Topics in Applied Stochastic Dynamics

by

Katarina Bodova

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied and Interdisciplinary Mathematics)
in The University of Michigan
2009

Doctoral Committee:

Professor Charles R. Doering, Co-Chair
Assistant Professor Anna Amirdjanova, Co-Chair
Professor Joseph Conlon
Assistant Professor Daniel B. Forger

ACKNOWLEDGEMENTS

First I would like to thank Peter Smereka, who helped me change my mind and turn my attention from financial mathematics to broader areas of applied math. Then I would like to thank Charlie Doering for giving me loads of advice and guidance over the last three years as well as for conveying the understanding of various physical phenomena to me many times mostly via color chalk. Both Danny Forger and Anna Amirdjanova deserve many thanks for answering my questions and introducing me to new fields in applied mathematics. Joe Conlon deserves thanks for the patience he showed over the past two years, during which I was TA for the class he taught. I cannot forget George Lin and Mohar Guha for fruitful mathematical discussions and my fellow graduate students, namely, Mark Iwen, Andy Stein, Arvind Baskaran, Liz Vivas, Johanna Mangahas, Alan Stapledon and Jose Gomez-Guerra for keeping me occupied with non-math related activities. Further thanks go to Martha Carlson and Emily Baecher, for giving me reasons to play Ultimate Frisbee. Finally, but most importantly, my deepest thanks belongs to Richard Kollár who was my greatest source of support over the past five years and more.

Chapter II is a joint work with Danny Forger that began as a project for a Math Biology class (Math 563) taught by Danny. Chapter III is a collaboration with Charlie Doering and George Lin.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

LIST OF ABBREVIATIONS

DDT dispersion-diffusion theory

FPE Fokker-Planck equation

HH Hodgkin-Huxley model

HT homogenization theory

HRBC homogenous Rayleigh-Bénard convection

ILT internal layer theory

LSHH Langevin-type stochastic HH model

MC Monte Carlo method

MCR Monte Carlo method with recycling

MSHH Markov-type stochastic HH model

RBC Rayleigh-Bénard convection

SE Schrődinger equation

SHH stochastic HH model

SLF stochastic Lyapunov function

S-L Sturm-Liouville problem

WKB Wentzel-Kramers-Brillouin approximation

CHAPTER I

Introduction

Many processes in nature exhibit a natural level of randomness due to various sources. In some cases, the level of randomness is low and deterministic models give quantitatively correct description of the problem. The purpose of this work is to study processes in which the effect of noise is not obvious or even counterintuitive. The core of the thesis is organized into three chapters, consisting of three different applications of stochastic dynamic modelling. The unifying goal of Chapters II, III and IV is to illustrate the counterintuitive roles and effects of noise on dynamics in diametrically different scenarios in which the origin and magnitude of randomness varies dramatically.

In Chapter II we study a system in which randomness comes from the finite nature of the underlying mechanism at a very small scale compared to the size of the system. Those systems are abundant in biological applications, ecological systems, interacting particle and chemical kinetics and other areas. The particular example studied in Chapter II is a model for a neuron whose basic biophysical function is to propagate electrical information within the neural network. In the underlying process the randomness arises from finiteness of the number of ion channels per unit area (density) in a membrane of an axon. The finiteness of the ion charges is neglected since it affects smaller time- and space- scales. The ion channel density varies for different

types of neurons and is one of the key factors that determines the functionality of the neuron. The computational study of the effects of channel noise, contained in Chapter II, consists of examination and comparison of two different stochastic approaches, based on a classical mathematical description by *Hodgkin and Huxley* (1952).

In many applications it is customary to represent a stochastic process, based on particle motion, by a deterministic and continuous description of the particle density. That reduces the complexity of the problem. The two approaches are different in nature and do not necessarily yield the same results. One typical example is a replacement of molecular diffusion by the diffusion equation. Chapter III contains a study of the fluid properties, in which the molecular diffusion is enhanced by stirring (advection) and at the same time constantly replenished and depleted by inhomogeneous sources and sinks. The mixing properties of the stirred fluid has been studied by using both asymptotic methods in the continuous advection diffusion equation by *Shaw et al.* (2007) (in the presence of sources and sinks) and underlying stochastic processes by *Majda and MacLaughlin* (1993) and *Fannjiang and Papanicolaou* (1994) (without sources and sinks). The two methods give two different predictions for the mixing efficiency of stirring. A way to resolve the disagreement between these two methods is to study the underlying stochastic processes and formulate a consistent description of the problem.

Even when process arising from some physical application is deterministic by its nature, some randomness is introduced during a numerical implementation of its mathematical model. Round-off error is often of a very small magnitude and intuitively it should not play any major role in determining the simulation results. However, in Chapter IV we construct a system of ordinary differential equations in two dimensions, that mimics the dynamical features of the Rayleigh-Bénard convection, where arbitrarily small random perturbation may change the system properties dramatically. For almost all initial conditions, the solutions of the deterministic model

diverge exponentially to infinity, while the solutions of the noise-perturbed system exhibit statistical oscillations. The technique used to prove existence of an invariant measure is based on a construction of auxiliary Lyapunov functions for stochastic systems, introduced by *Khasminskii* (1980). The properties of the constructed system are studied numerically and analytically as the first exit time problem.

CHAPTER II

Properties of a noise-induced firing and quiescence in a Hodgkin-Huxley model

2.1 Introduction

Neurons are essential building blocks of extremely complex, and not yet fully understood, systems of neural network. The complexity of a neural network is determined by the number and function of neurons and the number of synapses (connections) among them. For example, the human brain contains about 10^{11} individual neurons, each one connecting to about 7000 other neurons. Neurons are highly nonlinear, basically uni-directional devices, designed to process and transport electrical signals called action potentials (AP). Even though a signal can propagate in both directions inside the neuron, the physiology of synapses allows propagation in only one way. Received electrical signals enter a neuron through dendrites that form a branching structure around the cell body (soma). The signal is then modified and propagated in an axon. An axon is an excitable, cable-like device, covered by a membrane with ion channels that determine its electrical behavior. An electrical signal leaves axon through multiple axon terminals that synapse onto other neurons.

Neural electrical activity has been observed to contain a significant amount of variability, which derives from several sources. Presynaptic variability, affecting the

electrical signal before it enters a neuron, can arise from a variety of sources and propagate through this network. However, even if the same presynaptic stimulus is presented multiple times, the response of the neuron will still vary from trial to trial (*White et al.*, 2000; *Verveen and Derksen*, 1968). One reason for this is that the mechanism of synaptic transmission is itself random (*Aidley*, 1978). The mechanism of synaptic transmission is controlled by the stochastic opening and closing of synaptic channels, triggered by neurotransmitters that carry electrical signals inside the neuron. But even if the forementioned sources of randomness are eliminated, the electrical behavior of an axon is constructed from the random opening and closing of a finite number of ion channels. While other sources of variability merit study in their own right, here we will study variability of electrical responses within the neuron arising from the finite number of ion channels.

Ion channels are small electrically sensitive devices in the axonal membrane that open or close based on the electrical potential difference across the membrane. Ions can flow through a channel only if the channel is open. Channels are typically selective, i.e., only permeable to certain types of ions. The width of each channel is only slightly larger than an individual ion (*Aidley*, 1978), therefore, the ion transport through the channel is discrete in its nature. Moreover, ions are subject to molecular diffusion, which results in the random nature of opening and closing of each channel. Even though the current passing through an individual channel is rather small and the number of channels is very large, the fluctuations in the electrical output, due to random channel mechanism, might not be negligible. The impact of channel variability on the neuronal dynamics may be explained by strong nonlinearities in the channel mechanism, and signal propagation. Small fluctuations may cause a qualitative difference in the whole dynamics. See chapter IV for a mathematical example of a nonlinear system of ODE's, where arbitrarily small noise may lead to a qualitative difference in the dynamical properties of the system.

After Neher and Sakmann (*Neher and Sakmann*, 1976) invented the Nobel prize winning patch-clamp technique they found the first experimental evidence of a single channel variability. They measured the electrical noise by extracting a small area of the nerve's membrane. This isolated a tiny patch of a membrane in which the electrical current corresponding to a small number of channels could be measured. *Colquhoun and Sakmann* (1985) later examined single channel currents in a frog muscle and observed a random switching in the current corresponding to the random opening and closing of individual channels, see Fig. 2.1.

Figure 2.1: **Single-channel currents from acetylcholine receptors at the frog muscle end-plate.** Traces show the response to acetylcholine (ACh). From *Colquhoun and Sakmann* (1985).

Over 50 years ago, *Hodgkin and Huxley* (1952) developed a mathematical description of the behavior of one of the largest known axons in nature – the squid giant axon, based on the dynamics of ion channels. This model has been the basis of much research in computational neuroscience. The model did not incorporate the randomness of ion channel dynamics, probably since the neurons they were studying did not show much variability, due to the large numbers of channels. Nevertheless, some randomness can be seen even in the response of the squid giant axon, see *De Felice et al.* (1975). However, many neurons are small ($< 10\mu M$ in diameter), and have relatively low numbers of channels, see *Waxman et al.* (1989). Such neurons can show erratic behavior and random effects must be incorporated in the model in order to accurately describe experimental results.

There has been much work devoted to study stochastic versions of the Hodgkin-Huxley model. Two main approaches for incorporating channel variations are widely

used in the literature. In the first approach, random Gaussian noise terms are added directly inside the continuous HH equations. In *Fox* (1997), *Bazso et al.* (2003), *Schmidt et al.* (2006), *Ozer and Ekmekci* (2005) the noise is multiplied by a voltage and state dependent function, that reflects the intuitive feature that relative level of noise depends on the number of open channels. In *Saarinen et al.* (2006) the variance of the Gaussian term is taken to be constant.

The second approach uses a continuous time, discrete state-space Markov formulation where the state of the system is described by a number of open and closed channels at each time as in *Chow and White* (1996), *Rowat* (2007), *White et al.* (2000), *Schneidman et al.* (1998), *Skaugen and Walloe* (1979), *Clay and DeFelice* (1983), *Bruce* (2006) and the process can be faithfully simulated numerically (*Gillespie*, 1977). The state changes as channels open or close. The differences between multiple numerical algorithms for the Hodgkin-Huxley model are studied by *Mino et al.* (2002).

In this work, we demonstrate that many known phenomena (noise-induced quiescence and oscillations) are present in both stochastic formulations of the model. We perform exhaustive numerical simulations of each model for a wide range of noise parameters and show that the two models give the same results. We analyze the distribution of inter-spike intervals (ISI) that naturally depends not only on noise values, but also on the strength of the applied current. We perform a noise-sensitivity analysis based on the numerical results and show that certain types of channels are much less sensitive to random effects than others.

2.1.1 Hodgkin-Huxley model (HH)

The propagation of an impulse is a consequence of an ion transport through the ionic channels in the membrane. Channels are complicated protein structures of a typical width of about $10 - 15\mathring{A}$ that act as resistors. Each channel contains a

selectivity mechanism, and may be selective to a particular ion. The majority of channels in a brain are either Na^+, K^+ or Ca^{2+} permeable (mostly monovalent cations of diameter $< 6.5\mathring{A}$), and they are mainly permeable to this type of cation (even though in some cases this is not true, see *Dryer et al.* (1989)).

The driving force behind impulse propagation in the neural axon is the potential difference between the inside and outside of the axon. Once a channel opens, ions of the corresponding type are free to flow through the membrane until the electrical potential equilibrates. The flow lasts a very short time (less than 10ms) before the channel closes again.

Modelling channel dynamics and the consequent impulse propagation almost always involves the following simplifying assumptions:

- **Selectivity of channels**: Probability that a specific ion passes through a channel does not depend on the other ions.
- **Independence of channels**: Probability of channel opening and ion transport in a channel does not depend on other channels.
- **Stationarity of channels**: Channel densities are constant in time. Channels do not move and their number remains unchanged in time.
- **No myelination**: The axonal membrane is not myelinated.
- **Spatial homogeneity**: We consider a space-clamped axon.

Propagation of the signal inside the axon and influence of channel behavior on its

dynamics is described by a 4–dimensional system of ordinary differential equations

$$\begin{aligned} C_m \frac{dV}{dt} &= -g_{Na} m^3 h (V - V_{Na}) - g_K n^4 (V - V_K) - g_L (V - V_L) + I_{App}\,, & (2.1)\\ \frac{dm}{dt} &= \alpha_m (1 - m) - \beta_m m\,, & (2.2)\\ \frac{dn}{dt} &= \alpha_n (1 - n) - \beta_n n\,, & (2.3)\\ \frac{dh}{dt} &= \alpha_h (1 - h) - \beta_h h\,, & (2.4) \end{aligned}$$

where V denotes a potential difference between the interior and the exterior of an axon and time is measured in ms. The sodium channel consists of one m gate and three h gates, whereas the potassium channel consists of four n gates. Dynamical variables m and h thus denote probabilities that specific gate is in the open state. An ion channel is permeable to a given kind of ion only if all of its gates are open (for Na and K with probability $m^3 h$ and n^4 respectively). Action potential initiation is driven by sodium channels and prevented by activation of potassium channels. Rates for m, n and h in the HH dynamics are given by

$$\alpha_m = 0.1 \frac{25 - V}{e^{\frac{25 - V}{10}} - 1} \qquad \alpha_n = 0.01 \frac{10 - V}{e^{\frac{10 - V}{10}} - 1} \qquad \alpha_h = 0.07 e^{\frac{-V}{20}}$$

$$\beta_m = 4 e^{-\frac{V}{18}} \qquad \beta_n = 0.125 e^{-\frac{V}{80}} \qquad \beta_h = \frac{1}{e^{\frac{30 - V}{10}} + 1}$$

and parameter values for a squid axon (*Hodgkin and Huxley*, 1952) are

Constant	V_R	V_{Na}	V_K	V_L	g_{Na}	g_K	g_L	C_m
Value	0	115	-12	-10.613	120	36	0.3	1
Units	mV	mV	mV	mV	$mS\,cm^{-2}$	$mS\,cm^{-2}$	$mS\,cm^{-2}$	$\mu F\,cm^{-2}$

Table 2.1: **Parameter values for the HH model.**

2.2 Numerical modelling of stochastic channel dynamics

A simple view of a neuron as a unit that processes and modifies neuronal impulses may be rather oversimplified, but it reveals the fundamental property that

the processing inside the neuron is highly nonlinear in its nature and very rich to mathematically interesting phenomena. The corresponding mathematical Hodgkin-Huxley model contains this nonlinearity and exhibits various dynamical features, among them: nonlinear oscillations, bistability, scale separation, and chaotic behavior (*Peskin*, 2000). However, many aspects of the model may be studied only numerically, mainly due to the nonlinearity embedded in the underlying processes. While the numerical implementation of the deterministic Hodgkin-Huxley model is straightforward, there are multiple and principally different ways of implementing the model with noise. Each of the methods has some advantages and some limitations. Two computational methods, based on two different theoretical descriptions of the underlying stochastic processes, have been examined in literature: the Langevin and Markov descriptions.

Computational methods based on Langevin approach assume that the channel densities are sufficiently large so that a master equation is accurately approximated by a continuous model. The algorithm for the numerical simulation of the stochastic HH model, with the incorporated spatial dependance, is explored by *Fox* (1997) with the random terms being Gaussian with state-dependent second moments (multiplicative noise). The state-dependent Gaussian noise affects the inter-spike interval (ISI) distribution and the recovery period, see *Ozer and Ekmekci* (2005). The magnitude of noise is here parametrized by the membrane area, since the channel densities are assumed to be constant. In addition to the state-dependent channel noise, the capacitance fluctuations and their effect on the ISI properties are studied by *Schmidt et al.* (2006).

If the Gaussian channel noise has a constant, rather than state-dependent moments, the dynamical properties of the random channels will change. These differences in terms of spectrum and ISI distribution are discussed by *Bazso et al.* (2003). For a granule cell model the dynamics is driven by calcium ions. In such case, *Saarinen*

et al. (2006) observe that constant Gaussian noise with a small random extra-synaptic input leads to bursts and occasional spontaneous firing.

The most direct approach to study channel noise in axon is to use a kinetic model, based on Markov processes, that tracks opening and closing of individual channels. This method is particularly useful if the channel densities are low, i.e. there is a lot of variability. However, for large channel densities the method becomes computationally expensive.

The first developments of Markov model based computational methods are attributed to *Clay and DeFelice* (1983). The kinetic simulation shows that there are three major effects of channel variability: induction of spontaneous spiking activity (*Chow and White*, 1996), subthreshold oscillations and missing spikes for suprathreshold inputs (*Schneidman et al.*, 1998; *White et al.*, 2000). The properties of the spiking pattern in the presence of the channel noise are typically studied via the spiking frequency and its dependence on the applied current (*Skaugen and Walloe*, 1979) and via the inter-spike interval properties (*Chow and White*, 1996). In the latter the exponential decay of the ISI distribution is demonstrated and compared to the prediction from the Langevin model. However, the ISI distribution may be bimodal in some cases, see *Rowat* (2007). The author compares the ISI distribution obtained by the Markov based model with ISI distribution obtained by the Langevin kind model. The concentrations of Na and K are fixed throughout the article. With a suitable noise magnitude parameters the two distributions are shown to be nearly identical. In the case of variable input currents *Schneidman et al.* (1998) studies the reliability and precision of spike timing.

The comparison of different Markov-type and Langevin-type methods can be found in the work of *Bruce* (2006). The Langevin method is shown to be naturally faster and this is pronounced for large channel densities. On the other hand, comparing the firing efficiency (fraction of trials when the action potential is generated)

for different methods reveals that the Langevin method leads to some inaccuracies.

In this chapter, we aim to describe two implementations of the stochastic channel dynamics (Langevin and Markov models). The detailed properties of the stochastic dynamics for both models and their consistency are studied in the consequent chapter.

2.2.1 Langevin Stochastic Representation of the HH Model

The first and computationally simpler way to implement variability in the channel dynamics into a HH model is to input a Gaussian random noise directly into channel equations 2.2, 2.3 and 2.4. This approach will be referred to as a Langevin-type stochastic HH model (LSHH). Consider an electric potential equation (2.1) plus a random channel dynamics in the form

$$\frac{dm}{dt} = \alpha_m(1-m) - \beta_m m + g_m(t)\xi_m\,, \tag{2.5}$$

$$\frac{dn}{dt} = \alpha_n(1-n) - \beta_n n + g_n(t)\xi_n\,, \tag{2.6}$$

$$\frac{dh}{dt} = \alpha_h(1-h) - \beta_h h + g_h(t)\xi_h\,. \tag{2.7}$$

where $\xi_m(t)$, $\xi_n(t)$ and $\xi_h(t)$ are independent white noises. The random perturbation of HH is Gaussian with variances

$$g_m^2(t) = \sigma_M^2(\alpha_m(1-m) + \beta_m m) \tag{2.8}$$

$$g_n^2(t) = \sigma_N^2(\alpha_m(1-m) + \beta_m m) \tag{2.9}$$

$$g_h^2(t) = \sigma_H^2(\alpha_m(1-m) + \beta_m m) \tag{2.10}$$

and the "noise magnitudes" that are (unlike in the existing literature) chosen to be inversely proportional to the Na and K channel numbers, are

$$\sigma_m^2 = \frac{1}{3N_{Na}} \tag{2.11}$$

$$\sigma_n^2 = \frac{1}{N_K} \tag{2.12}$$

$$\sigma_h^2 = \frac{1}{N_{Na}} \tag{2.13}$$

where N_{Na} and N_K are the number of sodium and potassium channels, respectively. Note that even though there are three gate variables, there are only two parameters N_{Na} and N_K because the number of m gates is three-fold the number of h gates. The independent parameters of LSHH are therefore noise magnitudes σ_H and σ_N and the constant applied current level I_{app}.

The channel dynamics is fully described by the m, n and h equations and therefore no explicit noise is added to equation (2.1). Note that addition of a white noise to the electric potential equation corresponds to a random perturbation of the applied current. In this work we assume that applied current is constant in time.

The advantages of the direct model are that it is simple to implement numerically with Euler method or higher order methods, and the running time only depends on a time step and not on the number of channels. Unfortunately the noise structure in the Langevin approach may be oversimplified (Gaussian form, independence) and may therefore lead to skewed results.

2.2.2 Markov Chain Representation of the HH Model

The second approach assumes a more realistic structure of noise in the channel dynamics than the LSHH. In the Markov-type stochastic HH model (MSHH) stochastic opening and closing of ion channels is represented as memoryless chemical reactions occuring with transition rates that depend on $\alpha_i(V)$ and $\beta_i(V)$. The channel density,

determined by total number of gates of each type (M, N and H), is fixed and determines system's natural level of randomness. Gates can open or close. We keep track of the total number of open channels of each type ($m \leq M$, $n \leq N$ and $h \leq H$). At

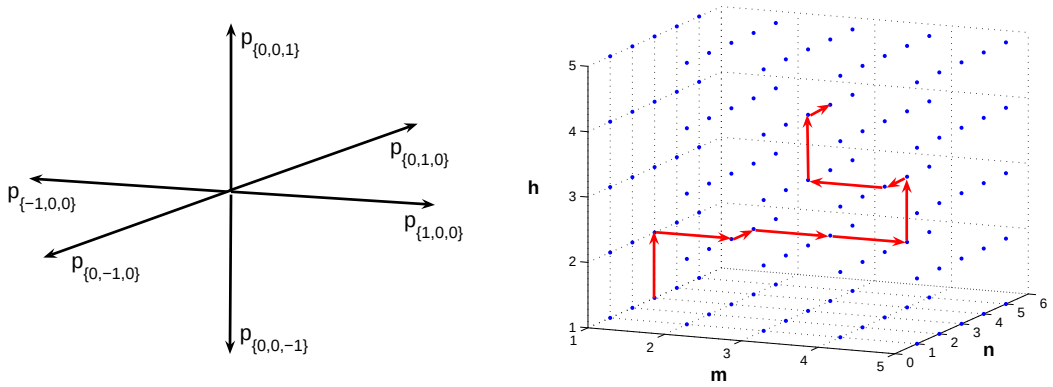

Figure 2.2: **Transition mechanism in the Markov chain representation of the Hodgkin-Huxley model.** (a) Six possible transitions from the state $[m, n, h]$. Exactly one of the gates can open or close at each moment. The probability of each transition is proportional to the corresponding nonlinear rate rescaled so that all possible transitions sum up to 1. (b) One possible realization of the $[m, n, h]$ process. Initial state is $[m, n, h] = [1, 3, 1]$, after 10 transitions the state changes to $[m, n, h] = [3, 4, 4]$.

each point of time where a transition occurs there are exactly six possible outcomes. Each of them has an assigned probability (that depends exclusively on the previous state of the system), see Fig. 2.2. All transition probabilities sum up to one at each transition point. The electric potential equation (2.1) is modified to

$$\frac{dV}{dt} = -g_{Na}\left(\frac{m}{M}\right)^3\left(\frac{h}{H}\right)(V - V_{Na}) - g_K\left(\frac{n}{N}\right)^4(V - V_K) - g_L(V - V_L) + I_{App} \quad (2.14)$$

where probabilities that gates are open are substituted by a relative number of open gates of each type, i.e. $\frac{m}{M}$, $\frac{n}{N}$, $\frac{h}{H}$. The physical condition $M = 3H$ is adopted. The time interval Δt until the next transition is a random number with distribution $\Delta t \sim \text{Exp}(\lambda)$ where λ is the sum of all transition rates as specified in step 2 of algorithm. The following numerical algorithm then describes opening and closing of

ion channels and the signal propagation dynamically in time.

ALGORITHM:

1. **Initialization:** Specify initial values of V, number of open gates m, n, h and number of closed gates $M-m$, $N-n$, $H-h$.

2. **Transition rates:** Calculate the effective total rate of transition

$$\lambda = (\alpha_m(M-m) + \beta_m m) + (\alpha_n(N-n) + \beta_n n) + (\alpha_h(H-h) + \beta_h h)\,.$$

   and individual transition probabilities $p_m^+ = \frac{\alpha_m(M-m)}{\lambda}$, $p_m^- = \frac{\beta_m m}{\lambda}$ and similarly for n and h.

3. **Time step:** Generate two random numbers $r_1, r_2 \sim U[0,1]$. The time to the next transition is:

$$\Delta t = -\frac{1}{\lambda}\log(r_1)\,.$$

4. **Integration:** Integrate the electric potential equation between time t and $t+\Delta t$ using Euler method.

5. **Transition:** Choose one of six possible reactions with a decision based on a randomly generated number r_2 where the reaction rates are

$$\begin{aligned} m-1 &\xleftarrow{\beta_M \cdot m} m \xrightarrow{\alpha_M \cdot (M-m)} m+1 \\ n-1 &\xleftarrow{\beta_N \cdot n} n \xrightarrow{\alpha_N \cdot (N-n)} n+1 \\ h-1 &\xleftarrow{\beta_H \cdot h} h \xrightarrow{\alpha_H \cdot (H-h)} h+1 \end{aligned}$$

6. **Loop:** Repeat steps (2)-(5) until the time reaches the terminal time.

Note that every sodium channel consists of three m-type gates and one h-type gates, therefore a restriction $M = 3H$ is incorporated into the initial condition in the

numerical scheme. The MSHH model, as opposed to LSHH model, contains randomness naturally. The level of variability is encoded in individual channel densities. If, say, $M >> 1$ then the single change of state m produces a change in relative number of open channels of size

$$\left| \frac{m \pm 1}{M} - \frac{m}{M} \right| = \frac{1}{M} .$$

Larger M leads to a smaller change in the above expression, i.e., smaller variability.

2.3 Numerical results

In this chapter we aim to explore dynamical properties of signal propagation in an the axon using both, previously described, stochastic algorithms. In both approaches we analyze the dependance of a spike propagation on (i) applied current (that is constant and deterministic) and on (ii) channel densities/noise magnitudes. The goal is to explore presence of spikes as well as statistical properties of the spiking sequence based on a wide range of parameters of the model. Also the noise sensitivity arguments are new as far as is known to us. The parameter values used in numerical simulations are as follows.

Markov model:

- Applied current: $I_{App} \in \{0, 0.5, 1, 1.5, \ldots, 12\}$
- Number of Na channels: $H = M/3 \in \{200, 500, 800, \ldots, 3800\}$
- Number of K channels: $N \in \{200, 500, 800, \ldots, 8000\}$
- Simulation length: $T = 20s$

Langevin model:

- Applied current: $I_{App} \in \{0, 0.5, 1, 1.5, \ldots, 12\}$
- Na channel noise magnitude: $\sigma_H = \sqrt{3}\sigma_M \in \{0, 0.001, 0.01, 0.02, \ldots, 0.25\}$

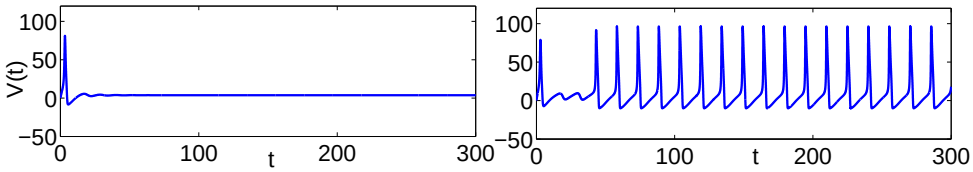

Figure 2.3: **Quiescence and oscillations in the deterministic HH model.** For $I_{App} < I_1$ deterministic HH model is in the quiescent state whereas for $I_{App} > I_2$ it regularly fires APs. Trajectory of HH in the sub-threshold interval $I_{App} = 6$ is on the first plot (no spiking, just an initial relaxation to a rest state) whereas trajectory past the Hopf bifurcation $I_{App} = 9$ is on the second plot. In the bistable region $I_1 < I_{App} < I_2$ trajectories may behave similar to either plot depending on the value of the initial condition.

- K channel noise magnitude: $\sigma_N \in \{0, 0.01, 0.02, \ldots, 0.25\}$
- Simulation length: $T = 200s$

2.3.1 Stochastic Hopf Bifurcation

One of the ways to understand the effect of channel noise on the dynamics of the HH model is to first understand the properties of the deterministic model and only then include variability. Our interest is focused on the influence of applied current on the HH dynamics, see *Rowat* (2007), *Schmidt et al.* (2006), *White et al.* (2000). For low applied currents neuron is in the rest state and it does not propagate any information, see Fig. 2.3 (a). We will call this state quiescent state. As the level of applied current passes a lower threshold $I_{App} = I_1 \approx 7.4$ the system goes through a saddle-node bifurcation of limit cycles where the stable and unstable branches are born, corresponding to the stable and unstable limit cycle. This leads to a bistable dynamics since both the rest state and the limit cycle are stable structures in the model at the same time. After passing an upper threshold $I_{App} = I_2 \approx 8.9$ the system undergoes a subcritical Hopf bifurcation (destabilization of the rest state by a collapse of the unstable limit cycle into the rest state). For high applied currents ($I_{App} > I_2$) no matter how the initial condition is chosen the axon exhibits periodic firing of

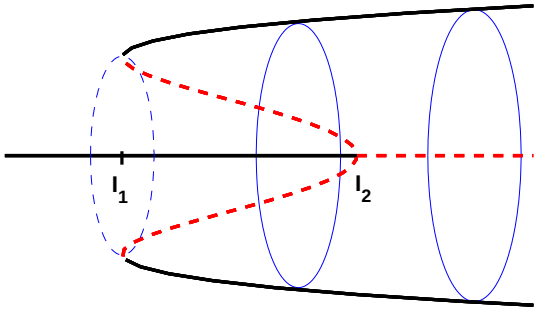

Figure 2.4: **Schematic figure of the bifurcation in HH model.** Stable fixed point if $I_{App} < I_1$, stable fixed point and two limit cycles (stable and unstable) born through saddle-node bifurcation at I_1 and destabilization of the fixed point via the Hopf bifurcation at I_2 ($I_{App} > I_2 > I_1$).

action potentials (AP), as on Fig. 2.3 (b). A schematic diagram of the saddle-node and subcritical Hopf bifurcation can be found on Fig. 2.4.

The subcritical Hopf bifurcation is also present in stochastically perturbed HH system but it is no longer possible to characterize it in the same way. The change of the system from a quiescent state to a periodic firing with a finite nonzero amplitude can be still observed but the transition is not localized in the parameter space anymore. One way to visualize this stochastic bifurcation is to plot the probability density function of the average trajectory location in the space. As the applied current grows the corresponding histogram in the V, m plane changes from a single peak, located in the neighborhood of a fixed point, to a smaller peak (around the fixed point) together with high concentration around the limit cycle, see Fig. 2.6. The probability mass, located in a neighborhood of the fixed point is redistributed to the region close to the limit cycle as I_{App} increases. Note that the height of the probability density function reflects also the trajectory speed, i.e. the trajectory always spends more time in the neighborhood of the fixed point and therefore there will be a peak there for all I_{App} values. The histogram is also plotted in different 2-dimensional subspaces of

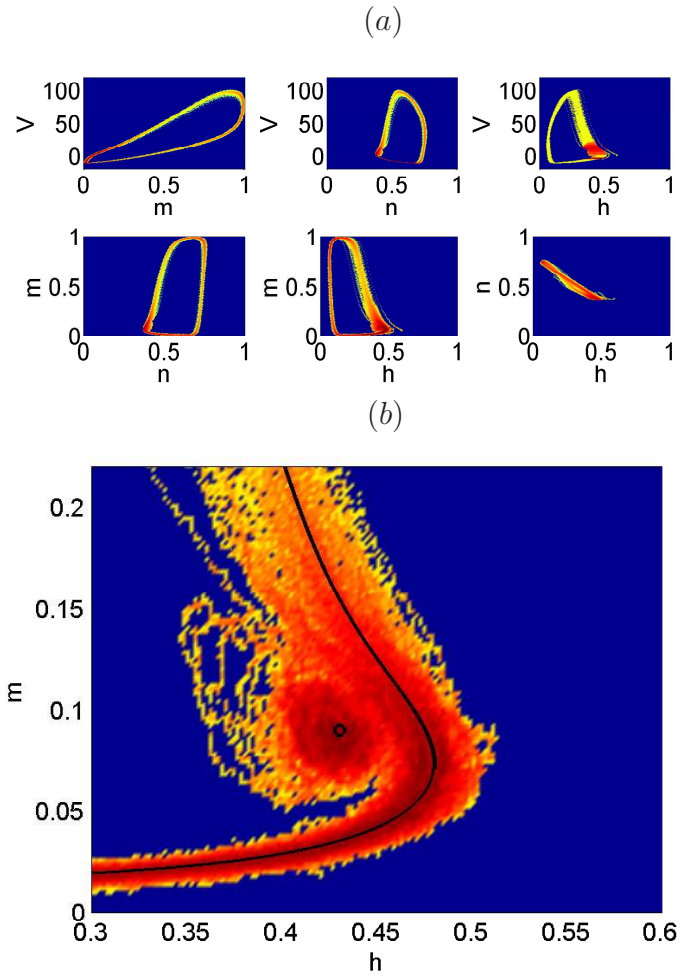

Figure 2.5: **Trajectory location of a Langevin type stochastic version of a HH model.** A set of multiple trajectories was computed numerically (using Langevin model) and 2-dimensional histogram of the trajectory position was plotted in different planes using a colorplot option of Matlab (initial condition is random). Values in the histogram are plotted on a logarithmic scale from the minimal value (blue) to the maximal value (dark-red). This experiment was performed for applied current where the deterministic HH is bistable (both quiescence and spiking can occur). (a) The figure in (n, h)-plane suggests approximately linear relationship between n and h. (b) The plot shows a detail of a near rest state behavior in the (m, h)-plane. Trajectories either oscillate around the limit cycle or they rotate in a close neighborhood of the fixed point. The noise is responsible for the switching between the two phases.

the 4-dimensional space (V, m, n, h) on Fig. 2.5 reflecting the effect of noise on these variables. The sodium gating variable m is known to be the fast variable whereas n

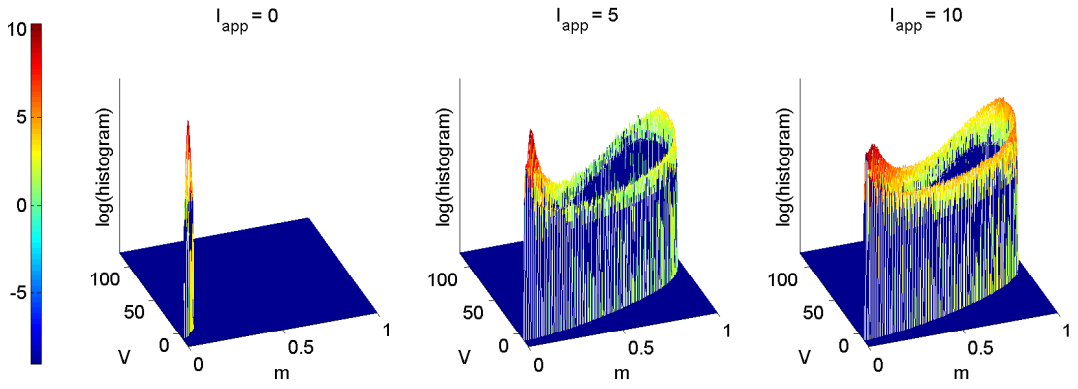

Figure 2.6: **Stochastic bifurcation in the Langevin model.** Probability density function of the trajectory location in the (V, m)-space changes from a single peak (concentrated around the rest state) to a peak plus a high concentration around the "deterministic" limit cycle. One can see that as I_{App} increases more mass is being redistributed from the rest state neighborhood to the limit cycle neighborhood. The fraction of mass around the fixed point thus decreases with I_{App} and the model of the neuron is more prone to spiking. The histogram values are on a logarithmic scale and therefore small differences in the density of trajectory location reflect rapid redistribution of probability.

and h are the slow variables *Peskin* (2000). Moreover, if the timescale of n and h are comparable and $h_\infty(v) + n_\infty(v) = c$ then there is approximately a linear relationship between $h(t)$ and $n(t)$. This relationship is observed on Fig. 2.5 where the (n, h) plane plot reflects the fact that the limit cycle lies approximately on a line with slope -1. Therefore most of the oscillatory dynamics should be hidden in the perpendicular space to (n, h), and thus (V, m)-plane should contain most of the information about the limit cycle properties.

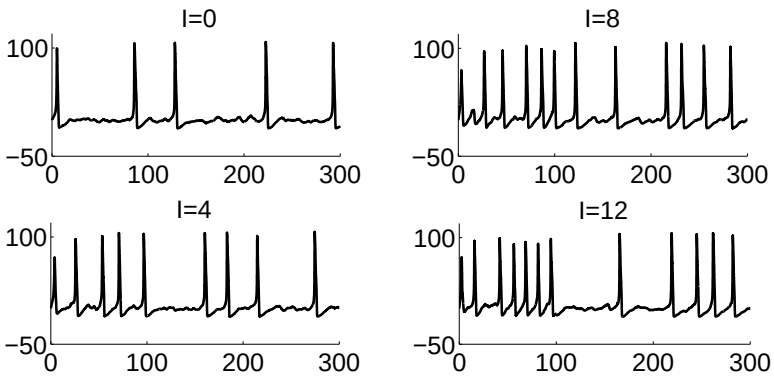

Figure 2.7: **Noise-induced and noise-supressed firing pattern in the stochastic HH model.** Trajectories $V(t)$ (in mV) against time (in ms) in a Markov model as the applied current varies ($N = 300$, $H = 300$, $M = 3H$). For small applied current $I_{App} = 0, 4$, the system oscillates, although in the deterministic HH only the rest state is stable. For large applied current $I_{App} = 8, 12$, spikes are occasionaly supressed.

2.3.2 Noise-Induced Oscillations and Quiescence

Stochastic numerical simulation of the HH model can exhibit spiking behavior also in the case when the corresponding deterministic model does not spike (*Rowat* (2007), *Schneidman et al.* (1998) for Markov chain models). Similarly, stochastic simulation can show temporary quiescence in case when the deterministic model fires regularly. On Fig. 2.7 we demonstrate that noise of relatively small magnitude can induce firing even in the case when no external current is applied. The firing in this case is rare and occurs in nonregular time intervals. On the other hand, even if the neuron is stimulated by a relatively large applied current, spikes can be temporarily suppressed in the spike train. Similar behavior is found in the LSHH model. There are certain parameter regions where the two behaviors (firing and quiescence) can be combined together in a state called coexistence. For moderate values of I_{App} (close to bistability of the HH model) one can find sequences containing various numbers of spikes alternating with small amplitude "oscillations" around the quiescent state when a trajectory enters a close neighborhood around the fixed point. The role of noise is to transition the trajectory from the neighborhood of the rest state to the

neighborhood of the limit cycle and conversely. These regions may or may not be attractive depending on whether the system is past the bifurcation. The region of bistability of the deterministic HH model $I_1 < I_{App} < I_2$ naturally exhibits coexistence in the stochastic Hodgkin-Huxley model (SHH). Both the fixed point and the limit cycle are stable in this case and therefore it is natural for a neuron with small noise to contain longer sequences of non-interrupted spikes and quiescence. However, even if the system is outside of the bistability region (low current, high current) coexistence in the stochastic model still occurs.

2.3.3 Distribution of Inter-Spike Intervals

Numerical results in the previous section suggest that the firing pattern of SHH "smoothly" depends on the external applied current unlike in the deterministic model where the transition from quiescent to spiking behavior is abrupt. The transition in the stochastic model is not localized in the applied current space; for low applied currents spiking is rare, whereas for high applied currents, the neuron spikes frequently with rare occurrence of quiescence. For small values of neuronal variability (low noise amplitudes or high channel density) the coexistence region gets narrower as the model transitions to the deterministic HH model. These spiking properties may be studied via inter-spike interval (ISI) that measures time between consecutive spikes.

We observe that randomness in the potassium dynamics has a major influence on the type of the ISI distribution. If the variation in potassium dynamics is small the distribution of ISI consists of multiple peaks, as shown on Fig. 2.8 for LSHH model and on Fig. 2.9 for MSHH model. The magnitude of these peaks seems to decay exponentially (linear decay on the logarithmic plot). This is a natural consequence of the alternation between two states: oscillations and quiescence. In particular, the ISI distribution reflects the fact that transitions between oscillations and quiescence occur at a small region in the phase space. Therefore the k-th peak of the distribution

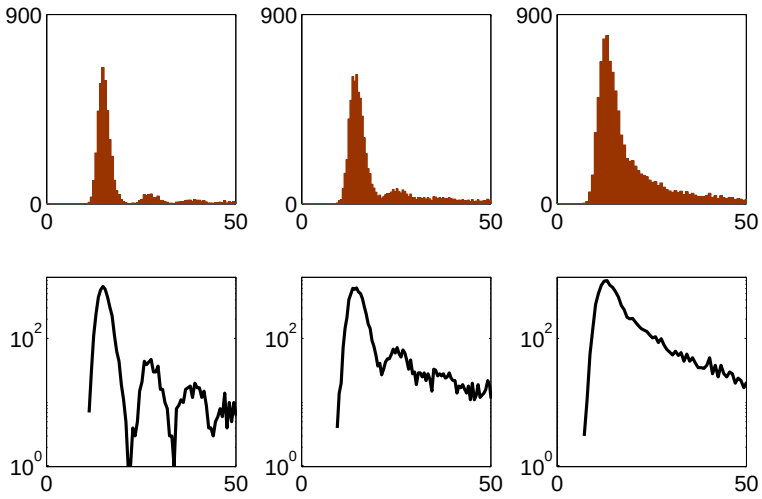

Figure 2.8: **Distribution of inter-spike intervals in the stochastic HH model.** Histograms of the ISI distribution for $I_{App} = 8$, $\sigma_H = 0.19$ and $\sigma_N = 0.05$ multimodal distribution (in the first column), $\sigma_N = 0.1$ (in the second column) and $\sigma_N = 0.2$ exponential tail distribution (in the third column).

corresponds to interval between two spikes if the trajectory enters the neighborhood of the quiescent state after the first spike and does not leave it for next $k-1$ small amplitude "oscillations". Naturally, the peaks decay as k grows since the probability of not exiting the quiescent region $k-1$ times in a row decays with k. One may study a very simple discrete model (reference) where the probability of spike equals p_s and the probability of quiescence equals $p_q = 1 - p_s$. The sequence of spikes and small amplitude "oscillations" can be then studied as a sequence of tosses of an unfair coin with constant (deterministic) time intervals between the tosses. The discrete distribution of ISI will then decay exponentially and will have peaks located at all natural numbers. By taking the time interval between tosses to be a random variable one may arrive to a similarly shaped distribution functions. Our observation shows that the continuous model behaves similarly to the coin model if the variation that controls the length of each spike/small amplitude "oscillation" is small. If the variation in the potassium dynamics is large then the variability in the duration of spikes/small amplitude "oscillations" is large as well. This causes individual modes in the distri-

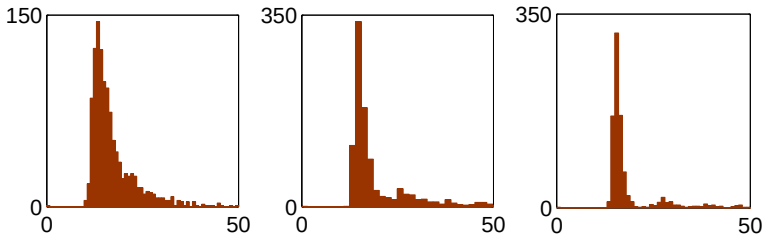

Figure 2.9: **Distribution of inter-spike intervals in the stochastic Langevin HH model.** Histograms of the ISI distribution (Markov model) for $I_{App} = 8$, $H = 2000$ and $N = 200$ exponential tail distribution (first column), $N = 2000$ (second column) and $N = 8000$ multimodal distribution (third column).

bution of the ISI to smear and the tail of the distribution decays exponentially. Both in the Langevin model and in the Markov model there are two distribution types observed: multi-peaked and single-peaked with exponential tail and the distribution shape is controlled by the potassium dynamics.

2.3.4 Spiking Frequency, Monotonicity and Sensitivity to noise

Our experiments show that occurrence and frequency of spikes in the subthreshold regime ($I_{App} < I_1$) may be enhanced by noise whereas in the superthreshold regime ($I_{App} > I_2$) spiking frequency tends to be decreased by noise. The most natural measure that reflects these observed properties is the average long-time spiking frequency. This quantity is deterministic and depends on the value of the applied current and channel densities (noise magnitudes). Numerical simulations for both LSHH and MSHH are performed for a sufficiently long time (we simulate 20s for MSHH and 200s for LSHH). The long term average spiking frequency then equals the fraction of total number of spikes in the given time interval to the total simulation time (in seconds).

Numerical results depicted on Fig. 2.10 (with relative or with fixed color scale) indicate some unexpected properties of the firing rate dependance on applied current and level of variability. We will sketch these properties with references to further

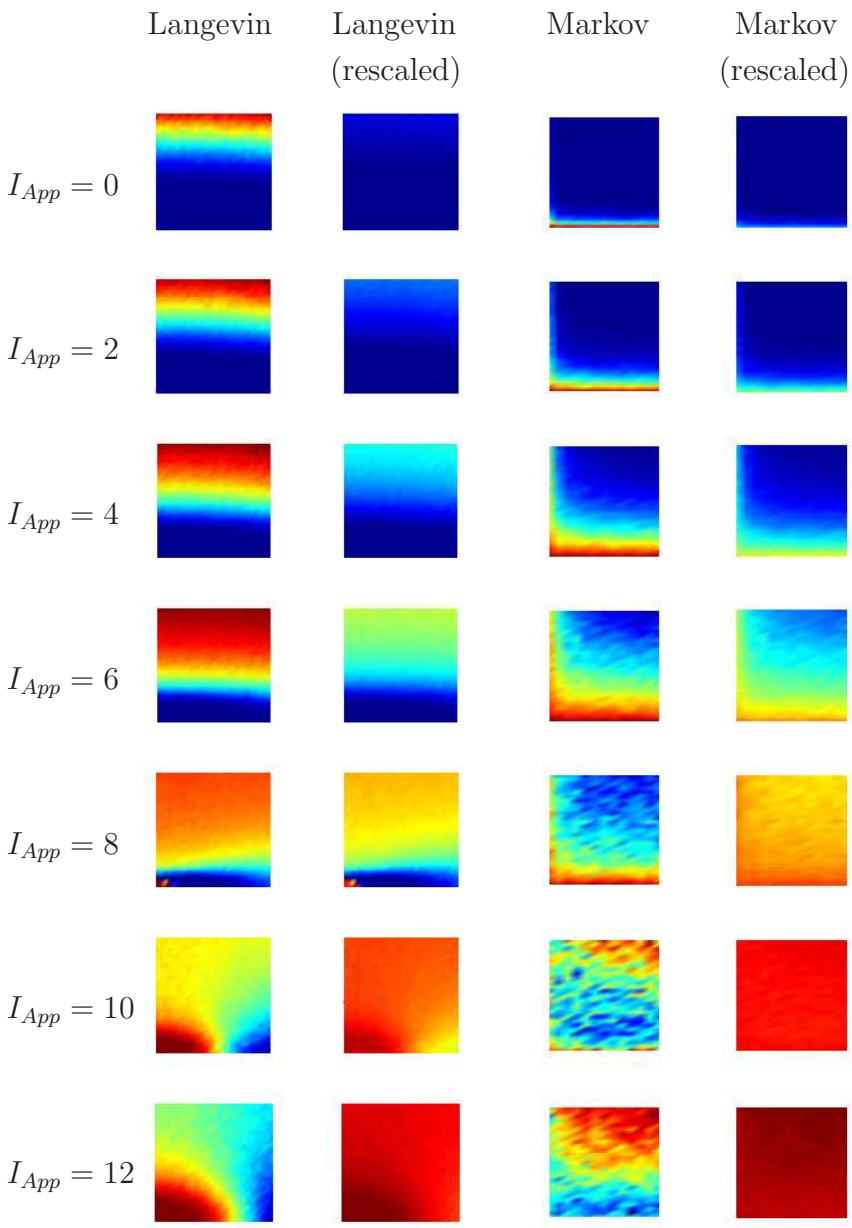

Figure 2.10: **Firing frequency dependence on applied current and noise.** Sequence of colorplots shows spiking frequency as a function of noise magnitudes σ_H on the horizontal axis and σ_M on the vertical axis (or channel densities M and H respectively). The color ranges from low (blue) to high (red). Plots is organized as follows (by column). First column: Langevin model, colorplots with scale adapted for each colorplot separately; second column: Langevin model with uniform scale $[0, 80]$ spikes/s; third column: Markov model, colorplots with scale adapted for each colorplot separately; last column: Markov model with uniform scale $[0, 80]$ spikes/s. Parameters from the range: $\sigma_M, \sigma_H \in [0, 0.25]$, $M \in [200, 8000]$ and $H \in [200, 3800]$. Particular values of the spiking frequency are further illustrated on Fig. 2.12 and Fig. 2.13.

numerical simulations.

1. Bifurcations in the stochastic model: The deterministic saddle-node bifurcation of limit cycles is depicted on Fig. 2.10 for $I_{App} = 8$, where the destabilization of the rest state forces HH to periodicaly fire action potentials with a high frequency (lower left corner of each LSHH figure corresponds to the deterministic limit). This spiking frequency is higher compared to the spiking frequency for SHH with any nonzero noise. The reason is that noise may suppress firing (and lead to quiescence) but it cannot enhance the firing above some level. Other way to understand this is by looking back at Fig. 2.5. It suggests that even though the trajectory is sensitive to noise everywhere in the 4 dimensional space, it can jump from the neighborhood of the fixed point to a neighborhood of the limit cycle and conversely only in the tiny region of the space (see the detailed plot) and the duration of the tour around the limit cycle is not very sensitive to noise. To summarize, noise may both enhance ($I_{App} < I_1$) and suppress ($I_{App} > I_1$) oscillations in different regimes. Also, the destabilization of the rest state at $I_{App} = I_2$ in the deterministic model introduces a high firing rate behavior that spreads away from the deterministic scenario towards the higher noise levels both in the Na and K channel dynamics as the applied current increases.

2. Potassium noise for a fixed subthreshold current: Increasing the variability of the potassium dynamics (an increase of σ_N or decrease of N) results in a higher firing rate in the subthreshold applied current scenario ($I_{App} < I_1$), see Fig. 2.10 and Fig. 2.11. This monotone pattern is observed up to the point of the saddle-node bifurcation of limit cycles, where the stable limit cycle of the deterministic system is born, see the bifurcation diagram Fig. 2.4. This potassium noise is capable of inducing action potentials by forcing the trajectory to enter some close neighborhood

of the limit cycle (that may be both stable or unstable).

3. Potassium noise for a fixed superthreshold current: Noise in K channels has an opposite role in the superthreshold and subthreshold regime. In the superthreshold regime the spiking frequency tends to decrease rather than increase as the potassium channel variability increases in the regime where the sodium variability is small, see Fig. 2.10 and Fig. 2.11. This is due to the fact that firing rate of the deterministic HH model past the Hopf bifurcation is higher than the noise-induced firing rate. Naturally the SHH approaches HH as noise magnitudes go to zero (lower left corner on Langevin colorplots and upper right corner on Markov colorplots). But since the firing rate depends smoothly on the noise level, as the potassium variability increases the firing rate must decrease from the deterministic firing rate to a noise-induced firing rate that is smaller.

4. Sodium noise for a fixed subthreshold current: Unlike in the case of potassium variability, the firing rate seems to be roughly independent of the sodium variability for subthreshold currents. This is reflected in a horizontal pattern in the colorplots Fig. 2.10 for the subthreshold current regime and in the low variation between blue curves on Fig. 2.11.

5. Sodium noise for a fixed superthreshold current: Past the supercritical Hopf bifurcation the independence of the firing rate on the sodium variability can be no more observed. The sodium channel variability results in a decrease of the firing rate. This effect is less dramatic than the effect of potassium noise.

6. Monotonicity due to applied current: As observed above, the firing rate properties strongly depend on the value of the applied current. Enhancement of

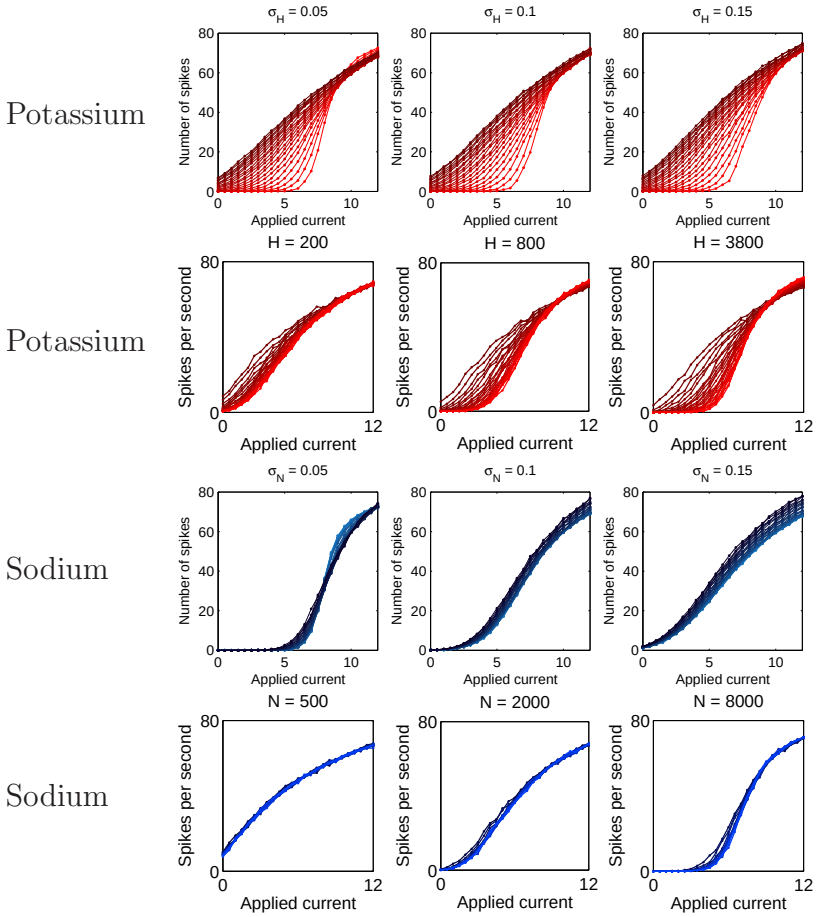

Figure 2.11: **Sensitivity of the firing rate on different levels of noise, rescaled.** Langevin model simulation in the first and third row, Markov model simulation in the second and last row. Each plot shows functional dependance of the spiking frequency on the applied current when either the noise (channel density) of sodium channels is fixed and functions for all potassium noise levels (channel density) are plotted or conversely. In all figures darker color corresponds to larger noise magnitude. Parameters are chosen for Langevin model to be $\sigma_N \in \{0, 0.01, \ldots, 0.25\}$ and $\sigma_H \in \{0, 0.001, 0.01, 0.02, \ldots, 0.25\}$. Channel densities in the Markov model are $N \in \{800, 1100, 1400, \ldots, 8000\}$ and $H \in \{800, 1100, 1400, \ldots, 3800\}$. The first two rows show that variability in the potassium dynamics influences the firing frequency considerably (red curves on each figure do not overlap). However, variability in the sodium dynamics does not produce almost any variation (blue curves in last two rows are almost identical).

firing in the case of subthreshold applied current transitions to suppression of firing in the case of the superthreshold current. A rather unexpected property we find is that firing rate depends monotonically on the applied current. No matter what noise parameters we choose the average long term firing rate will never decrease as the applied current increases. One way to visualize this is to plot spiking frequency

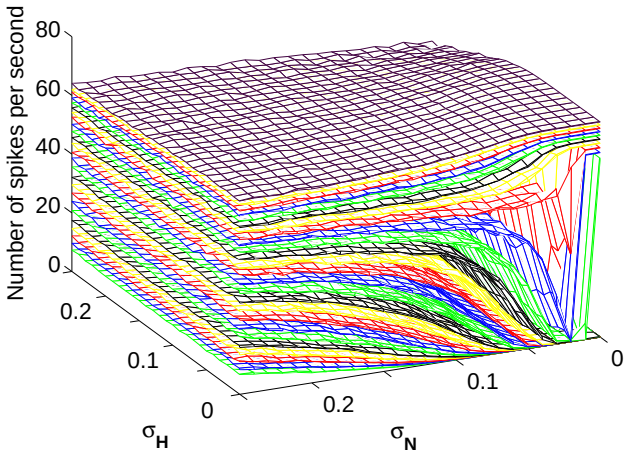

Figure 2.12: **Monotonicity of the firing frequency with respect to applied current (Langevin model).** Each value of applied current yields a spiking frequency surface (function of σ_N and σ_H). These surfaces do not overlap and they grow as the applied current grows. Values from bottom to top $I_{App} = 0, 0.5, 1, \ldots, 12$.

surfaces in the noise parameter plane each of which correspond to a different value of applied current. We observe that these surfaces do not intersect but rather form layers that divide the space, see Fig. 2.12 and Fig. 2.14. We hypothesize that if one computes these surfaces for all reasonable values of applied current and noise magnitudes ($0 \leq I_{App} \leq I_{max}$) infinitely accurately (infinite time of integration) they will completely fill a portion of the three dimensional rectangular box, i.e., there is an equivalence relation parametrized by applied current where each surface forms an equivalence class.

Monotonicity and smoothness of the average firing rate with respect to applied current yields to the following. If $F(\sigma_N, \sigma_H)$ is the average firing rate at noise level

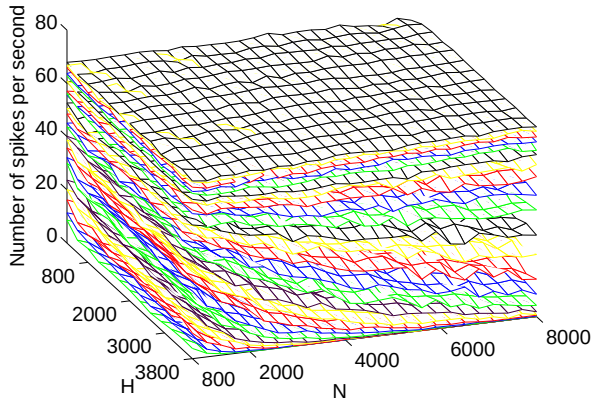

Figure 2.13: **Monotonicity of the firing frequency with respect to applied current (Markov model).** Each value of applied current yields a spiking frequency surface (function of N and H). These surfaces do not overlap (some overlap may be caused by large fluctuations) and they grow as the applied current grows. Values from bottom to top $I_{App} = 0, 0.5, 1, \ldots, 12$. Channel densities: $N \in \{800, 1100, 1400, \ldots, 8000\}$, $H \in \{800, 1100, 1400, \ldots, 3800\}$.

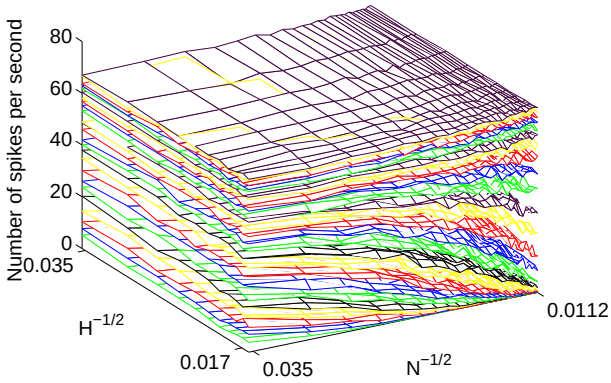

Figure 2.14: **Monotonicity of the firing frequency with respect to applied current (Markov model) in scaled coordinates.** Each value of applied current yields a spiking frequency surface (function of N and H). We plot the spiking frequency as a function of $1/\sqrt{N}$ and $1/\sqrt{H}$ that corresponds to the noise magnitudes σ_N and σ_H. These surfaces do not overlap (some overlap may be caused by large fluctuations) and they grow as the applied current grows. Moreover, the overall dependance in the Markov model resembles the Langevin model case. Values from bottom to top $I_{App} = 0, 0.5, 1, \ldots, 12$. Channel densities: $N \in \{800, 1100, 1400, \ldots, 8000\}$, $H \in \{800, 1100, 1400, \ldots, 3800\}$.

(σ_N, σ_H), then

$$F(\sigma_N, \sigma_H) = g(I_{App}) \tag{2.15}$$

where g is nondecreasing. On the one hand this implies that if one knows the channel densities and applied current, average firing rate can be read off from Fig. 2.12 or Fig. 2.13. On the other hand, given the average firing rate and applied current (observables in experiments), this yields a functional relation between channel density in sodium and potassium channels.

2.4 Conclusions

Perhaps the most important finding is that both stochastic numerical approaches (MSHH, LSHH) lead to qualitatively the same results. This implies that Markov model can be replaced by the computationally less expensive (particularly if channel densities are small) Langevin model. Regardless of the number of channels, we find that the firing rate always increases as the applied current increases (Fig. 2.12, Fig. 2.13 and Fig. 2.14). This suggests that some properties of the firing mechanism remain independent of the channel numbers. For low applied currents, we find that stochasticity induces neuronal firing (Fig. 2.11). The firing rate in the presence of low amplitude currents increases as the channel density decreases (more variability). Past a threshold applied current, the deterministic Hodgkin-Huxley equations show repetitive firing. This firing behavior is also captured in our stochastic simulations. However, unlike lower amplitude currents, lower channel numbers can cause slower rather than faster spiking in this case (Fig. 2.10).

We also find that noise from sodium channels has a smaller effect on the firing behavior of the neuron than noise from potassium channels. As the potassium channel density decreases, we find that the shape of the ISI distribution changes from a multimodal to an exponential-tailed function with each peak roughly symmetrical (Fig. 2.9,

Fig. 2.8). On the other hand as the sodium channel density changes both the shape of the ISI distribution and the spiking frequency remain almost unchanged. As the applied current decreases, the tail of the distribution gets heavier and spikes typically occur only after a long waiting time.

Our numerical findings suggest that even if channel densities are not known, they can be determined from the neuron's response to an applied current by observing the spiking activity. This may be done by comparing the experimental long term average firing frequency for different choices of applied current with the numerical data. This together with the distribution of inter-spike interval can be matched to Na and K channel densities.

Our studies show that noise can play significant and counterintuitive roles in determining the firing behavior of a neuron and lead to testable predictions of the real channel density based on the spiking frequency and shape of the inter-spike interval distribution.

We believe that the observed properties (suppression and enhancement of spikes, distribution shape, monotonicity with respect to applied current) are present in many other bistable systems that undergo a subcritical Hopf bifurcation and are subject to noise. A way to verify presence of these properties mathematically is to analyze a simpler canonical dynamical model and study the effect of random perturbation via analytical tools. The key feature that the model should possess is the nonuniform proximity of the fixed point to different parts of the limit cycle. The reason for this is that the noise-induced transitions occur at the place where this distance is small with large probability.

CHAPTER III

Mathematical models and measures of mixing

3.1 Introduction

Simple characterization of transport and mixing properties of complex flows, as on Fig. 3.1, is a great challenge. If important features could be quantified it would provide a tremendous insight in many applications. Mixing processes naturally arise in atmospheric and oceanic sciences and engineering applications (*Csanady*, 1973; *Dagan*, 1987) where the underlying flows may be simple laminar flows but are often as complicated as a fully developed turbulence. There has been continued interest in the qualitative and quantitative understanding of mixing processes, see *Fannjiang and Papanicolaou* (1994), *Taylor* (1922), *Taylor* (1953), *Richardson* (1926), and recently also in quantifying an ideal stirrer that mixes the fluid "the best".

One way to characterize the mixing properties of a fluid flow is to consider a motion of a passive substance (scalar) that is carried by the flow without effecting it. This advected substance may be thought of as temperature, salt, dye or other chemical composition markers. Even though this concept is purely theoretical, in many cases (typically if the concentration of the "passive tracer" is small) it is realistic and accurate.

Mixing of the passive scalar in any system occurs on several scales, the smallest being the molecular scale. The molecular diffusion (or thermal diffusion in case where

Figure 3.1: **River mixing.** Passau is a town in Lower Bavaria, Germany, located at the confluence of the Danube (North) and Inn (South) rivers on the German-Austrian border. The two rivers have water of very different color and at the point they meet they begin to mix together. Figure from www.maps.google.com

the scalar is a temperature) acts as an effective small scale stirring mechanism that may be enhanced by a large-scale stirrer due to the advection. The mixing effect of the stirring may eventually dominate "weak" molecular mixing.

There are two basic setups for the problem. Given a velocity field $\vec{u}(\vec{x},t)$ in two spatial dimensions satisfying a divergence-free condition $\nabla \cdot \vec{u} = 0$, consider a system of stochastic differential equations for the position of a particle $X(t)$ moving with the flow field

$$d\vec{X}(t) = \vec{u}(\vec{X},t)dt + \sqrt{2\kappa}\, d\vec{W}(t)\,, \quad \vec{X}(0) = \vec{x}_0\,. \tag{3.1}$$

with $\vec{W}(t)$ being the Wiener process. Capital letters are used for the particle position $\vec{X}(t)$ to highlight its stochastic nature. Or, consider the homogenous advection-

diffusion equation for the passive scalar concentration (or temperature)

$$\partial_t T + \vec{u} \cdot \nabla T = \kappa \Delta T \,, \quad T(\vec{x}, 0) = \delta(\vec{x} - \vec{x_0}), \tag{3.2}$$

where κ is a molecular (or thermal) diffusion coefficient. $\vec{X}(t)$ is the passive tracer particle position and $T(\vec{x}, t)$ is its probability distribution, i.e., the scalar concentration. The two setups are equivalent in the sense that the advection-diffusion equation (3.2) is a Fokker-Planck PDE for the stochastic process given by (4.2).

In many relevant applications the system is supplied by sources and sinks that act as a non-homogeneous additive term of the form $s(\vec{x}, t)$ in the equation (3.2). The inclusion of a source-sink distribution leads to additional complexity of the mixing problem and to potential inconsistencies in the existing theories. For completeness we will assume periodic boundary conditions for $T(\vec{x}, t)$.

In applications for "real" turbulent mixing, the velocity $u(\vec{x}, t)$ of the advective fluid in (3.2) should be determined as a solution of Navier-Stokes equations

$$\partial_t \vec{u} + \vec{u} \cdot \nabla \vec{u} = -\nabla p + \kappa \Delta \vec{u} + \vec{F} \,, \tag{3.3}$$

$$\nabla \cdot \vec{u} = 0 \,, \tag{3.4}$$

with appropriate boundary conditions. In this setting $p(\vec{x}, t)$ is the pressure field and $F(\vec{x}, t)$ is an appropriate external force. In many applications like convection, $\vec{F}$ depends on T. But here we study scenarios where passive tracer particles do not influence the flow and therefore the velocity of the advecting fluid does not depend on the concentration of the tracer.

3.1.1 Mixing Measures and Modelling Approaches

The fundamental question is how to characterize mixing and what aspects of mixing should be encoded in this choice. In reality, because of wide range of applications,

very different approaches and measures have been used to quantify the quality of mixing, sometimes giving answers that are inconsistent with each other. It is therefore essential to understand the differences and limitations of these theories and measures and to find a unifying mathematical model that may capture the different features.

Among some of the most mainstream approaches one may study the effectivity of mixing via

- single particle or particle pair dispersion in the absence of sources and sinks,
- enhancement of the flux due to the flow when a constant scalar gradient is imposed,
- the original advection-diffusion equation with arbitrary sources and sinks.

The particle dispersion approach and the flux-gradient approach are traditionally used to specify properties of mixing when there is a large scale separation between the flow and the scalar concentration, which sometimes occurs in long time limits, i.e., in homogenization theory (HT), see *Majda and MacLaughlin* (1993), *Majda and Kramer* (1999). Also, different measures have been established to characterize the effectivity of stirring (the choice of mixing measure is often but not always predetermined by the choice of method used), among them

- enhanced tracer particle dispersion,
- enhanced flux for a given gradient,
- enhanced scalar concentration variance suppression in the presence of sources and sinks.

The origins of the particle dispersion approach for turbulent flows reach back to *Taylor* (1922) and *Richardson* (1926). The most recent generalization to a multi-particle dispersion is due to (*Toschi and Bodenschatz*, 2009). More references may be found in the review article (*Faber and Vassilicos*, 2009).

In the mathematically natural approach equation (4.2) serves as a tool for calculating characteristics of the flow that may be then used to simplify the problem (3.2) in the absence of sources. In the simplified problem advection and diffusion terms are replaced by an "effective" diffusion matrix K_{ij}^{eff} satisfying

$$K_{ij}^{\text{eff}} \sim \frac{1}{2t}\mathbb{E}\left[(X_i(t) - X_i(0))(X_j(t) - X_j(0))\right] \tag{3.5}$$

as $t \to \infty$, where $\vec{X}(t)$ is a position of the passive particle at time t and $\mathbb{E}[\cdot]$ is an expectation taken over all possible realizations of $W(t)$. The approximation $\vec{u} \cdot \nabla - \kappa\Delta \to -\partial_i K_{ij}^{\text{eff}} \partial_j$ it then used to replace the advection-diffusion equation with the effective diffusion equation

$$\partial_t T = \partial_i K_{ij}^{\text{eff}} \partial_j T \,. \tag{3.6}$$

To quantify the effectivity of stirring the enhancement matrix $K_{ij}^{\text{eff}}/\kappa$ is computed. Its dependance on the flow structure in the large flow scale and long time approximation is a central theme of the homogenization theory. Note that the approach requires absence of sources and sinks in the problem. Although this technique seems to be transferable to a source-sink problem where (3.2) simplifies to

$$\partial_t T = \partial_i K_{ij}^{\text{eff}} \partial_j T + s \,. \tag{3.7}$$

as we will discuss later, this may not give the correct answer to the problem. The difficulty is that the recently entered material in the transient system does not in fact diffuse at a rate given by long time effective diffusivity K_{ij}^{eff} but is rather advected by a short time velocity field (the fresh material has not spent enough time in the system yet to see the long time diffusivity).

A variation of the homogenization approach used extensively in turbulence theory,

considers the advection-diffusion problem (3.2) with an imposed background gradient,

$$T = -Gx + \theta(\vec{x}, t)\,. \tag{3.8}$$

The perturbation concentration θ satisfies a similar advection-diffusion equation to (3.2) but with a source-like term:

$$\partial_t \theta + \vec{u} \cdot \nabla\theta = \kappa\Delta\theta + G(\hat{i} \cdot \vec{u})\,, \tag{3.9}$$

where $\hat{i}$ is a unit vector in the x-direction. The imposed gradient scale as well as the separation between the flow and scalar field scales are assumed to be large and boundary conditions are periodic. This technique, see review articles *Gollub et al.* (1991), *Warhaft* (2000), *Falkovich and Sreenivasan* (2006), has been extensively used to specify properties of the single particle distribution tail in a single-scale random velocity field in *Holzer and Siggia* (1995), *Shraiman and Siggia* (1994) and *Bourlioux and Majda* (2002). In many cases the flow enhances the molecular diffusion in the direction of the gradient forcing. The enhancement factor $K_{11}^{\text{eff}}/\kappa$ is a scalar that depends on the concentration field through a flux-gradient relationship

$$\frac{K_{11}^{\text{eff}}}{\kappa} = 1 + \frac{\langle ||\vec{\nabla}\theta||^2 \rangle}{G^2} \tag{3.10}$$

where $\langle\cdot\rangle$ is a time-space average over the periodic cell.

To avoid problems with transience of the effective diffusion matrix and at the same time to fully describe the mixing properties of the flow in the presence of sources and sinks the advection-diffusion equation

$$\partial_t \theta + \vec{u} \cdot \nabla\theta = \kappa\Delta\theta + s\,, \quad \theta(\vec{x}, 0) = \delta(\vec{x} - \vec{x_0}), \tag{3.11}$$

may be studied directly. In the recent advances the quality of stirring has been represented by the amount of suppressed variance of the scalar concentration field due to stirring, see *Rehab et al.* (2000), *Constantin et al.* (2008), *Doering and Thiffeault* (2006), *Plasting and Young* (2006), *Shaw et al.* (2007), *Thiffeault et al.* (2004) and in the past by *Danckwerts* (1952), *Edwards et al.* (1985). From the application's point of view this seems to be the most natural measure of mixing because the effect of stirring is to move passive tracer from places with high to low concentration and thus to smooth out the concentration field.

It is known that some flows (including turbulent flows) are mixing fluid differently on different length scales. The variance reduction mixing measure may be defined in terms of variance on different scales (*Thiffeault et al.*, 2004; *Schumacher et al.*, 2003; *Mathew et al.*, 2005; *Doering and Thiffeault*, 2006) by

$$\mathcal{E}_p = \frac{\kappa_p^{\text{eff}}}{\kappa} = \frac{< ||\nabla^p \theta_0||^2 >^{1/2}}{< ||\nabla^p \theta||^2 >^{1/2}} \tag{3.12}$$

where $\theta_0 = -\frac{1}{\kappa}\Delta^{-1}s$ solves the PDE problem in the absence of stirring. The equivalent diffusivity κ^{eff}, first introduced by *Thiffeault et al.* (2004) satisfies

$$\kappa_p^{\text{eff}} = \kappa \mathcal{E}_p = \sqrt{\frac{\langle (\Delta^p \Delta^{-1} s)^2 \rangle}{\langle \Delta^p \theta^2 \rangle}}\,, \tag{3.13}$$

and parametrizes the flow-enhanced mixing with a simple diffusion constant that achieves the same level of variance suppression.

Mixing efficiency $\mathcal{E}_0 = \kappa_0^{\text{eff}}/\kappa$ measures enhancement of mixing on intermediate scales in terms of a scalar variance suppression whereas $\mathcal{E}_1$ and $\mathcal{E}_{-1}$ measure variance suppression enhancement at small and large scales respectively. Without the loss of generality the source-sink distribution and the initial concentration are assumed to be spatially mean zero. The enhancement factor $\kappa^{\text{eff}}/\kappa$ naturally depends not only on how we stir the fluid but also on the non-homogenous "forcing" due to sources and

sinks.

3.1.2 Péclet Number Dependance

There has been a lot of effort spent to analytically quantify the properties of the mixing efficiency (enhancement factor) in the regime of strong flow for different approaches and using different measures. The Péclet number is a nondimensional quantity of the flow that measures the relative flow strength compared to the strength of molecular diffusion. We define the Péclet number by

$$Pe = \frac{U \cdot L_u}{\kappa} \tag{3.14}$$

where U is a velocity scale of the flow and L_u is a length scale of the flow. The two forementioned nondimensional mixing measures (enhancement matrix/number and variance suppression) then capture flow properties and may be considered as a function of also non-dimensional Péclet number.

$$\frac{\kappa^{\text{eff}}}{\kappa} = f(Pe) \tag{3.15}$$

It has been shown using a homogenization theory in the case of tracer particle dispersion measure or flux-gradient mixing measure without sources (*Majda and MacLaughlin*, 1993; *Fannjiang and Papanicolaou*, 1994) that any steady, spatially periodic flow satisfies an upper bound on the enhancement factor (its components) in the form

$$\frac{\kappa^{\text{eff}}}{\kappa} \leq 1 + Pe^2 \,. \tag{3.16}$$

As will be shown later, this bound can be saturated for a monochromatic sine flow in 2D (*Majda and Kramer*, 1999) This bound may be derived from the space-time

averaged advection-diffusion equation (3.9) multiplied by θ, using integration by parts

$$0 = \langle \theta(\kappa \Delta\theta + G(\hat{i} \cdot \vec{u})) \rangle = -\kappa \langle |\nabla\theta|^2 \rangle + G \langle \theta(\hat{i} \cdot \vec{u}) \rangle \tag{3.17}$$

with the use of Cauchy-Schwartz inequality in

$$\langle \theta(\hat{i} \cdot \vec{u}) \rangle \leq \langle |\nabla\theta|^2 \rangle^{\frac{1}{2}} \langle |\nabla^{-1}\vec{u}|^2 \rangle^{\frac{1}{2}} = U L_u \langle |\nabla\theta|^2 \rangle^{\frac{1}{2}} \tag{3.18}$$

where U is the rms velocity of the flow and L_u is a characteristic lengthscale of the flow

$$U^2 := \langle |\vec{u}|^2 \rangle, \qquad L_u^2 := \frac{\langle |\nabla^{-1}\vec{u}|^2 \rangle}{\langle |\vec{u}|^2 \rangle} \,. \tag{3.19}$$

and ∇^{-1} corresponds to multiplication by $||\vec{k}||^{-2}\vec{k}$ in Fourier space. On the other hand, it has been showed by *Thiffeault et al.* (2004) that for any flow in the presence of steady sources, that using the suppression variance mixing measure, the enhancement factor cannot grow faster than Pe'

$$\frac{\kappa^{\text{eff}}}{\kappa} \leq 1 + Pe' \tag{3.20}$$

where the source-depended Péclet number is defined by

$$Pe' = \frac{U \cdot L_s}{\kappa} \,. \tag{3.21}$$

The two results clearly show a conflict between different mixing measures. This conflict is partially caused by the inconsistency in the Péclet number definition. A partial resolution to the conflict can be seen in the following

$$\frac{\kappa^{\text{eff}}}{\kappa} \leq 1 + Pe' = 1 + \frac{U l_s}{\kappa} = 1 + \frac{l_s}{l_u} \cdot \frac{U l_u}{\kappa} = 1 + r \cdot Pe \tag{3.22}$$

where $r = L_s/L_u$ controls the separation between the source and flow scale. If the scale separation r is large ($L_u << L_s$) the enhancement coefficient may be as large as Pe^2. This suggests that the homogenization theory may be valid but only in the regime $r \geq \mathcal{O}(Pe)$.

A main challenge of this work is to form and analyze a single model that reconciles the particle dispersion modelling techniques with the multiscale mixing measures.

3.2 Dispersion-diffusion theory (DDT)

As we have noted above, the approximation of advection and diffusion by an effective diffusion matrix in (3.6) cannot be generally combined with the inclusion of the source-sink distribution due to the temporal inconsistency of the resulting problem. However, the nature of the passive scalar, advected by the fluid suggests that molecular equations (4.2) describe particle motion quite accurately. The main challenge is therefore to accommodate both the particle dispersion and the source-sink distribution in one theory in a temporarily consistent way. In order to do this the effective diffusivity matrix must contain temporal information. Following the same idea as sketched in review article by *Salazar and Collins* (2009) we define a time-dependent effective diffusivity as

$$K_{ij}(\vec{X}(t_0), t, t_0) := \frac{1}{2}\frac{d}{dt}\mathbb{E}\left[(X_i(t) - X_i(t_0))(X_j(t) - X_j(t_0))\right] . \tag{3.23}$$

Effective diffusivity matrix contains information about the initial position of a particle and the initial time when it was introduced to the system. Note that for steady or statistically stationary flow $\mathbf{K}(t_0, t, \vec{X}(t_0))$ does not depend on time t itself but rather on the time difference $t - t_0$. Because every particle possesses its own time benchmark t_0 we cannot treat the particles generated by sources at different times collectively.

For each particle the time-dependent effective diffusivity (3.23) may be used to

approximate the particle's position probability density $\rho(\vec{x}, t; \vec{x}_0, t_0)$ via

$$\frac{\partial}{\partial t}\rho(\vec{x}, t; \vec{x}_0, t_0) = \frac{\partial}{\partial x_i}K_{ij}(\vec{x}_0, t, t_0)\frac{\partial}{\partial x_j}\rho(\vec{x}, t; \vec{x}_0, t_0)\,, \tag{3.24}$$

where $\vec{x}(t_0) = \vec{x}_0$ and initially

$$\rho(\vec{x}, t; \vec{x}_0, t_0) = \delta(\vec{x} - \vec{x}_0). \tag{3.25}$$

The fundamental solution, i.e., the Green function of this problem on the periodic domain with period L in d spatial dimensions ($d = 2$ in our case) is

$$\rho(\vec{x}, t; \vec{x}_0, t_0) = \frac{1}{L^d}\sum_{\vec{k}\in\mathbb{Z}^d} e^{i\vec{k}\cdot(\vec{x}-\vec{x}_0)-\frac{1}{2}\vec{k}^T\mathbf{C}(\vec{x}_0,t,t_0)\vec{k}} \tag{3.26}$$

where $\mathbf{C}(\vec{x}_0, t, t_0)$ is a correlation matrix defined by

$$C_{ij}(\vec{x}_0, t, t_0) := \mathbb{E}\left[(X_i(t) - X_i(t_0))(X_j(t) - X_j(t_0))\right]\,. \tag{3.27}$$

In the nonhomogenous problem with sources and sinks we seek a solution of (3.11) using a superposition principle to account for the different time and position where each particle was introduced. After a long time, we approximate the solution of the d-dimensional advection-diffusion equation by

$$\theta_{\mathrm{DDT}}(\vec{x}, t) = \int_{-\infty}^{t} dt_0 \int_{[0,L]^d} dx_0 \rho(\vec{x}, t; \vec{x}_0, t_0)s(\vec{x}_0, t_0)\,. \tag{3.28}$$

This is called the dispersion-diffusion approximation. Note that θ_{DDT} itself does not satisfy an advection-diffusion equation even in the long-time limit because of the underlying spatial dependence of the effective diffusivity matrix. It is worth to mention that the dispersion-diffusion approximation captures the spatial and tempo-

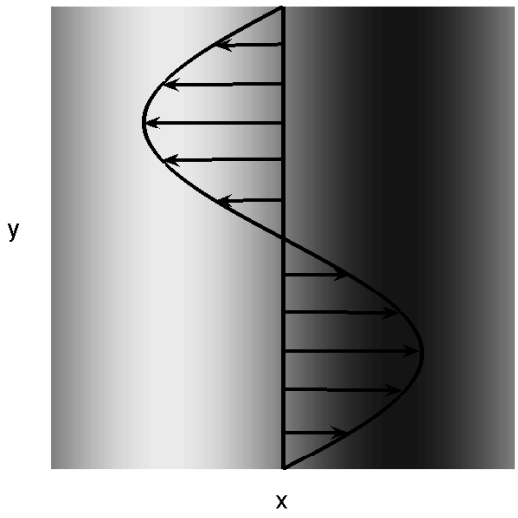

Figure 3.2: **Sine shear flow with single-scale distribution of sources and sinks.** The flow, depicted by black lines, is parallel to the horizontal axis, distribution of sources and sinks depend only on x.

ral information in the first two moments of the underlying stochastic tracer particle process exactly. This makes this approximation "best" among all particle dispersion-type methods based on matching the first two moments. In order to improve this method, one would probably need to do something like matching higher moments of the distribution.

3.2.1 Two scale problem: sine shear flow with sine sources

We will consider a two-scale problem (3.11) in $\mathbb{R}^2$ with a single-scale, uni-directional source

$$s(\vec{x}, t) = \sqrt{2} S \sin(k_s x) \tag{3.29}$$

with the wavenumber $k_s = 2\pi/L_s$ and a single-scale stirring flow

$$u(\vec{x}, t) = \hat{i}\sqrt{2} U \sin(k_u y) \tag{3.30}$$

as depicted on Fig. 3.2. The non-dimensional control parameters are

$$Pe = \frac{U}{k_u \kappa} \qquad \text{and} \qquad r = \frac{k_u}{k_s}\,. \tag{3.31}$$

This problem was studied by *Shaw et al.* (2007) in a regime of finite r and large Pe using asymptotic analysis of the internal layer theory (ILT). The nontrivial asymptotic scaling laws for the multiscale enhancement factor were found to be

$$\mathcal{E}_p \sim \begin{cases} r^{7/6} Pe^{5/6} & \text{if} \quad p = 0 \\ r^{1/2} Pe^{1/2} & \text{if} \quad p = 1 \\ r\, Pe / \log Pe & \text{if} \quad p = -1 \end{cases} \tag{3.32}$$

Note that this result contradicts the homogenization theory prediction (of Pe^2 scaling at all scales) and therefore HT and ILT approximations must have different regions of validity. The homogenization theory and internal-layer theory results are derived with the following assumptions made

- HT: $r \to \infty$ and Pe finite, and
- ILT: $Pe \to \infty$ and r finite.

As we will demonstrate, the transition between the two regimes occurs around $r = Pe$, that may supported by the following kinetic argument. In the homogenization theory regime, where the lengthscale separation between the passive tracer and flow is large, the advection-diffusion problem is approximated by a diffusion process. The typical passive tracer particle is pushed in one direction and travels along the streamline until its direction changes due to the molecular diffusion. This will move the particle onto a streamline of the opposite direction in a relatively long time L_u^2/κ. In this timeframe the effective diffusion covariance matrix is close to its long-time limit calculated by HT so the HT gives an accurate estimate of the mixing enhancement. On the other hand, when the source-sink distribution is nonzero, the scale separation may be violated and

the HT fails. The dominant timescale in the problem is then the minimum between the time to travel to a streamline of a different direction L_u^2/κ, and the time to travel from a source to a sink L_s/U, where the passive tracer particle exits the system. The balance of the two timescales occurs when

$$\frac{L_u^2}{\kappa} = \frac{L_s}{U} \qquad \Longleftrightarrow \qquad Pe = r\,. \tag{3.33}$$

The homogenization scaling $1 + Pe^2$ seems to form a universal upper bound for both regimes, but not necessarily a sharp bound for $Pe > r$. A dispersion-diffusion theory is designed with a hope that it will be able to reconcile the particle dispersion modelling techniques with the multiscale mixing measures and at the same time capture both the HT as well as the ILT predictions.

In the case of no stirring the nonhomogeneous diffusion equation (3.11) has a simple steady solution of form

$$\theta_0(x, y) = \frac{\sqrt{2}S}{k_s^2\kappa} \sin(k_s x)\,, \tag{3.34}$$

whereas in the presence of stirring the steady solution solves the non-homogenous, second order PDE

$$\sqrt{2}U \sin(k_u y)\partial_x\theta_\infty = \kappa\Delta\theta_\infty + \sqrt{2}S \sin(k_s x) \tag{3.35}$$

whose solution can be written in the form

$$\theta_\infty(x, y) = f(y) \sin(k_s x) + g(y) \cos(k_s x) \tag{3.36}$$

with $\frac{2\pi}{k_u}$-periodic functions f and g. The functions f and g may be found using numerical algorithms (spectral methods, finite differences) or asymptotical methods

as in *Shaw et al.* (2007). The numerical calculation yields operationaly exact results.

3.2.2 Effective diffusivity matrix derivation

Stochastic differential equations governing the motion of a passive tracer particle in the two-scale problem are described by the system

$$dX(t) = \sqrt{2}U\sin(k_u Y(t))dt + \sqrt{2\kappa}\,dW_X(t)\,, \tag{3.37}$$

$$dY(t) = \sqrt{2\kappa}\,dW_Y(t)\,, \tag{3.38}$$

where W_X and W_Y are independent Wiener processes. Since the dynamics of $Y(t)$ is decoupled from the dynamics of $X(t)$ the system may be written as

$$X(t) = x_0 + \sqrt{2}U\int_0^t \sin(k_u Y(t))dt + \sqrt{2\kappa}\,W_X(t)\,, \tag{3.39}$$

$$Y(t) = y_0 + \sqrt{2\kappa}\,W_Y(t)\,. \tag{3.40}$$

The effective diffusivity matrix may be derived in view of the calculation in *Majda and Kramer* (1999). In the y-direction,using standard properties of a Wiener process, one obtains

$$\mathbb{E}[(Y(t)-y_0)^2] = \mathbb{E}[(\sqrt{2\kappa}\,W_Y(t))^2] = 2\kappa t \tag{3.41}$$

The calculation becomes more complicated but straighforward in the x-direction

$$\mathbb{E}[(X(t)-x_0)^2] = 2\kappa t + 2U^2\int_0^t\int_0^t \mathbb{E}[\sin(k_u Y(s))\sin(k_u Y(s'))]dsds'\,. \tag{3.42}$$

The integrand is written as

$$\begin{aligned}
\mathbb{E}[\sin(k_u Y(s))\sin(k_u Y(s'))] &= \frac{1}{2}\mathbb{E}[\cos(k_u Y(s) - k_u Y(s')) - \cos(k_u Y(s) + k_u Y(s'))] \\
&= \frac{1}{2}\mathbb{E}[\cos(k_u\sqrt{2\kappa}\,(W_Y(s) - W_Y(s')))] \\
&- \frac{1}{2}\mathbb{E}[\cos(k_u\sqrt{2\kappa}\,(2y_0 + W_Y(s) + W_Y(s')))] \\
&= \frac{1}{2}\mathcal{R}\left\{\mathbb{E}[e^{ik_u\sqrt{2\kappa}\,(W_Y(s)-W_Y(s'))}]\right\} \\
&- \frac{1}{2}\mathcal{R}\left\{\cos(2k_u y_0)\mathbb{E}[e^{ik_u\sqrt{2\kappa}\,(W_Y(s)+W_Y(s'))}]\right\} \qquad (3.43)
\end{aligned}$$

where $\mathcal{R}$ stands for the real part. The sum or difference of Gaussian random variables is again Gaussian and in particular

$$\begin{aligned}
W_Y(s) - W_Y(s') &= W_Y(|s - s'|) \sim \mathcal{N}(0, |s - s'|) \qquad (3.44) \\
W_Y(s) + W_Y(s') &\sim \mathcal{N}(0, \tau) \qquad (3.45)
\end{aligned}$$

where the variance τ satisfies

$$\tau = Var(W_Y(s)) + Var(W_Y(s')) + 2Cov(W_Y(s), W_Y(s')) = s + s' + 2(s \wedge s') \qquad (3.46)$$

where the notation $s \wedge s' = \min\{s, s'\}$ is used. The following standard identity for Gaussian random variable ξ may be used to simplify expression in (3.43)

$$\mathbb{E}[e^{i\xi}] = e^{i\mathbb{E}[\xi]}e^{-\frac{1}{2}\mathbb{E}[(\xi-\mathbb{E}[\xi])^2]}\,. \qquad (3.47)$$

Further details may be found in *Øksendal* (1998). The mean in (3.43) simplifies to

$$\mathbb{E}[\sin(k_u Y(s))\sin(k_u Y(s'))] = \frac{1}{2}e^{-k_u^2\kappa\,|s-s'|} - \frac{1}{2}\cos(2k_u y_0)e^{-k_u^2\kappa\,(s+s'+(s\wedge s'))}\,. \qquad (3.48)$$

This expression can be easily integrated by using a symmetry of $|s-s'|$ and $s+s'+(s\wedge s')$ with respect to a line $s=s'$. We find the variance in the x-direction to be

$$\begin{aligned}
&\mathbb{E}[(X(t)-x_0)^2] = \qquad (3.49)\\
&= 2\kappa t + U^2 \int_0^t \int_0^t e^{-k_u^2\kappa|s-s'|} - \cos(2k_u y_0) e^{-k_u^2\kappa(s+s'+(s\wedge s'))} ds ds'\\
&= 2\kappa t + 2U^2 \int_0^t \frac{1}{k_u^2\kappa} - \frac{e^{-k_u^2\kappa s'}}{k_u^2\kappa} + \cos(2k_u y_0)\left(\frac{e^{-4k_u^2\kappa s'}}{3k_u^2\kappa} - \frac{e^{-k_u^2\kappa s'}}{3k_u^2\kappa}\right) ds'\\
&= 2\kappa t + 2U^2 \left(\frac{k_u^2\kappa t + e^{-k_u^2\kappa t} - 1}{k_u^4\kappa^2} - \cos(2k_u y_0)\left(\frac{3 - 4e^{-k_u^2\kappa t} + e^{-4k_u^2\kappa t}}{12k_u^4\kappa^2}\right)\right) \qquad (3.50)
\end{aligned}$$

Similar technique is used for finding the second mixed moment

$$\begin{aligned}
&\mathbb{E}[(X(t)-x_0)(Y(t)-y_0)] =\\
&= 2\sqrt{\kappa}U \int_0^t \mathbb{E}[\sin(k_u Y(s))\cdot W_Y(t)]ds + 2\kappa\,\mathbb{E}[W_X(t)\cdot W_Y(t)]\\
&= 2\sqrt{\kappa}U \int_0^t \mathbb{E}[\sin(k_u Y(s))\cdot (W_Y(s) + W_Y(t-s))]ds\\
&= 2\sqrt{\kappa}U \int_0^t \mathbb{E}[\sin(k_u y_0 + k_u\sqrt{2\kappa}\,W_Y(s))\cdot W_Y(s)]ds \qquad (3.51)
\end{aligned}$$

In order to evaluate the average above we use the following trick, consisting of evaluating the integral of the average with respect to variable of our choice first, and subsequently of differentiating it to obtain the desired result

$$\begin{aligned}
\mathbb{E}[\sin(k_u y_0 + k_u\sqrt{2\kappa}\,W_Y(s))W_Y(s)] &= -\frac{1}{k_u}\frac{d}{d\sqrt{2\kappa}}\mathbb{E}[\cos(k_u y_0 + k_u\sqrt{2\kappa}W_Y(s))]\\
&= -\frac{1}{k_u}\frac{d}{d\sqrt{2\kappa}}\mathcal{R}\left\{e^{ik_u y_0 + ik_u\sqrt{2\kappa}\,W_Y(s)}\right\}\\
&= -\frac{1}{k_u}\frac{d}{d\sqrt{2\kappa}}\cos(k_u y_0)e^{-\frac{1}{2}k_u^2 s\sqrt{2\kappa}^2}\\
&= \sqrt{2\kappa}\,k_u s\cos(k_u y_0)e^{-k_u^2\kappa s}\,. \qquad (3.52)
\end{aligned}$$

The time integration gives the final form of the second mixed moment

$$\begin{aligned}
\mathbb{E}[(X(t)-x_0)(Y(t)-y_0)] &= 2\sqrt{2}k_u\kappa U\cos(k_u y_0)\int_0^t s e^{-k_u^2\kappa s}ds \\
&= 2\sqrt{2}k_u\kappa U\cos(k_u y_0)\left[-\frac{se^{-k_u^2\kappa s}}{k_u^2\kappa}-\frac{e^{-k_u^2\kappa s}}{k_u^4\kappa^2}\right]_0^t \\
&= 2\sqrt{2}k_u\kappa U\cos(k_u y_0)\frac{1-e^{-k_u^2\kappa t}-k_u^2\kappa t e^{-k_u^2\kappa t}}{k_u^4\kappa^2} \\
&= \frac{2\sqrt{2}Ut}{k_u}\cos(k_u y_0)\left(\frac{1-e^{-k_u^2\kappa t}}{k_u^2\kappa t}-e^{-k_u^2\kappa t}\right). \quad (3.53)
\end{aligned}$$

Collecting the second moment results we calculate the correlation matrix from (3.27) as

$$\mathbf{C}(\vec{x_0},t)=\begin{pmatrix}\mathbb{E}[(X(t)-x_0)^2] & \mathbb{E}[(X(t)-x_0)(Y(t)-y_0)]\\ \mathbb{E}[(X(t)-x_0)(Y(t)-y_0)] & \mathbb{E}[(Y(t)-y_0)^2]\end{pmatrix}$$

but also the effective diffusivity matrix

$$\begin{aligned}
\mathbf{K}(\vec{x_0},t)&=\frac{1}{2}\frac{d}{dt}\mathbf{C}(\vec{x_0},t)=\begin{pmatrix}K_{11}(\vec{x}_0,t) & K_{12}(\vec{x}_0,t)\\ K_{21}(\vec{x}_0,t) & K_{22}(\vec{x}_0,t)\end{pmatrix}\\
&=\begin{pmatrix}1-e^{-k_u^2\kappa t}k_u^2\kappa-\cos(2k_u y_0)\left(\frac{4e^{-k_u^2\kappa t}-4e^{-4k_u^2\kappa t}}{12k_u^2\kappa}\right) & \sqrt{2}k_u\kappa t U\cos(k_u y_0)e^{-k_u^2\kappa t}\\ \sqrt{2}k_u\kappa t U\cos(k_u y_0)e^{-k_u^2\kappa t} & \kappa\end{pmatrix}
\end{aligned}$$

where in the limit $t\to\infty$ the effective diffusivity matrix has a diagonal form

$$\lim_{t\to\infty}\mathbf{K}(\vec{x_0},t)=\kappa\begin{pmatrix}1+Pe^2 & 0\\ 0 & 1\end{pmatrix} \quad (3.54)$$

with $Pe=\frac{U}{k_u\kappa}$. This is consistent with the homogenization theory result for the given flow (even without using the usual lengthscale separation assumption). Fig. 3.3

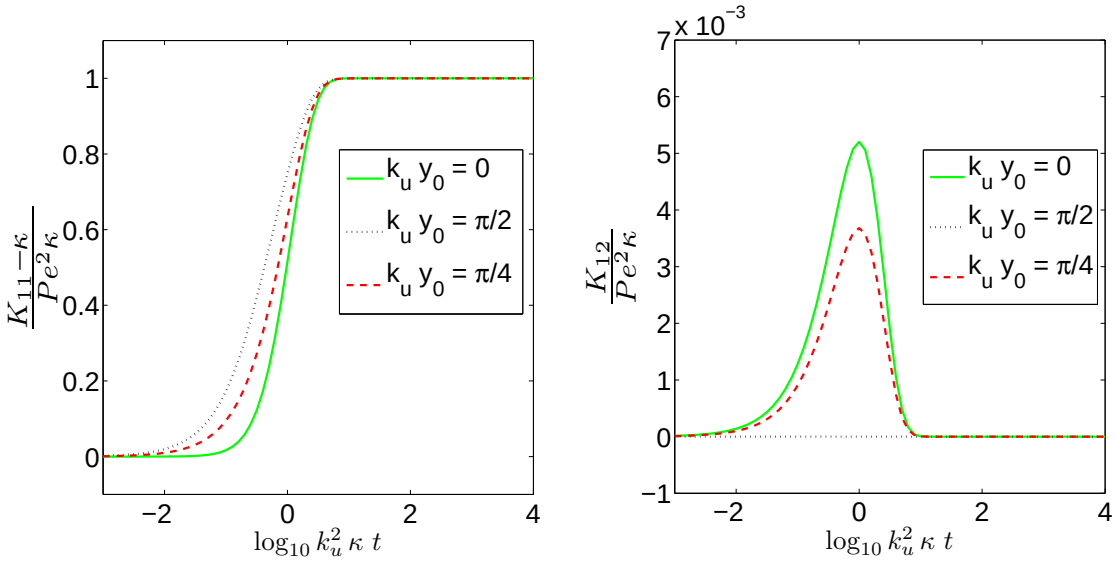

Figure 3.3: **Temporal evolution of $\frac{K_{11}-\kappa}{Pe^2\kappa}$ and $\frac{K_{12}}{Pe^2\kappa}$ against $k_u^2\kappa t$ for $k_u y_0 = 0, \pi/2, \pi/4$.**

shows the temporal evolution of K_{11} and K_{12} for different initial values y_0 where the renormalization shows the asymptotic Pe^2 scaling as a convergence to one. The homogenization theory treats the case when the lengthscale separation between the initial concentration field and stirring field is assumed to be large and the particles may experience the long-time effective diffusivity, as in (3.54). The homogenization theory prediction is simply

$$\theta_{HT} = \frac{s(x)}{\kappa k_s^2(1+Pe^2)}\,. \tag{3.55}$$

But at very high Péclet number in the temporal problem without a lengthscale separation between sources and the flow, many passive tracer particles experience only the short-time effective diffusivity since they spend relatively less time in the system before exiting through a sink. Therefore the full temporal dependance in the effective diffusivity matrix affects the mixing properties of the given flow, as in (3.28).

3.2.3 Multiscale mixing efficiencies in the dispersion-diffusion approximation

The DDT approximation of the solution to the two-scale problem satisfies (3.26) and (3.28). The source term may be written as

$$s(x) = \sqrt{2}S\sin(k_s x) = \frac{\sqrt{2}}{2i}S(e^{ik_s x} - e^{-ik_s x}) \tag{3.56}$$

In the next calculation we express the DDT approximation in the same form as (3.36). Given the wavevector $\vec{k} = (k_x, k_y k_u)$, $\theta_{DDT}(x,y)$ satisfies

$$\begin{aligned}
\theta_{DDT}(x,y) &= \lim_{t\to\infty}\int_0^t dt_0 \int_{[0,2\pi]^2} d\vec{x}_0 \rho(\vec{x},t;\vec{x}_0,t_0)s(\vec{x}_0) \\
&= \frac{k_u}{4\pi^2}\sum_{\vec{k}\in\mathbb{R}^2}\int_{[0,\frac{2\pi}{k_u}]\times[0,2\pi]} d\vec{x}_0 e^{i\vec{k}(\vec{x}-\vec{x_0})}s(x_0)\int_0^\infty dt e^{-\frac{1}{2}\vec{k}\mathbf{C}(\vec{x_0},t)\vec{k}^T}
\end{aligned}$$

that reduces to a difference

$$\begin{aligned}
\theta_{DDT}(x,y) &= \frac{\sqrt{2}k_u S}{4\pi i}e^{ik_s x}\sum_{k_y=-\infty}^{\infty} e^{ik_y k_u y}\int_0^\infty dt \int_0^{\frac{2\pi}{k_u}} e^{-ik_y k_u y_0 - \frac{1}{2}(k_s,k_y k_u)\mathbf{C}(y_0,t)(k_s,k_y k_u)^T} dy_0 \\
&- \frac{\sqrt{2}k_u S}{4\pi i}e^{ik_s x}\sum_{k_y=-\infty}^{\infty} e^{ik_y k_u y}\int_0^\infty dt \int_0^{\frac{2\pi}{k_u}} e^{-ik_y k_u y_0 - \frac{1}{2}(-k_s,k_y k_u)\mathbf{C}(y_0,t)(-k_s,k_y k_u)^T} dy_0
\end{aligned}$$

Note that due to the fact that the summands in the above expression are complex conjugates, the result is real. In particular,

$$\begin{aligned}
&\theta_{DDT}(x,y) = \\
&= \frac{\sqrt{2}k_u S}{2\pi}\mathcal{R}\left\{ e^{ik_s x}\sum_{k_y=-\infty}^{\infty} e^{ik_y k_u y}\int_0^\infty dt \int_0^{\frac{2\pi}{k_u}} e^{-ik_y k_u y_0 - \frac{1}{2}(k_s,k_y k_u)\mathbf{C}(y_0,t)(k_s,k_y k_u)^T} dy_0 \right\}
\end{aligned}$$

If we define

$$\hat{f}(y)+i\hat{g}(y) := \frac{\sqrt{2}k_u S}{2\pi}\sum_{k_y=-\infty}^{\infty} e^{ik_y k_u y}\int_0^{\infty} dt \int_0^{\frac{2\pi}{k_u}} e^{-ik_y k_u y_0 - \frac{1}{2}(k_s,k_y k_u)\mathbf{C}(y_0,t)(k_s,k_y k_u)^T} dy_0 \tag{3.57}$$

then the DDT approximation may be written as

$$\theta_{DDT}(x,y) = \hat{f}(y)\sin(k_s x) + \hat{g}(y)\cos(k_s x) \tag{3.58}$$

where

$$\begin{aligned}
\hat{f}(y) &= \frac{\sqrt{2}k_u S}{2\pi}\sum_{k_y=-\infty}^{\infty}\cos(k_y k_u y)\int_0^{\infty} dt\int_0^{\frac{2\pi}{k_u}}\cos(k_y k_u y_0)e^{-\frac{1}{2}(k_s,k_y k_u)\mathbf{C}(y_0,t)(k_s,k_y k_u)^T}dy_0 \\
\hat{g}(y) &= \frac{\sqrt{2}k_u S}{2\pi}\sum_{k_y=-\infty}^{\infty}\sin(k_y k_u y)\int_0^{\infty} dt\int_0^{\frac{2\pi}{k_u}}\cos(k_y k_u y_0)e^{-\frac{1}{2}(k_s,k_y k_u)\mathbf{C}(y_0,t)(k_s,k_y k_u)^T}dy_0
\end{aligned}$$

Note that the correlation matrix is an even function of y_0 since it has the form $\mathbf{C}(\vec{x_0},t) = \mathbf{C}(\cos(k_u y_0),\cos(2k_u y_0),t)$. This property together with the symmetry of sine function was used to simplify the above expressions for $\hat{f}$ and $\hat{g}$. Next we write $\hat{f}$ and $\hat{g}$ in the form

$$\hat{f}(y) = \frac{\sqrt{2}S}{Uk_s}\sum_{k_y=-\infty}^{\infty}\cos(k_y k_u y)I_{k_y}\,, \qquad \hat{g}(y) = \frac{\sqrt{2}S}{Uk_s}\sum_{k_y=-\infty}^{\infty}\sin(k_y k_u y)I_{k_y} \tag{3.59}$$

where the non-dimensional quantity I_{k_y} satisfies

$$I_{k_y} = \frac{k_u k_s U}{2\pi}\int_0^{\infty} dt\int_0^{\frac{2\pi}{k_u}}\cos(k_y k_u y_0)e^{-\frac{1}{2}(k_s,k_y k_u)\mathbf{C}(y_0,t)(k_s,k_y k_u)^T}dy_0\,, \quad k_y\in\mathbb{Z}\,. \tag{3.60}$$

To compute the mixing measures $\mathcal{E}_p$, $p = 0, \pm 1$ for the DDT approximation θ_{DDT} of the exact solution θ_∞ we use (3.57)-(3.60) with the Parseval's formula. The scalar variance of the steady unstirred solution satisfies

$$\langle \theta_0 \rangle = \frac{S^2}{\kappa^2 k_s^4} \tag{3.61}$$

We calculate the scalar variance on different scales in the presence of stirring. On the intermediate scale we have

$$\begin{aligned}
\langle \theta_{DDT}^2 \rangle &= \frac{1}{2}\langle f^2 + g^2 \rangle \\
&= \frac{2S^2}{U^2 k_s^2}\frac{k_u}{4\pi^2}\int_0^{2\pi/k_u} dy \int_0^{2\pi} dx \left(\sum_{k_y=-\infty}^{\infty} \cos(k_y k_u y) I_{k_y} \right)^2 + \left(\sum_{k_y=-\infty}^{\infty} \sin(k_y k_u y) I_{k_y} \right)^2 \\
&= \frac{S^2}{4U^2 k_s^2 \pi^2}\frac{1}{2}\left[2I_0^2 + \sum_{k_y=1}^{\infty} (I(k_y) + I(-k_y))^2 + \sum_{k_y=1}^{\infty} (I(k_y) - I(-k_y))^2 \right] \qquad (3.62) \\
&= \frac{S^2}{4\pi^2 U^2 k_s^2} \sum_{k_y=-\infty}^{\infty} I^2(k_y) \qquad (3.63)
\end{aligned}$$

On the small scale we use

$$\begin{aligned}
||\nabla \theta_{DDT}||_2^2 &= \left(\frac{\partial \theta_{DDT}}{\partial x} \right)^2 + \left(\frac{\partial \theta_{DDT}}{\partial y} \right)^2 \\
&= k_s^2 \left(f(y)\cos(k_s x) - g(y)\sin(k_s x) \right)^2 + \left(f'(y)\sin(k_s x) + g'(y)\cos(k_s x) \right)^2
\end{aligned}$$

to obtain

$$\left\langle \left(\frac{\partial \theta_{DDT}}{\partial x} \right)^2 + \left(\frac{\partial \theta_{DDT}}{\partial y} \right)^2 \right\rangle = \frac{k_s^2}{2}\langle f^2 + g^2 \rangle + \frac{1}{2}\langle f'^2 + g'^2 \rangle \,. \tag{3.64}$$

The symmetry of sine next results in

$$\frac{1}{2}\langle f'^2\rangle = \frac{S^2k_u^2}{8U^2k_s^2\pi^2}\sum_{k_y=1}^{\infty}k_y^2[I(k_y)-I(-k_y)]\,, \tag{3.65}$$

$$\frac{1}{2}\langle g'^2\rangle = \frac{S^2k_u^2}{8U^2k_s^2\pi^2}\sum_{k_y=1}^{\infty}k_y^2[I(k_y)+I(-k_y)]\,, \tag{3.66}$$

$$\frac{1}{2}\langle f'^2+g'^2\rangle = \frac{S^2k_u^2}{4U^2k_s^2\pi^2}\sum_{k_y=-\infty}^{\infty}k_y^2I^2(k_y)\,. \tag{3.67}$$

Then the previously derived (3.63) together with (3.64) leads to

$$\langle(\nabla\theta_{DDT})^2\rangle = \frac{S^2}{4U^2\pi^2}\sum_{k_y=-\infty}^{\infty}(1+r^2k_y^2)^pI^2(k_y) \tag{3.68}$$

This result applies on small scales in case $p=+1$, and by a similar calculation also on large scales in case $p=-1$. The multiscale mixing efficiencies take the form

$$\mathcal{E}_0 = \sqrt{\frac{\langle\theta_0^2\rangle}{\langle\theta_{DDT}^2\rangle}} = rPe\left(\sum_{k_y\in\mathbb{Z}}I_{k_y}^2\right)^{-\frac{1}{2}} \tag{3.69}$$

$$\mathcal{E}_{+1} = \sqrt{\frac{\langle\nabla\theta_0^2\rangle}{\langle\nabla\theta_{DDT}^2\rangle}} = rPe\left(\sum_{k_y\in\mathbb{Z}}(1+r^2k_y^2)I_{k_y}^2\right)^{-\frac{1}{2}} \tag{3.70}$$

$$\mathcal{E}_{-1} = \sqrt{\frac{\langle\nabla^{-1}\theta_0^2\rangle}{\langle\nabla^{-1}\theta_{DDT}^2\rangle}} = rPe\left(\sum_{k_y\in\mathbb{Z}}\frac{I_{k_y}^2}{1+r^2k_y^2}\right)^{-\frac{1}{2}} \tag{3.71}$$

Let us next explore the scaling of $\mathcal{E}_0$ in the homogenization regime. If $r=k_u/k_s\to\infty$, linear transformation $z=k_uy_0$ and $\tau=\frac{U^2k_s^2t}{\kappa k_u^2}$ in the nondimensional integral (3.60) leads to asymptotic dependance

$$\sum_{k_y=-\infty}^{\infty}I_{k_y} = \sum_{k_y=-\infty}^{\infty}\frac{r}{2\pi Pe}\int_0^{\infty}d\tau\int_0^{\frac{2\pi}{k_u}}\cos(k_yz)e^{-\frac{1}{2}(k_s,k_y)\mathbf{C}(k_s,k_y)^T}dz \sim I_0 = \frac{r}{Pe} \tag{3.72}$$

where similar calculation can be done for $\mathcal{E}_{-1}$. Then the following asymptotic estimate holds

$$\mathcal{E}_0 \sim \frac{rPe}{r/Pe} = Pe^2\,, \quad \text{as} \quad r \to \infty\,. \tag{3.73}$$

This suggests that the mixing efficiencies for the DDT approximation are consistent with HT in the regime of finite Pe and $r \to \infty$. Moreover, it has been shown by *Thiffeault et al.* (2004) that for any periodic flow with single-scale sources a universal upper bound holds

$$\mathcal{E}_0 \leq \sqrt{1 + 2r^2Pe^2}\,. \tag{3.74}$$

This bound is independent of the flow properties as it does not depend on the flow wave number. Its derivation as well as improvement to $\mathcal{E}_0 \leq \sqrt{1 + r^2Pe^2}$ for a sine flow in one direction can be found in Appendix D. On the other hand, it has been shown, using asymptotic methods *Shaw et al.* (2007), that finite r and $Pe \to \infty$ yields yet another law

$$\mathcal{E}_0 \sim r^{7/6}Pe^{5/6}\,, \quad \text{as} \quad Pe \to \infty\,. \tag{3.75}$$

and at small and large scales

$$\begin{aligned} \mathcal{E}_{+1} &\sim r^{1/2}Pe^{1/2}\,, \quad \text{as} \quad Pe \to \infty \\ \mathcal{E}_{-1} &\sim rPe\log Pe\,, \quad \text{as} \quad Pe \to \infty \end{aligned}$$

The three different asymptotic laws at intermediate scales cannot be valid at the same time and the only possible explanation is that the limits $r \to \infty$ and $Pe \to \infty$ do not commute and their order predetermines the asymptotic regime of the model. In the next section we provide the computational results capturing the mixing enhancement in the DDT model across the (r, Pe)-parameter space.

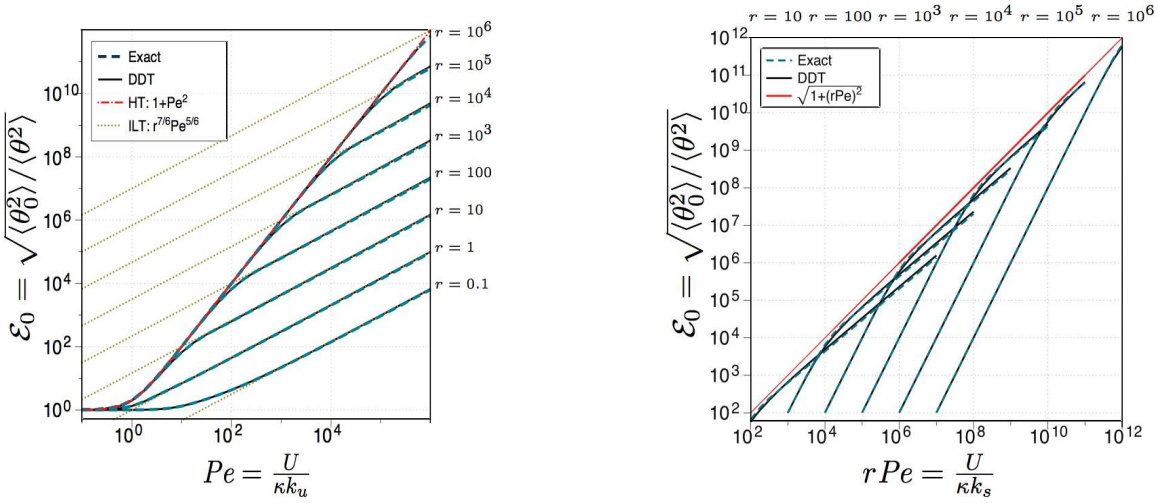

Figure 3.4: $\mathcal{E}_0$ **versus** Pe **and** rPe**.**

3.2.4 Numerical performance of dispersion-diffusion approximation

In our numerical results we use five different approximations to compute and compare the scalar field and the variance suppression mixing measure on various scales:

- **"Exact":** Numerically computed approximation of the exact solution (3.35) by spectral methods (dashed blue curves)
- **DDT:** Computation of the scalar field using (3.58)-(3.60) (with modes $|k_y| < 30$) and enhancement factors using (3.69)-(3.71) and (3.60) (solid black)
- **HT:** Computation of the enhancement factor scaling by (3.73) (dash-dotted red)
- **ILT:** Computation of the enhancement factor scaling by (3.75) (dotted yellow)
- **Bound:** Universal upper bound in terms of r and Pe, according to (3.74) (solid red)

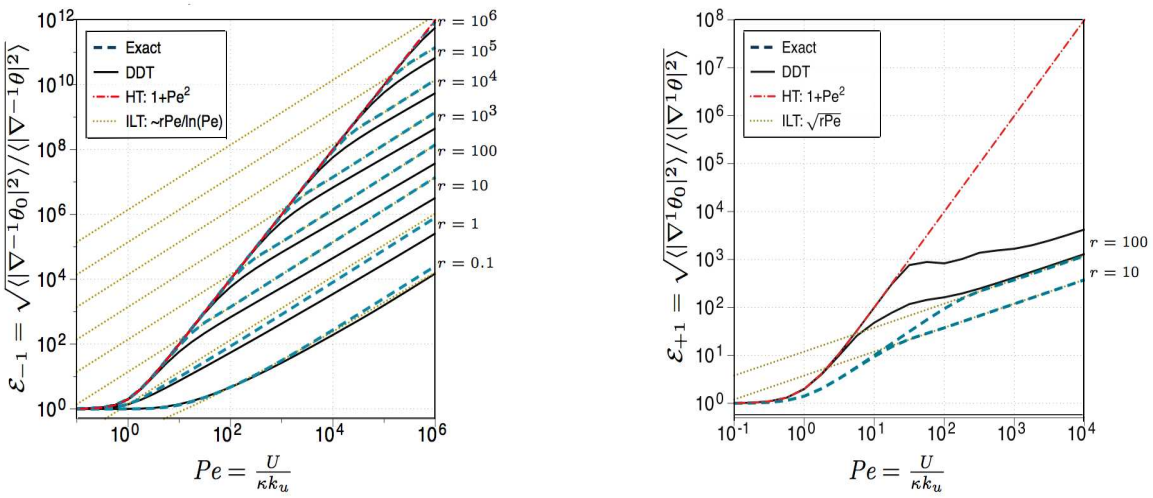

Figure 3.5: $\mathcal{E}_{-1}$ **and** $\mathcal{E}_1$ **versus** Pe**.**

Fig. 3.4 (left) illustrates accuracy of the different approaches (among them "exact", DDT, HT and ILT approximations) and their performance in capturing variance suppression on intermediate scale $\mathcal{E}_0$. The "exact" behavior shows that there are two regimes with a distinct r and Pe dependance: $\{r > Pe, r \to \infty\}$ and $\{r < Pe, Pe \to \infty\}$. While the dispersion-diffusion theory captures the variance suppression accurately in the whole (r, Pe)-space, the homogenization theory is only valid in the first region and the internal-layer theory in the second region. This reflects the noncomutativity of the large-r and large-Pe asymptotic regimes. The transition between the regimes occurs at $r = Pe$. Fig. 3.4 (right) shows that dispersion-diffusion theory respects the rigorous absolute bound (3.74) for the stationary monochromatic distribution that is independent of the choice of flow. The performance of the DDT method is contrasted with other methods on both the large scales (Fig. 3.5 on the left) and small scales (Fig. 3.5 on the right). As seen on Fig. 3.5, the DDT still captures the correct scaling of $\mathcal{E}_{-1}$ in both the HT regime ($r > Pe$) and in the ILT regime ($Pe > r$), even though the exact values differ slightly. However, on small scales, neither the HT nor the DDT predict the variance suppression correctly (ILT is accurate for $Pe > r$)

as seen on Fig. 3.5. Surprisingly, the exact scaling does not even coincide with the HT prediction in the regime $Pe < r \to \infty$. In the asymptotic regime for large r the numerical computation of I_{k_y} becomes very sensitive to numerical errors and the calculation of $\mathcal{E}_{+1}$ may be disrupted by these errors. The approximated density of the

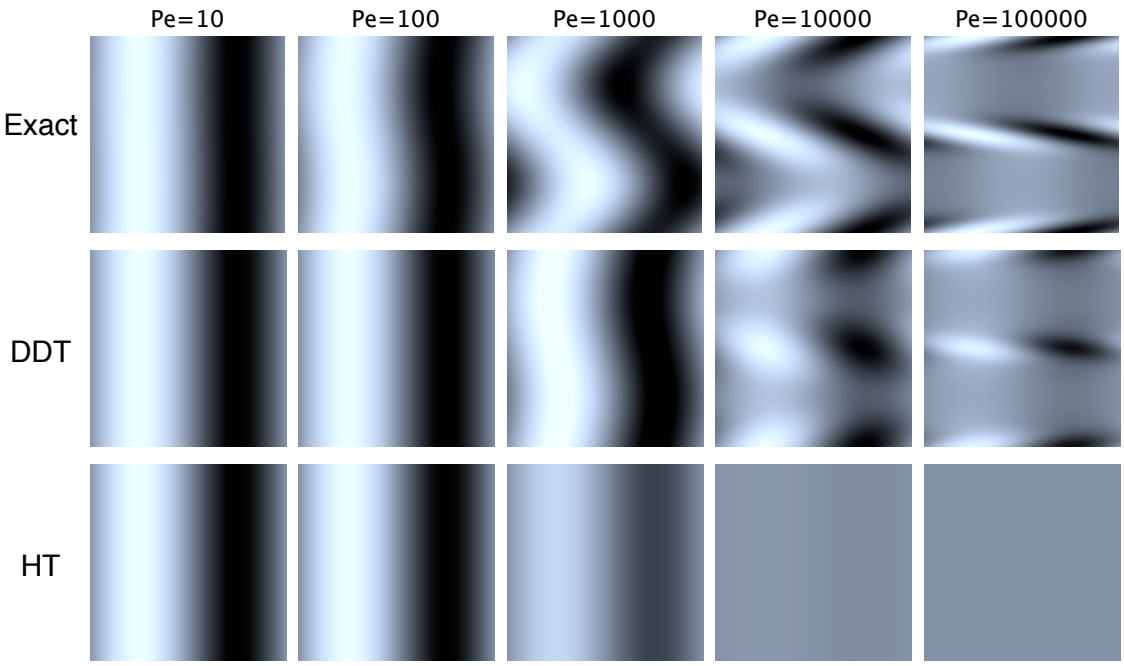

Figure 3.6: **Steady-state scalar fields for** $x \in [0, \frac{2\pi}{k_s}]$**,** $y \in [0, \frac{2\pi}{k_y}]$ **when** $r = 1000$**.**

passive scalar concentration can be compared and contrasted in the physical space $x \in [0, \frac{2\pi}{k_s}]$, $y \in [0, \frac{2\pi}{k_y}]$ between the three approaches: dispersion-diffusion approach, homogenization theory and "exact" solution. Fig. 3.6 shows the performance of DDT and HT, compared to the exact field plot, in capturing the small-scale features of the scalar concentration for $Pe = 10$, 100, 1000, 10000, 100000 and for fixed $r = 1000$. Within each column the fields are renormalized to the maximum magnitude of the exact solution and then plotted on the same grayscale (where -1 corresponds to black and 1 to white). As the stirring intensity increases with Pe, the internal layers form in places with the largest shear whereas the best mixing occurs at places with the strongest flow. Both the width and the angle of internal layers are remarkably well

captured by DDT on intermediate and large scales. However, HT does not contain any detailed structure and always resembles the source-sink distribution. While the DDT approximates the bulk properties of the concentration field very closely for all parameter values, HT greatly overestimates the concentration variance for $Pe > r$.

3.3 Conclusions

Multiple theories and computational techniques in the literature serve to quantify the flow enhancement of molecular diffusion. Their predictions are not always consistent. One of the fundamental questions answered in the chapter was how similar and different features of these theories can be reconciled into one model that produces the correct result for one of the mixing measures (scalar concentration variance suppression). We introduced a model utilizing the essence of the particle dispersion technique that respects the temporal structure of inhomogeneities introduced by sources and sinks. Our findings for a simple sine flow can be summarized as follows:

- **Effect of sources:** In the transient problem ($s = 0$) the time asymptotic bulk properties of the concentration field are dominated by the long-time dispersive behavior of the scalar, which ignores much of the structure of stirring. The approximation using the constant (in time) additive enhancement factor, as in homogenization theory, is accurate in its prediction of the Pe^2 scaling. However, in the presence of sources and sinks, the dominant contribution to the bulk variance comes from scalar particles that have been most recently introduced into the system. In such case, it is the balance between the lengthscales (flow and source scale), that determines the mixing properties of the given flow. Therefore the mixing properties naturally depend not only on Péclet number but also on the scale separation r.

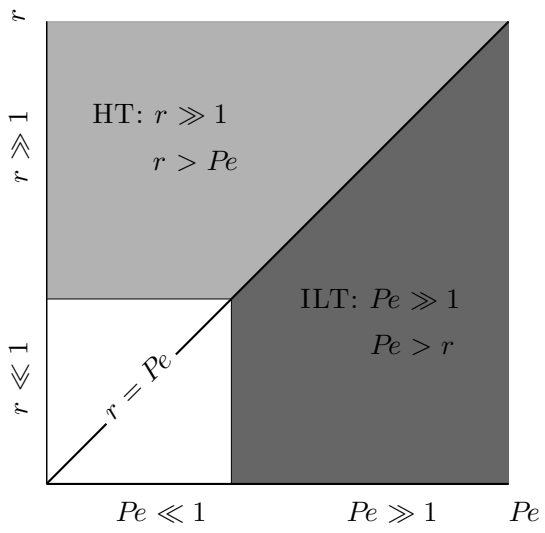

Figure 3.7: **Validity of mixing theories in the** (r, Pe) **phase space.**

- **Non-comutativity of asymptotic regimes:** In the nonhomogeneous problem ($s \neq 0$) the mixing properties depend on two nondimensional parameters: Pe and r. The two existing theories (HT and ILT) are both found to be valid but in different asymptotic regimes, specified by the parameters of the problem and illustrated on Fig. 4.4. This implies that there is no universal asymptotic scaling for the variance reduction mixing measure in the simple two-parameter model. It is therefore essential to know in which order the limits $r \to \infty$ and $Pe \to \infty$ are taken and the result itself depends on this choice.

- **Validity of DDT approach:** Unlike the HT and ILT theories, the dispersion-diffusion approach captures aspects of both theories and at the same time shows the non-commutativity of the limits. Its ability to capture the large-scale mixing properties encourages us to further study problems in which several lengthscales are present. The method would be even stronger if it allowed to correctly resolve the small scales as well. However, the numerical computations at this scale are highly sensitive to numerical errors.

So far, we have illustrated the DDT performance for a shear-flow problem with single-modal sources. This technique is potentially applicable to a wider class of problems, including those with more complicated flow or source structure. In particular, once the explicit formula for the time- and position-dependent efficiency matrix $\mathbf{C}(\vec{x}, t)$ is known, or at least its short-time properties, the problem with more general source-sink distribution may be studied. Among such examples we can mention turbulent flows like homogeneous and isotropic turbulence *Faber and Vassilicos* (2009), *Richardson* (1926) where the time-dependent covariance matrix may be modelled by

$$\mathbf{E}[(X_i(t) - X_i(0))(Y_i(t) - Y_i(0))] \sim (2\kappa t + U^2 t^2 + C_R \varepsilon t^3)\delta_{ij} \tag{3.76}$$

for displacements in the inertial range (or smaller), i.e., $L << U^3/\epsilon = L_u$, where the absolute constant $C_R \approx 1$ is called a Richardson constant and $\varepsilon = U^3/l_u$ is the mean energy dissipation rate. Prediction of the DDT using this covariance remains to be tested in direct numerical simulations, and ultimately in experiments.

CHAPTER IV

Noise–induced (statistical) stability in differential equations

4.1 Introduction: Rayleigh-Bénard Convection

It is a common observation that noise added to a dynamical system may destabilize it and produce large fluctuations. Clearly, this statement is true for many important applied models in biology, chemistry, physics and finance. One such model (Hodgkin-Huxley model) is studied in detail in Chapter II. However, an interesting question is whether noise may act in the opposite way and stabilize the system. In this chapter we endeavour to systematically construct and study an example relevant to an applied model arising in fluid dynamics for which almost all solution trajectories in the deterministic model escape to infinity, but an arbitrarily small stochastic perturbation of the system leads to stochastic oscillations and statistically "stable" patterns.

One of the classical models in fluid dynamics and a fundamental paradigm of nonlinear science is Rayleigh-Bénard convection (RBC), see *Kadanoff* (2001). A fluid confined between horizontal boundaries separated by height H is heated from below and cooled from above as on Fig. 4.1. The temperature difference between the top and bottom plates creates a buoyancy force that can cause the fluid to flow.

The problem can be described by nonlinear partial differential Boussinesq equations as in *Landau and Lifshitz* (1987), based on the approximation by *Oberbeck* (1879) and *Boussinesq* (1879). Qualitative properties of the fluid and heat flow are deter-

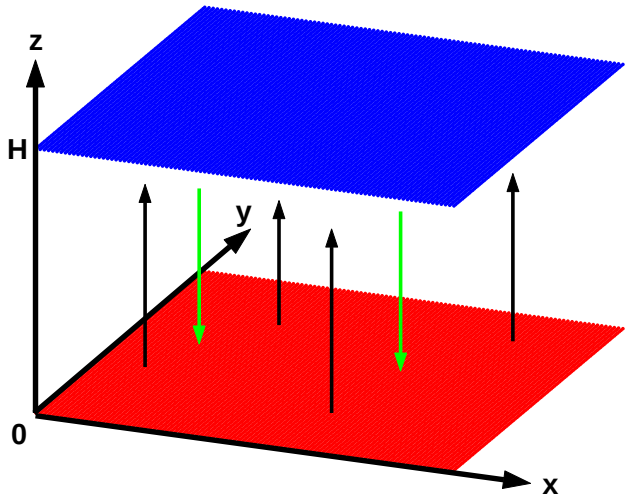

Figure 4.1: **Rayleigh-Bénard convection.**

mined to a great extent by two non-dimensional quantities (parameters in Boussinesq equations): the *Rayleigh number* and the *Prandtl number*. The dimensionless temperature difference ΔT between hot and cold boundary is conventionally indicated by a Rayleigh number, $Ra = \alpha g H^3 \Delta T (\nu\kappa)^{-1}$, where κ, α, ν and g are the coefficients of thermal diffusivity, thermal expansion, kinematic viscosity and the acceleration due to gravity, respectively. Prandtl number is given by the ratio of kinematic viscosity to thermal diffusion, $Pr = \frac{\nu}{\kappa}$.

A quantity of particular interest, that can be calculated from the fluid velocity and temperature field aposteriori, is the heat flux enhancement caused by the flow. It is usually measured by the dimensionless Nusselt number Nu given by the ratio of the total physical heat flux (including both conductive and convective heat flux) to the conductive heat flux in the absence of fluid motion, i.e. $Nu = c\kappa\Delta T/H$ where c is the specific heat.

The problem of thermal convection has been studied extensively from both mathematical and experimental point of view. A linearized analysis of the Boussinesq PDEs reveals that the conductive state is linearly stable for small values of the temperature forcing, measured by Ra. As the value of Ra exceeds a critical value R_c the first unstable mode of the linearized system emerges, leading to existence of an exact exponentialy growing solution of the linearized problem, see *Calzavarini et al.* (2006). In the weakly nonlinear regime close to this transition a convection cell pattern becomes a stable structure of the model. By increasing Ra even more the range of possible instabilities increases leading to chaotic behavior (qualitatively similar to that in the Lorentz system, which is a reduced model of convection), and eventualy to turbulence.

The conventional wisdom in fluid dynamics is that for large values of Ra the functional dependence between Nu and Ra is a power law, see review article *Ahlers et al.* (2009). In the *ultimate* regime, i.e. in the asymptotic limit $Ra \to \infty$, the scaling has been conjectured to be $Nu \sim Ra^{1/2}$ (*Kraichnan* (1962), *Spiegel* (1971)). It has been proven that such an ultimate regime forms a bound, i.e. $Nu \lesssim c_1 + c_2 Ra^{\alpha}$ where $\alpha = 1/2$ by *Busse* (2003), *Doering and Constantin* (1996), *Doering and Constantin* (1992), *Plasting and Kerswell* (2003). Much effort has been spent on laboratory experiments and more recently, in direct numerical simulations of Rayleigh–Bérnard convection in attempts to observe the ultimate regime. Rayleigh numbers up to 10^{17} have been enticed experimentally by *Niemiela and Sreenivasan* (2006) and up to 10^{14} have been resolved computationally by *Amati et al.* (2005), *Verzicco and Sreenivasan* (2008) but the scaling was found to be at most $Nu \sim Ra^{0.37}$, see *Stringano et al.* (2006). One difficulty arises from a fact that such a Rayleigh number might still be far from the ultimate regime; there is no measure telling how large Ra needs to be in order to be close to the ultimate regime – if it exists at all. Another difficulty is that the flow has qualitatively different properties close to the physical boundary:

thermal and viscous boundary layers get thinner (and thus more difficult to resolve computationally) for larger Rayleigh numbers.

One proposed way to eliminate the numerical simulation problem with boundary layers is to investigate the properties of a convective cell with periodic boundary conditions on every side, so–called homogenous Rayleigh-Bénard convection (HRBC). Two recent approaches were performed to study this problem, leading to different conclusions.

In the theoretical approach a critical value $Ra_c = (2\pi)^4$ was found such that above this (fairly small) value a family of *exact* exponentially growing solutions of the full nonlinear system exists, see *Calzavarini et al.* (2006). In the computational approach *Lohse and Toschi* (2003) transient exponentially growing solutions were observed by direct numerical simulation of the fully nonlinear problem sightly above the threshold $Ra > Ra_c$ for $Pr = 1$. When compared, the exponential growth factor matched the growth factor found analytically. Numerical observations reproduced in Fig. 4.2 show that after some seemingly random time these solutions are destabilized and a sudden collapse occurs making values drop down. This process keeps repeating in an unregular pattern and produces statistically steady turbulent heat transport.

The numerical simulators observed a $Nu = Ra^{1/2}$ scaling in the HRBC model and conjectured that the ultimate regime hypothesized by Kraichnan and Spiegel indeed occurs for large Ra. The analysts, on the contrary, were hesitant to draw the conclusion. In an effort to understand why the nonlinear model exhibits physically unfeasible solutions up to a point of collapse, and what the mechanism behind the collapse is, the authors of numerical work were asked to test their method. After setting a different level of precision in the numerical scheme (from the single to the double precision), both groups of researchers observe that result does not seem to change "much" as shown in Fig. 4.3. At higher precision the solution still grows exponentially up to a point of a sudden drop of values but the collapse is slightly

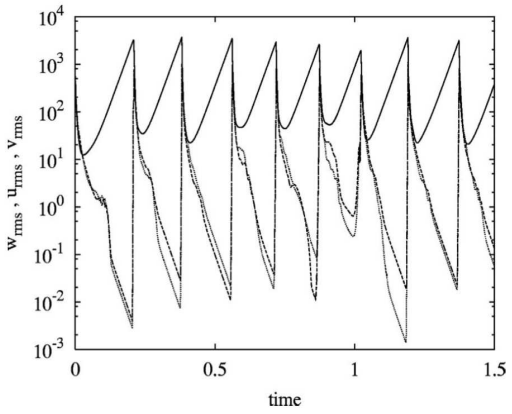

Figure 4.2: **Velocity components in time.** Linear-log plot of the spatial root mean squared (rms) value of the three velocity components and the thermal fluctuation in the direct numerical simulation at resolution 32^3, $w_{\rm rms}$ (solid), $u_{\rm rms}$ (dashed) and $v_{\rm rms}$ (dotted), see *Calzavarini et al.* (2006).

delayed. They concluded that the quantitative observation, i.e. $Nu \sim Ra^{1/2}$, is correct. The reason behind sudden collapses is not fully understood but it is likely that noise in the simulation from either the numerical method or round-off error may be triggering the collapse of the exponentially growing solution.

On the other hand, after observing the change between the simple and double precision calculations, one may hypothesize that (i) with smaller computational error one can track the exponential solution longer (note that one needs to do series of experiments for more than just two values of numerical precision to see whether the point of collapse converges), and more importantly, (ii) the numerical method is very sensitive to noise of very small magnitude such as round-off error (the magnitude of round-off error is usually too small to explain such a big change in the result). The computational result at any finite precision might therefore be qualitatively different from the prediction of the differential equations. It is thus reasonable to conjecture that the collapse of the solution is not a property of the model, but rather a consequence of a small artificial fluctuations that can play an important role when the domain of attraction to an exponential solution is small.

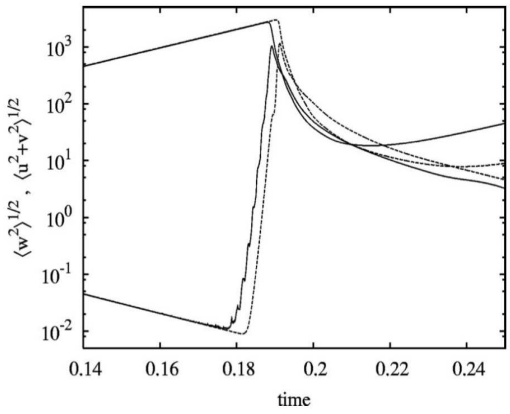

Figure 4.3: **Effect of numerical precision.** A comparison between floating point (solid) and double precision (dashed) calculations for the root mean squared velocity component w and $< u^2 + v^2 >^{1/2}$ vs time, with spatial resolution 32^3 and second order Adams-Bashford as time marching algorithm. The exponentially growing variable $< w^2 >$ collapses around time 0.19. This process is repeated as the simulation progresses, see *Calzavarini et al.* (2006).

Although the role of noise in PDEs or systems of ODEs is not fully understood, some particular questions have been answered already. For an ODE in one dimension, Scheutzow showed that if a deterministic equation has solutions that explode (i.e., goes to infinity) in a finite time for some initial condition then it necessarily explodes in a finite time with probability strictly greater than zero with an added white noise *Scheutzow* (1993). In later work *Scheutzow* (1995) an example of a two-dimensional system of ODEs was constructed where deterministic solutions explode in a finite time uniformly for all initial conditions, but for any arbitrarily small level of an additive white noise all solutions are nonexplosive with probability 1. We remark that the construction of that example is extremely complicated and the mechanisms at work are by no means clear. That is, it is not straightforward to identify the key properties of the system that lead to these results.

The main question of this work is to investigate whether arbitrary small level of an additive white noise can prevent trajectories from escaping to infinity. In this work

we present two planar deterministic dynamical systems with the property that almost every solution diverges to infinity exponentially in time. Both models are designed to capture key features of the homogenous convection problem, although no claim is made of a strict derivation. The models are then perturbed by a small magnitude Gaussian noise. The first model, similar to a 3-dimensional simplification of thermohaline convection by *Hughes and Proctor* (1990), "vacillates" [1] instead of producing statistical oscillations that do not diverge to infinity with probability one. The second model exhibits statistical oscillations. By using a theory of stochastic Lyapunov functions, see *Khasminskii* (1980), we show that adding arbitrary small white noise will produce a statistically steady stochastic evolution with an invariant probability measure. Next we develop an effective numerical method to find distribution of collapse times/values. Finally by asymptotic analysis of the corresponding first exit time problem we confirm that our numerical and analytical answers are consistent.

4.2 Topics from stochastic differential equations theory

This section is as a short digression to the topic of stochastic differential equations theory. The focus is on existence and uniqueness of an invariant measure for SDE's and the use of the stochastic Lyapunov function (SLF). We make use of a material by presented in spring of 2008 by Rafail Khasminskii at Wayne State University *Khasminskii* (2008). Mathematical proofs are omitted; interested reader can find them in the original book by *Khasminskii* (1980).

[1] vacillate = to swing indecisively from one course of action or opinion to another. In this context vacillation stands for seemingly random oscillations that wander off to infinity almost surely.

4.2.1 Regularity of solution

Consider a system of stochastic differential equations in $\mathbb{R}^n$

$$dx(t) = b(t, x(t))dt + \sigma(t, x(t))dW(t)\,, \qquad x(t_0) = x_0 \tag{4.1}$$

where $W(t)$ is a Brownian motion; and satisfying the regularity conditions for every x_1, x_2, $x \in \mathbb{R}^n$

$$|b(t, x)|^2 + ||\sigma(t, x)||^2 \leq K(1 + |x|^2)\,, \tag{4.2}$$

$$|b(t, x_2) - b(t, x_1)|^2 + ||\sigma(t, x_2) - \sigma(t, x_1)||^2 \leq K|x_2 - x_1|^2\,. \tag{4.3}$$

Then the solution exists at all times. Let

$$\tau_n = \inf\{t : |x(t)| > n\} \tag{4.4}$$

The sequence τ_n is increasing and thus its limit $\tau = \lim_{n\to\infty} \tau_n$ exists with probability one (it may be ∞). We say that a solution to (4.1) is *regular* if

$$P\{\tau < \infty\} = 0\,. \tag{4.5}$$

Define the stochastic Lyapunov operator L for any $V(t, x) \in C_1(\mathbb{R}_+) \cap C_2(\mathbb{R}^n)$ so that

$$LV(t, x) = \frac{\partial V}{\partial t}(t, x) + \sum_{i=1}^{n} b_i(t, x)\frac{\partial V}{\partial x_i}(t, x) + \frac{1}{2}\sum_{i,j=1}^{n} a_{i,j}(t, x)\frac{\partial^2 V}{\partial x_i \partial x_j}(t, x) \tag{4.6}$$

where $A(t, x) = \sigma(t, x)\sigma(t, x)^*$ is $n \times n$ covariance matrix with entries $a_{i,j}(t, x)$. The sufficient conditions for regularity of a solution of (4.1) are as follows.

Theorem IV.1. *Let $b(t, x)$ and $\sigma(t, x)$ be Lipschitz continuous and satisfy* (4.2), (4.3) *in $|x| < R$ for every $R > 0$. Assume there exists $V(t, x) \in C_1(\mathbb{R}_+) \cap C_2(\mathbb{R}^n)$ so*

that

1. $V(t,x) > 0$,

2. $LV(t,x) \leq CV(t,x)$ *for some* $C > 0$, *and*

3. $\inf_{t>0,|x|\geq R} V(t,x) = V_R \to \infty$ *as* $R \to \infty$.

Then any solution of stochastic system (4.1) *is regular.*

The proof consists of applying an Itô formula to a function $W(t,x) = V(t,x)e^{-C(t-t_0)}$ but with $t = \tau_n$. Taking expectation and the limit $n \to \infty$ together with condition 3 implies regularity. This is an extension of a similar statement for deterministic systems of ODE's.

Example IV.2. We show regularity of

$$\begin{cases} dx(t) = -x^3 dt + y^2 dt + dW_1(t) \\ dy(t) = x^2 dt - y^5 dt + dW_2(t) \end{cases} \tag{4.7}$$

Choose $V(x,y) = x^2 + y^2 + 1$. Then the inequality $ab \leq \frac{\gamma a^2}{2} + \frac{b^2}{2\gamma}$ for $\gamma = 1$ yields

$$\begin{aligned} LV(x,y) &= 2x(-x^3 + y^2) + 2y(x^2 - y^5) + 1 \\ &= -2x^4 - 2y^6 - 2xy^2 - 2x^2y + 1 \\ &\leq -2x^4 - 2y^6 + x^2 + y^4 + x^4 + y^2 + 1 \\ &\leq C(x^2 + y^2 + 1) \end{aligned}$$

for some $C > 0$. Therefore any solution is regular.

4.2.2 Recurrency

Let $\mathcal{D}$ be a compact domain in $\mathbb{R}^n$ and assume that the solution of (4.1) is regular and unique. Now define a time to reach domain $\mathcal{D}^c$

$$\tau_{\mathcal{D}} = \inf\{t : x(t) \in \mathcal{D}^c\} \tag{4.8}$$

There are two possible questions of interest: whether the time to reach $\partial\mathcal{D}$ is finite, and what is a solution of a Dirichlet problem $Lu = 0$ on a domain $\mathcal{D}$ with a boundary condition $u|_{\partial\mathcal{D}} = \phi(x)$. Both of these problems are connected with a recurrency of the stochastic dynamical system. The process $x(t)$, starting at point t_0 at $x_0 \in \mathcal{D}$ is *recurrent* with respect to $\mathcal{D}^c$ if $P^{t_0,x_0}\{\tau_{\mathcal{D}} < \infty\} = 1$ for all $x_0 \in \mathcal{D}$. Also, we call a process $x(t)$ starting at value t_0, x_0 *non-recurrent* if $P^{t_0,x_0}\{\tau_{\mathcal{D}} < \infty\} < 1$. The next theorem gives an answer to the first question again with the use of a stochastic Lyapunov function.

Theorem IV.3. *Suppose there is a Lyapunov function $V(t,x) \geq 0$ that satisfies*

1. *$V(t,x) \in C_1(\mathbb{R}_+) \cap C_2(\mathbb{R}^n)$ (smoothness), and*

2. *for all $x \in \mathcal{D}$ we have $LV(t,x) \leq -C$ where $C > 0$ is a constant*

then the process $x(t)$ starting at (t_0, x_0) is recurrent with respect to $\mathcal{D}^c$. And moreover

$$\mathbb{E}^{t_0,x_0}\tau_{\mathcal{D}} - t_0 \leq \frac{V(t_0,x_0)}{C}. \tag{4.9}$$

The proof requires use of Itô's lemma for $V(\tau_{\mathcal{D}}(T), X(\tau_{\mathcal{D}}(T)))$ where $\tau_{\mathcal{D}}(T) = \min\{\tau_{\mathcal{D}}, T\}$. One needs to take expectation and utilize property 2. To take the limit $T \to \infty$ one uses a Fatou lemma.

Example IV.4. What is $P\{\tau_{\mathbb{R}-\{0\}^c} < \infty\}$ for the one-dimensional Brownian motion $dx = d\xi$ with initial condition $x(0) = x_0 > 0$? In order to show that 1D Brownian

motion reaches origin at a finite time one needs to construct a SLF V such that

$$LV(t,x) = \frac{1}{2}\frac{\partial^2 V}{\partial x^2} \leq -C$$

for all $x > 0$. V must be both positive and concave for all $|x| > c \in \mathbb{R}$. Clearly, both conditions cannot be satisfied at the same time.

4.2.3 Conditions for stationary distribution

A solution of a time-homogeneous process

$$dx(t) = b(x(t))dt + \sigma(x(t))dW(t) \qquad x(t_0) = x_0\,, \tag{4.10}$$

is a Markov process, and therefore is characterized by a transition function $P(x,t,A) = P\{x^{t_0,x_0}(t) \in A\}$. A Markov process is stationary if there exists a measure $\hat{\mu}(A)$ such that

(i) $\hat{\mu}(A) = \int_{\mathbb{R}^n} \hat{\mu}(dx)P(x,t,A)$

(ii) $\int_{\mathbb{R}^n} \hat{\mu}(dx) = \hat{\mu}(\mathbb{R}^n) = 1$

If the condition (ii) is not satisfied but $\mu(\mathbb{R}^n) = \infty$ then the distribution is called an invariant measure but we omit the hat in that case. There exists a stationary density function $q(x) = \frac{d\hat{\mu}}{dx}$ corresponding to the distribution $\hat{\mu}(A)$ that satisfies a forward Kolmogorov equation, also called Fokker-Planck equation (FPE)

$$L^*(x)q(x) = 0\,, \quad \int_{\mathbb{R}^n} q(x)dx = 1$$

where $L^*(x)$ is the adjoint operator to the generator $L(x)$

$$L^*(x)q = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}(a_{i,j}(x), q(x)) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(b_i(x), q(x)) \tag{4.11}$$

The existence of the invariant measure and stationary initial distribution for a non-degenerate elliptic operators may be verified using the following.

Theorem IV.5. *Assume that*

(i) the matrix $A(x) = \sigma(x)\sigma^*(x)$ is positive definite (nondegenerate elliptic problem), and

(ii) the process $x(t)$ is recurrent with respect to some bounded domain.

Then the process (4.10) *has an invariant measure $\mu(A)$ (for arbitrary A). If in addition $\mathbb{E}^{t_0,x_0}(\tau_A) < \infty$ for all $x \in A^c$ then there exists a stationary distribution $\hat{\mu}(A)$: $\hat{\mu}(\mathbb{R}^n) = 1$.*

In the proof of the statement in one dimension the invariant measure is directly constructed. Since any point on the line can be reached with probability 1 one can construct cycles $0 \to L \to 0$ so that trajectory starting at 0 reaches value L at time τ_1 and then again value 0 at time τ_2 and so on. The value of $\mu(A)$ ($A \in [0, L]$) is then the fraction of time $\tau = \tau_1 + \tau_2$ that the trajectory spends in A. This measure can be shown to be invariant and if $\mathbb{E}(\tau) < \infty$ then it is also stationary. The proof in two dimensions requires a more elaborate construction of the cycles.

Remark IV.6. It is possible to show that positive recurrence of a non-degenerate elliptic problem implies uniqueness of the stationary distribution $\hat{\mu}$.

Example IV.7. Let us find the stationary distribution density for the process $x^{t_0,x_0}(t)$ described by the equation

$$dx(t) = (1 - 2x(t))dt + 2dW(t)\,, \quad x(0) = x_0\,.$$

The stochastic differential equation is a homogenous linear equation with a solution

$$x(t) = e^{-2t}\left[x_0 + \frac{e^{2t} - 1}{2} + 2\int_0^t e^{2s} dW(s)\right]$$

Mean $m(t)$ and standard deviation $\sigma(t)$ satisfy

$$\begin{aligned} m(t) &= x_0 e^{-2t} + \frac{1}{2}(1 - e^{-2t}) \\ \sigma^2(t) &= 4e^{-4t} \int_0^t e^{4s} ds = 1 - e^{-4t} \end{aligned}$$

Then the time-dependent probability density function is a gaussian $x(t) \sim \mathcal{N}(m(t), \sigma(t))$. The stationary distribution density for the problem is also a gaussian with the mean and standard deviation being the limits as $t \to \infty$ of $m(t)$ and $\sigma(t)$. Therefore

$$\hat{\mu}: \quad \hat{f}(x) = \frac{1}{\sqrt{2\pi}} exp\left[-\frac{(x - 1/2)^2}{2}\right]$$

4.2.4 Stochastic Lyapunov function theorem

In case the problem (4.10) is degenerate (for example if there is a noise term only in some of the equations) one needs to use a more general theorem. The following statement from *Khasminskii* (1980) applies to this general case.

Theorem IV.8 (SLF theorem)**.** *Let* (4.10) *be a time-homogenous problem where coefficients satisfy* (4.2) *and* (4.3) *in* $U_R = \{x \in \mathbb{R}^n : |x| \leq R\}$ *for every* R*. If there exists a function* $V(x) \in C_2(\mathbb{R}^n)$ *with the properties*

(i) $V(x) \geq 0$,

(ii) $\sup_{|x|>R} LV(x) = -A_R \to -\infty$ *as* $R \to \infty$.

and if the solution of (4.10) *is regular for at least one initial condition* $x_0 \in \mathbb{R}^n$ *then the solution of* (4.10) *is a stationary Markov process.*

In the proof of this theorem the stopped process $x(t)$ is considered only on U_R where the full statement is proved by taking a limit $R \to \infty$.

For a given SDE the existence of an invariant measure may be proved by finding an appropriate SLF $V(x)$. The function $V(x)$ may be considered as a stochastic

potential. In general, the search for a SLF is difficult especially if the underlying system has complicated dynamics. There is no general direct way or algorithm for finding a SLF but in some cases the system itself may hint on the form of the SLF. We will apply this theorem for a two dimensional SDE system that is degenerate in chapter 4.3.

4.3 Stochastic dynamical system approximations

4.3.1 Vaccilatory model – Model 1

Our aim is to construct a simple two-dimesional dynamical system that captures key features of the dynamics of HRBC, and that for arbitrary small stochastic perturbation changes its dynamics in nature. More complicated models with similar properties have been studied before. One of them is a three dimensional dynamical system for a thermohaline convection by *Hughes and Proctor* (1990) derived directly from fluid PDE's in a regime of small Rayleigh number. As a first step we propose a similar model but in two dimensions.

The key property to capture in the nonlinear model is the presence of exponentialy growing solutions. In order to mimic the HRBC, the model should contain at least one fixed point with one locally exponentialy growing and one locally exponentially decaying solution. The natural form of the linear part of the dynamical system is a saddle

$$\dot{x} = -x\,, \qquad \dot{y} = y\,. \tag{4.12}$$

In order to decrease the stability region of the exponentialy growing solutions (all of them approach y-axis) we add nonlinearity to the saddle model. The proposed

nonlinear model is

$$\dot{x} = -x(1-xy)\,, \tag{4.13}$$

$$\dot{y} = y - x^3\,. \tag{4.14}$$

The system is symmetric $(x, y) \to (-x, -y)$ and has three equilibria: saddle $(0, 0)$ and spiral sources $(\pm 1, \pm 1)$. The nullcline $xy = 1$ approaches the y-axis as y grows

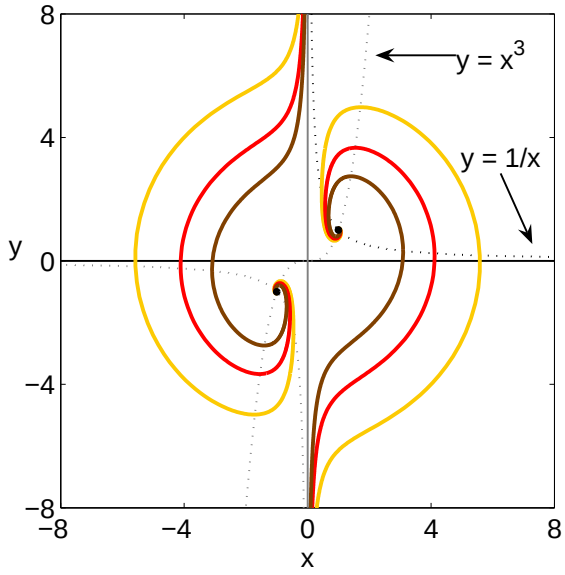

Figure 4.4: **Sample trajectories of the deterministic Model 1.** Nullclines of the map are $xy = 1$, $y = x^3$ and $x = 0$. Note that in the nonlinear system all solutions starting in quadrants I and III eventualy enter quadrants II and IV respectively.

and restricts the domain of attraction to the exponentialy growing solution. In other words, any exponentialy growing solution of this system is sensitive to a small perturbation in the x variable for y large. Therefore adding a white noise of a small magnitude to the x dynamics

$$dx = -x(1-xy)dt + \sigma dW\,, \tag{4.15}$$

$$dy = (y - x^3)dt\,. \tag{4.16}$$

should also result in destabilizing the exponentialy growing solution at some point. This leads to a stochastic dynamical system with a behavior similar to the fluid problem. To summarize, trajectories starting at almost any initial condition grow exponentialy to ∞ in the deterministic model. The only exceptions are when starting either directly at fixed points or at connections from/to fixed points. After adding a small noise in the x dynamics trajectories first climb up (or symmetricaly down) along the y-axis and after some random time the noise kicks them beyond the $xy = 1$ nullcline and throws them out of the region of attraction. This produces a collapse where y decreases and x temporarily increases but eventualy rapidly decreases. After this noise insensitive phase the trajectory starts to grow exponentialy again and the whole "cycle" repeats as on Fig. 4.5. There are two distinct parts of the trajectory: (1) noise sensitive part when y grows exponentialy fast and $|x| < \min\{1/|y|, c\}$ and (2) dominantly deterministic phase of trajectories $|x| > \min\{1/|y|, c\}$. The two different behaviors are separated by a jump – moment of exit from the atractive region. The jump occurs when $xy = 1$.

The repetition of exponential growth and subsequent collapse that we found in Model 1 is a feature that resembles the original RBC model. Our goal is to show that trajectories of this SDE model do not blow up to ∞ and that there exists an invariant probability measure for trajectories in this system. Given the "oscillatory" structure of the model this would show that statistical oscillations are a dominant feature in the model.

In reality, Model 1 only partialy satisfies the forementioned properties. It is true that starting from any given initial condition trajectory exits the noise-sensitive region in a random finite time with a finite expected value. Unfortunately, there does not exist any invariant measure for this problem and trajectories eventualy wander off to infinity. Surprisingly, the blow up occurs due to the properties of the dynamics in the noise insensitive region. Both claims will be justified in next two sections.

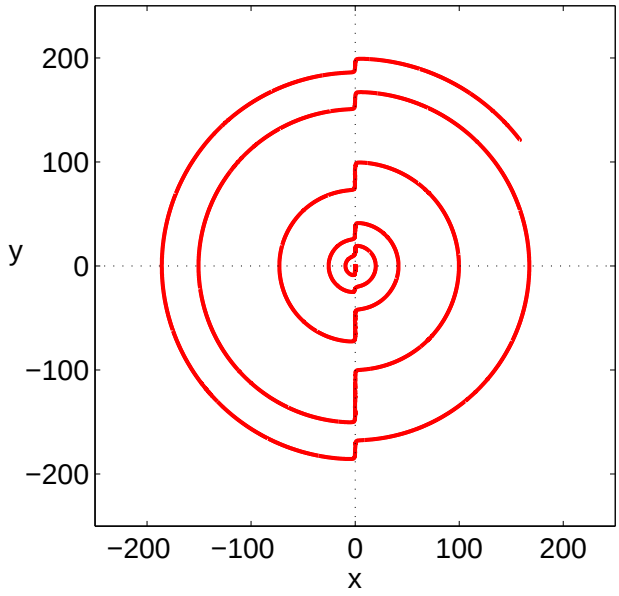

Figure 4.5: **Solutions of Model 1 with an added white noise of magnitude $\sigma = 1$.** Trajectories in a noisy region are climbing up/down the y axis. In the noise insensitive region trajectories seem to follow circles. Note the change of scale from the plot Fig. 4.4.

4.3.1.1 Noise sensitive region

As mentioned above trajectories can be in two regions, either noise sensitive region or deterministic region and these two are alternating in time. The focus of this section is to estimate the time it takes a trajectory to exit the first region. We show that starting in a noise sensitive region the solution exits the region on average by reaching some fixed time that depends on the noise magnitude σ. We prove the statement by showing that the system is recurrent (using theorem IV.3) with respect to some compact region around y-axis.

Theorem IV.9. *System* (4.15), (4.16) *is recurrent with respect to a region* $\Omega^c = \mathbb{R}^2 - \{(x, y) : |xy| \leq 1\,, |y| \geq 1\}$.

Proof. Choose a Lyapunov function to be a quadratic function

$$V(x, y) = a + bx - cx^2\,, \tag{4.17}$$

on $\Omega = \{(x,y) : |xy| \leq 1\,, |y| \geq 1\}$ where a, b and c are constants. Then for any $(x,y) \in \Omega$

$$\begin{aligned} LV(x,y) &= \partial_x V(x,y)\,(-x(1-xy)) + \frac{\sigma^2}{2}\partial_{x,x}V(x,y) && (4.18)\\ &= -(b-2cx)x(1-xy) - b\sigma^2 && (4.19)\\ &\leq -b\sigma^2 && (4.20) \end{aligned}$$

provided $0 < 2c < b$. Value of constant a is set so that $V(x,y) \geq 0$ on Ω and $c > 0$ is arbitrary. As an example one can take $(a,b,c) = (5,3,1)$. The inequality (D.1) implies that the average first exit time τ_Ω of a trajectory starting at $(x_0,y_0) \in \Omega$ is

$$\mathbb{E}(\tau_\Omega) \leq \frac{V(x_0,y_0)}{b\sigma^2} = \frac{a + bx_0 - cx_0^2}{b\sigma^2} \tag{4.21}$$

that does not depend on y_0. This gives a rather loose upper bound for the average exit time from Ω

$$\mathbb{E}(\tau_\Omega) \leq \frac{2}{\sigma^2}\,. \tag{4.22}$$

□

Although the upper bound does not use any information about the initial value y_0 it is consistent with our expectation that noise strength influences the time spent in the noise sensitive region.

4.3.1.2 Noise insensitive region

Numerical simulations of the noise perturbed Model 1 suggest that trajectories in the noise insensitive region move on a curve similar to a circle until they enter the noise sensitive part again. We will show that for Model 1 this region determines the properties of trajectories for large time. It is natural for us to keep track of the *exit value* – the height $|y_{\text{out}}|$ at which trajectory leaves the noise-sensitive region (when

it crosses the nullcline $xy = 1$) and the *entering value* – the height $|y_{\text{in}}|$ at which trajetory enters the noise-insensitive region next time (when $xy = -1$).

Definition IV.10. If the deterministic version of Model 1 has the property that for any $(x_{\text{out}}, y_{\text{out}}) = (1/y_{\text{out}}, y_{\text{out}})$ such that $|(x_{\text{out}}, y_{\text{out}})| > R$ for $R >> 1$ and

(i) $|y_{\text{out}}| > |y_{\text{in}}|$, then we call the map to be *contracting* in the noise insensitive region, or

(ii) $|y_{\text{out}}| < |y_{\text{in}}|$, then we call the map to be *expanding* in the noise insensitive region,

where $(x_{\text{in}}, y_{\text{in}})$ is the point of next entry of the slow region.

We will prove that Model 1 behaves as an expanding map in the deterministic region and thus there is no hope that there exists an invariant measure.

Theorem IV.11. *The deterministic system of ODE's* (4.13)-(4.14) *is an expanding map in the region defined by* $\Omega_+ = \mathbb{R}^2_{+x} - \{(x, y) : 0 < x \leq 1\,, |y| \leq 1/x\}$ *and symmetrically on* $\Omega_- = \mathbb{R}^2_{-x} - \{(x, y) : -1 \leq x < 0\,, |y| \leq -1/x\}$ *where* $\mathbb{R}^2_{+x} = \{(x, y) : x > 0\}$ *and* $\mathbb{R}^2_{-x} = \{(x, y) : x < 0\}$.

Based on the proof that may be found in Appendix A, the separation of time scales is responsible for the positive net growth of the distance from origin on the way from y_{out} to y_{in}. Note that by this argument the net growth is positive only if it is far away from the origin. Therefore any numerically found trajectory that starts close to the origin may "vaccilate" – statistically oscillate in some region around the origin and after reaching some point it will start to "monotonically" approach infinity. There are two destabilizing mechanisms in the model. On one hand the deterministic version of the map is expanding in the noise insensitive region. On the other hand the trajectory climbs away from the origin in the noisy region due to the direction of the slope field. Therefore all trajectories will eventualy blow up in a well controlled fashion yielding nonexistence of the invariant probability measure.

4.3.2 Resetting model – Model 2

For a model to have an invariant probability measure both noise sensitive and noise insensitive part of dynamics has to work in our favor. In the funnel-shaped region where stochastics has a dominant role the problem can be reformulated as a *first exit time* problem that can be analysed mathematically. On the other hand the noise insensitive part of the dynamics may be studied directly and must have certain contracting properties, i.e. trajectories that exit the noise dominant region at high value of y must enter it again at a smaller value of y. We design our second model so that trajectories revisit a small neighborhood of origin after every collapse (strong contraction towards origin). This way the dynamics effectively reduces to a repeated first exit value problem. The form of a modified model is

$$\dot{x} = -x(4 - xy + x^2)\,, \tag{4.23}$$

$$\dot{y} = \frac{y - x^3}{1 + x^2}\,. \tag{4.24}$$

Similarly to Model 1 this system has the symmetry $(x, y) \to (-x, -y)$ and has three equilibria: saddle $(0, 0)$ and oscillatory sources $\left(\pm\sqrt{\frac{1}{2} + \frac{\sqrt{17}}{2}}, \pm\sqrt{\frac{1}{2} + \frac{\sqrt{17}}{2}}^3\right)$. There exist heteroclinic connections from both sources to the origin and separatrices from ∞ to the origin. All trajectories eventualy escape to infinity at an exponential rate except when starting at fixed points, on the hetroclinic orbits or on the separatrices to the origin. After adding a small noise to the x-variable the behavior of the dynamics

$$dx = -x(4 - xy + x^2)dt + \sigma dW\,, \tag{4.25}$$

$$dy = \frac{y - x^3}{1 + x^2}dt\,. \tag{4.26}$$

consists of (1) slow phase in a noise sensitive region $|y| < |x| + 4/|x|$ and $|x| < c$; (2) jump from the slow phase; and (3) collapse of trajectory to a neighborhood of origin.

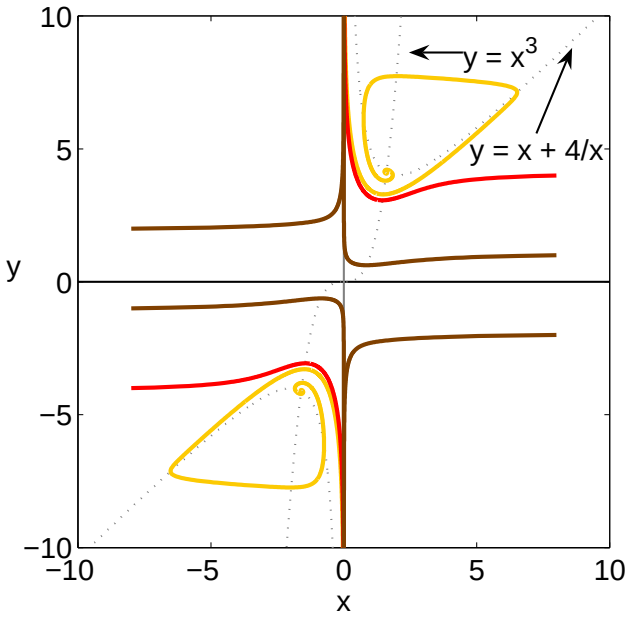

Figure 4.6: **Sample trajectories of the deterministic Model 2.** Nullclines are $4 - xy + x^2 = 0$, $x = 0$ and $y = x^3$.

4.3.2.1 Contraction of the map

To indicate the difference in the properties of Model 2 compared to Model 1, we examine the contraction properties of the chosen deterministic map, given by (4.23), (4.24), and show that any trajectory, starting outside of a narrow strip around the y-axis, will eventualy enter a fixed neighborhood of origin. Even though the full proof of statistical oscillations for this stochastic model (i.e. existence of an invariant probability measure) is provided in the next section, the essential property of the deterministic map is stated below. The dynamics of the noise-perturbed system naturaly divides into a noise-sensitive and -insensitive part. Here, we study the properties of the deterministic model in the region corresponding to the noise-insensitive region. In particular, we will decompose the state space into the following regions, as depicted on Fig. 4.8:

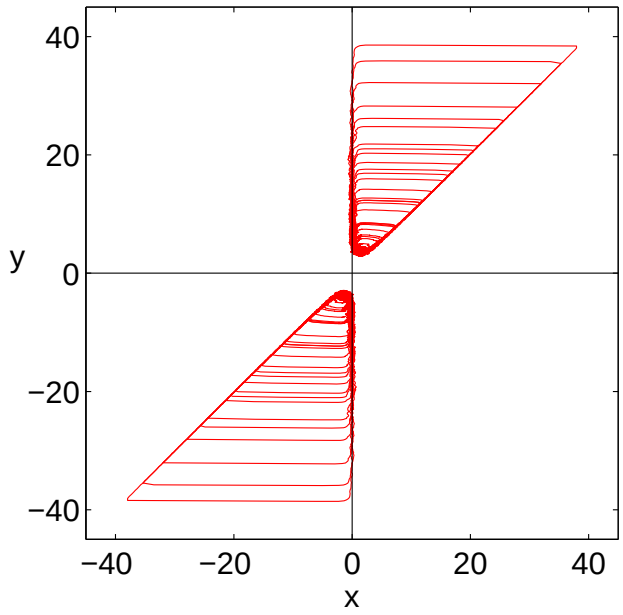

Figure 4.7: **Solutions of Model 2 with an added white noise of magnitude $\sigma = 1$.** In the noisy region trajectory stays close to the y axis. After crossing the approaching nullcline at a value y_{out} trajectory returns back to a neighborhood of the origin. Note the change of scale from Fig. 4.7.

- **Growing $x(t)$:** $A = \{(x,y) : \frac{d|x|}{dt}(x,y) > 0\} \cap \{(x,y) : |x| \geq 2\}$,
- **Decaying $x(t)$:** $B = \{(x,y) : \frac{d|x|}{dt}(x,y) < 0\} \cap \{(x,y) : x(y-x) > 0\}$,
- **Rectangular region around origin:** $R = \max\{|x|, |y|/2\} \leq 2$,
- The rest.

Theorem IV.12 (Contraction of Model 2.). *Let* $A = \{(x,y) : \frac{d|x|}{dt}(x,y) > 0\} \cap \{(x,y) : |x| \geq 2\}$ *be defined as in Fig. 4.8. The deterministic dynamical system* (4.23)-(4.24) *has a property that if* $(x(0), y(0)) \in A$ *then for some* $\tau < \infty$ $(x(\tau), y(\tau)) \in R$.

The proof of the claim is split into two simpler parts. In the first part we show that a trajectory, starting in region A will reach region B in finite time whereas in the second part we show that a trajectory in region B will enter rectangle R in a finite time. The proof is given in Appendix B.

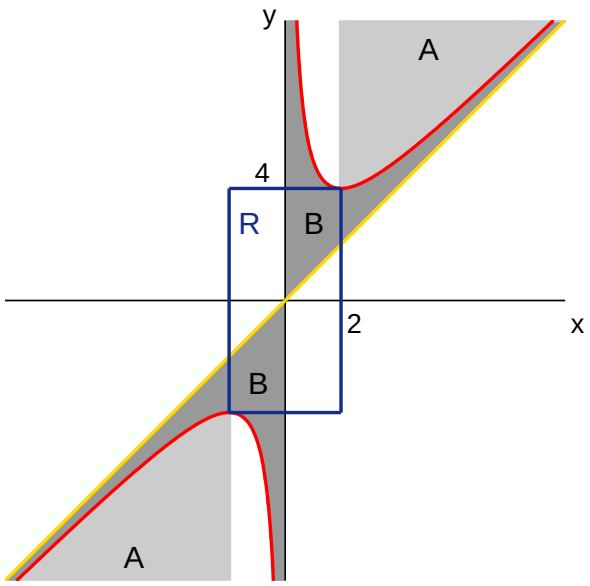

Figure 4.8: **Regions of the phase space:** A, B, R. Trajectory from A eventualy enters B and after a finite time it enters the rectangle R.

4.3.3 Existence of invariant measure in Model 2

By construction, the deterministic version of Model 2 does not have any Lyapunov function. This is because almost all its trajectories blow up exponentially fast. We must design SLF for the noise-perturbed system in such a way that the diffusion term in the generator dominates the rest. Clearly, trajectories are most sensitive to noise in the noise-sensitive part of region B down the y-axis and the key is to construct the SLF in this region. Moreover, we want to find SLF for arbitrarily small noise magnitude σ. But since the diffusion term in the generator is proportional to the magnitude of the noise itself, the form of the SLF needs to depend on σ as well. Note that SLF may be defined arbitrarily on any given compact region without violating the SLF theorem (since the key condition must be valid for $|x| > R$ as $R \to \infty$).

Our goal is to find a simple SLF function that respects the structure of the problem (symmetry, fast/slow phase). The construction consists of three steps.

1. Fast region: $\hat{V}(x,y) = (x-y)^2$ satisfies SLF condition on $\Omega_F = \{(x,y) \in \mathbb{R}^2 : |x| \geq C\,, |y| \geq 1\}$ where $C > 1$.

2. Slow region: $f(x,\sigma)\hat{V}(x,y)$ satisfies SLF condition on $\Omega_S = \{(x,y) \in \mathbb{R}^2 : |x| \leq C\}$ where $C > 1$ is the same as above. In addition, it can be patched with constant multiples of $\hat{V}(x,y)$ on Ω_F so that the result will be C^2. SLF will be defined separately for $y > 1$ and by parity for $y < -1$. It will depend on the noise strength σ.

3. All the remaining "rough" parts of the function ($|y| \leq 1$) will be patched together to produce an explicit C^2 function $V(x,y,\sigma)$ that satisfies SLF condition on $\mathbb{R}^2$ for arbitrarily small noise strength σ.

4.3.3.1 Fast region Ω_F

In the fast region $\Omega_F = \{(x,y) \in \mathbb{R}^2 : |x| \geq C\}$ where $C > 1$ is set, trajectories of Model 2 are not very sensitive to noise. We choose $\hat{V}(x,y)$ to be one of the simplest positive definite functions – quadratic, that satisfies SLF condition in this region. That is, $\hat{V}(x,y) = (x-y)^2$. Then

$$\begin{aligned} L\hat{V}(x,y) &= \left(\frac{2}{1+x^2} - 2x^2\right) y^2 && (4.27)\\ &+ \left(8x + 4x^3 - \frac{2x}{1+x^2} - \frac{2x^3}{1+x^2}\right) y && (4.28)\\ &+ \sigma^2 - 8x^2 - 2x^4 + \frac{2x^4}{1+x^2}\,. && (4.29) \end{aligned}$$

This nonnegative quadratic functional is concave down in variable y for $x \geq C > 1$. For technical pirposes we analyse two cases: $x \in [C, M]$ (bounded interval), and $x \in [M, \infty]$ for sufficiently large M. First observe that if $x \in [C, M]$ simple inequalities

imply that

$$L\hat{V}(x,y) \leq \left(\frac{2}{1+M^2} - 2\right) y^2 + 16M^3|y| + \sigma^2 - 10M^2 \tag{4.30}$$

and if $y \to \infty$ indeed $L\hat{V}(x,y) \to -\infty$ and the SLF condition is satisfied. Next we use concavity of $L\hat{V}(x,y)$ with respect to variable y to find its maximum on the remaining interval $[M,\infty)$ by setting $\frac{\partial L\hat{V}}{\partial y}(x,y^*) = 0$ and solving for y^*.

$$y^* = \frac{2x^5 + 5x^3 + 3x}{2x^4 + 2x^2 - 2} \sim x \text{ for } x >> 1 \tag{4.31}$$

The result y^* can be furthemore substituted into $\mathcal{L}\hat{V}(x,y)$.

$$\begin{aligned} L\hat{V}(x,y^*) &= \frac{-4x^{12} + O(x^8)}{2x^{10} + O(x^8)} && (4.32) \\ &\sim -2x^2 \text{ for } x >> 1 && (4.33) \end{aligned}$$

For $x \in [M,\infty)$ function $L\hat{V}(x,y)$ is bounded from above by some negative polynomial and therefore goes to $-\infty$ anytime $|\mathbf{x}| \to \infty$. In fact $L\hat{V}(x,y) \to \infty$ anytime when $|(x,y)| \to \infty$ since it is a polynomial in y as well. Therefore $\hat{V}(x,y)$ satisfies SLF condition on $[C,\infty)$. The SLF on the interval $[C,\infty)$ is a constant multiple of $\hat{V}(x,y)$ where the constant can be specified to satisfy smoothness properties.

4.3.3.2 Slow region Ω_S

Sensitivity of the dynamics on the noise magnitude is a dominant feature of the model in the slow phase. In the slow region $|x| \leq C > 1$ the function $\hat{V}(x,y) = C(x-y)^2$ does not satisfy properties of a SLF mainly because it does not even depend on the noise magnitude σ. In order to modify it we define SLF on a half-strip

region $|x| \leq C$, $y > 1$ (and by symmetry for $y < -1$) to have a form

$$V(x, y, \sigma) = f(x, \sigma)\hat{V}(x, y) = f(x, \sigma)(x - y)^2 \tag{4.34}$$

where a function $f \in C^2$ is to be specified and we set $C = 2$. Note that the smoothness property of a SLF requires us to impose boundary conditions on $f(x, \sigma)$ for $x = C = 2$.

$$f''(-2, \sigma) = f''(2, \sigma) = 0 \text{ and } f'(-2, \sigma) = f'(2, \sigma) = 0\,, \tag{4.35}$$

but also $f(-2, \sigma), f(2, \sigma) > 0$. After applying the generator of Model 1 to function $V(x, y, \sigma)$ we obtain

$$\begin{aligned} LV(x, y, \sigma) &= A_3(x, f, f', f'')y^3 + A_2(x, f, f', f'')y^2 && (4.36)\\ &+ A_1(x, f, f', f'')y + A_0(x, f, f', f'') && (4.37)\\ &= x^2y^3f' + y^2\left(-2x^2f - 4xf' - 3x^3f' + \frac{1}{2}\sigma^2 f'' + \frac{2f}{1 + x^2}\right) && (4.38)\\ &+ yA_1(x, f, f', f'') + A_0(x, f, f', f'') && (4.39) \end{aligned}$$

A necessary condition (for $V(x, y)$ to be a SLF) is that $yf' < 0$ implying that f needs to be decreasing in the positive half-strip. Since the deterministic version of "Model 2" blows up exponentially (and for that reason it does not have a stochastic Lyapunov function) it must be the stochastic term in the generator $\frac{1}{2}\sigma^2\partial_{xx}V(x, y, \sigma)$ that kills all the terms that blow up and makes function $LV(x, y, \sigma)$ approach $-\infty$. This leads to the observation that $V(x, y, \sigma)$ must be concave in x for $x \approx 0$ and also that its second derivative (with respect to x) must dominate positive terms in $LV(x, y, \sigma)$. The construction is divided into three steps, for more details, see Appendix C.

- Construct $f''(x)$ as a piecewise linear function that is equal to a large negative constant in a ε-neighborhood of $x = 0$. The simplest candidate is schematically plotted on Fig. 4.3.3.2. There are 5 parameters that we control: ε, K, L, X

and Y.

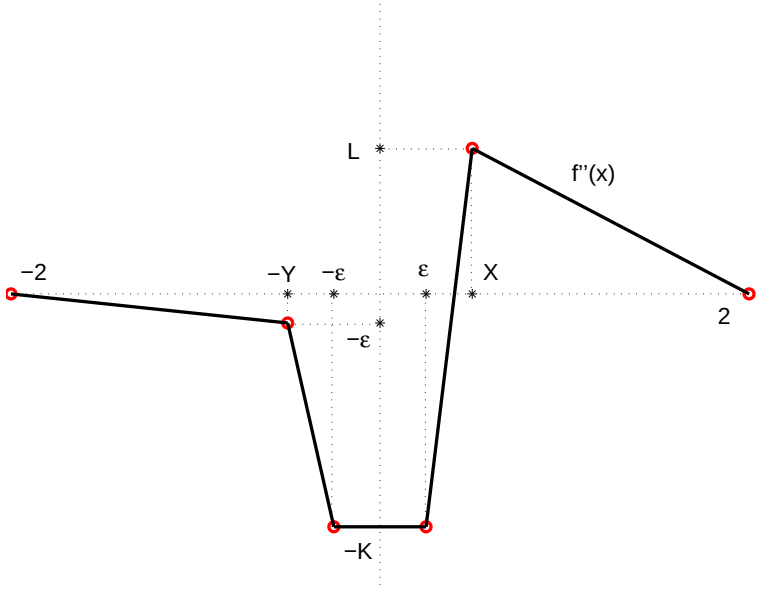

Figure 4.9: **Piecewise defined function $f''(x,\sigma)$ in the slow region.** Values of K and ε both depend exclusively on the magnitude of the noise strength σ. K is chosen big enough and ε small enough to satisfy SLF properties.

- Integrate the function $f''(x)$ twice to find $f(x)$ with the use of boundary conditions $f''(-2) = f''(2) = f'(-2) = f'(2) = 0$, $f(-2) = A > 0$ and $f(2) = B > 0$. Note that constants A and B are linked together by a fundamental theorem of calculus. As we verify by calculation in Appendix C this yields

$$L = \varepsilon + \frac{X+Y+2\varepsilon}{2-\varepsilon}, \tag{4.40}$$

$$A = B + \frac{1}{6}(-2\varepsilon^3 + K(2+Y)(X+Y) + \varepsilon^2(2K-X-Y) \tag{4.41}$$

$$+ \; 2\varepsilon(4+X+Y+K(2+X+Y))). \tag{4.42}$$

- Choose $B = \varepsilon$ and $X = Y = 2\varepsilon$. Lyapunov condition implies that the coefficient $A_2(x)$ needs to be negative. This sets a positive upper bound on ε as a function

of noise level σ. The SLF condition becomes

$$-6\varepsilon^3 + \varepsilon^2(8 + 22K - 6x - 9x^3) + \varepsilon(3K(8 + 6x + 9x^3)) + \tag{4.43}$$

$$+2(10 + 6x + 9x^3)) + 3K(6x^2 + 6x^4 - \sigma^2) < 0\,. \tag{4.44}$$

that leads to restrictions on K and ε based on the magnitude of $\sigma \leq 1$. The condition is satisfied if $K = \frac{1}{\varepsilon}$ and $\varepsilon(\sigma) \leq \frac{1}{16}\sigma^2$ (for example $\varepsilon = \frac{1}{16}\sigma^2$).

The coefficient $A_2(x)$ then by construction satisfies

$$A_2(x) < -c\sigma^2 K \text{ for any } x \in [-\varepsilon, \varepsilon)\,, \tag{4.45}$$

for some positive $c \in \mathbb{R}$ and in the limit $y \to \infty$,

$$V(x, y, \sigma) \to A_3(x)y^3 + A_2(x)y^2 \leq -(xy)^2|yf'(x)| - c\sigma^2 K y^2 \to -\infty \tag{4.46}$$

for any $x \in [-2, 2]$. Moreover the stochastic Lyapunov function connects in a $\mathcal{C}^2$ manner to $V(x, y) = A\hat{V}(x, y)$ for $x \leq -2$ and to $V(x, y) = B\hat{V}(x, y)$ for $x \geq 2$. Note that for $y < -1$ we need to define $V(x, y, \sigma) = V(-x, -y, \sigma)$ and $V(x, y) = B\hat{V}(x, y)$ for $x \leq -2$ and to $V(x, y) = A\hat{V}(x, y)$ for $x \geq 2$.

Below is the summary of all boundary conditions for $f(x,\sigma)$ in the upper half-strip ($|x| \leq 2$, $y \geq 1$)

$$\begin{aligned} f(-2,\sigma) &= A = -\varepsilon^3 + \frac{14K+4}{6}\varepsilon^2 + \frac{12K+14}{6}\varepsilon && (4.47)\\ f'(-2,\sigma) &= 0 && (4.48)\\ f''(-2,\sigma) &= 0 && (4.49)\\ f(2,\sigma) &= B = \varepsilon && (4.50)\\ f'(2,\sigma) &= 0 && (4.51)\\ f''(2,\sigma) &= 0 && (4.52) \end{aligned}$$

and by symmetry in the lower half-strip ($|x| \leq 2$, $y \leq -1$).

The last thing to do is to smoothly connect $A\hat{V}(x,y)$ with $B\hat{V}(x,y)$ to avoid a sharp transition of $V(x,y)$ along the x-axis. This can be done in a straightforward way by using a fifth-order polynomial as a smoothing function close to the x-axis.

4.3.3.3 Final patching

Here we define the function $V(x,y)$ on $|y| < 1$ for $x > 2$ (denoted by $+$) and on $|y| < 1$ for $x < -2$ (denoted by $-$) in order for it to have a smooth connection with the rest. We will take $V(x,y)$ of a form $V(x,y) = g_{\pm}(y)\hat{V}(x,y)$ where $g_{\pm}(y)$ satisfy boundary conditions (smoothness requirements)

$$\begin{aligned} g_+(1) &= \varepsilon \quad \text{and} \quad g_+(-1) = A && (4.53)\\ g_-(1) &= A \quad \text{and} \quad g_-(-1) = \varepsilon && (4.54) \end{aligned}$$

and also $g'_{\pm}(y) = g''_{\pm}(y) = 0$ for $y = \pm 1$. Polynomial function of the lowest order that satisfies all stated requirements is

$$g_{+}(y) = A - (A - \varepsilon)\frac{\int_{-1}^{y}(u+1)^2(u-1)^2 du}{\int_{-1}^{1}(u+1)^2(u-1)^2 du} \tag{4.55}$$

$$g_{-}(y) = \varepsilon + (A - \varepsilon)\frac{\int_{-1}^{y}(u+1)^2(u-1)^2 du}{\int_{-1}^{1}(u+1)^2(u-1)^2 du} \tag{4.56}$$

We need to make sure that as $|x| \to \infty$ then $LV(x,y) \to -\infty$ (y is bounded). The stochastic system in the described region has an approximate form

$$dx = (-x^3 + \mathcal{O}(x^2 y))dt + \sigma d\xi \tag{4.57}$$

$$dy = (-x + \mathcal{O}(1+y))dt \tag{4.58}$$

therefore

$$LV(x,y) = -2x^4 g(y) + \delta(x^3, y, g, g') \tag{4.59}$$

where $\delta(x^3, y, g, g') \sim \mathcal{O}(x^3)$ and y, $g(y)$ and $g'(y)$ are all bounded by a constant. Since the leading term is negative (and grows to $-\infty$ with x) $V(x,y)$ satisfies SLF condition. Now is the construction a SLF for Model 1 almost complete. The function $V(x,y)$ is a C^2 function in the following form, schematicaly plotted on Fig. C.1 where

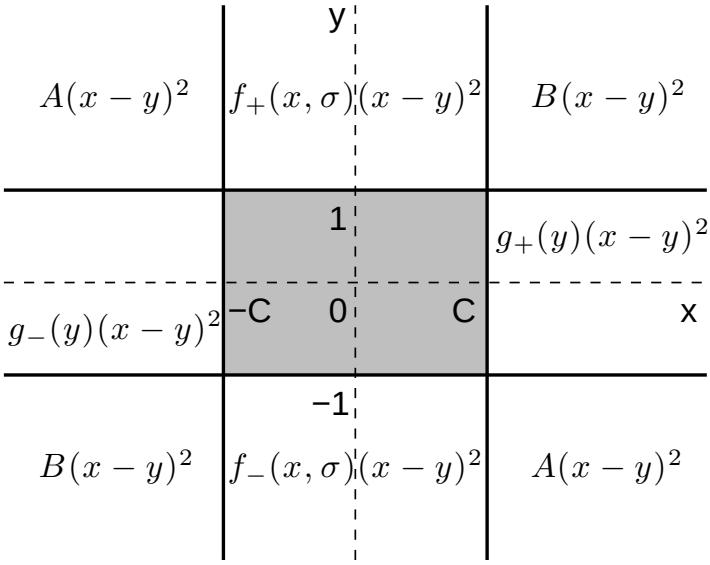

Figure 4.10: **Schematic picture of the stochastic Lyapunov function in $\mathbb{R}^2$.**

$\hat{V}(x,y) = (x-y)^2$, where $\varepsilon(\sigma)$ depends on the noise strength:

$$V(x,y) = \begin{cases} \varepsilon\hat{V}(x,y) & \text{if } x > 2 \text{ and } y > 1, \\ A\hat{V}(x,y) & \text{if } x < -2 \text{ and } y > 1, \\ \varepsilon\hat{V}(x,y) & \text{if } x < -2 \text{ and } y < -1, \\ A\hat{V}(x,y) & \text{if } x > 2 \text{ and } y < -1, \\ f(x,\varepsilon)\hat{V}(x,y) & \text{if } |x| \le 2 \text{ and } y \ge 1, \\ f(-x,\varepsilon)\hat{V}(x,y) & \text{if } |x| \le 2 \text{ and } y \le -1, \\ g_+(y)\hat{V}(x,y) & \text{if } x \ge 2 \text{ and } |y| \le 1, \\ g_-(y)\hat{V}(x,y) & \text{if } x \le -2 \text{ and } |y| \le 1, \\ \text{arbitrary} & \text{if } |x| < 2 \text{ and } |y| < 1 \end{cases} \tag{4.60}$$

Remark IV.13. In order to satisfy $V(x,y) > 0$ for all $(x,y) \in \mathbb{R}^2$ we add an arbitrary

positive constant to the function V. The properties of $LV(x, y)$ will not depend on the choice of this constant.

Remark IV.14. For simplicity the stochastic Lyapunov function was based on a function $V(x, y) = (x - y)^2$. This polynomial does not allways grow to infinity as $|(x, y)| \to \infty$. One can modify the construction and base the stochastic Lyapunov function on a different polynomial $V'(x, y) = (x - y)^2 + y^2$. The growth condition will then be satisfied and the whole construction will be completely analogous (only a few constants will change).

Remark IV.15. Regularity of the process $(x(t), y(t))$ is trivially satisfied for every initial condition. This is due to the fact that Theorem (IV.3) requires finding a stochastic Lyapunov function with a weaker condition that we have already constructed. Process is therefore regular for all initial conditions.

4.4 First exit time problem

The dynamics of (4.23), (4.24) is very sensitive to noise in the slow region $\{|y| \leq |x| + 1/|x|\} \cap \{|x| \leq c\}$ whereas in the rest of the phase space ("weak" region) noise does not alter the trajectory considerably. Noise-driven collapse together with the system's dynamical properties force trajectories to repeatedly visit a small rectangle around the origin. Therefore we can say that the system has short memory, e.g. it effectively resets once it enters the neighborhood of origin. The trajectory may also travel very far away along the y-axis (in the slow region) before returning back towards origin. The dynamics is therefore dominated by the slow noisy region and we can analyse the dynamical system in the whole real plane by a *first exit time* problem in the noisy region with a carefuly chosen initial condition. However, it seems that the initial condition does not have major effect on the outcome and for simplicity we choose it to be fixed $(x, y) = (0, 1)$. Alternatively, one may use a random initial

condition (say, $y = 1$ and $x \in [-4, 4]$ according some distribution). Using Monte-Carlo simulation of multiple trajectory paths we can find statistical properties of the first hitting time/value and a time-depended distribution of position $x(t)$ conditioned on not having yet hit the boundary (conditional distribution).

Formally we can write the first exit value problem as the first time in which the dynamical system

$$x\text{-dynamics:} \qquad dx = -x(4 - xy + x^2)dt + \sigma dW \tag{4.61}$$

$$y\text{-dynamics:} \qquad dy = \frac{y - x^3}{1 + x^2}dt \tag{4.62}$$

with initial conditions

$$\text{initial conditions:} \qquad x(0) = x_0\,, \quad y(0) = 1\,, \tag{4.63}$$

hits the nullcline $4 - xy + x^2 = 0$

$$\text{exit condition:} \qquad 0 = 4 - xy + x^2\,. \tag{4.64}$$

In order to find statistical properties of the first exit value in the asymptotic regime $y \to \infty$ (and also $\sigma \to \infty$) we can make two approximations. We first set initial condition to be $x_0 = 0$ and let the system forget this initial condition before the statistics is measured. In the second step we approximate the x and y dynamics for $x \approx 0$ by

$$dx = -x(4 - xy)dt + \sigma dW\,, \tag{4.65}$$

$$dy = ydt\,. \tag{4.66}$$

The dynamics is then reduced to a one-dimensional non-autonomous SDE

$$dx = -x(4 - xe^t)dt + \sigma dW\,. \tag{4.67}$$

The nonlinear first exit time problem with constraints 4.67, 4.63 and 4.64 will be studied both numericaly using an effective Monte Carlo simulation and analyticaly in the regime $t \to \infty$ (or $y \to \infty$) using an asymptotic analysis in terms of the corresponding Fokker-Planck equation.

4.4.1 Monte Carlo method with recycling (MCR)

Standard Monte Carlo method (MC) is a simple numerical method that may be used to compute properties of the first exit value problem above. Unfortunately we are interested in large y (respectively long time) limit. The biggest disadvantage of MC in this context is the fact that during most of the trials trajectory does not even reach the specified time/value because it exits the noise-sensitive region before that time. In addition, the region is narrowing as time progresses at an exponential rate and with a set value of noise level σ trajectories are progressively less likely to stay inside the region as time goes on (also the step size must be finer as the region gets narrower). The probability that a trajectory stays in the region is very small for long time/large y and the events are therefore called *rare*. It is extremely computationaly expensive to answer even the simpliest possible questions such as: "What does the survival probability $S(t)$ look like as $t \to \infty$ (or $y \to \infty$)?"

The easiest solution is to design a modified MC that reduces the number of trajectories that exit given region, i.e., to use an importance sampling method. Such methods have been useful in applications, for example to find harmonic measure for critical percolation Ising clusters, see *Adams et al.* (2008). We develop a new computationaly effective method, suitable for the first exit value problem.

The idea behind the method is to "recycle" otherwise irrelevant trajectories. Therefore we will call it MCR. In terms of the implementation, MCR will simultaneously keep track of many trajectories. Once a trajectory hits the boundary of the region it is replaced by a different, randomly chosen trajectory from the sample. Consequently the survival probability $S(t)$ is updated and the process continues. Numerical algorithm of the MCR is

1. **Initialization:** Start with a population of N trajectories at $t^{(0)} = 0$ with the same initial condition

$$(x_n^{(0)}, y_n^{(0)}) = (x, 1) \tag{4.68}$$

   where the initial value x is either randomly chosen from some distribution or $x = 0$. Set the inital values of the survival probability $S(t)$ at time t and the probability $P(t)$ of having hit the boundary by time t

$$S(t^{(0)}) = 1 - P(t^{(0)}) = 1\,. \tag{4.69}$$

2. **Time step:** For all trajectories $n \in \{1, \ldots, N\}$ use an Euler scheme to move forward in time

$$\begin{aligned} t^{(i+1)} &\rightarrow t^{(i)} + \Delta t \\ x_n^{(i+1)} &= x_n^{(i)} - x_n^{(i)}(4 - x_n^{(i)} y_n^{(i)})\Delta t + \sigma\sqrt{\Delta t} W^{(i)} \\ y_n^{(i+1)} &= y_n^{(i)} + y_n^{(i)} \Delta t \end{aligned}$$

   with adjustable Δt.

3. **Exit of trajectory:** If $x_n^{(i)} \geq 4e^{-t^{(i)}}$, replace trajectory $(x_n^{(i)}, y_n^{(i)})$ by an other, randomly chosen from the remaining $N - 1$ trajectories. Update survival probability $S(t^{(i)} + \Delta t) = \frac{N-1}{N} S(t^{(i)})$.

4. **Loop:** Repeat steps (2)-(3) until the terminal time.

The MCR algorithm is able to compute survival probabilities of very small magnitudes. It is important to choose a number of sample trajectories N large enough in order to visit majority of possible outcomes at each time. Also, the time step needs to be adjusted in such a way that the probability of two trajectories exiting at the same time step is negligible. In practice we need to decrease the time step at the same rate as the nullcline approaches the y-axis. This happens at an exponential rate and the calculation therefore exponentialy slows down in time.

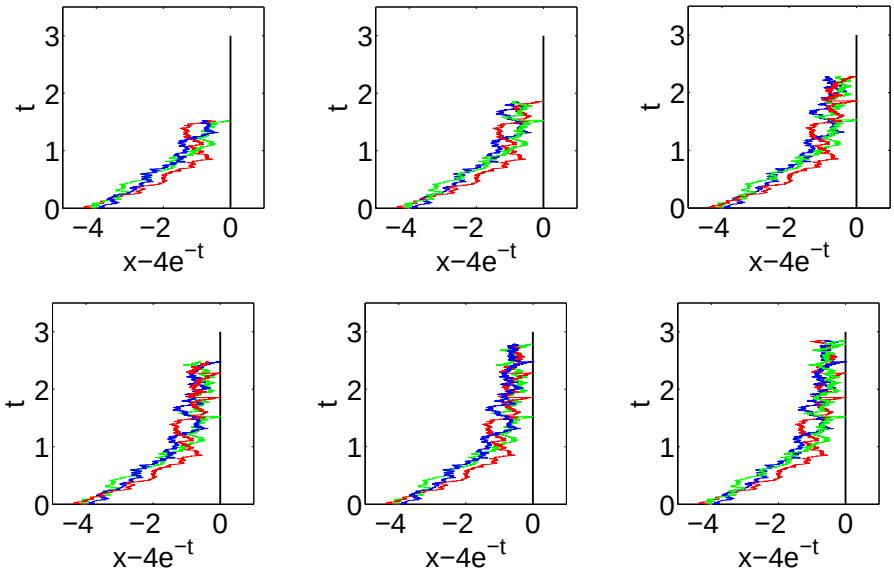

Figure 4.11: **Monte Carlo method with recycling for three sample trajectories.** Simulation is plotted at six different times: at a time of first, second, ... up to the sixth hit of the boundary with $\sigma = 1$. Each time a boundary is hit trajectory is replaced by a different one from the surviving set.

4.4.1.1 Application to first exit time problem

In order to turn the first exit value problem with an exponentialy fast moving boundary to a simpler problem with a constant boundary the transformation $\bar{x} =$

$x - 4e^{-t}$ is used. The first exit time problem constraints then become

$$x\text{-dynamics:} \qquad d\bar{x} = (16e^{-t} + 7\bar{x} + \bar{x}^2 e^t)dt + \sigma dW \tag{4.70}$$

$$\text{init. condition:} \qquad \bar{x}(0) = -4\,, \tag{4.71}$$

$$\text{exit condition:} \qquad \bar{x}(t) = 0\,. \tag{4.72}$$

First we illustrate the algorithm by applying it to the first exit value problem (4.70), (4.71), (4.72) for only three trajectories. In the Fig. 4.11 one can observe the numerical simulation until terminal time $T = 3$ at a time of first, second, ..., up to sixth hit of the boundary and at the terminal time.

The algorithm provides information about both the decay of the survival probability and about the distribution of trajectories that have not hit the boundary yet. For the numerical analysis we choose terminal time $T = 10$ (corresponding to a value $y = 4e^{10} \approx 88106$) and number of sample trajectories $N = 2000$. By numerically applying the MCR algorithm to the first exit time problem with constraints (4.63), (4.64) and (4.67) we numerically justify the following two hypotheses in the asymptotic regime $t \to \infty$:

- The survival probability $S(t)$ has a double exponential asymptotic growth: $S(t) \sim e^{-k\sigma^\delta e^{\alpha t}}$ with appropriate k, δ and α.
- The mean $m(t)$ of the conditional distribution of $\bar{x}(t)$ asymptotically approaches exponential function of the form: $m(t) = Be^{-\beta t}$

The survival probability is naturally a decreasing function of time and y. The simulation results on Fig. 4.12 suggest that $\log|\log S(t)|$ is approaching a linear function. Results on Fig. 4.12 for different values of noise parameter σ imply that the slope does not depend on σ whereas the shift depends linearly on some power of σ. This

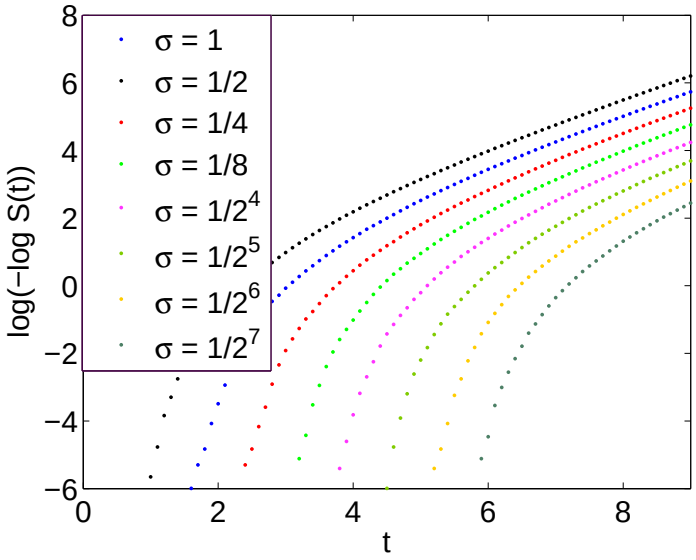

Figure 4.12: **Survival probability $S(t)$ approaches a double exponential function as $t \to \infty$.** This information is contained in the graph of $\log(-\log S(t))$ that appears to converge to a linear function for any σ. The slope of the function corresponds to α and vertical shift corresponds to a linear function $logk + \delta log(\sigma)$.

numericaly confirms the hypothesis that $S(t)$ satisfies

$$S(t) \sim e^{-k\sigma^{\delta}e^{\alpha t}} \tag{4.73}$$

in the large time regime. To estimate values of coefficients α, k, δ for different σ we fit numerical data to a function of the form (4.73). Best fit in L^2 space for this rule is a solution of the minimization problem

$$(k, a) = \operatorname{argmin} \sum_{i=i^*}^{M} (\log(-\log S(i)) + \log k - \delta \log \sigma - \alpha t^{(i)})^2 \,. \tag{4.74}$$

Note that the fitting function is chosen to match the values only in the asymptotic region, i.e. after some time (specified by index i^*) For $\sigma = 1$, $\frac{1}{2}$ and $\frac{1}{3}$ the minimization problem yields Results are expected to be less accurate for smaller σ because the asymptotic region is reached at larger times. Fig. 4.12 agrees with the assumed

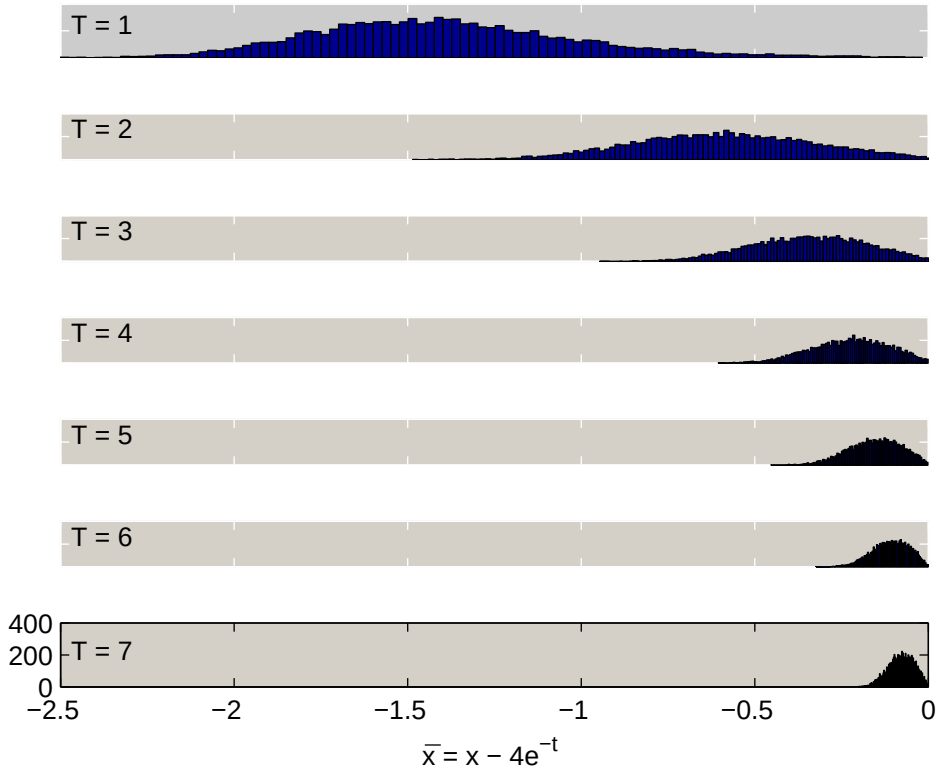

Figure 4.13: **Histogram of trajectories in the MCR simulation at time** $T = 1, 2, \ldots, 8$**.** Distribution is being squeezed towards the origin as time increases.

asymptotic form for $S(t)$, with $\alpha \approx \delta \approx 2/3$.

Next we explore the properties of the conditional distribution of $\bar{x}$ (shifted x) at time t conditioned on not having hit the boundary yet. The distribution is represented in the MCR simulation by an unconditional distribution of trajectories at a given time. The Fig. 4.13 shows histograms of numericaly computed random trajectories at times $t = 1, 2, \ldots, 7$. Clearly, for $t = 0$ the distribution is a delta function positioned at

Constant	α	δ	k
$\sigma = \mathbf{1}$	0.6659	0.6434	0.0373
$\sigma = \mathbf{1/2}$	0.6562	0.6434	0.0555
$\sigma = \mathbf{1/4}$	0.6710	0.6329	0.1100

Table 4.1: **Fitted parameter values for the decay of S(t).**

$\bar{x} = -4$. The distribution function is being squeezed towards $\bar{x} = 0$ as time progresses. The rate at which the contraction happens can be estimated from the time evolution of the mean $\bar{x}$ position. The mean $m(t)$ for a fixed boundary exit value problem is a

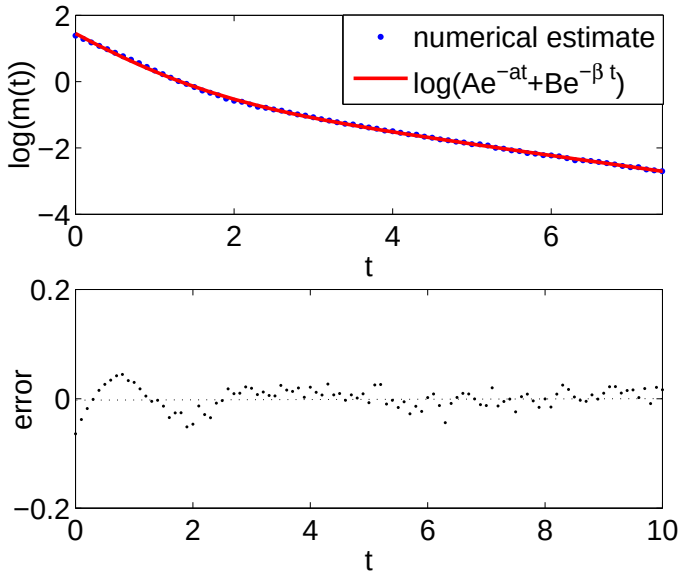

Figure 4.14: **Mean of the conditional distribution of $z(t)$ as a function of time.** Function can be closely approximated by a sum of two exponential functions with a random nonsystematic error.

monotone negative function that increases to 0 exponentialy as the time progresses. There is an initial transition from one rate to a different exponential rate of decay. It is possible to accurately fit profile of $m(t)$ to a function of the form

$$\log m(t) = \log(Ae^{-at} + Be^{-\beta t}) \tag{4.75}$$

by solving a nonlinear minimization problem

$$(A, B, a, \beta) = \operatorname{argmin} \sum_{i=1}^{N} (\log m(i) - \log(Ae^{-at^{(i)}} + Be^{-\beta t^{(i)}}))^2 \tag{4.76}$$

where the criterion of good fit is the L^2 norm. The fit remarkably resembles the original function and the error does not seem to be systematic (at least not in the

Constant	A	B	a	β
$\sigma = \mathbf{1}$	3.3683	0.8161	1.4744	0.3372

Table 4.2: **Fitted parameter values for the particle mean m(t).**

asymptotic regime). Also the value of β coefficient is very close to the $\frac{1}{3}$. Therefore we hypothesize that asymptoticaly

$$m(t) \sim e^{-\beta t} \quad \text{where} \quad \beta = \frac{1}{3}\,. \tag{4.77}$$

A natural way to confirm that (4.73) and (4.77) are true is to rescale the histograms Fig. 4.13 in a following way

$$z = e^{\beta t}\bar{x} \tag{4.78}$$

$$u(t,\bar{x}) = e^{-k\sigma^{\delta}e^{\alpha t}} v(t,z) \tag{4.79}$$

where $u(t,\bar{x})$ was the unconditional pdf on the interval $(-\infty, 0)$ and $v(t,z)$ is the conditional pdf after the normalization. If both hypotheses are correct then as time grows to ∞ density function $v(t,z)$ should converge to a limiting profile. And indeed, the rescaled probability density function of the conditional distribution of $v(t,z)$ at time $t = 3, \ldots t = 7$, plotted on Fig. 4.15, numerically confirms the hypothesis.

4.4.2 Asymptotic Methods for the first exit time problem

As shown in \$4.3.3 trajectories of this system exhibit statistical oscillations due to the existence of the probabilistic invariant measure. No matter how small the noise magnitude σ is, the dynamics in the noise sensitive region is dominated by the random term. Here, the SDE system may be reduced to a simpler, 1-dimensional stochastic differential equation due to the fact that $x << 1$ and $y >> 1$. The problem is formulated as a reduced first passage time problem (4.63), (4.64), (4.67) with the absorbing boundary at $x(t)y_0 = 4e^{-t}$ (transition between the noise regulated and

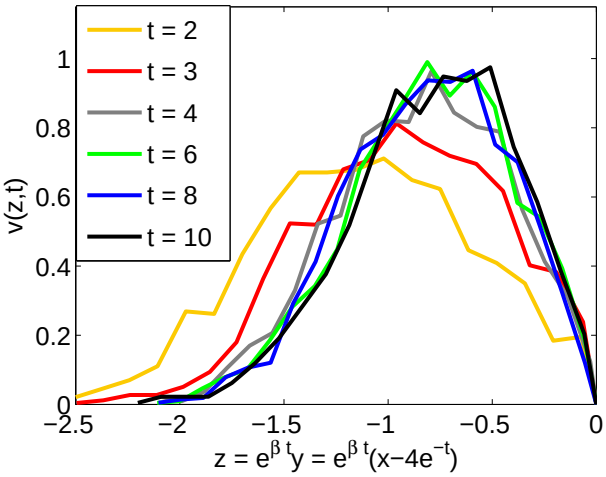

Figure 4.15: **Rescaled probability density function** $v(t,z)$ **at times** $T = 3,4,\ldots,7$**.** Functions converge to a time invariant pdf that captures behavior in the asymptotic regime $t = \infty$.

deterministic regions). After choosing $y_0 = 1$ the problem becomes

$$dx = (-4x + x^2 e^t)dt + \sigma dW_t\,, \tag{4.80}$$

$$x(t) = 4e^{-t} \qquad \text{absorbing boundary,} \tag{4.81}$$

This problem has been studied numerically using the MCR method in \$4.4.1 and the suitable scaling laws have been proposed and numerically verified. Next, we will use these to arrive at a better theoretical understanding of the problem. The conditional probability density function $u(t,x)$ of $x(t)$ that has not yet reached the boundary satisfies the FPE with a moving boundary

$$u_t = \left((4x - x^2 e^t)u\right)_x + \frac{\sigma^2}{2} u_{xx}\,, \tag{4.82}$$

$$0 = u(t, 4e^{-t})\,. \tag{4.83}$$

The problem can be transformed to a stationary boundary problem by a linear in x transformation $\bar{x} = x - 4e^{-t}$. Specifically

$$u_t = u_t + u_{\bar{x}}\bar{x}_t = u_t + 4e^{-t}u_{\bar{x}}\,, \tag{4.84}$$

$$u_{\bar{x}} = u_x\,, \tag{4.85}$$

$$u_{\bar{x}\bar{x}} = u_{xx}\,. \tag{4.86}$$

The FPE with a fixed boundary at 0 has the form

$$u_t = -(4 + 2\bar{x}e^t)u - (4\bar{x} + \bar{x}^2e^t + 4e^{-t})u_{\bar{x}} + \frac{\sigma^2}{2}u_{\bar{x}\bar{x}}\,, \tag{4.87}$$

$$0 = u(t,0)\,. \tag{4.88}$$

Based on numerical results the survival probability $S(t)$ has a simple form in the asymptotic limit $t \to \infty$,

$$S(t) = \int_{-\infty}^{0} u(t,\bar{x})d\bar{x} \sim e^{-k\sigma^\delta e^{\alpha t}} \tag{4.89}$$

with k, δ and α are suitable growth constants. Also, numerics suggests that as $t \to \infty$ the profile of $u(t,\bar{x})$ is contracting at an exponential rate

$$u(t,\bar{x}) \sim e^{\beta t}u(t,e^{\beta t}\bar{x}) \tag{4.90}$$

Due to the above information ((4.73) and (4.77)) we formulate an ansatz for the asymptotic form of $u(t,z)$ for $t \to \infty$.

$$u(t,\bar{x}) = e^{-k\sigma^\delta e^{\alpha t}}e^{\beta t}v(t,e^{\beta t}\bar{x}) = e^{\beta t - k\sigma^\delta e^{\alpha t}}v(t,e^{\beta t}\bar{x}) \tag{4.91}$$

Note that

$$S(t) = \int_{-\infty}^{0} u(t,\bar{x})d\bar{x} = e^{-k\sigma^\delta e^{\alpha t}} \int_{-\infty}^{0} e^{\beta t} v(t, e^{\beta t}\bar{x})d\bar{x} = e^{-\sigma^k e^{\alpha t}} \int_{-\infty}^{0} v(t,\hat{x})d\hat{x} \tag{4.92}$$

Therefore $v(t,\bar{x})$ is the conditional probability density function of $\bar{x}$ in the noise dominated region (it is normalized to give $\int_{-\infty}^{0} v(t,\bar{x}) = 1$). If we denote $z = e^{\beta t}\bar{x}$ then

$$\begin{aligned} u_t &= e^{\beta t - k\sigma^\delta e^{\alpha t}} \left((\beta - \alpha k \sigma^\delta e^{\alpha t})v + v_t + \beta z v_z \right) , & (4.93)\\ u_{\bar{x}} &= e^{\beta t - k\sigma^\delta e^{\alpha t}} (e^{\beta t} v_z) , & (4.94)\\ u_{\bar{x}\bar{x}} &= e^{\beta t - k\sigma^\delta e^{\alpha t}} (e^{2\beta t} v_{zz}) . & (4.95) \end{aligned}$$

Subtitution into the Fokker–Planck equation gives

$$v_t = \left(-\beta + \alpha k \sigma^\delta e^{\alpha t} \right) v - \left((4z + z^2 e^{(1-\beta)t} + 4e^{(\beta-1)t})v \right)_z + \frac{\sigma^2}{2} e^{2\beta t} v_{zz} . \tag{4.96}$$

For t large we can use the numerical evidence that $\alpha, \beta \in (0,1)$ to obtain effective long time behavior as

$$v_t = e^{\alpha t}(\alpha k \sigma^\delta v) - e^{(1-\beta)t}(z^2 v)_z + e^{2\beta t}\frac{\sigma^2}{2} v_{zz} . \tag{4.97}$$

The stationary distribution, if it exists, has the property that $\partial_t = 0$. But the right hand side of the PDE also depends on time and as $t \to \infty$, so the terms with the largest exponential growth rate dominate. In fact, we may rewrite (4.97) using $r = \max\{\alpha, 1-\beta, 2\beta\}$ as

$$e^{-rt} v_t = e^{-(r-\alpha)t}(\alpha k \sigma^\delta v) - e^{-(r-1+\beta)t}(z^2 v)_z + e^{-(r-2\beta)t}\frac{\sigma^2}{2} v_{zz} . \tag{4.98}$$

If the stationary distribution exists, then it satisfies the ODE where the left hand side of (4.98) is replaced by 0 and right hand side only contains terms without exponential factors. The existence of invariant measure is then closely connected to the magnitude of growth constants α and β. There are three possible cases for constant β: (i) $\beta > \frac{1}{3}$; (ii) $\beta < \frac{1}{3}$; and (iii) $\beta = \frac{1}{3}$ and then after elimination of conflicting cases there are two possible cases for constant α: (1) $2\beta \neq \alpha$ and (2) $2\beta = \alpha$. The cases eliminate in the following way:

- If $\beta > \frac{1}{3}$ the nonlinear term vanishes. There are no solutions of the linear ODE that are both normalizable and go through the origin.

- If $\beta < \frac{1}{3}$ the diffusion term is negligible and there is no solution that is both normalizable and that goes through the origin.

$\Rightarrow$ For the invariant distribution to exists it must be true that $\beta = \frac{1}{3}$.

- If $2\beta \neq \alpha$ the solution that satisfies boundary conditions is not normalizable.

$\Rightarrow$ For the invariant distribution to exists it must be true that $\alpha = \frac{2}{3}$, $\beta = \frac{1}{3}$.

The only way how we can obtain a normalizable solution satisfying boundary conditions $v(z = 0) = 0$ and $\lim_{z\to-\infty} v(z) = 0$ is that all exponential growth coefficients in the FPE are the same. That implies $\alpha = \frac{2}{3}$ and $\beta = \frac{1}{3}$. This is consistent with numerically estimated values of constants $\alpha \approx 0.725$ and $\beta = 0.337$. The problem becomes

$$0 = \alpha k \sigma^{\delta} v - (z^2 v)_z + \frac{\sigma^2}{2} v_{zz} \,. \tag{4.99}$$

This is a linear eigenvalue problem with an eigenvalue $\lambda = \alpha k \sigma^{\delta} = \frac{2}{3} k \sigma^{\delta}$ that we can compute numerically. We are interested in a positive normalizable solution, e.g. the eigenfunction (or the ground state), corresponding to the smallest positive eigenvalue of the related operator. There are two ways of finding the proper eigenvalue

- using a numerical method to satisfy boundary conditions $v(0) = v(\infty) = 0$,
- solve initial value problem with boundary conditions at $z = 0$ and find λ such that $v(\infty) = 0$ is satisfied (simplified shooting algorithm)

To find the boundary condition on $v_z(0)$ we can assume that $\int_{-\infty}^{0} v(z)dz = 1$ (by linearity). Then integration of the ODE results in

$$0 = \lambda \int_{-\infty}^{0} v(z)dz + \int_{-\infty}^{0} [-z^2 v(z) + \frac{\sigma^2}{2} v_z(z)]_z dz \tag{4.100}$$

$$v_z(0) = -\frac{2\lambda}{\sigma^2} = -\frac{2\alpha k}{\sigma^{2-\delta}} \tag{4.101}$$

There is only one eigenvalue for which the eigenfunction is positive and integrable at the same time.

Normal form of the Sturm-Liouville problem (S-L) problem. We first transform the eigenvalue problem (4.99) to a self-adjoint S-L problem by an exponential transformation

$$v = we^{\frac{z^3}{3\sigma^2}} \tag{4.102}$$

$$v_z = \left(w_z + \frac{1}{\sigma^2} z^2 w\right) e^{\frac{z^3}{3\sigma^2}} \tag{4.103}$$

$$v_{zz} = \left(w_{zz} + \frac{2}{\sigma^2} z^2 w_z + \frac{2}{\sigma^2} z^2 w + \frac{1}{\sigma^4} z^4 w_z\right) e^{\frac{z^3}{3\sigma^2}} \tag{4.104}$$

The term with w_z vanishes and we obtain a singular Sturm-Liouville problem on interval $(-\infty, 0]$

$$\alpha k \sigma^{\delta} w = -\frac{\sigma^2}{2} w_{zz} + \left(\frac{1}{2\sigma^2} z^4 + z\right) w \quad \textbf{(Formulation 1)} \tag{4.105}$$

with $\lambda = \alpha k \sigma^{\delta}$ being the eigenvalue of the corresponding regular S-L problem. Note that initial conditions remain unchanged

$$w(0) = 0 \qquad \text{and} \qquad w_z(0) = v_z(0) = -\frac{2\lambda}{\sigma^2}\,. \tag{4.106}$$

Scaling argument, normalization. The dependence of eigenvalue λ on a noise strength σ can be verified by a scaling argument. Linear transformation $z = \sigma^{2/3}\bar{z}$ yields

$$2\lambda\sigma^{-2/3} w = -w_{\bar{z}\bar{z}} + \left(\bar{z}^4 + 2\bar{z}\right) w\,. \tag{4.107}$$

Since $\rho = 2\lambda\sigma^{-2/3}$ is an eigenvalue of the operator that does not depend on a noise strength σ, the eigenvalue itself should be independent of σ as well. This implies $2\lambda\sigma^{2/3} = 2\alpha k \sigma^{\delta - 2/3} = 2\alpha k$, where $\delta = 2/3$ and the eigenvalue $\rho = 2\alpha k = 4/3k$. The scaled **SL!** (**SL!**) problem is

$$\frac{4}{3} k w = -w_{\bar{z}\bar{z}} + \left(\bar{z}^4 + 2\bar{z}\right) w \quad \textbf{(Formulation 2)} \tag{4.108}$$

This SL problem still depends on the noise magnitude σ via the boundary conditions

$$w(0) = 0 \qquad \text{and} \qquad w_{\bar{z}}(0) = w_z(0)\frac{dz}{d\bar{z}} = v_z(0)\sigma^{2/3} = -\frac{2\alpha k}{\sigma^{2/3}}\,. \tag{4.109}$$

The eigenfunctions of the operator in **Formulation 1** $w(t, z, \sigma)$ is related to the eigenfunctions of the noise-independent operator in **Formulation 2** $w(t, \bar{z})$ by

$$w(t, \bar{z}) = w(t, \frac{z}{c}) = w(t, \frac{z}{\sigma^{2/3}}) = w(t, z, \sigma)\,. \tag{4.110}$$

4.4.2.1 Asymptotic scaling argument.

In this section we use asymptotic methods to verify the smallest eigenvalue and the form of the ground state of problem **Formulation 1**. First, we rescale the independent variable z and introduce a new variable $\bar{z} = (\frac{1}{2\sigma^2})^{\frac{1}{4}} z$. The result after dropping bars is

$$\frac{\sigma}{\sqrt{8}} w_{zz} + (\lambda - V(z))w = 0 \quad \textbf{(Formulation 3)} \tag{4.111}$$

where

$$V(z) = z^4 + (2\sigma^2)^{\frac{1}{4}} z \tag{4.112}$$

Let us denote

$$\phi(z) = \lambda - V(z) = \lambda - (z^4 + (2\sigma^2)^{\frac{1}{4}} z) \tag{4.113}$$

The ODE we obtained has a form of a time-independent Schrődinger equation (SE) with a potential $V(z)$ for $z \leq 0$ and $V(z) = \infty$ for $z > 0$ (this implies $w(0) = 0$). The parameter σ is small and as $\sigma \to 0$ the potential asymptotically approaches potential of an anharmonic oscillator $V(z) = z^4$. The physical interpretation of the problem suggests that the smallest eigenvalue of the problem is positive. Otherwise there would need to be an exponentialy growing pattern in the time evolution of the survival probability. But the survival probability of a first exit time problem must be non-increasing in time. Therefore there are only two distinct regions: (i) $z \in (\infty, z)$ where $\lambda > V(z)$; and (ii) $(z, 0)$ where $\lambda < V(z)$. We can approximate the solution of SE on both intervals.

If $\lambda > V(z)$ the solution is oscillatory and we use the ansatz

$$
\begin{aligned}
w &= f e^{i\left(\frac{8}{\sigma^2}\right)^{1/4} g} && (4.114)\\
w_y &= \left(f' + i\left(\frac{8}{\sigma^2}\right)^{1/4} f g'\right) e^{i\left(\frac{8}{\sigma^2}\right)^{1/4} g} && (4.115)\\
w_{yy} &= \left(\frac{8}{\sigma^2}\right)^{1/4} \left(\left(\frac{8}{\sigma^2}\right)^{-1/4} f'' + 2if'g' + ifg'' - f(g')^2\right) e^{i\left(\frac{8}{\sigma^2}\right)^{1/4} g(z)} && (4.116)
\end{aligned}
$$

The differential equations for the real and imaginary part are

$$
\begin{aligned}
0 &= \frac{\sigma}{\sqrt{8}} f'' - f(g')^2 + \phi f && (4.117)\\
0 &= 2f'g' + fg'' && (4.118)
\end{aligned}
$$

The noise magnitude σ is small and if we assume that f does not oscillate too fast one can use the Wentzel-Kramers-Brillouin approximation (WKB) approximation and neglect the f'' term. The parameter σ shows up also in the potential term that we keep in the equation (with power 1/2 rather then 1). More careful calculation of the asymptotic solution follows in the next section. The solution of above equations is

$$
\begin{aligned}
g(z) &= \int_{z_1}^{z} \sqrt{\phi(z')} dz' && (4.119)\\
f(z) &= \left(\frac{\phi(z_1)}{\phi(z)}\right)^{1/4} w(z_1) && (4.120)
\end{aligned}
$$

where z_1 is an integration constant. So far we have only found the form of the solution

in the case $\lambda > V(z)$:

$$w_+(z) \sim \frac{K}{\phi(z)^{1/4}} e^{i\left(\frac{8}{\sigma^2}\right)^{1/4} \int_{z_1}^{z} \sqrt{\phi(z')} dz'} \tag{4.121}$$

$$= \frac{a}{\phi(z)^{1/4}} \sin\left[\left(\frac{8}{\sigma^2}\right)^{1/4} \int_{z_1}^{z} \sqrt{\phi(z')} dz'\right] \tag{4.122}$$

$$+ \frac{b}{\phi(z)^{1/4}} \cos\left[\left(\frac{8}{\sigma^2}\right)^{1/4} \int_{z_1}^{z} \sqrt{\phi(z')} dz'\right] \tag{4.123}$$

If $\lambda < V(z)$ one can find similar general form for the solution $w_-(z)$. In this interval we obtain exponentialy growing and decaying solutions

$$w_-(z) \sim \frac{A}{\phi(z)^{1/4}} e^{\left(\frac{8}{\sigma^2}\right)^{1/4} \int_{z_1}^{y} \sqrt{\phi(z')} dz'} \tag{4.124}$$

$$+ \frac{B}{\phi(z)^{1/4}} e^{-\left(\frac{8}{\sigma^2}\right)^{1/4} \int_{z_1}^{z} \sqrt{\phi(z')} dz'} \tag{4.125}$$

Let us for simplicity denote the argument of the sine by U. In order to find all eigenvalues of the problem we need to connect the solution at a "turning point" $\hat{z}$ (such thet $\phi(\hat{z}) = 0$). We will follow the matching process discussed in J. D. Murray – Asymptotic Analysis (1984). Required decay at $z = -\infty$ forces $B = 0$. Connection formula has a form

$$a = \frac{1}{\sqrt{2}}\left(\frac{B}{2} - A\right) \tag{4.126}$$

$$b = \frac{1}{\sqrt{2}}\left(\frac{B}{2} + A\right) \tag{4.127}$$

and after simplification $a = -b$. The boundary condition at $z = 0$ then gives

$$\sin(U) = \cos(U) \tag{4.128}$$

where

$$U = \left(\frac{8}{\sigma^2}\right)^{1/4} \int_{z_1}^{z} \sqrt{\phi(y')} dz' \tag{4.129}$$

This yields that $U = \frac{\pi}{4} + k\pi$. The quantization of eigenvalues for this problem has a form

$$\left(\frac{8}{\sigma^2}\right)^{1/4} \int_{\hat{z}}^{0} \sqrt{\phi(z')} dz' = \frac{\pi}{4} + k\pi \tag{4.130}$$

With a potential $V(z)$ the quantization condition for eigenvalues becomes

$$\left(\frac{8}{\sigma^2}\right)^{1/4} \int_{\hat{z}}^{0} \sqrt{\lambda - z'^4 - (2\sigma^2)^{1/4} z'} dz' = \frac{\pi}{4} + k\pi \tag{4.131}$$

where $\hat{z}$ is the unique negative solution of a polynomial equation

$$\hat{z}^4 + (2\sigma^2)^{1/4} \hat{z} = \lambda \tag{4.132}$$

The ground state corresponds to the smallest eigenvalue, for $k = 0$. The ground state is the only eigenfunction that remains nonnegative for all $z \in (-\infty, 0]$. The asymptotic approximation of the smallest eigenvalue can be found for σ small if we take $V(z) \approx z^4$. The eigenvalue can be explicitly found from the quantization equation (turning point is $\hat{z} = -\lambda^{1/4}$) since

$$\int_{-\lambda^{1/4}}^{0} \sqrt{\lambda - z'^4} dz' = \lambda^{3/4} \int_{-1}^{0} \sqrt{1 - z'^4} dz' = \lambda^{3/4} \frac{\sqrt{\pi}\Gamma(\frac{1}{4})}{8\Gamma(\frac{7}{4})} \tag{4.133}$$

The approximation for λ is

$$\lambda \approx \left[\frac{3}{4}\sqrt{\pi}\left(\frac{\sigma^2}{8}\right)^{1/4} \frac{8\Gamma(\frac{7}{4})}{\sqrt{\pi}\Gamma(\frac{1}{4})}\right]^{4/3} \sim \sigma^{2/3} \tag{4.134}$$

where the scaling law is consistent with the intuitive answer found by the scaling

argument $\lambda = \alpha k \sigma^{\delta} = \frac{2}{3} k \sigma^{\frac{2}{3}}$.

4.4.2.2 Numerical results

In this section we compare numerical findings for the first exit problem (4.63), (4.64), (4.67), using two different methods. The goal is to estimate the probability density function of the position of the trajectory in the noise-sensitive region. The first approach is to simulate the problem using MCR method and to arrive to a time-dependent density function (in rescaled coordinates), as on Fig. 4.15. In the limit $t \to \infty$ this converges to the invariant probability density function for the problem. Second approach is a combination of analysis and numerics. Here we numerically solve the rescaled Fokker-Planck equation, given by **Formulation 2** with boundary conditions $w(0) = 0$ and $w_{\hat{z}}(0) = -\frac{2\alpha k}{\sigma^{2/3}}$. We choose the value of parameter k so that the solution satisfies $w(-\infty) = 0$. Note that there is only one such value. Then we transform the solution from $w(\hat{z})$ back to $v(z)$ to obtain the ground state of the S-L problem (4.99). This ground state, after normalization to $\int_{-\infty}^{0} w(z)dz = 1$, should be the same as the invariant probability density function from the first approach. Also, one can evaluate the asymptotic solution found by using a WKB approximation as sketched above. As Figure 4.16 suggests, the solution of the Fokker-Planck equation captures the trend in MCR results as time grows. Since the time $t = 7$ is still not very large, the match is not perfect. Note that time $t = 7$ in the MCR method corresponds to $y = e^7 \approx 10^3$ for which the time step required is the original time step reduced by the factor of 10^3. The slowdown of the numerical calculation due to the exponential approach of the exit boundary is the major difficulty of the numerical method.

The WKB analysis predicts, based on the quantization condition (4.134), the value of constant k as follows

$$k = \frac{3}{2}\lambda\sigma^{-2/3} \approx \left[\frac{3}{4}\sqrt{\pi}\left(\frac{1}{8}\right)^{1/4}\frac{8\Gamma(\frac{7}{4})}{\sqrt{\pi}\Gamma(\frac{1}{4})}\right]^{4/3} \sim 1.31184 \tag{4.135}$$

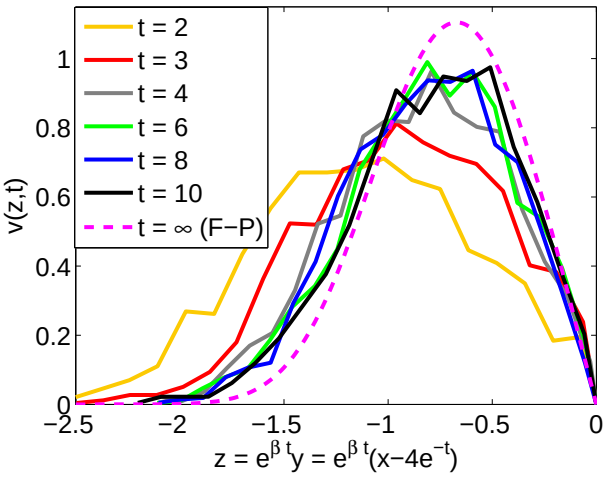

Figure 4.16: **Comparation of MCR and Fokker-Planck solutions.** Solid lines are probability density functions of the conditional survival probability distribution inside the noise-sensitive region found by MCR (for times $t = 3, \ldots 7$), dashed line is the numericaly found and normalized eigenfunction of the stationary Fokker-Planck equation corresponding to the first exit time problem. The dashed line seems to capture the limit of solid lines as time grows to ∞.

On the other hand the numerics yields that in order to satisfy $w(-\infty) = 0$ we need $k \sim 0.5535$. The following is a review of parameter predictions by the numerical MCR method and by analytic approach for $\sigma = 1$: We remark that MCR method

Constant	α	δ	k	β
Monte Carlo with recycling	0.6659	0.6434	0.0373	0.3372
Fokker-Planck equation	2/3	2/3	1.3118	1/3
Numerics of Fokker-Planck	(2/3)	(2/3)	0.5535	1/3

Table 4.3: **Parameter values for different methods.**

and analysis via the Fokker-Planck equation give almost identical values of α (double exponential decay rate of the survival probability), β (exponential contraction rate on the horizontal axis) and δ (noise-sensitivity coefficient in the decay rate of the survival probability). On the other hand, the three different ways of evaluating the prefactor k give all different result.

4.5 Conclusions

The noise-driven collapse of exponentially growing solutions in the RBC model is phenomenologically studied using the simple two-dimensional system of differential equations, that mimic the essential features of the RBC dynamics. Perhaps the most significant outcome is that Gaussian noise perturbation of arbitrarily small magnitude can induce collapses of trajectories towards origin. The stochastic Lyapunov function theory is used to prove existence of an invariant measure for the underlying stochastic process for arbitrary small noise.

The properties of the process, i.e. when does the collapce occur, are studied (both numerically and analytically) by the first exit time problem in a noise-sensitive region. Using numerical methods, based on Monte Carlo type algorithms, we estimate the survival probability $S(t)$ and mean trajectory position $m(t)$ in the noise-sensitive region and formulate two hypotheses

$$S(t) \sim e^{-k\sigma^{\delta}e^{\alpha t}} \qquad \text{and} \qquad m(t) \sim Be^{-\beta t} \qquad \text{as} \quad t \to \infty$$

with appropriate values of constants α, δ, k and β. In the analytic approach, we use the hypotheses to study the the appropriate Fokker-Planck equation. The two approaches give remarkably similar parameter values α, δ and β (the parameter k differs). Also, the solution of the FP equation (eigenfunction that corresponds to the smallest eigenvalue) is compared to the numerically found probability density function close to the y-axis. As $t \to \infty$, the numerical solution appears to converge to the solution of the Fokker-Planck equation.

One of the originally intended goals was to show, that the simple SDE model reproduces a scaling similar to $Nu \sim Ra^{\alpha}$ in the RBC model. The system with

additional parameter r (analog of Rayleigh number) has a form

$$\begin{aligned} dx &= -x(4 - xy + x^2)dt + \sigma dW\,, \\ dy &= \frac{ry - x^3}{1 + x^2}dt\,. \end{aligned}$$

Unfortunately, as r grows the model undergoes a series of bifurcations, changing the dynamics of the problem. Therefore this particular model cannot be used to illustrate the behavior in the limit $r \to \infty$.

CHAPTER V

Conclusions

To understand the effect of randomness, caused by a discrete nature of a system on small scale, we studied a classical Hodgkin-Huxley model in Chapter II. Two different stochastic representations were examined and compared. The mechanism of signal propagation, that is driven by the opening and closing of Na and K channels was demonstrated to be highly sensitive to the variability (channel density) of sodium channels as opposed to being almost completely insensitive to the variability in potassium channels. The computations, performed for a large set of parameter values, demonstrated the effect of noise on a bistable system that goes through a subcritical Hopf bifurcation. The simplest deterministic quantity, that captures the effects of variability is the average firing rate of the neuron. The numerical experiments showed that noise is responsible for the induction of action potentials in the subthreshold region, whereas it is also responsible for suppression of the action potentials in the superthreshold region. But for any level of noise the firing rate appeared to increase monotonically with the input current. The robustness of the observed features was confirmed as both numerical approaches led to qualitative agreement. In order to obtain a quantitative agreement an additional understanding of the relationship between the two stochastic models is necessary.

Next we turned our attention to the problem of mixing. Several mixing mea-

sures found in literature were described and an apparent conflict among existing theories (homogenization theory and internal-layer theory) was outlined. In an effort to resolve the conflict we presented a new approach (dispersion-diffusion approach), based on the particle dispersion modelling with a time consistent implementation and studied it for a simple shear flow. In the first step of the construction, explicit calculation of the covariance matrix for the nonlinear stochastic differential system was presented. This matrix was then used to find a Green's function of the homogenous problem (no sources and sinks). Finally, the solution of the nonhomogenous problem was expressed as a superposition of the source with the Green's function. Using the variance suppression mixing measure we computed the concentration fields and the mixing efficiency for the numerical approximation of the exact solution, for homogenization theory (HT) and internal-layer theory (ILT) and finally also for the dispersion-diffusion theory (DDT). The DDT captured HT in the regime $r > Pe$ and ILT in the regime $Pe > r$. The DDT matched perfectly with the exact solution of the advection-diffusion equation. We explored the non-comutativity between the limits $Pe \to \infty$ and $r \to \infty$ for the simple test problem.

In effort to understand the observed repetitive random pattern in homogeneous Rayleigh-Bénard convection, described in *Calzavarini et al.* (2006), we constructed two simpler, 2-dimensional systems of ODEs in Chapter IV, possessing the same features as the homogeneous Rayleigh-Bénard convection. In the first system (Model 1) the nonlinearity in the dynamics was showed not to be strong enough to force trajectories to stay in some bounded region. However, the second system (Model 2) had a property that trajectories always return to a neighborhood of the origin. We proved that arbitrarily small perturbation in the horizontal component of Model 2 leads to an existing invariant measure, even though the system without noise diverges to infinity at an exponential rate. This was proved using the stochastic Lyapunov function theorem by *Khasminskii* (1980). Once we knew that trajectories almost

surely do not blow up we approximated the problem by the first exit time problem from the noise-sensitive region. We developed a fast computational method, based on Monte Carlo methods with importance sampling and showed that probability of exiting by time t behaves asymptotically for large times as $P(t) \sim 1 - e^{-k\sigma^{\delta}e^{\alpha t}}$. The mean position of the trajectory that has not yet exited the noise-sensitive region was found to be asymptotically $m(t) \sim Be^{-\beta t}$. This ansatz was used to derive an ordinary differential equation for the distribution of the trajectory position as $t \to \infty$. We noted that the result accurately approximates the distribution found by the Monte Carlo method for relatively large times. The model was originally designed to explain not only the effects of small noise on its dynamics but also to investigate the regime when the rate of exponential blowup grows to infinity. Unfortunately, this model problem was unable to do so because of the bifurcation as the rate parameter increased.

APPENDICES

APPENDIX A

Expanding property in Model 1

Proof. We first transform the deterministic dynamical system into polar coordinates by

$$x = r\cos\theta\,, \qquad y = r\sin\theta$$

This leads to a dynamical system for r and θ

$$\begin{aligned} \dot{r} &= -r\cos 2\theta\,, \\ \dot{\theta} &= \sin 2\theta - r^3\cos^2\theta\,. \end{aligned}$$

Given that the trajectory enters the region Ω_+ at a value r_{out} we may specify the angle $\theta_{\text{out}} \in [0, \frac{\pi}{2}]$ at which it entered by a exit condition $xy = r^2\sin\theta\cos\theta = 1$ (where x is small). The resulting θ_{out} satisfies

$$\theta_{\text{out}} = \frac{1}{2}\arcsin\frac{2}{r^2} = \mathcal{O}\left(\frac{\pi}{2} - \frac{1}{r^2}\right) \quad \text{as} \quad r \to \infty$$

and the entering angle $\theta_{\text{in}} = -\theta_{\text{out}}$. Moreover it has a property that $\theta_{\text{out}} \to 0$ as $r \to \infty$. First note that for sufficiently large $r > R$ the angle θ decreases on Ω_+ and

if R is large the effective form of Model 1 in polar coordinates is

$$\dot{r} = -r\cos 2\theta\,, \tag{A.1}$$

$$\dot{\theta} = -r^3\cos^2\theta\,. \tag{A.2}$$

Monotonicity of θ implies that there exists a bijection $\tau : t \to \theta$. Using τ we can view radius as a function of θ. Then

$$r_{\rm in} = r_{\rm out} + \int\limits_{\theta_{\rm out}}^{\theta_{\rm in}} \frac{dr/dt}{d\theta/dt} d\theta = r_{\rm out} + \int\limits_{\theta_{\rm in}}^{\theta_{\rm out}} \frac{-r(\theta)\cos 2\theta}{r(\theta)^3\cos^2\theta - \sin 2\theta} d\theta$$

Observe that for $\theta \in (\frac{\pi}{4}, \frac{\pi}{2}) \cup (-\frac{\pi}{2}, \frac{\pi}{2})$ the value of r increases whereas for $\theta \in (-\frac{\pi}{4}, \frac{\pi}{4})$ it decreases. Using symmetry of sin and cos functions the net increase of r is

$$\begin{aligned}
\Delta r &= 2\left[\int\limits_{\frac{\pi}{4}}^{\theta_{\rm out}} \frac{-r(\theta)\cos 2\theta}{r(\theta)^3\cos^2\theta - \sin 2\theta} d\theta - \int\limits_{0}^{\frac{\pi}{4}} \frac{r(\theta)\cos 2\theta}{r(\theta)^3\cos^2\theta - \sin 2\theta} d\theta\right] \\
&\geq 2\left[\int\limits_{\frac{\pi}{4}}^{\theta_{\rm out}} \frac{-\cos 2\theta}{\cos^2\theta}\frac{1}{r(\theta)^2} d\theta - \int\limits_{0}^{\frac{\pi}{4}} \frac{\cos 2\theta}{\cos^2\theta}\frac{1}{(r(\theta)-1)^2} d\theta\right]
\end{aligned}$$

The key information is that for large $r_{\rm out}$ the rate at which θ decreases is much greater than the rate ar which r changes (compare coefficients r and r^3 in (A.1)-(A.2)). The range of angles for which this statement is true can be extracted from

$$\frac{dr}{d\theta} = \frac{-r(\theta)\cos 2\theta}{\sin 2\theta - r(\theta)^3\cos^2\theta} \sim \frac{\cos 2\theta}{r(\theta)^2\cos^2\theta} = -\frac{1}{r(\theta)^2}(\tan^2\theta - 1)$$

For $\theta \in [-\frac{\pi}{4}, \frac{\pi}{4}]$ the magnitude $|dr/d\theta| \sim \mathcal{O}(1/r^2)$ and the dominant contribution is when $\tan^2\theta > \mathcal{O}(r^2)$. But

$$\tan^2\theta \sim r^2 \quad \text{when} \quad \frac{\pi}{2} - \theta \sim \frac{1}{r}$$

The biggest contribution of $dr/d\theta$ is for $\theta \leq \mathcal{O}(\frac{\pi}{2} - \frac{1}{r})$. As shown in (A), $\theta_{\text{out}} \sim \mathcal{O}(\frac{\pi}{2} - 1/r^2)$ implying that the interval of biggest contribution is $[\theta_{\text{out}}, \frac{k}{r})$ (contained in the first integral of expression for Δr). Because r grows on this interval, this makes $\Delta r > 0$. The proof is analogous for trajectories that enter the deterministic region Ω_- due to symmetry of the problem. $\square$

APPENDIX B

Contraction property in Model 2

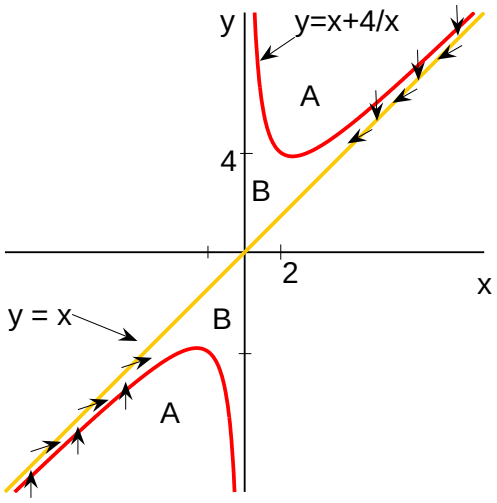

Figure B.1: Invariance of region B under the flow.

Lemma B.1. *Let* $B = \{(x, y) : \frac{d|x|}{dt}(x, y) < 0, x(y - x) > 0\}$. *Then any trajectory that starts in region* A *enters* B *in a finite time.*

Proof. The proof is done for the deterministic system and consists of three parts. Only the first quadrant case will be studied (the proof in the third quadrant is analogous). First we show that if $(x(0), y(0)) \in A$ then $x(t)$ has to satisfy $x(\tau_1) = 2$ at a finite

time τ_1. If this is not true then certainly $x(t)$ is bounded and monotone increasing, therefore $\lim_{t\to\infty} x(t) = x_* \leq 2$ and $\lim_{t\to\infty} y(t) = \infty$. But this implies

$$\lim_{t\to\infty} \dot{x} = -4x_* - {x_*}^3 + {x_*}^2 \lim_{t\to\infty} y = \infty$$

that contradicts with assumed finite limit of $x(t)$. By smoothness of the map $x(\tau_1) = 2$ at $\tau_1 \in \mathbb{R}$. Next we will argue that trajectory will also cross the nullcline $y = x^3$ in a finite time τ_2 due to a faster growth of $x(t)$ compared to $y(t)$. Assume that $\lim_{t\to\infty} y(t) = \infty$ and $\lim_{t\to\infty} x(t) = \infty$ but at the same time $y(t) < x(t)^3$ for all t (note that the x-limit cannot be different by the previous contradiction argument). Now compare the slopes

$$\begin{aligned} \dot{x} &= -x(4 - xy + x^2) > -4x + y(x^2 - 1) \sim y(x^2 - 1) \\ \dot{y} &\sim \frac{y}{x^2} - x \end{aligned}$$

But since $x > 2$ this implies that $\dot{x} >> \dot{y} > 0$ (contradiction with not crossing the nullcline). The trajectory therefore crosses $y = x^3$ in a finite time τ_2. The last claim is that once trajectory has crossed y-nullcline $y = x^3$ then it will also cross x-nullcline $y = 4/x + x$ in a finite time τ_3. This will be done by a direct proof. Observe that if $(x(\tau_3), y(\tau_3)) \in A - \{(x, y) : y < x^3\}$ then x will increase and y will decrease until it potentialy crosses the x-nullcline. The crossing must occur in a finite time because either the growth of x or the decay of y is bounded away from 0. In conclusion, trajectory starting in A will enter B in a finite time. □

Lemma B.2. *Trajectories of the flow given by* (4.23)-(4.24) *starting in region* $B \cap \{(x, y) : |x| > 2\}$ *may leave the region only by crossing* $x = 2$. *Moreover, at the point of crossing* $|y|/2 \leq 2$.

Proof. To prove the lemma it is sufficient to show that the flow on both boundaries

$y = x$ and $y = x + 4/x$ of region B is pointing inside the region. First, if $x = y$ then the flow has a form

$$\begin{aligned} \dot{x} &= -4x \\ \dot{y} &= \frac{x - x^3}{1 + x^2} \end{aligned}$$

But $0 > \dot{y} > \dot{x}$ for any $x > 0$ and therefore the flow points inside B. Second, if $y = x + 4/x$ then

$$\begin{aligned} \dot{x} &= 0 \\ \dot{y} &= \frac{x + 4/x - x^3}{1 + x^2} \end{aligned}$$

where $\dot{y} < 0$ for any $x > 2$. Since for $|x| = 2$ the crossing poit satisfies $|y| < |x| + 4/|x| \leq 2 + 2 = 4$. □

Proof. **Contraction of Model 2.** The Lemmas above imply that if $(x(0), y(0)) \in A$ then in a finite time $(x(\tau), y(\tau)) \in B$ and since $y(t)$ decays monotonicaly at the rate that is bounded away from 0 then in finite time $\max\{|x(\tau)|, \frac{|y(\tau)|}{2}\} \leq 2$. □

APPENDIX C

Details of SLF calculation

Define piecewise linear function $f(x)$ to be

$$f''(x) = \begin{cases} -\frac{\varepsilon}{2-Y}(x+2) & \text{for } x \in [-2, -Y) \\ -\frac{K-\varepsilon}{Y-\varepsilon}(x+Y) - \varepsilon & \text{for } x \in [-Y, -\varepsilon) \\ -K & \text{for } x \in [-\varepsilon, \varepsilon) \\ \frac{K+L}{X-\varepsilon}(x-\varepsilon) - K & \text{for } x \in [\varepsilon, X) \\ \frac{2-x}{2-X}L & \text{for } x \in [X, 2) \end{cases}$$

and set $X = Y = 2\varepsilon$. Then integration with respect to x plus the boundary condition

$f'(-2) = 0$ gives us

$$f'(x) = \begin{cases} -\frac{\varepsilon(2+x)^2}{4(1-\varepsilon)} & \text{for } x \in [-2, -2\varepsilon) \\ \varepsilon(\varepsilon - 1 + x) + \frac{x^2}{2} - K(2\varepsilon + 2x - \frac{x^2}{2\varepsilon}) & \text{for } x \in [-2\varepsilon, -\varepsilon) \\ -\frac{\varepsilon}{2}(2-\varepsilon) - \frac{K}{2}(3\varepsilon + 2x) & \text{for } x \in [-\varepsilon, \varepsilon) \\ \frac{6\varepsilon^2 - 2\varepsilon^4 + 2\varepsilon^3 x + 2\varepsilon x^2 - \varepsilon^2(2+x)^2}{2\varepsilon(2-\varepsilon)} + & \\ + K\frac{8\varepsilon^3 - x\varepsilon(8-5x) + 2x^2 - \varepsilon^2(4+8x)}{2\varepsilon(2-\varepsilon)} & \text{for } x \in [\varepsilon, 2\varepsilon) \\ -\frac{\varepsilon(2+6K-\varepsilon)(2-x)^2}{4(2-\varepsilon)(1-\varepsilon)} & \text{for } x \in [2\varepsilon, 2) \end{cases}$$

where K and L are be chosen such that the boundary condition $f'(2) = 0$ is satisfied

$$L = \varepsilon + K\frac{X + Y + 2\varepsilon}{2 - \varepsilon}\,.$$

After one more integration we obtain $f(x)$ in a form of piecewise defined third-order polynomial in x. Formulas are presented just for the sake of completeness of the

function definition:

$$f(x) = \begin{cases} \frac{\varepsilon(2+\varepsilon-3\varepsilon^2)}{1-\varepsilon}K+ & \\ +\frac{\varepsilon(20-12\varepsilon-28\varepsilon^2+12\varepsilon^3-12x-6x^2-x^3)}{12(1-\varepsilon)} & \text{for } x \in [-2,-2\varepsilon) \\ \frac{10\varepsilon^3+12\varepsilon^2(1-x)-6\varepsilon x^2-x^3}{6\varepsilon}K+ & \\ +\frac{10\varepsilon^2+4\varepsilon^3-2\varepsilon^4-6\varepsilon^2x+6\varepsilon^3x+3\varepsilon^2x^2+\varepsilon x^3}{6\varepsilon} & \text{for } x \in [-2\varepsilon,-\varepsilon) \\ \frac{1}{6}K(11\varepsilon^2+\varepsilon(12-9x)-3x^2)+ & \\ +\frac{1}{6}(10\varepsilon+4\varepsilon^2-3\varepsilon^3-6\varepsilon x+3\varepsilon^2x) & \text{for } x \in [-\varepsilon,\varepsilon) \\ \frac{-16\varepsilon^4+8\varepsilon^3(1+3x)-12\varepsilon^2(x^2+x-2)+\varepsilon x^2(12-5x)-2x^3}{6\varepsilon(2-\varepsilon)}K+ & \\ +\frac{4\varepsilon^5-\varepsilon^4(12+6x)+\varepsilon^3(-2+18x+3x^2)+\varepsilon^2(20-12x-6x^2-x^3)+2\varepsilon x^3}{6\varepsilon(2-\varepsilon)} & \text{for } x \in [\varepsilon,2\varepsilon) \\ -\frac{\varepsilon(x-2)^3}{2(\varepsilon-2)(\varepsilon-1)}K+ & \\ +\frac{\varepsilon(12\varepsilon^2+\varepsilon(-44+12x-6x^2+x^3)+(40-24x+12x^2-2x^3))}{12(\varepsilon-2)(\varepsilon-1)} & \text{for } x \in [2\varepsilon,2) \end{cases}$$

Constants A and B are chosen to satisfy boundary conditions $f(-2) = A$ and $f(2) = B = \varepsilon$ leading to a relationship between A and B (again by using a fundamental theorem of calculus)

$$\begin{aligned} A &= B + \frac{1}{6}(-2\varepsilon^3 + K(2+Y)(X+Y) + \varepsilon^2(2K-X-Y) \\ &+ 2\varepsilon(4+X+Y+K(2+X+Y)))\,, \end{aligned}$$

or using $X = Y = 2\varepsilon$ values

$$A = -\varepsilon^3 + \left(\frac{4}{3}+3K\right)\varepsilon^2 + \left(\frac{7}{3}+2K\right)\varepsilon\,.$$

At the end we plug $f(x)$, $f'(x)$ and $f''(x)$ into the coefficient function $A_2(x)$ to find conditions on remaining parameters K and ε that will guarantee $A_2(x) < 0$ and consequently the validity of the stochastic Lyapunov function. The coefficient

function $A_2(x)$ for $x \in [-\varepsilon, \varepsilon)$ will have a form

$$\begin{aligned} A_2(x) &= 3Kx^4 + \frac{1}{6}(18\varepsilon - 9\varepsilon^2 + 27\varepsilon K)x^3 + 3Kx^2 \\ &+ \frac{1}{6}(12\varepsilon - 6\varepsilon^2 + 18\varepsilon K)x \\ &+ \frac{1}{6}(20\varepsilon + 8\varepsilon^2 - 6\varepsilon^3 + 24\varepsilon K + 22\varepsilon^2 K - 3K\sigma^2) \\ &= (4\varepsilon - \frac{1}{2}\sigma^2)K + \mathcal{O}(\varepsilon(1 + \varepsilon K)) \end{aligned}$$

For small values of noise level σ we can certainly choose $\epsilon(\sigma) < \frac{1}{16}\sigma^2$ small enough and $K \sim \mathcal{O}(\frac{1}{\varepsilon})$ so that the coefficient $A_2(x)$ satisfies

$$A_2(x) < -\frac{1}{4}\sigma^2 K \text{ for any } x \in [-\epsilon, \epsilon)$$

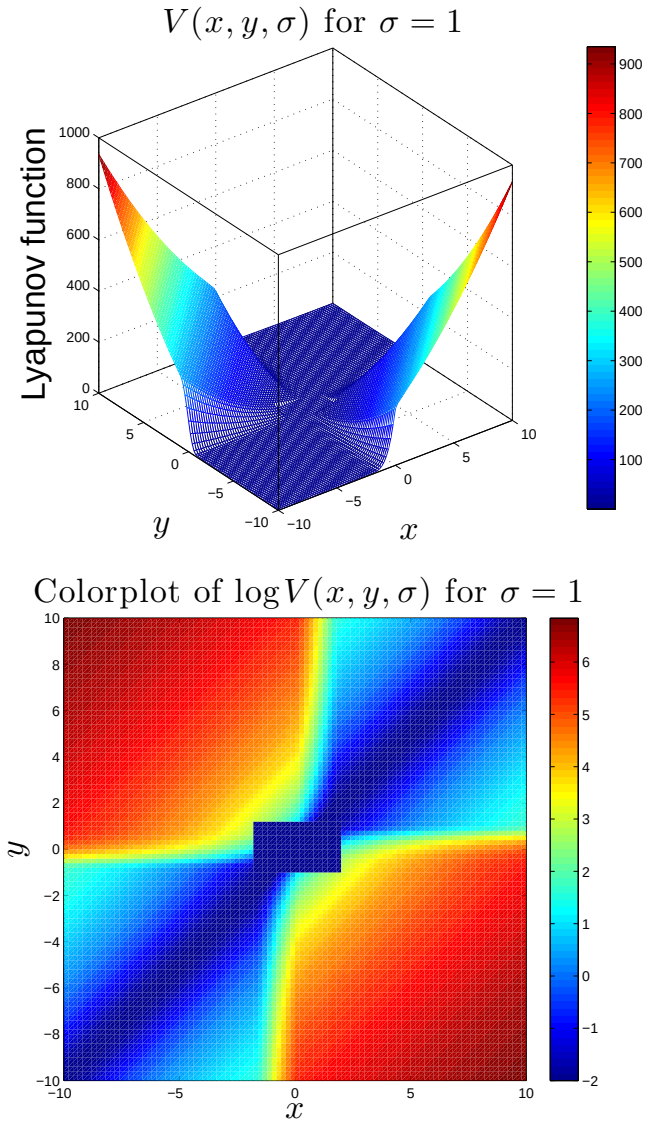

Figure C.1: **Stochastic Lypunov function** $V(x, y, \sigma)$**.** A positive constant $c = e^{-2}$ is added to the function to satisfy $V(x, y, \sigma) > 0$ for all $\mathbb{R}^2$.

APPENDIX D

Flow-independent upper bound on mixing efficiency $\mathcal{E}_0$

The advection-diffusion equation for the mixing problem with arbitrary stirring flow and spatially inhomogeneous steady sources given by $s = \sqrt{2}S \sin k_s x$ is

$$\dot{\theta} + \vec{u} \cdot \vec{\nabla}\theta = \kappa \Delta \theta + s \,.$$

Multiplying the equation by a test function $\phi(\vec{x})$ and space-time averaging yields

$$\langle \vec{u} \cdot \vec{\nabla}\theta \cdot \phi \rangle = \langle \kappa \Delta \theta \cdot \phi \rangle + \langle s \cdot \phi \rangle \,,$$

and after integrating by parts and using Cauchy-Schwartz inequality one obtains

$$\langle s \cdot \phi \rangle = -\langle (\kappa \Delta \phi + u \cdot \nabla \phi)\theta \rangle \leq \langle (\kappa \Delta \phi + u \cdot \nabla \phi)^2 \rangle^{1/2} \cdot \langle \theta^2 \rangle^{1/2} \,. \tag{D.1}$$

Inequality (D.1) may be used to find an upper bound for the mixing efficiency at intermediate scales since $\theta_0 = \frac{1}{\kappa}\langle \Delta^{-1} s \rangle$

$$\mathcal{E}_0^2 \;=\; \frac{\langle \theta_0^2 \rangle}{\langle \theta^2 \rangle} \leq \frac{\langle \Delta^{-1} s \rangle^2 \langle (\kappa \Delta \phi + u \cdot \nabla \phi)^2 \rangle}{\kappa^2 \langle (s \cdot \phi)^2 \rangle}$$

In particular, if the test function is chosen to be $\phi = \sqrt{2}S \sin k_s x$ the flow-independent bound simplifies to

$$\begin{aligned}
\mathcal{E}_0^2 &\leq \frac{2S^2\langle \sin^2 k_s x\rangle}{\kappa^2 k_s^4} \cdot \frac{2\kappa^2 k_s^4 S^2 \langle \sin^2 k_s x\rangle + \langle (u\cdot\nabla\phi)^2\rangle}{4S^4\langle \sin^2 k_s x\rangle^2} \\
&= 1 + \frac{\langle (u\cdot\nabla\phi)^2\rangle}{\kappa^2 k_s^4 S^2} = 1 + \frac{\langle 2\cos^2 k_s x \cdot u_x^2\rangle}{\kappa^2 k_s^2} \\
&\leq 1 + 2\frac{U^2}{\kappa^2 k_s^2} = 1 + 2r^2 Pe^2 \qquad \text{(D.2)}
\end{aligned}$$

where $U = \langle ||(u_x, u_y)||_2\rangle$. In case the fluid is stirred by a shear flow in the direction perpendicular to x, e.g., $u = \hat{i}\sqrt{2}U \sin k_u y$, a tighter upper bound can be obtained by

$$\begin{aligned}
\mathcal{E}_0^2 &\leq = 1 + \frac{\langle 2\cos^2 k_s x \cdot u_y^2\rangle}{\kappa^2 k_s^2} \\
&\leq 1 + \frac{U^2}{\kappa^2 k_s^2} = 1 + r^2 Pe^2 \qquad \text{(D.3)}
\end{aligned}$$

BIBLIOGRAPHY

BIBLIOGRAPHY

Adams, D. A., L. M. Sander, and R. M. Ziff (2008), Harmonic Measure for Percolation and Ising Clusters Including Rare Events, *Phys. Rev. Let.*, *101*, 144,102.

Ahlers, G., S. Grossmann, and D. Lohse (2009), Heat transfer and large scale dynamics in turbulent Rayleigh-Bénard convection, *Rev. Mod. Phys.*, *81*, 503.

Aidley, D. J. (1978), *The physiology of excitable cells*, Cambridge University Press, Cambridge.

Amati, G., K. Koal, F. Massaioli, and K. R. Sreenivasan (2005), Turbulent thermal convection at high Rayleigh numbers for a constant-Prandtl-number fluid under Boussinesq conditions, *Phys. Fluids*, *17*, 121,701.

Bazso, F., L. Zalanyi, and G. Csardi (2003), Channel noise in Hodgkin-Huxley model neurons, *Physica A*, *325*, 165–175.

Bourlioux, A., and A. J. Majda (2002), Elementary models with probability distribution function intermittency for passive scalars with a mean gradient, *Phys. Fluids*, *2*, 881–897.

Boussinesq, J. (1879), *Theorie Analytique de la Chaleur*, vol. 2, Gauthier-Villars, Paris.

Bruce, I. C. (2006), Implementation issues in approximate methods for stochastic Hodgkin-Huxley models, *Annals Biomed. Eng.*, *35*, 315–318.

Busse, F. H. (2003), The sequence-of-bifurcations approach towards understanding turbulent fluid flow, *Surv. Geophys.*, *24*, 269–288.

Calzavarini, E., C. R. Doering, J. D. Gibbon, D. Lohse, A. Tanabe, and F. Toschi (2006), Exponentially growing solutions in homogeneous Rayleigh–Bérnard convection, *Phys. Rev. E*, *73*, 035,301.

Chow, C. C., and J. A. White (1996), Spontaneous action potentials due to channel fluctuations, *Biophys. J.*, *71*, 3013–3021.

Clay, J. R., and L. J. DeFelice (1983), Relationship between membrane excitability and single channel open-close kinetics, *Biophys. J.*, *42*, 151–157.

Colquhoun, D., and B. Sakmann (1985), Fast events in single-channel currents activated by acetylcholine and its analogues at the frog muscle end-plate, *J. Physiol.*, *369*, 501–557.

Constantin, P., A. Kiselev, L. Ryzhik, and A. Zlatoš (2008), Diffusion and mixing in fluid flow, *Annals of Math*, *168*, 643–674.

Csanady, G. T. (1973), Turbulent diffusion in the environment, *Geophysics and Astrophysics Monographs*, *3*.

Dagan, G. (1987), Theory of solute transport by groundwater, *Ann. Rev. of Fluid Mech.*, *19*, 183–215.

Danckwerts, P. V. (1952), The definition and measurement of some characteristics of mixers, *Appl. Sci. Res.*, pp. 279–296.

De Felice, L. J., F. Conti, and E. Wanke (1975), Potassium and Sodium Current Noise in the Membrane of the Squid Giant Axon, *J. of Phys.*, *248*, 45–82.

Doering, C. R., and P. Constantin (1992), Energy dissipation in shear driven turbulence, *Phys. Rev. Lett.*, *69*(11), 1648–1651.

Doering, C. R., and P. Constantin (1996), Variational bounds on energy dissipation in incompressible flows. iii. convection, *Phys. Rev. E*, *53*(6), 5957–5981.

Doering, C. R., and J.-L. Thiffeault (2006), Multiscale mixing efficiencies for steady sources, *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, *74*(2), 025,301.

Dryer, S. E., J. T. Fujii, and A. R. Martin (1989), A Na+-activated K+ current in cultured brain stem neurones from chicks., *J. Physiol.*, *410*, 283–296.

Edwards, A. C., W. D. Sherman, and R. E. Breidenthal (1985), Turbulent mixing in tubes with transverse injection, *AIChE*, *31*, 516–518.

Faber, T., and J. C. Vassilicos (2009), Turbulent pair separation due to multicsale stagnation point structure and its time assymetry in two-dimensional turbulence, *Phys. Fluids*, *21*, 015,106.

Falkovich, G., and K. R. Sreenivasan (2006), Lessons from hydrodynamic turbulence, *Phys. Today*, *59*, 43–49.

Fannjiang, A., and G. C. Papanicolaou (1994), Convection enhanced diffusion for periodic flows, *SIAM J. App. Math*, *54*, 333–408.

Fox, R. F. (1997), Stochastic versions of the Hodgkin-Huxley Equations, *Biophys. J.*, *72*, 2068–2074.

Gillespie, D. T. (1977), Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.*, *81*, 2340 – 2361.

Gollub, J. P., J. Clarke, M. Gharib, B. Lane, and O. N. Mesquita (1991), Fluctuations and transport in a stirred fluid with a mean gradient, *Phys. Rev. Lett.*, *67*(25), 3507–3510.

Hodgkin, A. L., and A. F. Huxley (1952), A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. of Phys.*, *119*, 500–544.

Holzer, M., and E. Siggia (1995), Erratum: "turbulent mixing of a passive scalar", *Physics of Fluids*, *7*(6), 1519–1519.

Hughes, D. W., and M. R. E. Proctor (1990), A low-order model of the shear instability of convection: chaos and the effect of noise, *Nonlinearity*, *3*(1), 127–153.

Kadanoff, L. P. (2001), Turbulent heat flow: Structures and scaling, *Physics Today*, *54*(8), 34–39.

Khasminskii, R. Z. (1980), *Stochastic stability of differential equations*, Sijthoff and Nordhoff.

Khasminskii, R. Z. (2008), *Stochastic stability of differential equations*, Lecture Notes, Wayne State University.

Kraichnan, R. H. (1962), Turbulent thermal convection at arbitrary prandtl number, *Physics of Fluids*, *5*(11), 1374–1389.

Landau, L. D., and E. M. Lifshitz (1987), *Fluid Mechanics*, Pergamon.

Lohse, D., and F. Toschi (2003), Ultimate state of thermal convection, *Phys. Rev. Lett.*, *90*(3), 034,502.

Majda, A. C., and R. M. MacLaughlin (1993), The effect of mean flows on enhanced diffusivity in transport by incompressible periodic velocity fields, *Stud. Appl. Math*, *89*, 245–279.

Majda, A. J., and P. R. Kramer (1999), Simplified models for turbulent diffusion: Theory, numerical modelling, and physical phenomena, *Phys. Reports*, *314*, 237–574.

Mathew, G., I. Mezic, and L. Petzold (2005), A multiscale measure for mixing, *Physica D*, *221*, 23–46.

Mino, H., J. T. Rubinstein, and A. White (2002), Comparison Algorithms for the Simulation of Action Potentials with Stochastic Sodium Channels, *Annals Biomed. Eng.*, *30*, 578–587.

Neher, E., and B. Sakmann (1976), Single channel currents recorded from membrane of denervated frog muscle cells, *Nature*, *260*, 799–802.

Niemiela, J., and K. R. Sreenivasan (2006), The Use of Cryogenic Helium for Classical Turbulence: Promises and Hurdles, *J. Low Temp. Phys.*, *143*, 163–212.

Oberbeck, A. (1879), Űber die wărmeleitung der flűssigekeiten bei berűcksichtigung der strőmungen infolge von temperaturdifferenzen, *Ann. Phys. Chem.*, *7*, 271.

Øksendal, B. (1998), *Stochastic differential equations*, 2 ed., Springer, Berlin.

Ozer, M., and N. H. Ekmekci (2005), Effect of channel noise on the time-course of recovery from inactivation of sodium channels, *Physics Letters A*, *338*(2), 150 – 154.

Peskin, C. S. (2000), Mathematical Aspects of Neurophysiology, handwritten lecture notes, http://www.math.nyu.edu/faculty/peskin/.

Plasting, S. C., and R. R. Kerswell (2003), Improved upper bound on the energy dissipation rate in plane Couette flow, *J. Fluid Mech.*, *477*, 363–379.

Plasting, S. C., and W. R. Young (2006), A bound on scalar variance for the advection-diffusion equation, *Journal of Fluid Mechanics*, *552*, 289–298.

Rehab, H., R. A. Antonia, L. Djenidi, and J. Mi (2000), Characteristics of fluorescein dye and temperature fluctuations in a turbulent near-wake, *Exps. Fluids*, *28*, 462–470.

Richardson, L. F. (1926), Atmospheric diffusion shown on a distance-neighbour graph, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *110*(756), 709–737.

Rowat, P. (2007), Interspike interval statistics in the stochastic Hodgkin-Huxley model: Coexistence of gamma frequency bursts and highly irregular firing, *Neural Comp.*, *19*, 1215–1250.

Saarinen, A., M. L. Linne, and O. Y. Harja (2006), Modeling single neuron behavior using stochastic differential equations, *Neurocomp.*, *69*, 1091–1096.

Salazar, J. P. L. C., and L. R. Collins (2009), Two-particle dispersion in isotropic turbulent flows, *Annu. Rev. Fluid Mech.*, *41*, 405–432.

Scheutzow, M. (1993), Stabilization and destabilization by noise in the plane, *Stochastic Analysis and Applications*, *11*, 97–113.

Scheutzow, M. (1995), An integral inequality and its application to a problem of stabilization by noise, *Journal of Mathematical Analysis and Applications*, *193*, 200–208.

Schmidt, G., I. Goychuk, and P. Hanggi (2006), Capacitance fluctuations causing channel noise reduction in stochastic Hodgkin-Huxley systems, *Phys. Biol.*, *3*, 248–254.

Schneidman, E., B. Freedman, and I. Segev (1998), Ion Channel stochasticity may be critical in determining the reliability and precision of spike timing, *Neural Comp.*, *10*, 1679–1703.

Schumacher, J., K. R. Sreenivasan, and P. K. Yeung (2003), Schmidt number dependence of derivative moments for quasistatic straining motion, *J. Fluid Mech.*, *479*, 221–230.

Shaw, T. A., J.-L. Thiffeault, and C. R. Doering (2007), Stirring up trouble: Multi-scale mixing measures for steady scalar sources, *Physica D*, *231*, 143–164.

Shraiman, B. I., and E. D. Siggia (1994), Lagrangian path integrals and fluctuations in a random flow, *Phys. Rev.*, *49*, 2912–2927.

Skaugen, L., and L. Walloe (1979), Firing behavior in a stochastic nerve membrane model based upon the Hodgkin-Huxley equations, *Acta Physiol. Scand.*, *107*, 343–363.

Spiegel, E. A. (1971), Convection in stars, *Annu. Rev. Astron. Astrophys.*, *9*, 323–352.

Stringano, G., G. Pascazio, and R. Verzicco (2006), Turbulent thermal convection over grooved plates, *J. Fluid Mech.*, *557*, 307–336.

Taylor, G. I. (1922), Diffusion by continuous movement, *Proc. Lond. Math Soc.*, *20*, 196–212.

Taylor, G. I. (1953), Dispersion of Soluble Matter in Solvent Flowing Slowly through a Tube, *Proc. R. Soc. Lond.*, *219*, 186–203.

Thiffeault, J.-L., C. R. Doering, and J. D. Gibbon (2004), A bound on mixing efficiency for the advection-diffusion equation, *J. Fluid Mech.*, *521*, 105–114.

Toschi, F., and E. Bodenschatz (2009), Lagrangian properties of particles in turbulence, *Annu. Rev. Fluid Mech.*, *41*, 375–404.

Verveen, A. A., and H. E. Derksen (1968), Fluctuation phenomena in nerve membrane, *Proc. IEEE*, *56*, 906–916.

Verzicco, R., and K. R. Sreenivasan (2008), A comparison of turbulent thermal convection between conditions of constant trmperature and constant heat flux, *J. Fluid Mech.*, *595*, 203–219.

Warhaft, Z. (2000), Passive scalars in turbulent flows, *Ann. Rev. Fluid Mech.*, *32*, 203–240.

Waxman, S. G., J. A. Black, J. D. Kocsis, and J. M. Ritchie (1989), Low density of sodium channels supports action potential conduction in axons of neonatal rat optic nerve, *Neurobiology*, *86*, 1406–1410.

White, J. A., J. T. Rubinstein, and A. R. Kay (2000), Channel noise in neurons, *Trends Neuroscience*, *23*, 131–137.