving the partner cell time to initiate its cycle. Mathematical modeling showed that such a system of communication works robustly also under noise.

## 4.9 Conclusion

In this chapter, a stochastic oscillator model that mimics the dynamic properties of the core circadian clock in cultures of zebrafish cells was presented. The model accurately recapitulates the rhythmic properties and entrainment of luminescence recordings under a range of lighting conditions, only losing accuracy when simulating the effects of exposure to light-dark cycles with periods significantly longer or shorter than 24 hours. Furthermore, the model can characterize luminescence recordings in terms of changes in single-cell core clock regulation. In particular, the model was successfully used to quantify the effects of several pharmacological treatments by correctly grouping compounds with similar effects. The model was also used to predict the distinct effect of two pharmacological treatments that appeared to have a similar effect at the population level but a distinct impact on the single-cell regulatory mechanisms. Such assessment of the degree of synchronization of single-cell oscillators in a cell population in culture cells would involve technically challenging and time-consuming imaging that is not trivial to perform as part



**Figure 4.24:** Fungal hyphae exhibit anti-phase locking (entrainment) at a short distance. (A) The mathematical model consists of two stochastic oscillators connected through a common medium. (B) At a long distance, the oscillators generate independent sustained oscillations. (C) At a short distance, the oscillators lock to each other in anti-phase. This figure is adapted from a previous publication [3].

of a large-scale screen. Within this context, the here-presented modeling approach provides the possibility to make rapid predictions about the behavior of individual cell clocks from population-level luminescence recordings. These predictions could then be followed up by more refined imaging-based assays of single-cell dynamics on a much smaller set of samples. Therefore, the presented mathematical model should contribute a rapid and scalable tool for interpreting the effects of pharmacological treatments and genetic modifications on the circadian clock at the cellular level.

# 5 Conclusion

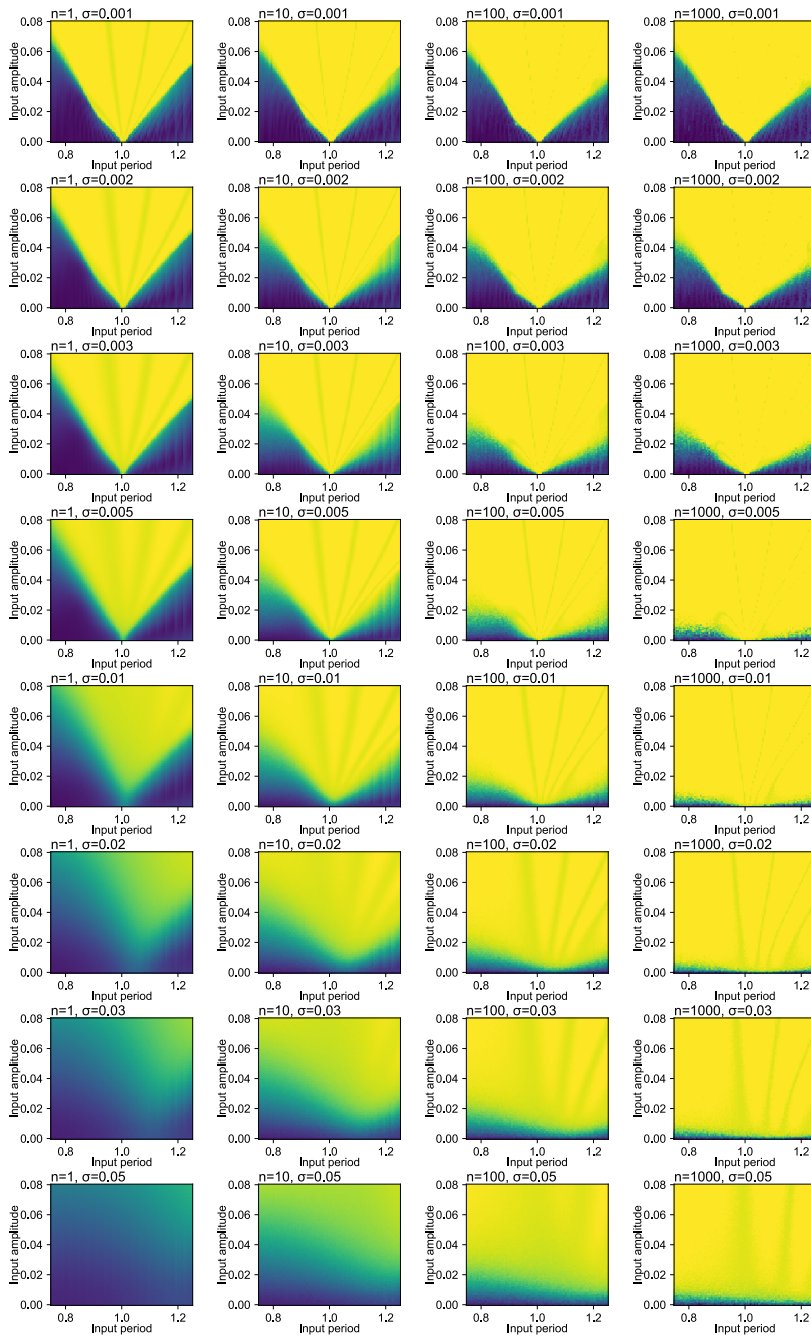
Biological oscillators are inherently noisy and yet they have a great ability to synchronize to the changing environmental conditions, such as the day-night cycle. It is, however, not clear what design principles make the oscillators robust to noise or to what degree they might even utilize noise to enhance their entrainment performance. This thesis considered those questions at the theoretical level, by numerically analyzing computational models, and also in a practical application, by analyzing experimental data of biological stochastic oscillators using the developed methods. In a nutshell, the major contributions presented in this thesis are:

1. Comparison and evaluation of the numerical methods for noise approximation concerning their accuracy and simulation time (Section 2.1).
2. Definition, comparison, and evaluation of various metrics of entrainment in the context of the quantification of stochastic oscillations (Section 2.2).
3. New theoretical findings showing that noise allows for population-level entrainment to a wider range of input signal periods and amplitudes (Section 2.3).
4. Implementing T-cycle phase response curve as an alternative to the phase response curve when the analysis based on oscillation peaks cannot be performed because of fast population-level desynchronization (Section 2.4.2).
5. New theoretical findings showing that noise allows for a larger response to external stimuli (Section 2.4.1) and faster recovery from jet lag (Section 2.4.3).
6. Demonstration that the noise-improved entrainment emerges for a larger class of generic limit cycle oscillators, not only biological or bio-inspired (Section 2.5).
7. Showing that population heterogeneity can also enhance entrainment similarly to intrinsic stochasticity, but lead to a more complex dynamical behavior that is challenging to quantify (Section 2.6).
8. Novel step-wise fitting algorithm that fits stochastic population models in a step-wise manner. This saves computational time and enables the exploration of larger parameter ranges in comparison to direct trajectory fitting (Section 2.7.2).

9. Software implementations of the described methods, which enhances the thesis reproducibility and utilization of the methods for future research (Chapter 3).
10. Novel normalization approach to the luminescence recordings of pharmacological treatments, which normalize individual treatments with respect to the control treatment and allows thus precise comparison of treatments from separate experimental repetitions (Section 4.3).
11. Novel model to capture precisely the entrainment dynamics of the zebrafish clock with the minimum of parameters (Sections 4.4 and 4.5).
12. Demonstration that a model can be used to quantify pharmacological treatments from population-level recordings at the level of individual cells (Section 4.6).
13. New predictions made with the model regarding the effects of the pharmacological treatments on the stochasticity of the individual cells (Section 4.7).
14. Novel model demonstrating dialog-like information exchange between two cells utilizing a single signaling channel (Section 4.8).

Considering the contributions listed above, the thesis met all objectives outlined in Section 1.4. Nevertheless, with regard to the open questions outlined in Section 1.3, there are still many further directions that should be investigated in future work. At the theoretical level, the presented thesis considered mainly populations of uncoupled identical stochastic oscillators. The future work should thus consider population heterogeneity, as briefly outlined in Section 2.6, and coupling among the individual oscillators in the population. The theoretical findings could be also supported by biological experiments using zebrafish luminescence reporter assays (introduced in Section 4.2). Ultimately, single-cell imaging could be employed to verify experimentally the presented model predictions on the relation among individual stochastic oscillators and their population-level mean. Beyond the research on biological oscillators, the results of this thesis are potentially relevant also to applications in the fields of biotechnology, control engineering, and automation. For example, the presented modeling approach could be used to create an automatic artificial light control to achieve optimal plant growth in greenhouses. Control of a population of noisy oscillator units could also prove important in the domains of movement generators for robots or semiconductor superlattices.

# **A Supplementary material for Chapter 2**



**Figure A.1:** Arnold tongues used to calculate the mean phase coherence for Figure 2.17. The color indicates mean phase coherence in the range 0 (dark blue) - 1 (bright yellow).

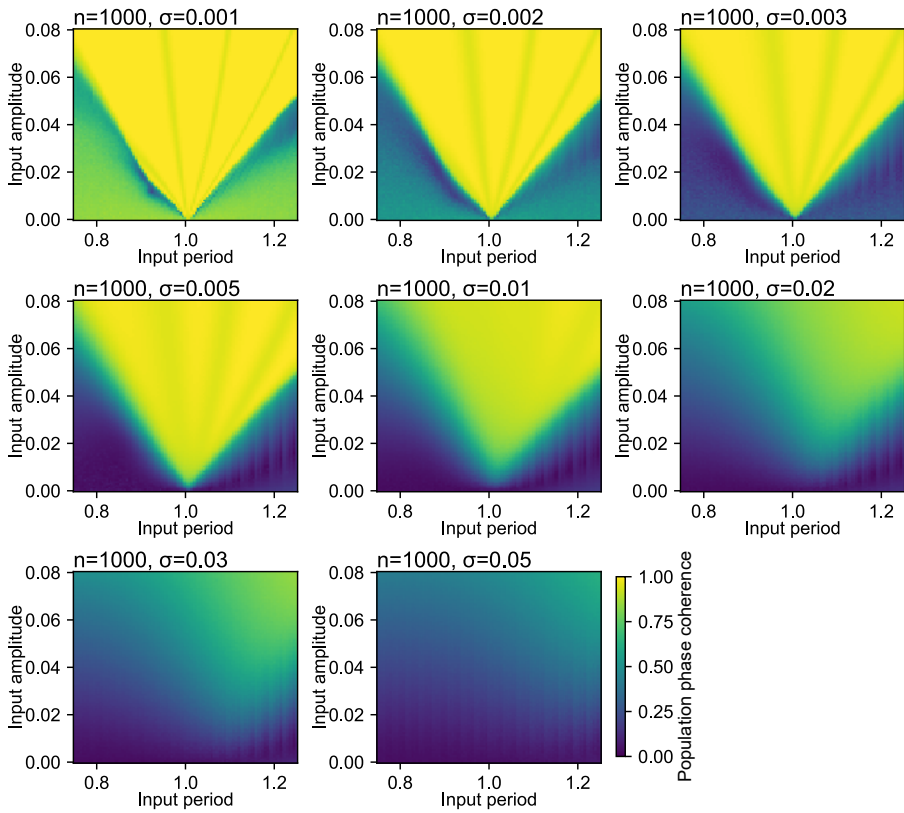
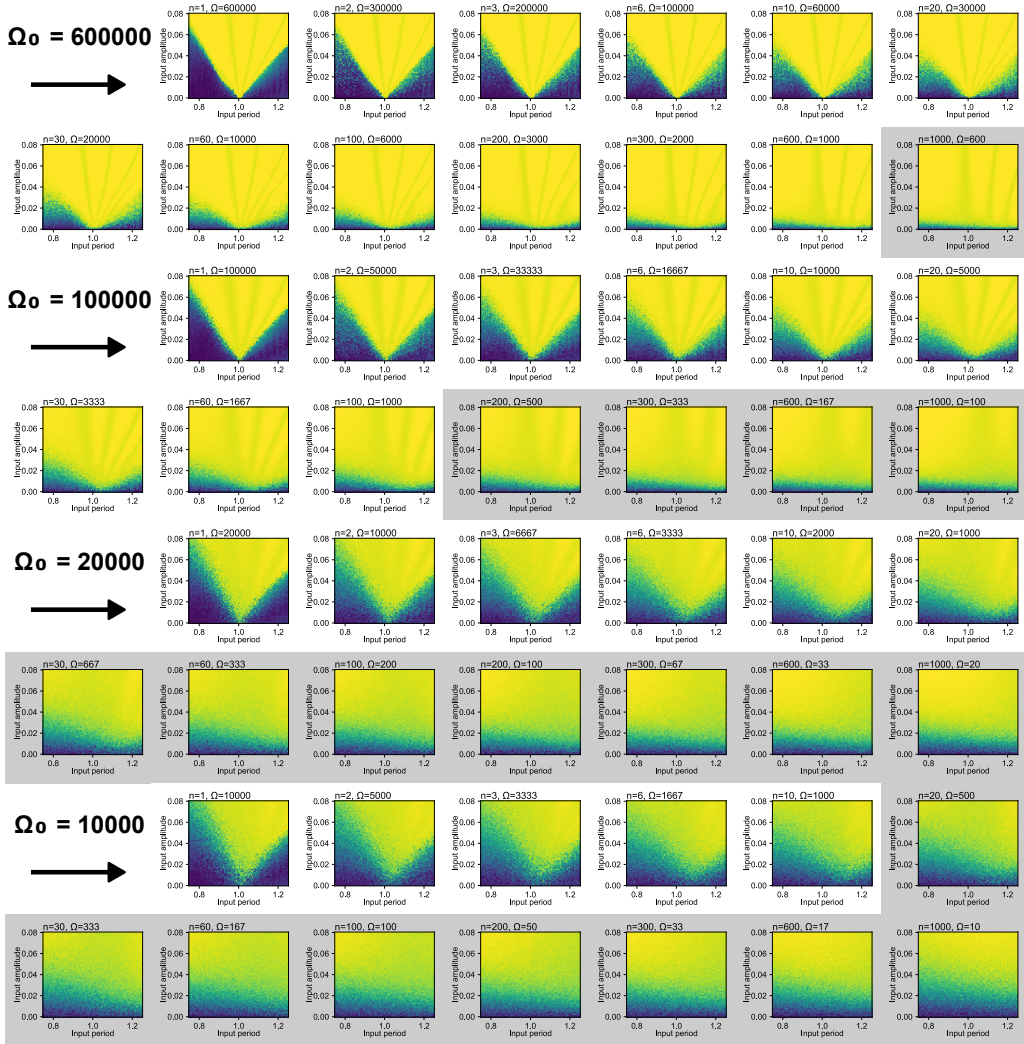
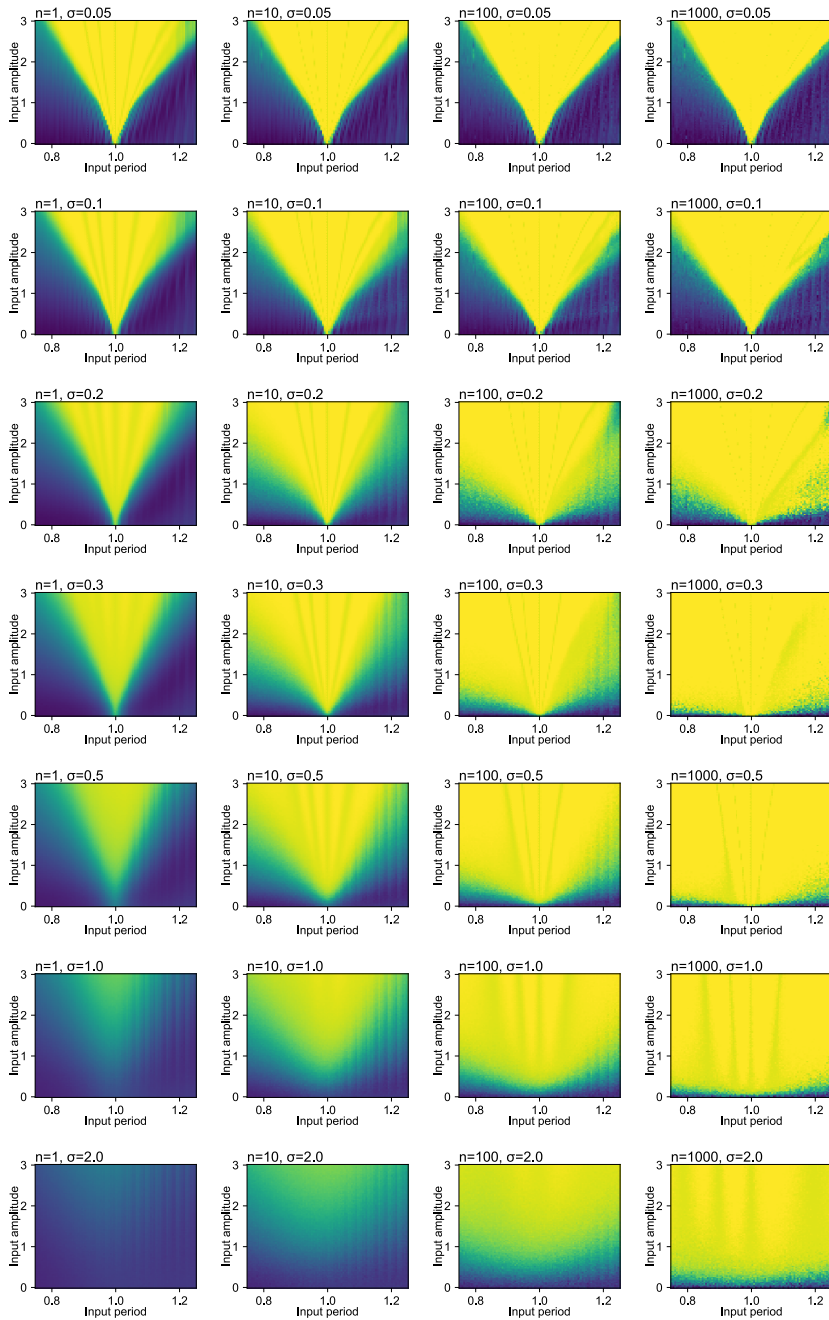


Figure A.2: Arnold tongues used to calculate the mean population phase coherence for Figure 2.18.

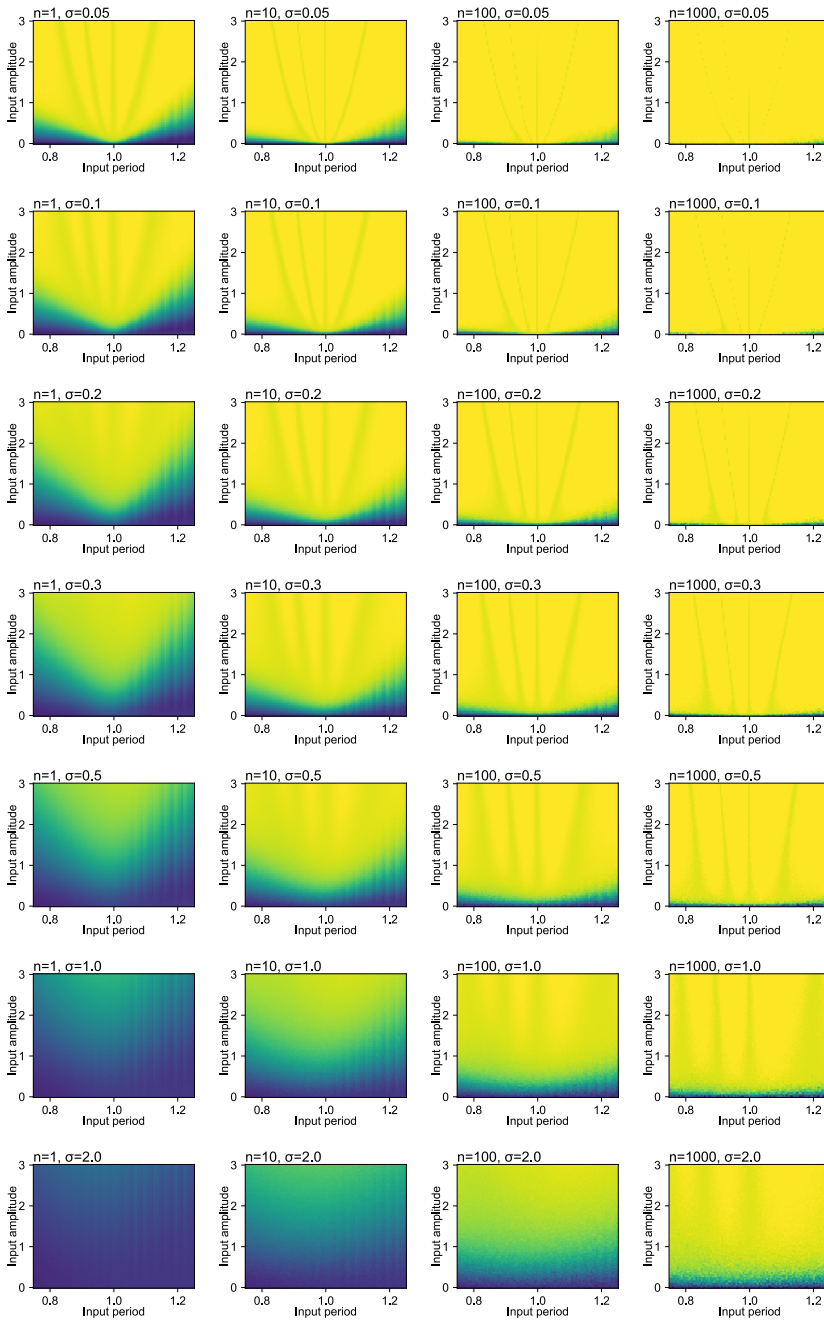


**Figure A.3:** Arnold tongues used to calculate the mean phase coherence for Figure 2.19. In gray, for low values of  $\Omega$  Gillespie's Stochastic Simulation Algorithm (SSA) instead of the Chemical Langevin Equation (CLE) was used as described in Section 2.1.2. The color indicates mean phase coherence in the range 0 (dark blue) - 1 (bright yellow).

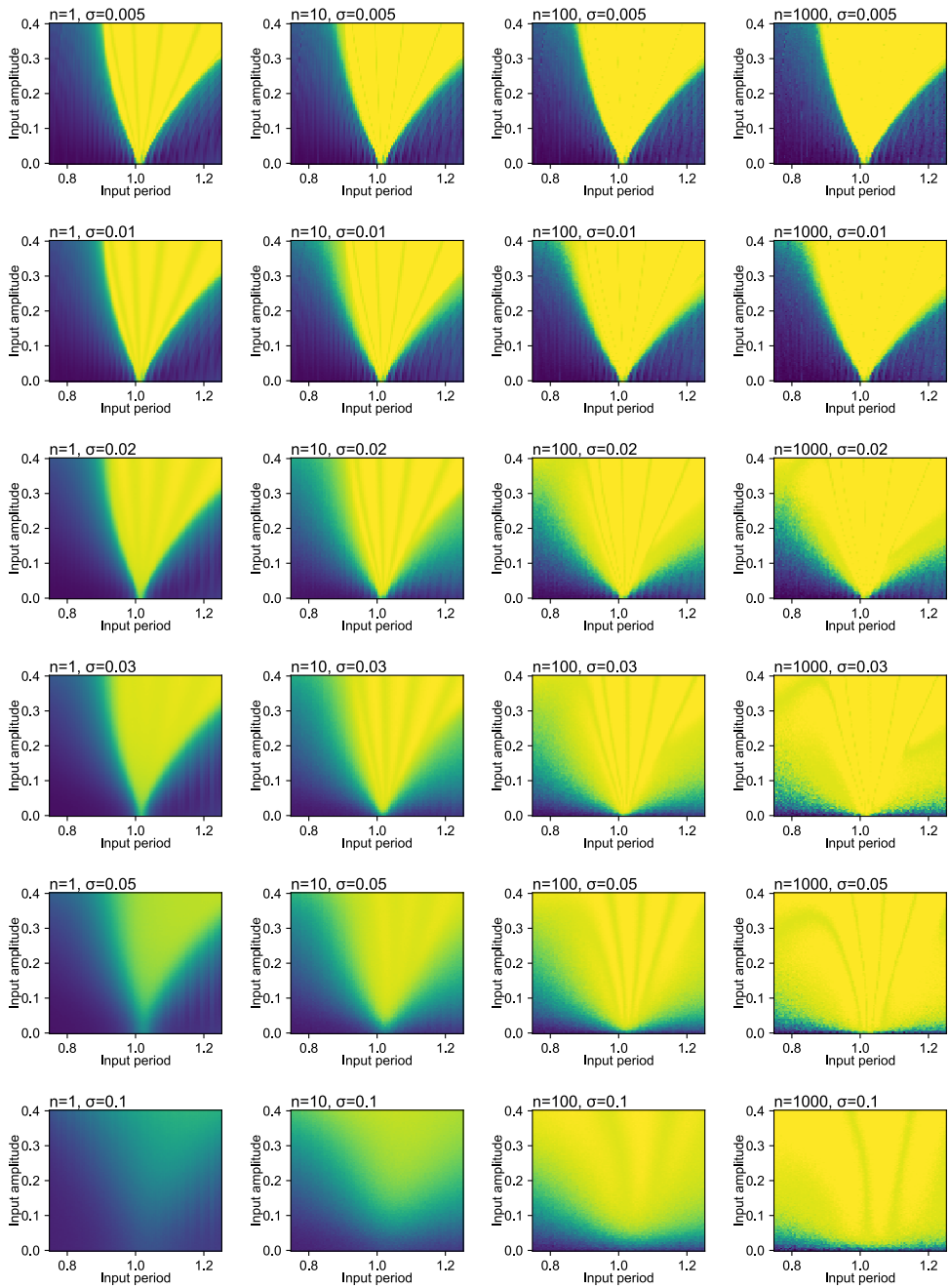




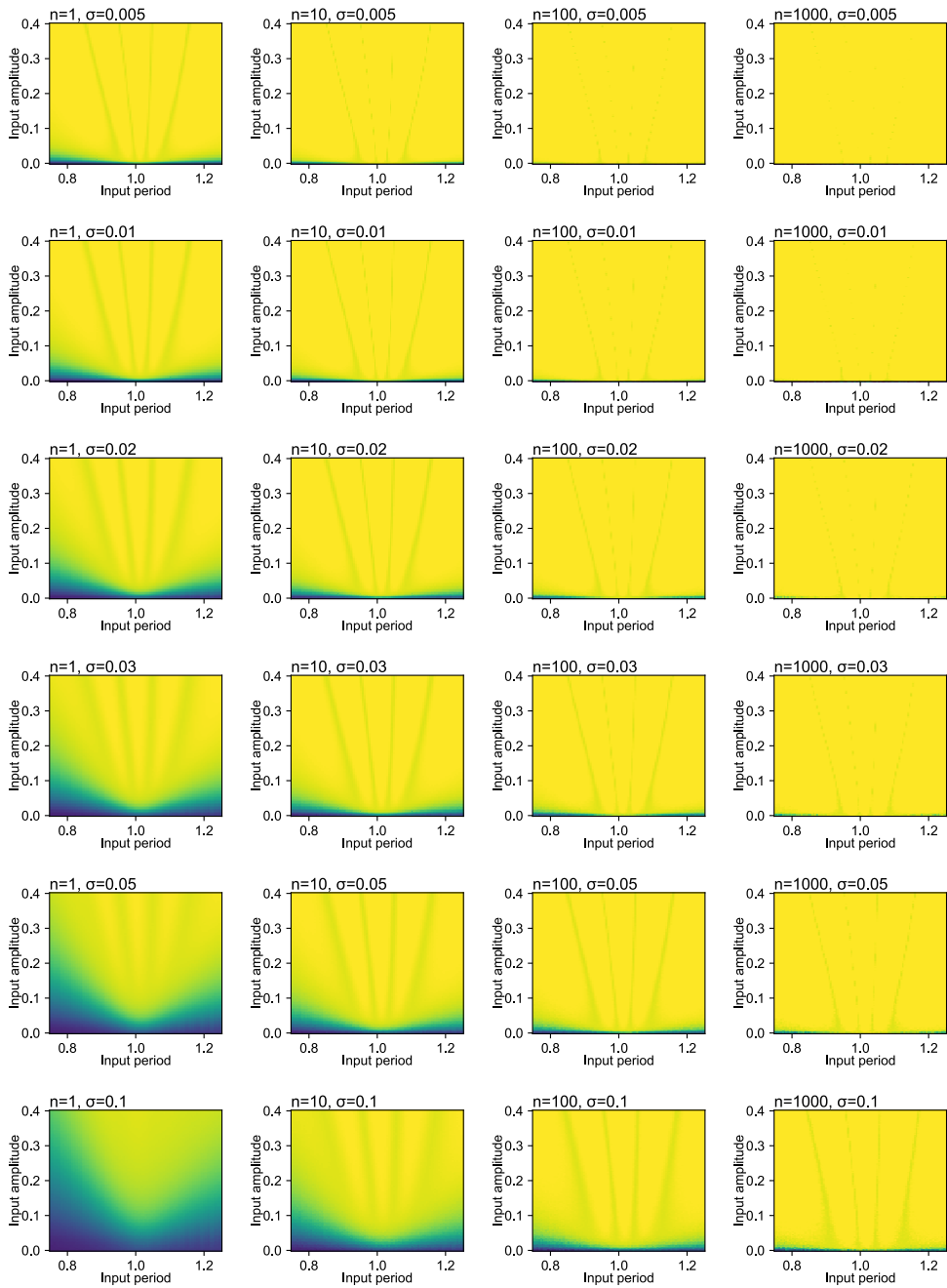
**Figure A.4:** Arnold tongues used to calculate the mean phase coherence for Figure 2.28A. The color indicates mean phase coherence in the range 0 (dark blue) - 1 (bright yellow).



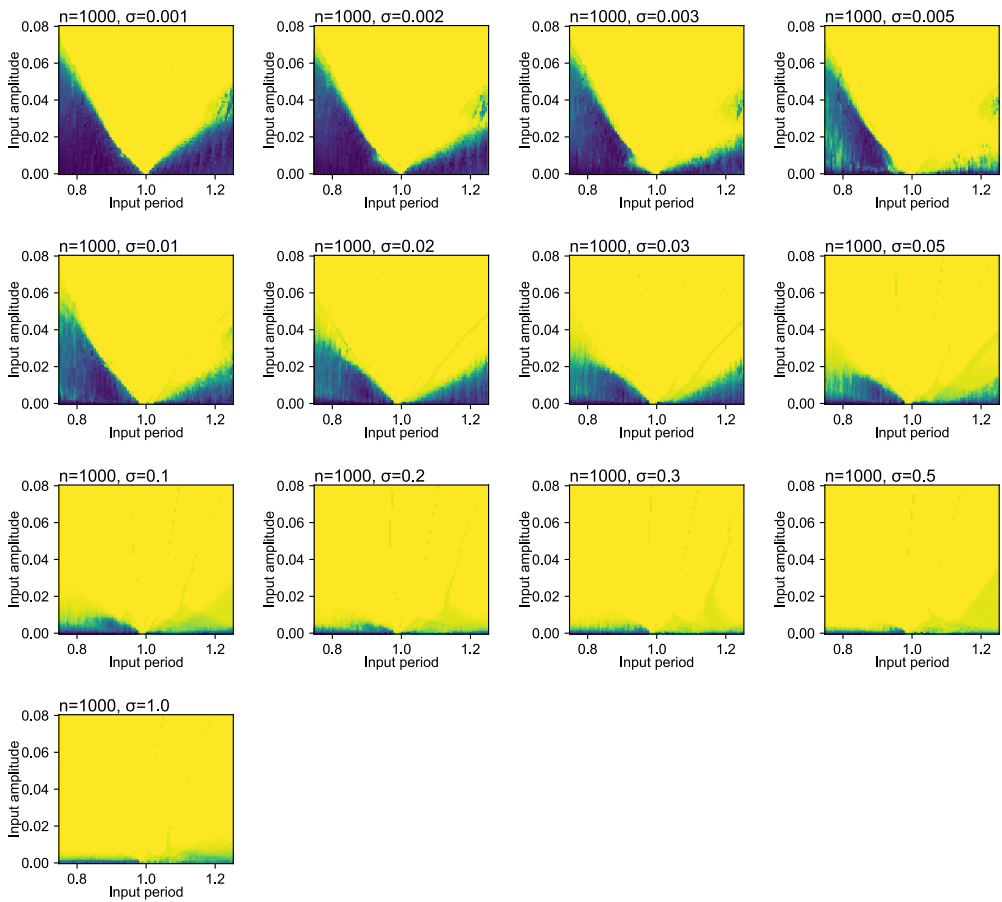
**Figure A.5:** Arnold tongues used to calculate the mean phase coherence for Figure 2.28B. The color indicates mean phase coherence in the range 0 (dark blue) - 1 (bright yellow).



**Figure A.6:** Arnold tongues used to calculate the mean phase coherence for Figure 2.29A. The color indicates mean phase coherence in the range 0 (dark blue) - 1 (bright yellow).



**Figure A.7:** Arnold tongues used to calculate the mean phase coherence for Figure 2.29B. The color indicates mean phase coherence in the range 0 (dark blue) - 1 (bright yellow).



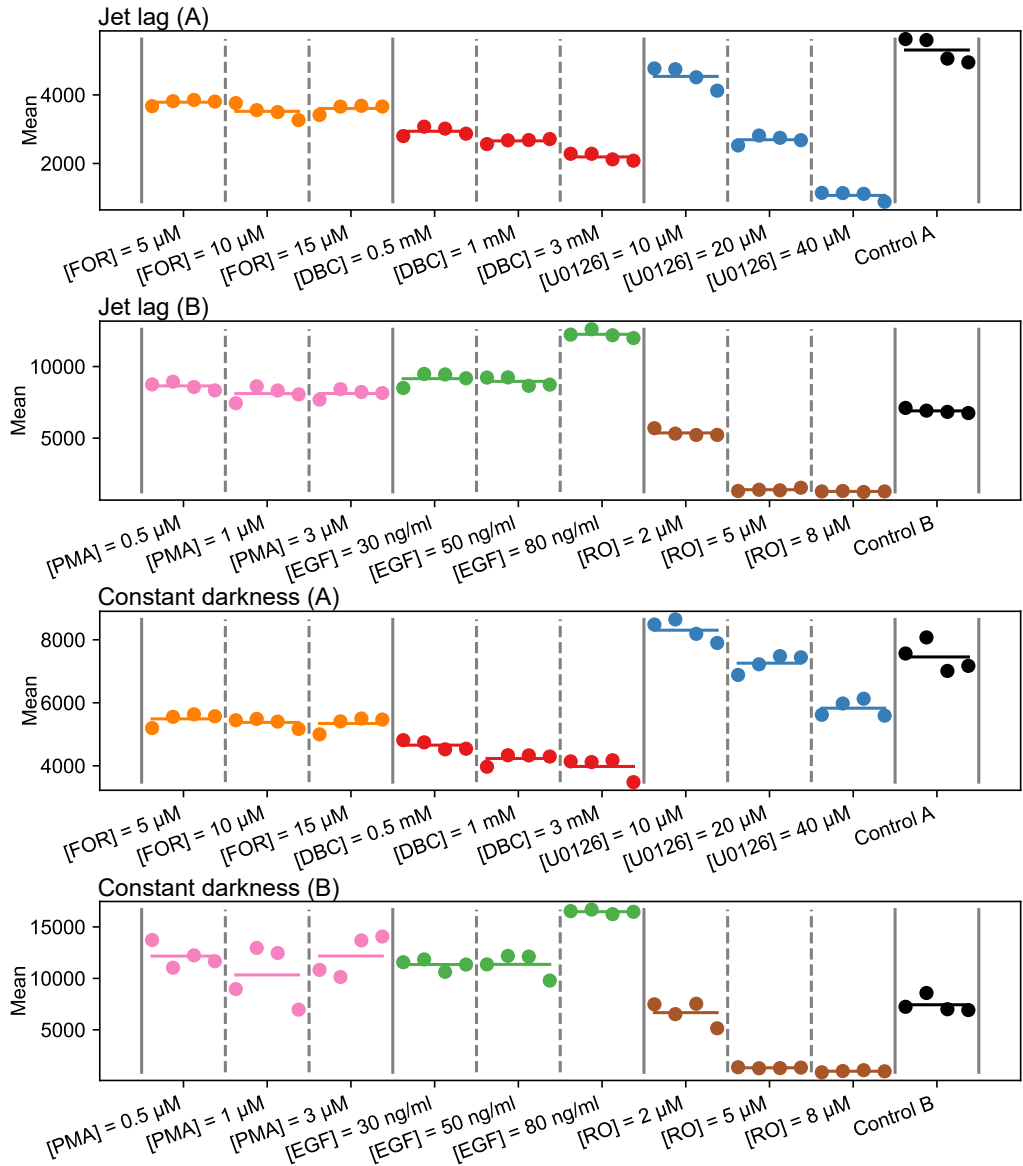
**Figure A.8:** Arnold tongues used to calculate the mean phase coherence for Figure 2.37. The color indicates mean phase coherence in the range 0 (dark blue) - 1 (bright yellow).



## B Supplementary material for Chapter 4

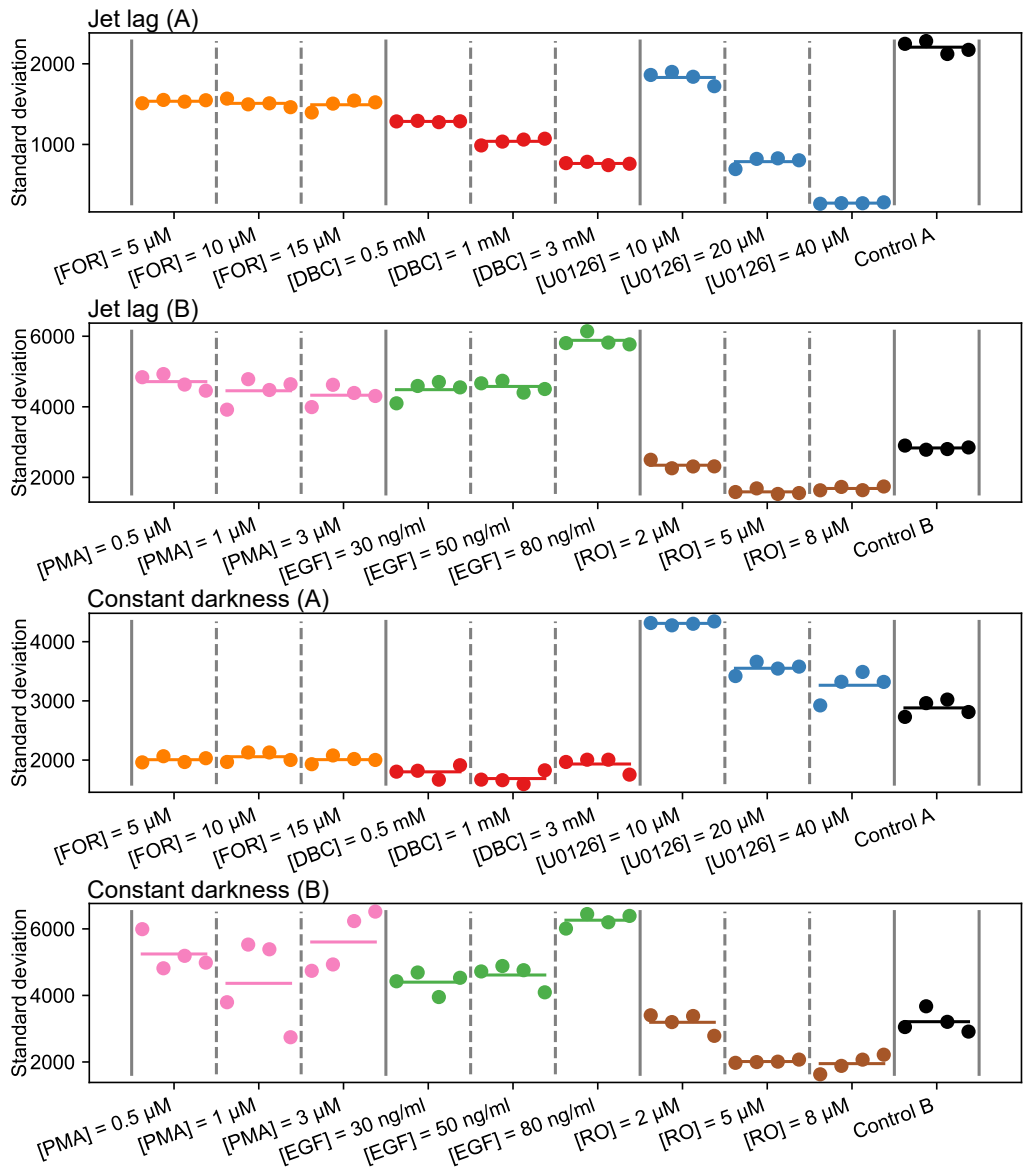
Treatment	$A$	$d_M$	$d_P$	$d_R$	$i$	$\sigma$
Untreated cells	9.68	0.16	0.16	0.16	0.05	0.04
DMSO control A	9.70	0.17	0.17	0.17	0.03	0.10
DMSO control B	11.84	0.16	0.16	0.16	0.02	0.09
[FOR] = 5 $\mu\text{M}$	5.11	0.13	0.13	0.34	0.01	0.04
[FOR] = 10 $\mu\text{M}$	4.92	0.12	0.13	0.35	0.01	0.04
[FOR] = 15 $\mu\text{M}$	5.17	0.13	0.12	0.35	0.01	0.04
[DBC] = 0.5 mM	4.77	0.05	0.27	0.27	0.01	0.02
[DBC] = 1 mM	3.89	0.04	0.35	0.26	0.00	0.01
[DBC] = 3 mM	2.40	0.01	0.90	0.18	0.00	0.04
[U0126] = 10 $\mu\text{M}$	7.74	0.15	0.14	0.24	0.02	0.08
[U0126] = 20 $\mu\text{M}$	3.84	0.17	0.08	0.41	0.04	0.12
[U0126] = 40 $\mu\text{M}$	1.29	0.06	0.33	0.50	0.60	0.02
[EGF] = 30 ng/ml	12.66	0.98	0.67	0.02	0.06	0.05
[EGF] = 50 ng/ml	12.62	0.90	0.73	0.02	0.06	0.09
[EGF] = 80 ng/ml	13.39	1.68	0.61	0.02	0.11	0.07
[PMA] = 0.5 $\mu\text{M}$	16.49	0.33	0.12	0.11	0.07	0.30
[PMA] = 1 $\mu\text{M}$	13.65	1.50	0.42	0.02	0.05	0.04
[PMA] = 3 $\mu\text{M}$	13.93	1.72	0.36	0.02	0.07	0.05
[RO] = 2 $\mu\text{M}$	15.41	0.09	0.05	0.06	0.00	0.17
[RO] = 5 $\mu\text{M}$	13.03	0.37	1.83	0.00	22.50	0.00
[RO] = 8 $\mu\text{M}$	13.11	0.39	1.74	0.00	40.50	0.00

Table B.1: Fitted parameter values for individual treatments.

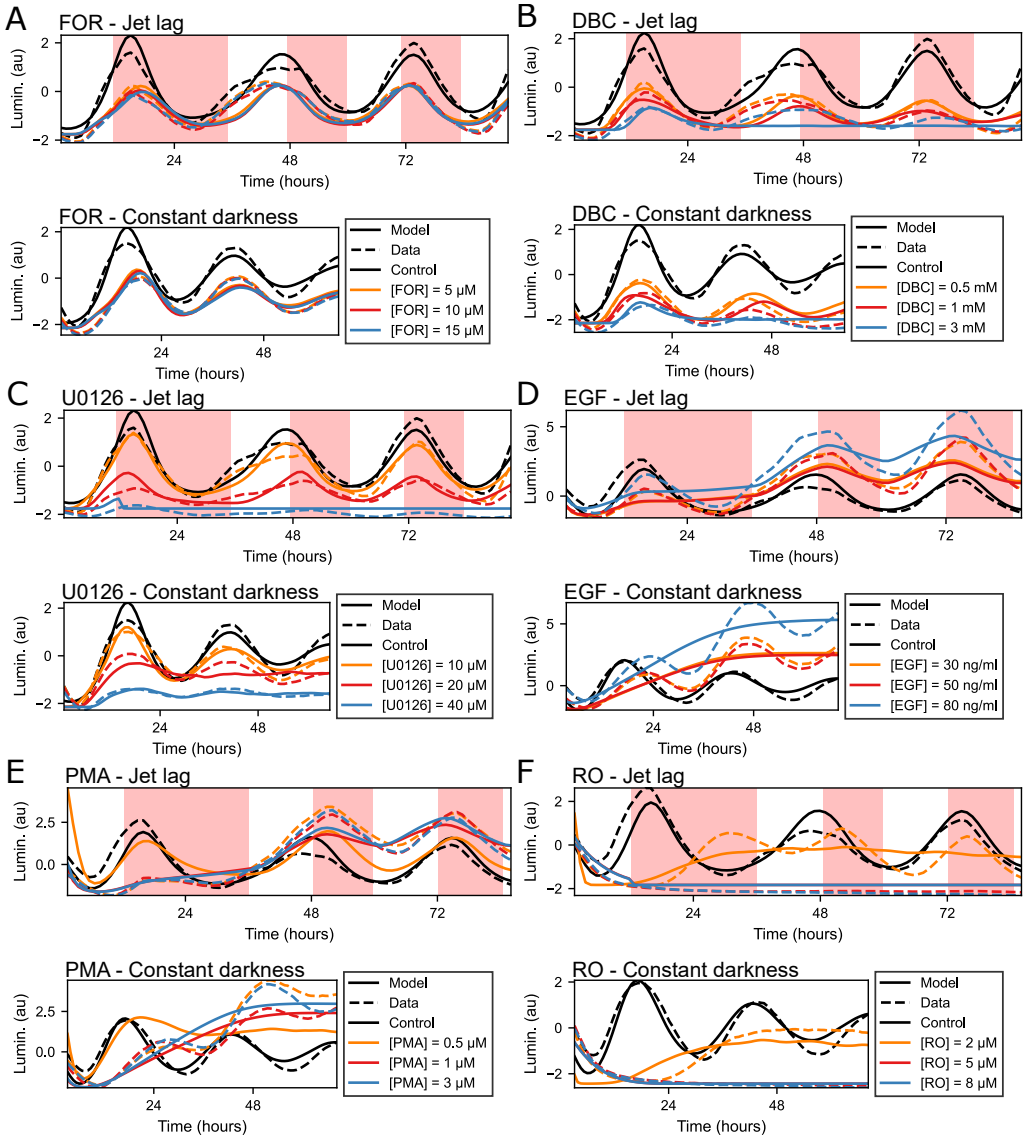


**Figure B.1:** Mean value for the pharmacological treatment recordings. Dots indicate mean values for the individual recordings and horizontal line indicate the mean of means for the specific compound and dose.

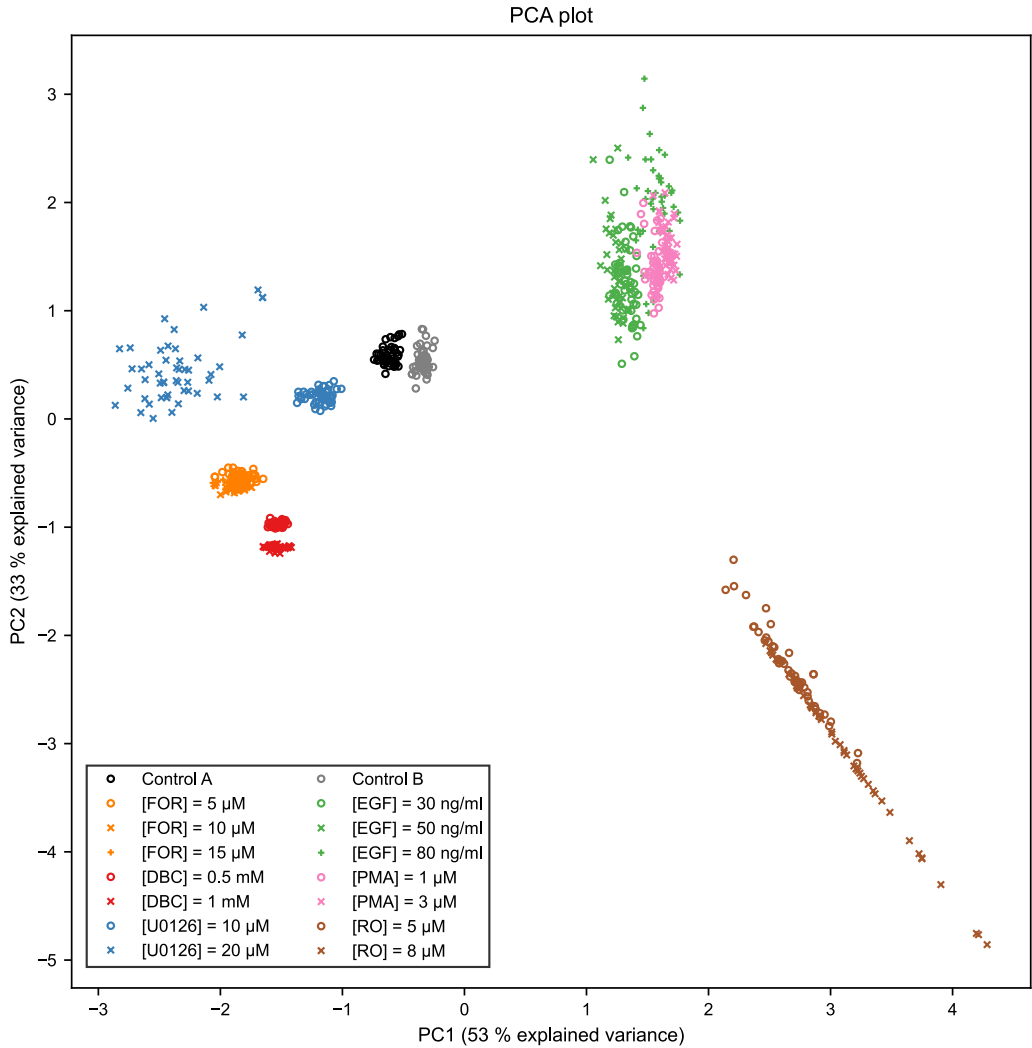




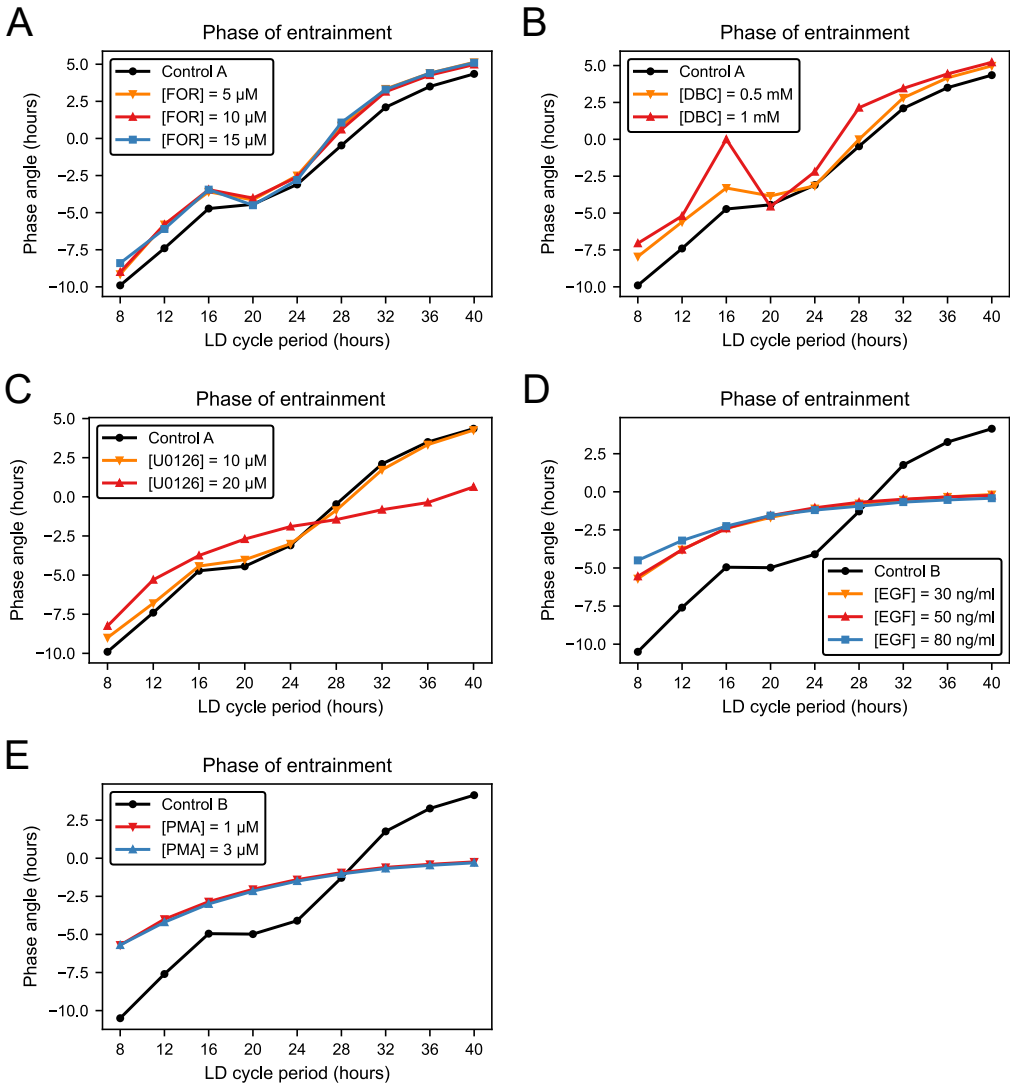
**Figure B.2:** Standard deviation for the pharmacological treatment recordings. Dots indicate standard deviations for the individual recordings and horizontal line indicate the mean of standard deviations for the specific compound and dose.



**Figure B.3:** Model fits to all pharmacological treatments. (A) Model fit to treatments with forskolin (FOR). (B) Model fit to treatments with dibutyl cAMP (DBC). (C) Model fit to treatments with U0126. (D) Model fit to treatments with epidermal growth factor (EGF). (E) Model fit to treatments with phorbol-12-myristate-13-acetate (PMA). (F) Model fit to treatments with ro-318220 (RO). This figure is adapted from the previous publication [2].



**Figure B.4:** Individual values in the PCA analysis. This plot is equivalent to Figure 4.18 showing individual points for each compound concentration. Each point represents one member of the final population obtained from running a differential evolution algorithm to find the best parameter fit for the data. The population size was 50. This figure is adapted from the previous publication [2].



**Figure B.5:** The phase of entrainment depends on the light-dark cycle period ( $T$ ) and administrated compound. The phase angle is calculated as the time interval between the light onset and the oscillation peak normalized by  $T/24$ . Negative values indicate that the oscillation peak is delayed with respect to the light onset. (A) forskolin (FOR) (B) dibutyryl cAMP (DBC) (C) U0126 (D) epidermal growth factor (EGF) (E) phorbol-12-myristate-13-acetate (PMA). This figure is adapted from the previous publication [2].

# List of Figures

1.1	Negative feedback loop . . . . .	2
1.2	Limit cycle oscillator . . . . .	4
1.3	Hopf bifurcation . . . . .	6
1.4	Goodwin model . . . . .	7
1.5	Inhibition functions . . . . .	8
1.6	Entrainment to square signal . . . . .	9
1.7	Arnold tongue and onion . . . . .	10
1.8	Phase response curves . . . . .	11
1.9	Stochastic simulation . . . . .	12
1.10	Stochastic resonance . . . . .	13
1.11	Noise-induced oscillator . . . . .	14
1.12	Population of oscillators . . . . .	15
1.13	Population-level desynchronization . . . . .	15
1.14	Stochastic population entrainment . . . . .	16
2.1	Bifurcation diagram for the Kim-Forger model . . . . .	20
2.2	Numerical simulation of the Kim-Forger model . . . . .	20
2.3	Chemical reaction network . . . . .	21
2.4	SSA and CLE time series . . . . .	24
2.5	SSA and CLE simulation times . . . . .	25
2.6	Integration step . . . . .	25
2.7	SSA and CLE phase-plane histograms . . . . .	26
2.8	Population-level simulation using SSA and CLE . . . . .	27
2.9	Period estimation by autocorrelation . . . . .	28
2.10	Winding number . . . . .	30
2.11	Peak detection . . . . .	30
2.12	Circular cross-correlation . . . . .	31
2.13	Phase coherence . . . . .	33
2.14	Population phase coherence . . . . .	33
2.15	Comparison of the entrainment metrics . . . . .	34
2.16	Arnold tongues for varying population size . . . . .	36
2.17	Mean phase coherence as a function of population size and noise intensity . . . . .	36
2.18	Mean population phase coherence as a function of noise intensity . . . . .	37

2.19	Mean phase coherence as a function of the total volume and population size . . . .	38
2.20	Phase response curve estimation . . . . .	39
2.21	Noise increases phase response . . . . .	40
2.22	T-cycle phase response curve estimation . . . . .	41
2.23	Noise increases slope of T-PRC . . . . .	42
2.24	Jet lag experiment. . . . .	42
2.25	Noise allows faster recovery from jet lag. . . . .	43
2.26	Limit cycle and noise-induced regimes of the amplitude-phase model . . . . .	45
2.27	Limit cycle and noise-induced regimes of the Van der Pol model . . . . .	46
2.28	Mean phase coherence for the amplitude-phase model . . . . .	48
2.29	Mean phase coherence for the Van der Pol model . . . . .	49
2.30	Phase response curves for the amplitude-phase model . . . . .	49
2.31	Phase response curves for the Van der Pol model . . . . .	50
2.32	Jet lag for the amplitude-phase model . . . . .	50
2.33	Jet lag for the Van der Pol model . . . . .	51
2.34	Superposition of two Arnold tongues . . . . .	52
2.35	Population-level trajectories for a system of two heterogeneous oscillators . . . . .	53
2.36	Superposition of three, five, and seven Arnold tongues . . . . .	54
2.37	Mean phase coherence for heterogeneous population of 1000 oscillators . . . . .	54
2.38	Generated benchmark data . . . . .	56
2.39	Differential evolution . . . . .	58
2.40	Metrics used for pre-estimation with deterministic model . . . . .	60
2.41	Results of the first optimization step . . . . .	61
2.42	Damped sine fit . . . . .	62
2.43	Damping rate is monotonically increasing function of noise intensity . . . . .	63
2.44	Results of the second optimization step . . . . .	63
2.45	Results of the third optimization step . . . . .	64
2.46	Final fit to the generated data . . . . .	65
2.47	Optimization convergence . . . . .	66
2.48	Model efficiency coefficient (toy example) . . . . .	67
3.1	Workflow for a simple population simulation . . . . .	73
3.2	Output of the simple example . . . . .	73
3.3	Workflow for a simple population simulation . . . . .	75
3.4	Output of the Arnold tongue example . . . . .	75
3.5	Workflow for data fitting . . . . .	77
3.6	Peak properties . . . . .	79
4.1	Circadian clock negative feedback loop . . . . .	82
4.2	Bioluminescence reporter assay . . . . .	83
4.3	Light protocols . . . . .	84
4.4	Z-score normalization . . . . .	86
4.5	Standard Z-score for controls from different plates . . . . .	87

---

4.6	Example of original data for pharmacological treatments . . . . .	88
4.7	Example of normalized data for pharmacological treatments . . . . .	88
4.8	Standard Z-score fails for pharmacological treatments . . . . .	89
4.9	Mathematical model of the zebrafish circadian clock . . . . .	90
4.10	Population simulation . . . . .	92
4.11	Fit to fitting data . . . . .	93
4.12	Fit to validation data . . . . .	94
4.13	Fit to different T periods . . . . .	94
4.14	Fitted phase response curve . . . . .	95
4.15	Step-wise fitting of the pharmacological treatments . . . . .	96
4.16	Fit to pharmacological treatments (set A) . . . . .	96
4.17	Fit to pharmacological treatments (set B) . . . . .	98
4.18	Principal component analysis . . . . .	99
4.19	Estimated parameter values for the pharmacological treatments . . . . .	101
4.20	Comparison of DBC and U0126 (time traces) . . . . .	102
4.21	Comparison of DBC and U0126 (probability density function) . . . . .	103
4.22	Comparison of DBC and U0126 (transcription activation and noise intensity) . . . . .	103
4.23	Comparison of DBC and U0126 (phase of entrainment) . . . . .	104
4.24	Fungal hyphae exhibit anti-phase locking at a short distance . . . . .	106
A.1	Arnold tongues for Figure 2.17 . . . . .	110
A.2	Arnold tongues for Figure 2.18 . . . . .	111
A.3	Arnold tongues for Figure 2.19 . . . . .	112
A.4	Arnold tongues for Figure 2.28A . . . . .	113
A.5	Arnold tongues for Figure 2.28B . . . . .	114
A.6	Arnold tongues for Figure 2.29A . . . . .	115
A.7	Arnold tongues for Figure 2.29B . . . . .	116
A.8	Arnold tongues for Figure 2.37 . . . . .	117
B.1	Mean value for the pharmacological treatment recordings . . . . .	120
B.2	Standard deviation for the pharmacological treatment recordings . . . . .	121
B.3	Model fits to all pharmacological treatments . . . . .	122
B.4	Estimated parameter values . . . . .	123
B.5	T-cycle PRCs for pharmacological treatments . . . . .	124





# List of Tables

1.1	Examples of biological oscillators . . . . .	2
2.1	Reaction model . . . . .	22
2.2	Fitting results for the generated data . . . . .	64
2.3	Search ranges . . . . .	65
3.1	Overview of implemented methods . . . . .	70
4.1	Recorded bioluminescent reporter assay data . . . . .	85
4.2	Goodness of fit for the pharmacological treatments . . . . .	97
B.1	Fitted parameter values . . . . .	119



# List of Publications

## Journal articles

- [1] Vojtech Kumpost, Lennart Hilbert, and Ralf Mikut. Noise facilitates entrainment of a population of uncoupled limit cycle oscillators. *Journal of the Royal Society Interface*, 20: 20220781, 2023.
- [2] Vojtech Kumpost, Daniela Vallone, Srinivas Babu Gondi, Nicholas S. Foulkes, Ralf Mikut, and Lennart Hilbert. A stochastic oscillator model simulates the entrainment of vertebrate cellular clocks by light. *Scientific Reports*, 11:14497, 2021.
- [3] Valentin Wernet, Vojtech Kumpost, Ralf Mikut, Lennart Hilbert, and Reinhard Fischer. Synchronization of oscillatory growth prepares fungal hyphae for fusion. *bioRxiv*, 2022.

## Conference posters and talks

- [4] Vojtech Kumpost. Poster: Intrinsic noise facilitates population synchronization of uncoupled cellular oscillators to external signals. *13th Conference on Dynamical Systems Applied to Biology and Natural Sciences*, Virtual, 2022.
- [5] Vojtech Kumpost. Poster: Elucidating mechanisms underlying experimentally induced changes in the zebrafish circadian clock. *Annual Meeting of the Society for Mathematical Biology*, Virtual, 2020.
- [6] Vojtech Kumpost. Talk: Molecular noise facilitates population synchronization of uncoupled cellular oscillators to external signals. *CompSysBio: Advanced Lecture Course on Computational Systems Biology*, Aussois, France, 2021.



# Bibliography

- [7] Gene F. Franklin, J. David Powell, and Abbas Emami-Naeini. *Feedback control of dynamic systems*. Pearson, 7th edition, 2014.
- [8] Jean-Marc Ginoux and Christophe Letellier. Van der Pol and the history of relaxation oscillations: toward the emergence of a concept. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(2):023120, 2012.
- [9] Auke Jan Ijspeert. Central pattern generators for locomotion control in animals and robots: a review. *Neural Networks*, 21(4):642–653, 2008.
- [10] Colin S Pittendrigh. Temporal organization: reflections of a Darwinian clock-watcher. *Annual Review of Physiology*, 55:17–54, 1993.
- [11] Brian O’Rourke, Brian M Ramza, and Eduardo Marban. Oscillations of membrane current and excitability driven by metabolic oscillations in heart cells. *Science*, 265(5174):962–966, 1994.
- [12] Mustafa Khammash. An engineering viewpoint on biological robustness. *BMC Biology*, 14(1):1–11, 2016.
- [13] Béla Novák and John J Tyson. Design principles of biochemical oscillators. *Nature Reviews Molecular Cell Biology*, 9(12):981–991, 2008.
- [14] Yuri Lazebnik. Can a biologist fix a radio?—Or, what i learned while studying apoptosis. *Cancer Cell*, 2(3):179–182, 2002.
- [15] Harley H McAdams and Adam Arkin. It’s a noisy business! Genetic regulation at the nanomolar scale. *Trends in Genetics*, 15(2):65–69, 1999.
- [16] Irina Mihalcescu, Weihong Hsing, and Stanislas Leibler. Resilient circadian oscillator revealed in individual cyanobacteria. *Nature*, 430(6995):81–85, 2004.
- [17] Galit Lahav, Nitzan Rosenfeld, Alex Sigal, Naama Geva-Zatorsky, Arnold J Levine, Michael B Elowitz, and Uri Alon. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nature Genetics*, 36(2):147–150, 2004.

- [18] Avigdor Eldar and Michael B Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467(7312):167–173, 2010.
- [19] Marc Lefranc. Entraining oscillations in the NF- $\kappa$ B signaling system: with a little help from noise. *Cell Systems*, 3(6):507–508, 2016.
- [20] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC Press, 2019.
- [21] Yoshiaki Futakata and Tetsuya Iwasaki. Entrainment to natural oscillations via uncoupled central pattern generators. *IEEE Transactions on Automatic Control*, 56(5):1075–1089, 2010.
- [22] Jeffrey D Plautz, Maki Kaneko, Jeffrey C Hall, and Steve A Kay. Independent photoreceptive circadian clocks throughout drosophila. *Science*, 278(5343):1632–1635, 1997.
- [23] Takuya Matsuo, Shun Yamaguchi, Shigeru Mitsui, Aki Emi, Fukuko Shimoda, and Hitoshi Okamura. Control mechanism of the circadian clock for timing of cell division in vivo. *Science*, 302(5643):255–259, 2003.
- [24] John Jalife. Mutual entrainment and electrical coupling as mechanisms for synchronous firing of rabbit sino-atrial pace-maker cells. *The Journal of Physiology*, 356(1):221–243, 1984.
- [25] Martin Bier, Barbara M Bakker, and Hans V Westerhoff. How yeast cells synchronize their glycolytic oscillations: a perturbation analytic treatment. *Biophysical Journal*, 78(3):1087–1093, 2000.
- [26] Ryan A Kellogg and Savaş Tay. Noise facilitates transcriptional control under dynamic inputs. *Cell*, 160(3):381–392, 2015.
- [27] Alba Jiménez, Ying Lu, Ashwini Jambhekar, and Galit Lahav. Principles, mechanisms and functions of entrainment in biological oscillators. *Interface Focus*, 12(3):20210088, 2022.
- [28] Steven H. Strogatz. *Nonlinear dynamics and chaos*. CRC Press, 2nd edition, 2015.
- [29] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500, 1952.
- [30] Jie Huang and Donald L Turcotte. Are earthquakes an example of deterministic chaos? *Geophysical Research Letters*, 17(3):223–226, 1990.

- 
- [31] Didier Gonze. Modeling circadian clocks: roles, advantages, and limitations. *Open Life Sciences*, 6(5):712–729, 2011.
- [32] John H Abel and Francis J Doyle III. A systems theoretic approach to analysis and control of mammalian circadian dynamics. *Chemical Engineering Research and Design*, 116: 48–60, 2016.
- [33] Peter C St John, Stephanie R Taylor, John H Abel, and Francis J Doyle III. Amplitude metrics for cellular circadian bioluminescence reporters. *Biophysical Journal*, 107(11): 2712–2722, 2014.
- [34] Ute Abraham, Adrián E Granada, Pål O Westermark, Markus Heine, Achim Kramer, and Hanspeter Herzel. Coupling governs entrainment range of circadian clocks. *Molecular Systems Biology*, 6(1):438, 2010.
- [35] Daniel Kaplan and Leon Glass. *Understanding nonlinear dynamics*. Springer Science & Business Media, 1995.
- [36] Isao T Tokuda, Christoph Schmal, Bharath Ananthasubramaniam, and Hanspeter Herzel. Conceptual models of entrainment, jet lag, and seasonality. *Frontiers in Physiology*, 11: 334, 2020.
- [37] Balth Van der Pol. LXXXVIII. On “relaxation-oscillations”. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):978–992, 1926.
- [38] Namiko Mitarai, Uri Alon, and Mogens H Jensen. Entrainment of noise-induced and limit cycle oscillators under weak noise. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(2):023125, 2013.
- [39] Megan E Jewett and Richard E Kronauer. Refinement of a limit cycle oscillator model of the effects of light on the human circadian pacemaker. *Journal of Theoretical Biology*, 192 (4):455–465, 1998.
- [40] Krzysztof Grudziński and Jan J Żebrowski. Modeling cardiac pacemakers with relaxation oscillators. *Physica A: Statistical Mechanics and its Applications*, 336(1-2):153–162, 2004.
- [41] Brian P. Ingalls. *Mathematical modeling in systems biology*. MIT Press, 2013.
- [42] John J Tyson, Attila Csikasz-Nagy, and Bela Novak. The dynamics of cell cycle regulation. *Bioessays*, 24(12):1095–1109, 2002.
- [43] Brian C Goodwin. Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation*, 3:425–437, 1965.

- [44] Luis A Sanchez. Global asymptotic stability of the Goodwin system with repression. *Nonlinear Analysis: Real World Applications*, 10(4):2151–2156, 2009.
- [45] Jianzhi Cao and Haijun Jiang. Stability and Hopf bifurcation analysis on Goodwin model with three delays. *Chaos, Solitons & Fractals*, 44(8):613–618, 2011.
- [46] Didier Gonze and Peter Ruoff. The Goodwin oscillator and its legacy. *Acta Biotheoretica*, 69(4):857–874, 2021.
- [47] John S Griffith. Mathematics of cellular control processes I. Negative feedback to one gene. *Journal of Theoretical Biology*, 20(2):202–208, 1968.
- [48] Jae Kyoung Kim and Daniel B Forger. A mechanism for robust circadian timekeeping via stoichiometric balance. *Molecular Systems Biology*, 8:630, 2012.
- [49] Jae Kyoung Kim. Protein sequestration versus Hill-type repression in circadian clock models. *IET Systems Biology*, 10(4):125–135, 2016.
- [50] Jae Kyoung Kim, Zachary P Kilpatrick, Matthew R Bennett, and Krešimir Josić. Molecular mechanisms that regulate the coupled period of the mammalian circadian clock. *Biophysical Journal*, 106(9):2071–2081, 2014.
- [51] Jae Kyoung Kim and John J Tyson. Misuse of the Michaelis–Menten rate law for protein interaction networks and its remedy. *PLOS Computational Biology*, 16(10):e1008258, 2020.
- [52] Martin Clayton. What is entrainment? Definition and applications in musical research. *Empirical Musicology Review*, 7:49–56, 2012.
- [53] Keith B Doelling, M Florencia Assaneo, Dana Bevilacqua, Bijan Pesaran, and David Poeppel. An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences USA*, 116(20):10113–10121, 2019.
- [54] T Pham Dinh, J Demongeot, P Baconnier, and G Benchetrit. Simulation of a biological oscillator: the respiratory system. *Journal of Theoretical Biology*, 103(1):113–132, 1983.
- [55] Mogens H Jensen and Sandeep Krishna. Inducing phase-locking and chaos in cellular oscillators by modulating the driving stimuli. *FEBS Letters*, 586(11):1664–1668, 2012.
- [56] Bharath Ananthasubramaniam, Christoph Schmal, and Hanspeter Herzel. Amplitude effects allow short jet lags and large seasonal phase shifts in minimal clock models. *Journal of Molecular Biology*, 432(12):3722–3737, 2020.



- 
- [57] Christoph Schmal, Daisuke Ono, Jihwan Myung, J Patrick Pett, Sato Honma, Ken-Ichi Honma, Hanspeter Herzel, and Isao T Tokuda. Weak coupling between intracellular feedback loops explains dissociation of clock gene dynamics. *PLOS Computational Biology*, 15(9):e1007330, 2019.
- [58] Raphaela Heussen and David Whitmore. The importance of stochastic effects for explaining entrainment in the zebrafish circadian clock. *Computational and Mathematical Methods in Medicine*, 2015, 2015.
- [59] Jean-Christophe Leloup and Albert Goldbeter. Toward a detailed computational model for the mammalian circadian clock. *Proceedings of the National Academy of Sciences USA*, 100(12):7051–7056, 2003.
- [60] Vladimir Igorevich Arnold. Small denominators, I: mappings of the circumference into itself. *American Mathematical Society Translations: Series 2*, 46:213, 1965.
- [61] Mogens H Jensen, Per Bak, and Tomas Bohr. Transition to chaos by interaction of resonances in dissipative systems. I. Circle maps. *Physical Review A*, 30(4):1960, 1984.
- [62] M Zalalutdinov, Keith L Aubin, Manoj Pandey, Alan Taylor Zehnder, Richard H Rand, Harold G Craighead, Jeevak M Parpia, and Brian H Houston. Frequency entrainment for micromechanical oscillator. *Applied Physics Letters*, 83(16):3281–3283, 2003.
- [63] Christoph Schmal, Jihwan Myung, Hanspeter Herzel, and Grigory Bordyugov. A theoretical study on seasonality. *Frontiers in Neurology*, 6:94, 2015.
- [64] Carl Hirschie Johnson. Forty years of PRCs-what have we learned? *Chronobiology International*, 16(6):711–743, 1999.
- [65] Leon Glass and A. T. Winfree. Discontinuities in phase-resetting experiments. *American Journal of Physiology*, 246(2):R251–R258, 1984.
- [66] Michael R Guevara, Alvin Shrier, and Leon Glass. Phase resetting of spontaneously beating embryonic ventricular heart cell aggregates. *American Journal of Physiology-Heart and Circulatory Physiology*, 251(6):H1298–H1305, 1986.
- [67] Tetsuro Funato, Yuki Yamamoto, Shinya Aoi, Takashi Imai, Toshio Aoyagi, Nozomi Tomita, and Kazuo Tsuchiya. Evaluation of the phase-dependent rhythm control of human walking using phase response curves. *PLOS Computational Biology*, 12(5):1–23, 2016.
- [68] Daiki Tamura, Shinya Aoi, Tetsuro Funato, Soichiro Fujiki, Kei Senda, and Kazuo Tsuchiya. Contribution of phase resetting to adaptive rhythm control in human walking based on the phase response curves of a neuromusculoskeletal model. *Frontiers in Neuroscience*, 14:17, 2020.

- [69] Darren J Wilkinson. Stochastic modelling for quantitative description of heterogeneous biological systems. *Nature Reviews Genetics*, 10(2):122–133, 2009.
- [70] Maria Luisa Guerriero, Alexandra Pokhilko, Aurora Piñas Fernández, Karen J Halliday, Andrew J Millar, and Jane Hillston. Stochastic properties of the plant circadian clock. *Journal of the Royal Society Interface*, 9(69):744–756, 2012.
- [71] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976.
- [72] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- [73] Daniel T Gillespie. The chemical Langevin equation. *The Journal of Chemical Physics*, 113(1):297–306, 2000.
- [74] Desmond J Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43(3):525–546, 2001.
- [75] Didier Gonze, José Halloy, and Albert Goldbeter. Deterministic versus stochastic models for circadian rhythms. *Journal of Biological Physics*, 28(4):637–653, 2002.
- [76] Didier Gonze, José Halloy, and Albert Goldbeter. Robustness of circadian rhythms with respect to molecular noise. *Proceedings of the National Academy of Sciences USA*, 99(2):673–678, 2002.
- [77] Didier Gonze and Albert Goldbeter. Circadian rhythms and molecular noise. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16(2):026110, 2006.
- [78] José MG Vilar, Hao Yuan Kueh, Naama Barkai, and Stanislas Leibler. Mechanisms of noise-resistance in genetic oscillators. *Proceedings of the National Academy of Sciences USA*, 99(9):5988–5992, 2002.
- [79] Michele Monti, David K Lubensky, and Pieter Rein Ten Wolde. Optimal entrainment of circadian clocks in the presence of noise. *Physical Review E*, 97(3):032405, 2018.
- [80] Guanyu Wang and Charles S Peskin. Entrainment of a cellular circadian oscillator by light in the presence of molecular noise. *Physical Review E*, 97(6):062416, 2018.
- [81] Benjamin Lindner, Jordi Garcia-Ojalvo, Alexander Neiman, and Lutz Schimansky-Geier. Effects of noise in excitable systems. *Physics Reports*, 392(6):321–424, 2004.

- 
- [82] Mark D McDonnell and Lawrence M Ward. The benefits of noise in neural systems: bridging theory and experiment. *Nature Reviews Neuroscience*, 12(7):415–425, 2011.
- [83] Luca Gammaitoni, Peter Hänggi, Peter Jung, and Fabio Marchesoni. Stochastic resonance. *Reviews of Modern Physics*, 70(1):223, 1998.
- [84] Mark D McDonnell and Derek Abbott. What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology. *PLoS Computational Biology*, 5(5):e1000348, 2009.
- [85] Lawrence M Ward, Sam M Doesburg, Keiichi Kitajo, Shannon E MacLean, and Alexa B Roggeveen. Neural synchrony in stochastic resonance, attention, and consciousness. *Canadian Journal of Experimental Psychology*, 60(4):319, 2006.
- [86] A Aldo Faisal, Luc PJ Selen, and Daniel M Wolpert. Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292–303, 2008.
- [87] James J Collins, Thomas T Imhoff, and Peter Grigg. Noise-enhanced information transmission in rat SA1 cutaneous mechanoreceptors via aperiodic stochastic resonance. *Journal of Neurophysiology*, 76(1):642–645, 1996.
- [88] Toshio Mori and Shoichi Kai. Noise-induced entrainment and stochastic resonance in human brain waves. *Physical Review Letters*, 88(21):218101, 2002.
- [89] Zhonghuai Hou and Houwen Xin. Internal noise stochastic resonance in a circadian clock system. *The Journal of Chemical Physics*, 119(22):11508–11512, 2003.
- [90] Ming Yi and Ya Jia. Light-noise-induced suprathreshold circadian oscillations and coherent resonance in *Drosophila*. *Physical Review E*, 72(1):012902, 2005.
- [91] Jingkui Wang, Marc Lefranc, and Quentin Thommen. Stochastic oscillations induced by intrinsic fluctuations in a self-repressing gene. *Biophysical Journal*, 107(10):2403–2416, 2014.
- [92] Ankit Gupta, Benjamin Hepp, and Mustafa Khammash. Noise induces the population-level entrainment of incoherent, uncoupled intracellular oscillators. *Cell Systems*, 3(6):521–531, 2016.
- [93] Octavio Mondragón-Palomino, Tal Danino, Jangir Selimkhanov, Lev Tsimring, and Jeff Hasty. Entrainment of a population of synthetic genetic oscillators. *Science*, 333(6047):1315–1319, 2011.
- [94] Simon C Thain, Anthony Hall, and Andrew J Millar. Functional independence of circadian clocks that regulate plant gene expression. *Current Biology*, 10(16):951–956, 2000.

- [95] Charna Dibner, Ueli Schibler, and Urs Albrecht. The mammalian circadian timing system: organization and coordination of central and peripheral clocks. *Annual Review of Physiology*, 72:517–549, 2010.
- [96] Amanda-Jayne F Carr and David Whitmore. Imaging of single light-responsive clock cells reveals fluctuating free-running periods. *Nature Cell Biology*, 7(3):319–321, 2005.
- [97] Frank Fan and Keith V Wood. Bioluminescent assays for high-throughput screening. *Assay and Drug Development Technologies*, 5(1):127–136, 2007.
- [98] Florian Dörfler and Francesco Bullo. Synchronization in complex networks of phase oscillators: A survey. *Automatica*, 50(6):1539–1564, 2014.
- [99] Samuel Bernard, Didier Gonze, Branka Čajavec, Hanspeter Herzel, and Achim Kramer. Synchronization-induced rhythmicity of circadian oscillators in the suprachiasmatic nucleus. *PLOS Computational Biology*, 3(4):e68, 2007.
- [100] Neda Bagheri, Stephanie R Taylor, Kirsten Meeker, Linda R Petzold, and Francis J Doyle III. Synchrony and entrainment properties of robust circadian oscillators. *Journal of The Royal Society Interface*, 5(suppl\_1):S17–S28, 2008.
- [101] Ethan D Buhr, Seung-Hee Yoo, and Joseph S Takahashi. Temperature as a universal resetting cue for mammalian circadian oscillators. *Science*, 330(6002):379–385, 2010.
- [102] John Hongyu Meng and Hermann Riecke. Synchronization by uncorrelated noise: interacting rhythms in interconnected oscillator networks. *Scientific Reports*, 8(1):1–14, 2018.
- [103] Sungwon An, Rich Harang, Kirsten Meeker, Daniel Granados-Fuentes, Connie A Tsai, Cristina Mazuski, Jihee Kim, Francis J Doyle III, Linda R Petzold, and Erik D Herzog. A neuropeptide speeds circadian entrainment by reducing intercellular synchrony. *Proceedings of the National Academy of Sciences USA*, 110(46):E4355–E4361, 2013.
- [104] Jing Chen. Stochastic gene expression. In *Case Studies in Systems Biology*. Springer, 2021.
- [105] David Schnoerr, Guido Sanguinetti, and Ramon Grima. The complex chemical Langevin equation. *The Journal of Chemical Physics*, 141(2):07B606\_1, 2014.
- [106] Lawrence F Shampine, Skip Thompson, JA Kierzenka, and GD Byrne. Non-negative solutions of ODEs. *Applied Mathematics and Computation*, 170(1):556–569, 2005.
- [107] Andrew Thwaites, Ian Nimmo-Smith, Elisabeth Fonteneau, Roy D Patterson, Paula Buttery, and William D Marslen-Wilson. Tracking cortical entrainment in neural activity:

- auditory processes in human temporal cortex. *Frontiers in Computational Neuroscience*, 9:5, 2015.
- [108] Kai Hormann and Alexander Agathos. The point in polygon problem for arbitrary polygons. *Computational Geometry*, 20(3):131–144, 2001.
- [109] Till Roenneberg, Serge Daan, and Martha Merrow. The art of entrainment. *Journal of Biological Rhythms*, 18(3):183–194, 2003.
- [110] Steven H Strogatz. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1-4):1–20, 2000.
- [111] Malte Schröder, Marc Timme, and Dirk Witthaut. A universal order parameter for synchrony in networks of limit cycle oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(7):073119, 2017.
- [112] Arthur T Winfree. Integrated view of resetting a circadian clock. *Journal of Theoretical Biology*, 28(3):327–374, 1970.
- [113] Jonas Buchli, Ludovic Righetti, and Auke Jan Ijspeert. Engineering entrainment and adaptation in limit cycle systems. *Biological Cybernetics*, 95(6):645–664, 2006.
- [114] Ruiting Zhang, Tiejun Xiao, and Zhonghuai Hou. Internal noise enhanced oscillation in a delayed circadian pacemaker. *Biophysical Chemistry*, 158(1):54–60, 2011.
- [115] Pål O Westermark, David K Welsh, Hitoshi Okamura, and Hanspeter Herzel. Quantification of circadian rhythms in single cells. *PLOS Computational Biology*, 5(11):e1000580, 2009.
- [116] J Hizanidis, Alexander G Balanov, Andreas Amann, and Eckehard Schöll. Noise-induced oscillations and their control in semiconductor superlattices. *International Journal of Bifurcation and Chaos*, 16(06):1701–1710, 2006.
- [117] Aurore Woller, Didier Gonze, and Thomas Erneux. The Goodwin model revisited: Hopf bifurcation, limit-cycle, and periodic entrainment. *Physical Biology*, 11(4):045002, 2014.
- [118] Peter S Swain, Michael B Elowitz, and Eric D Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences USA*, 99(20):12795–12800, 2002.
- [119] Dao-Guang Wang, Shaobing Wang, Bo Huang, and Feng Liu. Roles of cellular heterogeneity, intrinsic and extrinsic noise in variability of p53 oscillation. *Scientific Reports*, 9(1):1–11, 2019.

- [120] Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths. *Synchronization: a universal concept in nonlinear sciences*. Cambridge University Press, 2001.
- [121] Satyajith Amaran, Nikolaos V Sahinidis, Bikram Sharda, and Scott J Bury. Simulation optimization: a review of algorithms and applications. *Annals of Operations Research*, 240(1):351–380, 2016.
- [122] Jeff Hasty, Milos Dolnik, Vivi Rottschäfer, and James J Collins. Synthetic gene network for entraining and amplifying cellular oscillations. *Physical Review Letters*, 88(14):148101, 2002.
- [123] Rainer Storn and Kenneth Price. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.
- [124] Tarik Eltaeib and Ausif Mahmood. Differential evolution: a survey and analysis. *Applied Sciences*, 8(10):1945, 2018.
- [125] Petru A Simionescu. A collection of bivariate nonlinear optimisation test problems with graphical representations. *International Journal of Mathematical Modelling and Numerical Optimisation*, 10(4):365–398, 2020.
- [126] James CW Locke, Andrew J Millar, and Matthew S Turner. Modelling genetic networks with noisy and varied experimental data: The circadian clock in *Arabidopsis thaliana*. *Journal of Theoretical Biology*, 234(3):383–393, 2005.
- [127] Peter C St John and Francis J Doyle III. Quantifying stochastic noise in cultured circadian reporter cells. *PLOS Computational Biology*, 11(11):e1004451, 2015.
- [128] J Eamonn Nash and Jonh V Sutcliffe. River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3):282–290, 1970.
- [129] Pavel Kraikivski. Computational software. In *Case Studies in Systems Biology*. Springer, 2021.
- [130] Stefan Hoops, Sven Sahle, Ralph Gauges, Christine Lee, Jürgen Pahle, Natalia Simus, Mudita Singhal, Liang Xu, Pedro Mendes, and Ursula Kummer. Copasi—a complex pathway simulator. *Bioinformatics*, 22(24):3067–3074, 2006.
- [131] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017.

- 
- [132] Christopher Rackauckas and Qing Nie. DifferentialEquations.jl—a performant and feature-rich ecosystem for solving differential equations in Julia. *Journal of Open Research Software*, 5(1), 2017.
- [133] Carrie L Partch, Carla B Green, and Joseph S Takahashi. Molecular architecture of the mammalian circadian clock. *Trends in Cell Biology*, 24(2):90–99, 2014.
- [134] Till Roenneberg and Russell G Foster. Twilight times: Light and the circadian system. *Photochemistry and Photobiology*, 66(5):549–561, 1997.
- [135] Tara A LeGates, Diego C Fernandez, and Samer Hattar. Light as a central modulator of circadian rhythms, sleep and affect. *Nature Reviews Neuroscience*, 15(7):443–454, 2014.
- [136] Akhilesh B Reddy and John S O’Neill. Healthy clocks, healthy body, healthy mind. *Trends in Cell Biology*, 20(1):36–44, 2010.
- [137] Gad Vatine, Daniela Vallone, Yoav Gothilf, and Nicholas S Foulkes. It’s time to swim! Zebrafish and the circadian clock. *FEBS Letters*, 585(10):1485–1494, 2011.
- [138] David Whitmore, Nicholas S Foulkes, and Paolo Sassone-Corsi. Light acts directly on organs and cells in culture to set the vertebrate circadian clock. *Nature*, 404(6773):87–91, 2000.
- [139] Aurélio Balsalobre, Francesca Damiola, and Ueli Schibler. A serum shock induces circadian gene expression in mammalian tissue culture cells. *Cell*, 93(6):929–937, 1998.
- [140] Aurélio Balsalobre, Steven A Brown, Lysiane Marcacci, François Tronche, Christoph Kellendonk, Holger M Reichardt, Günther Schütz, and Ueli Schibler. Resetting of circadian time in peripheral tissues by glucocorticoid signaling. *Science*, 289(5488):2344–2347, 2000.
- [141] Daniela Vallone, Srinivas Babu Gondi, David Whitmore, and Nicholas S. Foulkes. E-box function in a period gene repressed by light. *Proceedings of the National Academy of Sciences USA*, 101(12):4106–4111, 2004.
- [142] TK Tamai, AJ Carr, and D Whitmore. Zebrafish circadian clocks: Cells that see light. *Biochemical Society Transactions*, 33(5):962–966, 2005.
- [143] Gad Vatine, Daniela Vallone, Lior Appelbaum, Philipp Mracek, Zohar Ben-Moshe, Kajori Lahiri, Yoav Gothilf, and Nicholas S Foulkes. Light directs zebrafish period2 expression via conserved D and E boxes. *PLOS Biology*, 7(10):e1000223, 2009.

- [144] Philipp Mracek, Cristina Santoriello, M Laura Idda, Cristina Pagano, Zohar Ben-Moshe, Yoav Gothilf, Daniela Vallone, and Nicholas S Foulkes. Regulation of per and cry genes reveals a central role for the D-box enhancer in light-dependent gene expression. *PLOS One*, 7(12):e51278, 2012.
- [145] Srinivas Babu Gondi. *Circadian expression and promoter analysis of the zebrafish Period-4 gene*. PhD thesis, Eberhard-Karls-Universität Tübingen, 2004.
- [146] Nicolas Cermakian, Matthew P Pando, Carol L Thompson, Anna B Pinchak, Christopher P Selby, Laura Gutierrez, Dan E Wells, Gregory M Cahill, Aziz Sancar, and Paolo Sassone-Corsi. Light induction of a vertebrate clock gene involves signaling through blue-light receptors and MAP kinases. *Current Biology*, 12(10):844–848, 2002.
- [147] Jun Hirayama, Norio Miyamura, Yoshimi Uchida, Yoichi Asaoka, Reiko Honda, Kenji Sawanobori, Takeshi Todo, Takuro Yamamoto, Paolo Sassone-Corsi, and Hiroshi Nishina. Common light signaling pathways controlling DNA repair and circadian clock entrainment in zebrafish. *Cell Cycle*, 8(17):2794–2801, 2009.
- [148] Philipp Mracek, Cristina Pagano, Nadine Fröhlich, M Laura Idda, Ines H Cuesta, Jose Fernando Lopez-Olmeda, F Javier Sánchez-Vázquez, Daniela Vallone, and Nicholas S Foulkes. ERK signaling regulates light-induced gene expression via D-box enhancers in a differential, wavelength-dependent manner. *PLOS One*, 8(6):e67858, 2013.
- [149] Bruno C R Ramos, Maria Nathália C. M. Moraes, Maristela O Poletini, Leonardo H R G Lima, and Ana Maria L Castrucci. From blue light to clock genes in zebrafish ZEM-2S cells. *PLOS One*, 9(9):e106252, 2014.
- [150] Niko Komin, Adrian C Murza, Emilio Hernández-García, and Raul Toral. Synchronization and entrainment of coupled circadian oscillators. *Interface Focus*, 1(1):167–176, 2011.
- [151] Didier Gonze, Samuel Bernard, Christian Waltermann, Achim Kramer, and Hanspeter Herzel. Spontaneous synchronization of coupled circadian oscillators. *Biophysical Journal*, 89(1):120–129, 2005.
- [152] Daniel B Forger. Signal processing in cellular clocks. *Proceedings of the National Academy of Sciences USA*, 108(11):4281–4285, 2011.
- [153] John S O’Neill and Akhilesh B Reddy. The essential role of cAMP/Ca<sup>2+</sup> signalling in mammalian circadian timekeeping. *Biochemical Society Transactions*, 40(1):44–50, 2012.
- [154] Ilona Brändlin, Susanne Hübner, Tim Eiseler, Marina Martinez-Moya, Andreas Horschinek, Angelika Hausser, Gisela Link, Steffen Rupp, Peter Storz, Klaus Pfizenmaier, et al. Protein kinase C (PKC)  $\eta$ -mediated PKC $\mu$  activation modulates ERK and JNK signal pathways. *Journal of Biological Chemistry*, 277(8):6490–6496, 2002.



- [155] Jae Kyoung Kim, DB Forger, M Marconi, D Wood, A Doran, T Wager, C Chang, and KM Walton. Modeling and validating chronic pharmacological manipulation of circadian rhythms. *CPT: Pharmacometrics & Systems Pharmacology*, 2(7):1–11, 2013.
- [156] John H Abel, Ankush Chakrabarty, Elizabeth B Klerman, and Francis J Doyle III. Pharmaceutical-based entrainment of circadian phase via nonlinear model predictive control. *Automatica*, 100:336–348, 2019.
- [157] Dae Wook Kim, Cheng Chang, Xian Chen, Angela C Doran, Francois Gaudreault, Travis Wager, George J DeMarco, and Jae Kyoung Kim. Systems approach reveals photosensitivity and PER2 level as determinants of clock-modulator efficacy. *Molecular Systems Biology*, 15(7):e8838, 2019.