

Reverse-time inference of biological dynamics

Dissertation

for the award of the degree
“Doctor rerum naturalium”

Division of Mathematics and Natural Sciences of the Georg August University
Göttingen within the doctoral program GGNB of the Georg August University School of
Science (GAUSS)
submitted by

Nicolas Wolfgang Lenner

from Regensburg

Göttingen 2019

Thesis committee:

- Prof. Dr. Fred Wolf (Thesis supervisor and 1st reviewer)
Max Planck Institute for Dynamics and Self-Organization
- Prof. Dr. Jörg Großhans (2nd reviewer)
Faculty of Medizin, University of Göttingen
- Prof. Dr. Jörg Enderlein (Thesis cosupervisor)
Third Institute of Physics, University of Göttingen

Further members of the examination board:

- Prof. Dr. Stefan Klumpp,
Third Institute of Physics, University of Göttingen
- Prof. Dr. Ulrich Parlitz,
Max Planck Institute for Dynamics and Self-Organization
- Prof. Dr. Michael Wibral,
Faculty of Biology and Psychology, University of Göttingen

Date of oral examination: November 13th, 2019

Contents

Contents	1
1 Preface	3
1.1 Synopsis	3
1.2 Content	4
2 Introduction	7
2.1 Stochastic dynamics in biological systems	7
2.2 Inference of stochastic dynamics in biological systems	9
2.3 A revolution in the imaging technology	11
2.4 Parallel dynamics in complex systems	14
2.5 Reverse Time Stochastic Processes	15
3 Reverse-time inference of biological dynamics evolving towards target-states	17
3.1 Introduction	17
3.2 Terminal pseudo forces and the mixed nature of TSA ensembles	18
3.3 Reverse time Fokker-Planck and SDE for the TSA ensemble	20
3.4 The TSA ensemble close to the target state	22
3.5 Inferring actomyosin turnover from the terminal dynamics of cytokinesis	24
3.6 Discussion	25
4 The mathematical theory of target state alignment	27
4.1 Stochastic dynamics setting	27
4.2 Time reversal and alignment	28
4.3 Normalized TSA dynamics	43
4.4 Exactly solvable reverse time dynamics	45
4.5 TSA dynamics close to the target state	52
4.6 Small L and weak noise approximation – Force and Noise driven TSA dynamics .	57
5 Testing TSA model inference and distinction on a simple model of cytokinesis	69
5.1 A model of cytokinetic ring constriction with distinct effective force laws	69
5.2 Reverse-time ensemble path inference	72
6 Understanding cell contact constriction in epithelial morphogenesis through data driven reverse-time inference	75
6.1 Introduction	75
6.2 Biophysical models	81
6.3 From modeling to inference – a paradigm shift	83
6.4 Models of increasing complexity	88
6.5 Methods	103
7 Generalizations of the theory of target state alignment	107
7.1 Heterogeneity in the realizations of the strength γ of the force law	107
7.2 Generalization to multiplicative noise	109

7.3	Ensemble equivalence of forward and reverse time ensembles	123
7.4	Insights on meander processes	127
8	Conclusion	131
	Bibliography	133
	Acknowledgements and CV	143
	Acknowledgements	143

Preface

1.1 Synopsis

Emergent biological phenomena can often be characterized as the endpoint of directional development. Such directional processes with functionally important target states are spread over all biologically relevant scales. A prominent example are decision processes where the outcome is easy to observe as an action, but the underlying stochastic biological dynamics that lead to the observed behavior are hard to infer. Another example is the fixation of a gene or phenotype in the course of evolution. Both examples have in common that initial conditions of the underlying processes are experimentally either in-accessible or ill defined. Typically such dynamics can not be traced back to a single observable event but emerge from a background of complex ongoing dynamics. While their onset is often not detectable, such dynamics aggregate over time into a collective phenomenon with its signal most apparent close to the target state. If such a collective phenomenon can be summarized in a measurable mesoscopic collective variable, then at least close to the target state the formulation of an effective dynamics approach is possible. This time dependent variable sheds light on the relevant mechanisms which trigger the observed phenomena. Due to the intrinsic stochasticity of any time dependent biological observable, a single time dependent observation is not sufficient to characterize the effective dynamics. Instead, a set of independently recorded sample paths is required to form an ensemble, evaluate its statistics, and characterize the process under study. For the here studied class of directional effective dynamics with well defined target states, I propose to analyze the ensemble in its natural frame of reference, i.e. to align all sample paths to the target state and infer the dynamics in reverse time. The target state then becomes the initial condition of the newly formed reverse-time ensemble.

Target state alignment (TSA) allows to determine whether and when such a system can be characterized by a single collective dynamical variable. Target state alignment however also implies that sample paths are selected based on a condition. This conditioning of the dynamics inevitably introduces a bias in the statistics and therefore leads to the inference of spurious forces. The here derived theory allows to separate these spurious forces from the true underlying biological dynamics.

The characterization as an ensemble allows for the first time to predict statistical properties of experimentally acquired TSA ensembles. I derive the universal low-noise and short term behavior of TSA ensembles. In addition to the intrinsic stochasticity of each sample path realization, biological processes are never identical copies of each other. They differ slightly in the underlying microscopic machinery which is summarized in the collective variable. The TSA theory allows to quantify this heterogeneity between different realizations independently of the intrinsic stochasticity of the dynamic. The third source of stochasticity relevant for the study of biological dynamics is the measurement noise. Often excluded in theoretical studies,

I show how various known techniques to access the measurement error directly from the data can be combined with the theory of TSA ensembles. The theory of target state alignment thus establishes a transparent mathematical foundation to quantitatively study directional biological processes.

In collaboration with the Großhans lab, I apply this theory to study cell contact constriction of intercalating cells during *Drosophila* germband extension. At this developmental stage and directly before the onset of germband extension, the fly embryo is composed of a fairly homogeneous single layered football shaped epithelium. The then starting germband extension is a key process in the reshaping of the developing fly embryo. It is a prototypical process of axis elongation which establishes the head-tail (anterior-posterior) body axis of a developing organism. This directional tissue elongation is driven by cell intercalation, that is, cells seemingly squeeze in between each other effectively shortening the tissue in one direction while elongating the other direction. Cell intercalation in *Drosophila* is driven by junction shrinkage of anterior posterior cell interfaces, its resolution into a vertex of four neighboring cells, and the subsequent junction formation in orthogonal direction. Here I study the process of junction shrinkage as a directional process with the target state defined as the 4-vertex. The tissue is treated as an effective two-dimensional mesh of junctions linked by vertices. The collective variable is the contact length of the constriction cell contact in this 2 dimensional representation.

The TSA ensemble approach allows to infer an effective dynamical model of junction shrinkage. I apply the reverse time inference scheme to an ensemble of ~ 1000 constriction events. To my knowledge this is the first time that the full time dependent statistics of such a non-stationary non-equilibrium biological process are used to systematically test different model classes of increasing complexity. If such an effective model can be inferred, it constrains the modeling space of more involved biophysical models. If it is possible to infer such a simple effective model which predicts all observed statistics every more complex model must at least have the same predictive power.

The most simple dynamical model I find describes cell contact constriction as elastic on short time scales and viscous on the timescale of constriction. It predicts all accessible statistics including mean, variance and two-time covariance. The inferred elastic fluctuations occur around an assumed equilibrium position. Presumably due to junction remodeling this equilibrium position then changes over time. The increasing speed with which the equilibrium position changes can probably be linked to feedback mechanism reported in the literature.

1.2 Content

Biological processes are often well represented within the mathematical framework of stochastic dynamics. In chapter 2, I provide a quick review on time-dependent stochastic processes and their inference from biological data. Due to a revolution in the fluorescence based imaging technology, today, the development of whole organisms can be imaged in 3d and in single cell resolution. After a brief review on fluorescence microscopy, I conclude the introductory chapter with a recapitulation on the stochastic dynamics literature on which my derivation of the TSA framework builds.

TSA ensemble analysis provides a powerful tool for model analysis, inference and distinction. It allows to study directional dynamics irrespective of any unknown initial conditions and thus without perturbation of the system under study. In chapter 3, I introduce the key concepts necessary to understand the mathematical construction of TSA ensembles. I discuss the universal limit of force and noise driven dynamics close to the target state.

Chapter 4 can be read as an independent building block where the key results of reverse time TSA inference, introduced in chapter 3, are derived rigorously step by step. I provide analytical

results for reverse time TSA ensemble dynamics and compare them to forward simulations after target state alignment.

With the intuition formed by these results, I show in chapter 5 the applicability of this approach to a simple model of cytokinetic ring constriction. I provide the mathematical framework to melt the TSA approach with an ensemble path inference scheme with emphasis on numerical stability.

Chapter 6 starts with an introduction on cell intercalation as an important building block of the evolutionary preserved phenomenon of convergent extension. It is followed by a review on the current biological understanding of contact constriction during germband extension and how such dynamics are currently modeled using biophysical intuition. The intuition about TSA ensembles invoked in chapter 3 is then used to infer the most simple effective model of cell contact constriction during germband elongation in *Drosophila* morphogenesis. Guided by the experimental evidence the theory is stepwise extended to explore whether a multiplicative noise model, memory in the noise or heterogeneity in the dynamics is capable to explain the observed time-dependent statistics. This chapter concludes with preliminary results on myosin dynamics during contact constriction and how these dynamics can quantitatively be linked to the junction dynamics.

In its scope, chapter 7 goes beyond the results used in chapter 3 - 6. I show, that Langevin bridges, i.e. stochastic dynamics with delta initial and final conditions can equivalently be studied in forward and reverse time. While identical as an ensemble they differ however in their dynamics. I further show how meander processes, i.e. processes, which up to time t do not return to their initial conditions, can be constructed additively from Langevin bridges. While probably known to some experts, both connections have to my knowledge not been linked in a publication. When ever possible, analytical results for reverse time TSA dynamics are obtained and compared to forward simulations after target state alignment.

Introduction

2.1 Stochastic dynamics in biological systems

Biological systems are comprised of a multitude of interacting components spanning several orders of magnitude. Starting from the very small scales of individual molecules, via protein complexes to cells, membranes and whole tissues, one observes noisy fluctuations. Rooted in thermal fluctuations on the small scales, are the cause of fluctuations on larger scale often active, energy consuming processes and the response of their environment. Independent of the details and the exact mechanisms, a prevalent observation for biological systems is a non negligible noise level on all biological relevant scales. In biological systems constitutive components are typically densely packed. Individual molecules and higher order structures do not move independently, but continuously interact with each other. These dynamical interactions lead to an almost instantaneous dissipation of accelerated dynamics at the typical time scales of observations. Dynamics are thus almost exclusively observed in the overdamped limit.

This heuristic description of biological systems can be formalized in the framework of statistical physics. In this framework each particle is characterized with its position and momentum. The whole system is summarized with its Hamiltonian, comprised of the kinetic and potential contributions of all particles. The Hamilton equations of motion describe the dynamics.

While a biological system with all its details is in principle described by its Hamiltonian, are the dynamics of interest, typically summarizable as a collective phenomenon, well represented by one or a few degrees of freedom. This idea has been exploited in a mathematically involved formalisms which separates the effective dynamics of interest from the residual system^{9,115,117,185}. This residual degrees of freedom then act as a stochastic bath in which the effective dynamics evolve. For most biological systems it is furthermore well justified to assume that accelerating forces have dissipated on the time scale of the observation due to the dominant viscous damping of the surrounding. Assuming in a last simplifying step that the noisy fluctuations occur on a smaller timescale than the effective dynamics of interest, one arrives at a simple representation of effective stochastic biological dynamics. This equation is called an (overdamped) Langevin equation

$$d\hat{L}(t) = f(\hat{L}) dt + \sqrt{D} dW_t , \quad (2.1)$$

here stated for a single dynamical degree of freedom \hat{L} . It is composed of a deterministic drift $f(\hat{L})$ and a stochastic forcing dW_t term, which for simplicity, is independent of \hat{L} . The stochastic forcing is called a Wiener increment dW_t and defines a zero mean $\langle dW_t \rangle = 0$, delta correlated $\langle dW_t dW_{t'} \rangle = \delta(t - t')$ Gaussian process. The noise strength is given by the diffusion constant D . With the Langevin equation, a sample path representation of stochastic biological dynamics, reduced to the core elements of an effective force and a stochastic component, is achieved.

In the limit of infinitely many sample paths, generated by Eq. (2.1) and equipped with a common initial distribution, the resulting ensemble can equivalently be described by the (forward) Fokker-Planck equation (FPE)

$$\frac{\partial}{\partial t} P(\widehat{L}, t | \widehat{L}_0, t_0) = -\frac{\partial}{\partial \widehat{L}} f(\widehat{L}) P(\widehat{L}, t | \widehat{L}_0, t_0) + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t | \widehat{L}_0, t_0), \quad (2.2)$$

where $P(\widehat{L}, t | \widehat{L}_0, t_0)$ denotes a with time evolving probability density conditioned on the initial condition (\widehat{L}_0, t_0) . The first term denotes the drift and the second the diffusion. Both terms are in their effect identical to the respective terms in the Langevin equation. For example setting $f(\widehat{L}) = 0$ and assuming vanishing probabilities at $\pm\infty$ one finds

$$P(\widehat{L}, t | \widehat{L}_0, t_0) = \frac{1}{\sqrt{2\pi Dt}} e^{-\frac{(\widehat{L} - \widehat{L}_0)^2}{2Dt}}, \quad (2.3)$$

which is the continuous representation of diffusion under Gaussian white noise.

The evolution of the transition probability can equivalently be obtained from the backward Fokker-Planck equation

$$\frac{\partial}{\partial t} P(\widehat{L}_f, t_f | \widehat{L}, t) = -f(\widehat{L}) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) - \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}_f, t_f | \widehat{L}, t), \quad (2.4)$$

here denoted for transitions to the final state (\widehat{L}_f, t_f) . It is important to note that solving the backward Fokker-Planck equation yields the very same transition probability as the forward Fokker-Planck equation. It does not(!) describe dynamics in reverse time.

In principle, both the Langevin and Fokker-Planck equation can capture the two recurrent themes of biological dynamics. That is, biological dynamics are typically (i) out of equilibrium and (ii) they change over time. To study the temporal evolution of an ensemble, within the Fokker-Planck approach a common initial reference time must be available to which all sample paths align. With well defined initial conditions temporal ensemble properties such as mean and variance can be determined. However, even in in-vitro systems, it is often a technical challenge to synchronize the observed dynamics (or find a common reference onset for independently recorded sample paths), to yield an ensemble of observations with identical temporal onset.

Giving up on a temporal description, but not on the non-equilibrium properties of biological systems, a natural simplifying assumption replaces the time dependent evolution of the FPE by a constant flux. The classical realization of this state is provided by constant flux boundary conditions, where the influx and outflux of the system are identical. A for the main text important generalization of this approach, lifts the confinement between two boundaries and replaces the insertion boundary by a source term with insertion probability $P^{\text{in}}(\widehat{L})$. Under this assumption the FPE Eq. (2.2) with absorbing boundary at \widehat{L}_{ts} simplifies to

$$\begin{aligned} -\lambda P^{\text{in}}(\widehat{L}) &= -\frac{\partial}{\partial \widehat{L}} \left(f(\widehat{L}) Q(\widehat{L}) \right) + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} Q(\widehat{L}) \\ Q(\widehat{L}_{\text{ts}}) &= 0, \end{aligned} \quad (2.5)$$

where $Q(\widehat{L})$ denotes the non-equilibrium stationary state (NESS) density and $\lambda = \frac{D}{2} \frac{\partial}{\partial \widehat{L}} Q(\widehat{L}_{\text{ts}})$ ensures that influx and outflux are balanced. The differential operator on the right hand side of Eq. (2.5) is identical to the right hand side of the forward FPE, while the temporal evolution on the left hand side is replaced by a constant in time. For the NESS density $Q(\widehat{L})$ an analytical expression has been obtained¹⁸². We will recapitulate this solution and its derivation in the main text. An interesting interpretation of this approach is provided in the Langevin, that is

sample path picture. Sample paths which start at \mathcal{A} progress over time until they terminate at \mathcal{B} and are instantaneously reinserted at \mathcal{A} . This approach thus only gives up on a global wall clock time to which all sample paths align but fully preserves the non-equilibrium character of many biological systems. In the next section, I will briefly discuss how the stochastic dynamics approach recapitulated in this section, underlies both classical and modern inference schemes used to characterize biological single trajectory dynamics.

2.2 Inference of stochastic dynamics in biological systems

The inference of stochastic dynamics dates back to the first quantitative diffusion experiments. Two conceptually different approaches have been suggested. In 1908, Perrin tracked individual diffusing particles to test Einsteins⁴² and Smoluchowskis¹⁷⁴ theory of Brownian motion^{122,128}, using the ensemble mean square displacement (MSD)

$$\langle \widehat{L}^2(t) \rangle = \int \widehat{L}^2 P(\widehat{L}, t) d^d \widehat{L} \quad (2.6)$$

for trajectories aligned to a common initial condition, here chosen as $(\widehat{L}_0, t_0) = (0, 0)$ and for d dimensions. For “normal” diffusion, the ensemble MSD can be calculated directly from Eq. (2.6) using the solution of the diffusion equation Eq. (2.11) generalized to d dimensions. One finds

$$\langle \widehat{L}^2(t) \rangle = dDt, \quad (2.7)$$

where we follow the notation convention used in the stochastic process literature such as Gardiner⁵⁵. Note, that often, and especially in the experimental biophysics literature, a factor 2 occurs in front of the diffusion constant D .

Six years after Perrin, Nordlund characterized the diffusion of single particles from their trajectories^{121,122}, using the time averaged MSD

$$\overline{\delta(\Delta)} = \lim_{T \rightarrow \infty} \frac{1}{T - \Delta} \int_0^{T-\Delta} (\widehat{L}(t + \Delta) - \widehat{L}(t))^2 dt \quad (2.8)$$

for lag intervals Δ and total observation time T . For normal diffusion, and in the limit of $T \rightarrow \infty$, ensemble average and time average exactly coincide. The ergodicity assumption of classical equilibrium statistical physics is fulfilled.

In summary, the first approach evaluates diffusion from its ensemble properties, the second from the temporal evolution of individual time dependent observations. With recursion to the stochastic dynamics theory reviewed in section 2.1, this illustrates the two discussed different representations, either as Fokker-Planck or Langevin equation.

Interestingly, in many biological systems, the measured ensemble MSD does not linearly depend on t ^{68,114,143}. Instead, one finds a powerlaw t^α , with α in the range $0 < \alpha < 1$. Such “anomalous” diffusing processes are classified as sub-diffusive. Unlike the ensemble MSD, the time averaged MSD in such systems typically still depend linearly on the lag Δ ¹²². Several models have been suggested to explain this statistical phenomenon of ergodicity breaking, including spatial heterogeneity or temporal correlations in the diffusion term^{113,114}.

Moment based approaches have further been used to classify Langevin dynamics in the overdamped limit as defined in Eq. (2.1). The classical version of this approach identifies the force with the ensemble averaged spatial change $\Delta \widehat{L}$ in the drift

$$f(\widehat{L}) = \lim_{\Delta t \rightarrow 0} \frac{\langle \Delta L \rangle}{\Delta t} \quad (2.9)$$

within one time step Δt , and for sample paths passing through \widehat{L} ^{53,72,156}. Assuming that the force law does not change over time and the dynamics are memory free, the average is with respect to all increments which originate at \widehat{L} independent of t .

The diffusion constant D is similarly defined as averaged short term MSD

$$D = \lim_{\Delta t \rightarrow 0} \frac{\langle (\Delta L)^2 \rangle}{\Delta t} \quad (2.10)$$

of all observed incremental steps^{53,72,156}. In this review section, we only consider spatially homogeneous diffusion with no dependence on \widehat{L} and thus extend the average over all starting positions \widehat{L} . Importantly, the expression for the diffusion constant is only valid in the limit $\Delta t \rightarrow 0$. For processes observed in discrete time, corrections can be obtained using maximum likelihood based estimators as detailed below.

In general, these classical time and ensemble average approaches face several problems when applied to biological systems⁹⁵. (i) First and foremost, time averages are clearly not suitable under with time change conditions as they often occur in vivo, i.e. they fail if the system is non-stationary. (ii) While time averages introduce temporal coarse graining, are ensemble averages, of the kind defined in Eq. (2.9) and Eq. (2.10), based on spatial coarse graining. Only if enough sample paths enter the interval $[\widehat{L}, \widehat{L} + \Delta\widehat{L}]$, can Eq. (2.9) and Eq. (2.10) faithfully be evaluated. The chosen bin width $\Delta\widehat{L}$ defines the coarse graining. (iii) The limit assumption $\Delta t \rightarrow 0$ is often inadequate to describe data sampled in discrete timesteps. For processes observed in discrete time, the description used in Eq. (2.9) and Eq. (2.10) breaks down and corrections of $\mathcal{O}(\Delta t)$ become important.

In the framework of likelihood based inference, all these problems can formally be addressed. For example, corrections for the expressions Eq. (2.9) and Eq. (2.10) can be derived under the assumption of Gaussian transition probabilities, where the increments are distributed according to

$$P(\Delta\widehat{L}) = \frac{1}{\sqrt{2\pi D\Delta t}} e^{-\frac{(\Delta\widehat{L} - f(\widehat{L})\Delta t)^2}{2D\Delta t}}. \quad (2.11)$$

This expression can directly be read off from the Langevin equation Eq. (2.1), bringing the force term to the left side of the equation. The resulting form implies, that fluctuations around the drift corrected increment are drawn from a Gaussian. With the further assumption of N_{ind} independent measurements, the likelihood function \mathcal{L} , is the product over independent Gaussian increment transition probabilities. For numerical stability typically the logarithm is evaluated

$$\log \mathcal{L} = - \sum_{j=1}^{N_{\text{ind}}} \frac{(\Delta\widehat{L}^{(j)} - f(\widehat{L}^{(j)})\Delta t)^2}{2D\Delta t} - \frac{N_{\text{ind}}}{2} \log(2\pi D\Delta t). \quad (2.12)$$

Assuming that all processes are memory free, the likelihood function can be extended to include all increments at all measured times. Taking the derivative with respect to D and f to evaluate the maxima, the resulting maximum likelihood estimators are¹⁰⁸

$$\widehat{f}(\widehat{L}) = \frac{\sum_{j=1}^{N_{\text{ind}}} \Delta\widehat{L}^{(j)}}{N_{\text{ind}}\Delta t} \quad (2.13)$$

and

$$\widehat{D} = \frac{\sum_{j=1}^{N_{\text{ind}}} \left(\Delta\widehat{L}^{(j)} - \widehat{f}(\widehat{L}^{(j)})\Delta t \right)^2}{N_{\text{ind}}\Delta t}. \quad (2.14)$$

While the estimator for the force is identical to the classical estimate stated in Eq. (2.10), the equation for the diffusion constant Eq. (2.9) is corrected by the estimator of force induced drift $\hat{f}(\hat{L})$.

For single trajectory measurements, the likelihood inference scheme can be generalized to include the full history of each sample path. The respective stochastic process is defined by its transition probability $P(\hat{L}_n, t_n | \hat{L}_{n-1}, t_{n-1}, \dots, \hat{L}_0, t_0)$. To simplify the analysis it is often assumed that the current state of a stochastic process only depends on the previous position. The dynamics are thus assumed to be memory free and the Markov property

$$P(\hat{L}_n, t_n | \hat{L}_{n-1}, t_{n-1}, \dots, \hat{L}_0, t_0) = P(\hat{L}_n, t_n | \hat{L}_{n-1}, t_{n-1}) \quad (2.15)$$

holds. The probability of an observed sample path with N data points can then be written as the product $\prod_{i=1}^N P(\hat{L}_i, t_i | \hat{L}_{i-1}, t_{i-1})$ over all transition probabilities. Studying n_{ens} independently recorded sample paths of equal duration $N\Delta t$, the likelihood

$$\mathcal{L} = \prod_{j=1}^{n_{\text{ens}}} \prod_{i=1}^N P(\hat{L}_i^{(j)}, t_i | \hat{L}_{i-1}^{(j)}, t_{i-1}) \quad (2.16)$$

is evaluated as the product over both.

For small enough forces the Gaussian assumption for the transition probability is often well justified. With increasing Δt , the transition probability in general becomes skewed and starts to deviate from a pure Gaussian. Especially Ait-Sahalia has proposed a powerful approach, which allows to systematically correct for deviations from the Gaussian reference^{2,3}. For sparse sampled data numerical techniques which allow to approximate the incremental transition probabilities have been proposed¹³. In practice, the numerical evaluation of transition probabilities is still computationally demanding. If the estimation pipeline is already computationally highly demanding, the numerical evaluation of transition probabilities can exactly be what breaks the inference scheme⁹⁵.

While the classical approaches were designed to efficiently find estimates for maximum likelihood parameters, it is now possible to sample the full likelihood. This provides access to the moments of the likelihood and thus allows to determine confidence intervals for the inferred parameter.

The likelihood approach however neglects prior knowledge about the inference objective. Sole goal of this “frequentist” inference approach is to provide as good estimates as possible on the “true” underlying parameter. This is contrasted with Bayesian inference, where measurements are updates of our beliefs, given reasonable prior distributions. The resulting posterior distribution $P(\Theta|D) \approx P(D|\Theta)P(\Theta)$ is given as the likelihood $\mathcal{L} = P(D|\Theta)$ times the prior of the parameter $P(\Theta)$, divided by a normalization. In practice the difference between both methods is often that the Bayesian approach is explicit about its assumptions, while the likelihood approach either implicitly includes prior knowledge, using regularizing terms, or assumes flat (constant) priors, where both methods become indistinguishable. In the limit of larger sample data sets, maximum posterior parameters converge to the maximum likelihood parameters^{14,25,89,175}.

2.3 A revolution in the imaging technology

Over the last two decades a revolution in the capabilities of life imaging and massive parallel data acquisition has occurred. Early single trajectory or particle measurements were often only possible under the well defined in vitro circumstances and for only a few sample paths^{121,128}. Unlike these classical experiments, today in vivo single particle tracking in $\propto 1\text{nm}$ resolution in

small volumes is possible⁷, and, on larger scales, the development of whole organisms can be imaged in single cell resolution and with minimal experimental interference. Movies in single cell resolution have been recorded for the embryonic development of *Drosophila*^{92,165}(Fig. 2.1a) and the early stages of zebrafish⁸³(Fig. 2.1b) and *C.elegans*²⁹(Fig. 2.1c) development. For the visualization of such complex and rapidly changing biological systems, large improvements in the available optical microscopy imaging technology were necessary.

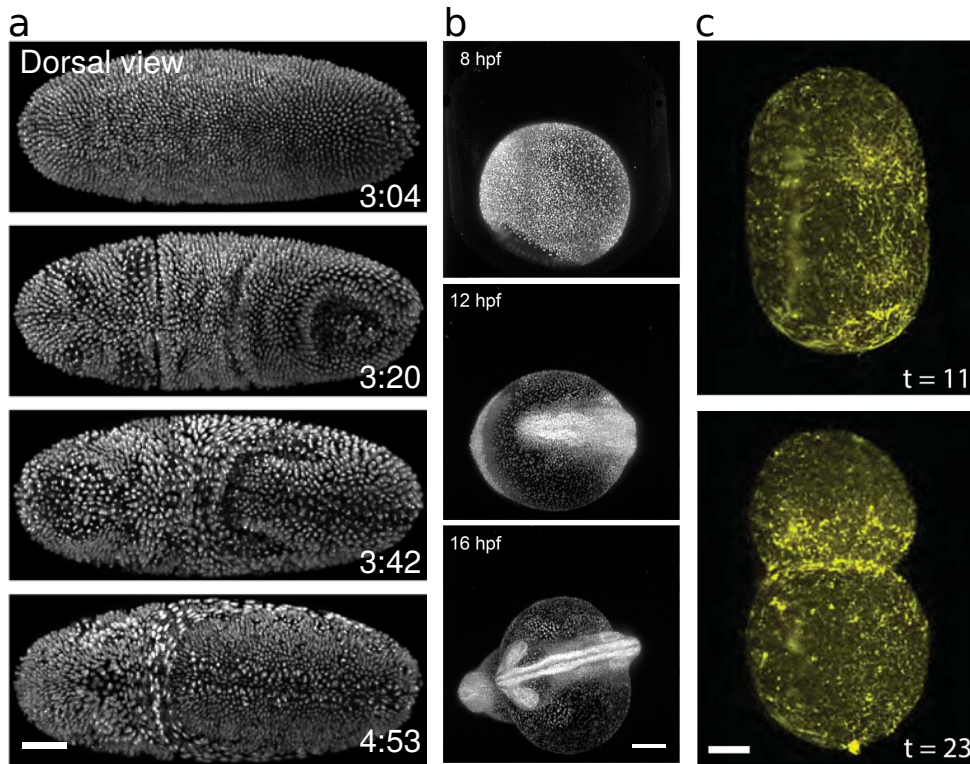


Figure 2.1: Progress in the microscopy techniques enables time-lapse imaging of developmental processes in 3d and all cell resolution. (a): Time-lapse imaging of nuclei in early *Drosophila* development, recorded with a multiview lightsheet setup⁴(reproduced from Amat et. al.⁴). The time inset denotes hours and minutes after egg laying. Scale bar, $50\mu\text{m}$. **(b):** Early zebrafish development, visualized with a GFP-Histone marker¹²⁷ and recorded with multi-view lightsheet microscopy with recordings taken every 3min and over a course of 12h¹⁷⁶(reproduced from Weber et. al.¹⁷⁶). The time inset denotes hours post fertilization (hpf). Scale bar, $150\mu\text{m}$. **(c):** First division of *C. elegans* oocyte, recorded with a lattice lightsheet and with respect to its actin dynamics²⁹(reproduced from Chen et. al.²⁹). Time in minutes with $t = 0\text{min}$ placed during pseudocleavage ingression. Scale bar, $5\mu\text{m}$.

The imaging revolution began with the discovery of the green fluorescent protein (GFP), and the ability to specifically tag proteins of interest^{26,101}. Early experiments illuminated the entire fluorescence tagged system, while simultaneously recording from it¹²². This wide field microscopy termed technique, faces two problems when recording living systems¹⁶⁴. (ii) Illuminating all focal planes simultaneously yields a diminished resolution both due to background noise and light scattering from out of focus planes and directions. (i) For time lapse recordings, illuminating the whole system simultaneously leads to fast photo-bleaching and faces the problem of photo-toxicity. The first problem is addressed in confocal microscopy, which due to an illumination pinhole in front of the light source and the confocal pinhole in front of the camera, only illuminates and records from fluorophores in one spot of the focal plane^{122,164}. In confocal laser scanning microscopy (LSM), this spot is then scanned over the probe^{162,164}. Due to the high scanning frequencies, the laser intensity must be high to excite enough fluorophores during

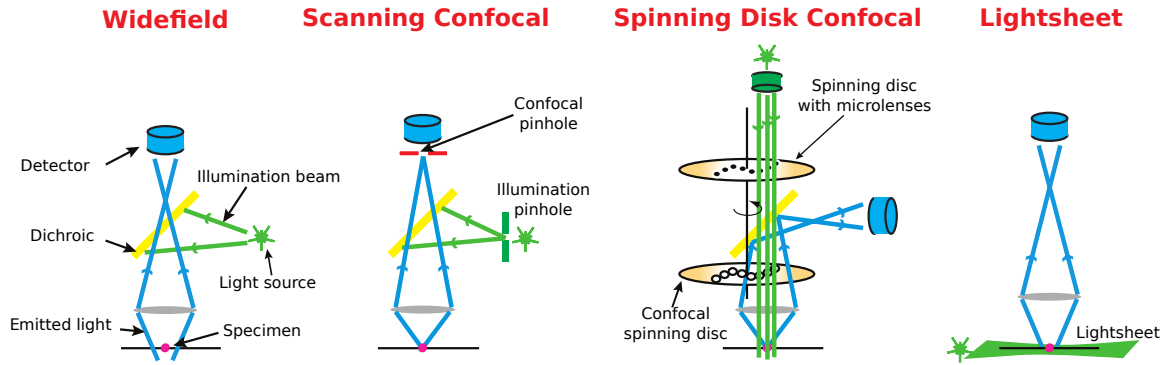


Figure 2.2: Schematic overview of different fluorescence based microscopy techniques. Shown are the illumination (green) and recording (blue) beams including the name giving and essential components (Figure adapted from Stephens and Allan¹⁶²). (Left): Widefield microscopy records from the full specimen simultaneously¹²². (Middle left): In laser scanning confocal microscopy (LSCM) a laser beam is scanned over the probe and a confocal pinhole prevents out of focus light from entering the detector^{162,164}. (Middle right): In spinning disc confocal microscopy (SDCM), a Nipkow disk parallelizes the concept of a confocal pinhole from LSCM. Often, a second in sync spinning disc, placed in the incoming lightpath and equipped with microlenses, is used to realize SDCM as a fully parallelized version of a LSCM⁶¹. (Right): Selective plane imaging microscopy (SPIM) replaces the illumination beam with a lightsheet orthogonal to the detection pathway, which only illuminates the current focal plane¹³². Lightsheets have been realized optically⁷⁶, as digital laser scanning beam⁸³, or as an interfering lattice of bessel beams²⁹. Today, they are often realized with two orthogonal lightsheets to allow for true 3d recordings^{92,132,165}.

the short illumination time per spot⁶¹. While the resolution is increased compared to widefield microscopy, photo-toxicity and photo-bleaching stays a relevant problem^{61,164}.

Spinning disc confocal microscopy (SDCM) addresses the problem of slow single beam scanning and its high illumination intensities per spot^{61,162,164}. In SDCM the single excitation and emission pinhole in the LSM is replaced by up to 1000 pinholes on a rotating disc, which, in current designs, serve both as illumination and confocal pinholes⁶¹. Instead of illuminating one spot at a time up to 1000 spots are simultaneously illuminated and recorded from⁶¹. This allows for a drastic increase in the obtained frame rates. Compared to single beam LSM, it allows to illuminate each individual spot longer, and thus with less intensity, which is believed to reduce both bleaching and photo-toxicity¹⁶⁴. To enhance the resolution of spinning disc microscopy, and avoid losing most of the incoming light at the spinning disc, the incoming light is often guided through another disc equipped with micro-lenses which rotates in sync with the pinhole disc⁶¹. Modern spinning disc microscopes are thus parallel confocal microscopes⁶¹.

In general, confocal imaging techniques can be combined with two-photon approaches^{75,152}. This technique exploits the spectroscopic properties of the used fluorophores. Instead of one-photon excitation, fluorophores are excited by 2 (or more) photons, with longer wavelength¹²². This switch to longer wavelengths then allows for a deeper penetration into the tissue with less scattering and as it seems increased resolution^{122,152}. However, as the likelihood of simultaneously photo absorption is relatively low, high laser peak intensities are required in two-photon microscopy, which is associated with increased photo-toxicity^{83,122}.

The relatively low light efficiency of confocal microscopy and in consequence, resulting increased bleaching and photo-toxicity, has been addressed with selective plane imaging microscopy (SPIM)⁷⁶. In this already 1902 by Zsigmondy invented¹⁵⁵ and decades later recovered technique^{75,76,173}, a light-sheet illuminates the probe perpendicular to the optical axis, and thus almost only the plane in focus. The pinhole used in confocal microscopy to confine the recorded light in z -direction becomes obsolete⁷⁵. Additionally, the whole plane defined by the light-sheet can be recorded simultaneously. Sequentially illuminating and recording from several planes

has been used to reconstruct 3d images^{76,83}. Combining two perpendicular light-sheets, it is now possible to record a developing organism in 3d and with equal resolution in all spatial directions^{92,132,165}.

The currently standard SPIM microscopy relies on light-sheets with a relatively thick Gaussian intensity profile in z -direction^{29,132}. This problem has been addressed with lattice light-sheet microscopy. This technique replaces the lightsheet by a lattice of Bessel beams, which through interference, yield an adjustable light-sheet²⁹. This constructive version of a light-sheet now allows to combine super resolution techniques with 3d all cell long term imaging in one conceptual setup^{54,102}.

2.4 Parallel dynamics in complex systems

With unprecedented temporal and spatial resolution, complex biological systems can now be studied in-vivo and with minimal experimental interference. These qualitatively and quantitatively new data sets raise new questions for model inference and data analysis. One of the key tasks is to identify meaningful collective variables that summarize the complex observed dynamics. In systems based on fluorescence tagging, as described above, the choice of such collective variables is already partially pre-determined by the current knowledge about the systems and its constitutive elements. In developing organisms, such collective variables could be properties of cells, such as their outlines, apical cell area or volume, or sub-cellular properties such as protein or RNA concentrations.

To distinguish between random fluctuations and the “true” underlying dynamics a single observation of such a collective variable is never enough. Only ensemble properties such as mean, variance and covariance are able to faithfully represent the full dynamics. To acquire such ensemble properties, researchers have in the past typically repeated the experiment a few dozen times to determine the (often stationary) mean. We here assume the perspective that, with the revolution in the imaging technology, complex, with time changing, biological systems can now be analyzed with the same rigor as diffusing single particle trajectories. With the parallel recording of all cells in a developing organism, it is now possible to study repetitions of the same process in different cells and occurring at different times. This in principle allows to form an ensemble comprised of both repeated experiments, and repetitions of the same process within one organism. The number of observations multiplies, and higher order statistics can be evaluated, which allows to distinguish between different models.

For stationary systems, the assembly of such ensembles is straightforward and all individually recorded and independent observations can be used in the averaging procedure. In practice, and for living systems, with time changing asynchronous dynamics, it is surprisingly difficult to define these ensembles. Such systems are typically neither stationary nor in equilibrium. Well defined collective observables often only crystallize over time, while true onsets of the dynamics are buried under a plethora of parallel ongoing processes. Without well defined initial conditions and a common reference time, classical ensemble analysis is prevented.

Although, a clear onset is rarely detectable, developmental processes seem to be directed towards functionally important target states – mandatory steps in a complex scheme that leads to an ever increased complexity of the developing organism. With ill defined initial conditions, these states of convergence provide a natural point of data alignment. Ensembles are then defined in reverse time, using their time to completion as natural system time. Data analysis in reverse time then allows to study biological dynamics in their natural frame of reference, and without perturbation by artificially induced synchronization, to obtain common initial conditions.

2.5 Reverse Time Stochastic Processes

Time reversal of a stochastic process is a difficult task. We here discuss the case of the Langevin equation as defined in Eq. (2.1). Setting the noise to zero ($D \rightarrow 0$), time reversal for this deterministic dynamics is achieved by a simple sign flip of the temporal increment $dt \rightarrow -d\tau$, where $\tau = t_f - t > 0$ measures the reverse time starting from the time after which the process is reversed t_f . To distinguish reverse time dynamics from forward dynamics we denote variables in the forward process as \widehat{L} and in the reverse time dynamics with L . In this notation, the final condition of the forward process (\widehat{L}_f, t_f) is now the initial condition (L_0, t_0) of the reverse time process.

In a stochastic dynamics setting, the operation of time reversal faces two complicating factors. (i) Due to the intrinsic stochasticity of each sample path, the final condition of a stochastic process is not a fixed position, but a random variable. The initial conditions of the reverse time process must be drawn from the final distribution of the forward process. (ii) In addition to the sign flip in front of the drift term, for the noise free case, the entropy production on the single trajectory level must be counteracted. Due to the stochasticity in the dynamics, information about the initial conditions is lost, as the single trajectory entropy¹⁴⁸ increases over time. Time reversed dynamics must therefore evolve uphill along the spatial gradient of the single trajectory entropy. Both aspects, the distribution of final values (i) and the reversal of the entropy production (ii), are captured in the reverse time Langevin equation⁵

$$dL(\tau) = \left(-f(L) + D \frac{\partial}{\partial L} \log \left(P^{\text{fw}}(L, T - \tau | L_f, 0) \right) \right) d\tau + \sqrt{D} dW_\tau, \quad (2.17)$$

where $P^{\text{fw}}(\widehat{L}, t | \widehat{L}_0, t_0)$ denotes the solution of the forward Fokker Planck equation Eq. (2.2) evaluated in reverse time τ , with $T = t_f - t_0$ defined as the duration of the forward process before time reversion. The condition is with respect to the initial condition of the forward process. The first term is the sign flipped forward force, the second a “guiding force” which compensates the entropy production of the forward process. For $\tau \rightarrow T$, the guiding force starts to dominate and ensures that reverse time sample paths approach the initial conditions. The Wiener increment dW_τ obeys the same gaussian statistics as the forward process. A step by step derivation of Eq. (2.17) is provided in chapter 4.

We note, that the reversal of a stochastic process has also been introduced in the stochastic thermodynamics literature^{30,103}. Here, however, not the time is reversed, but a general control parameter is used to reverse an external (time-dependent) protocol¹⁴⁹. While in principle, this approach should be complementary to the here taken Langevin approach towards time reversal, its power lies in a different objective, that is, the characterization of out-of-equilibrium ensemble properties. It is based upon the realization that macroscopic thermodynamic quantities such as heat, work and entropy can also be characterized on the mesoscopic scale^{148,150}, and that these quantities are, even for strongly out-of-equilibrium systems, exactly related to equilibrium system ensemble properties via fluctuation theorems¹⁴⁹.

For the task at hand, the here assumed Langevin approach seems most suitable to characterize and synthesize reverse time ensembles dynamics. This approach is rooted in the wider context of constraint (forward) Langevin dynamics, which i.a. is suggested for the sampling of rare events¹⁰⁴. One example of constrained stochastic dynamics are bridge processes, where the initial and final position of a stochastic process are kept fixed, i.e. the initial and final distribution of a stochastic process are drawn from delta-distributions centered at \widehat{L}_0 and \widehat{L}_f . Until recently, such a bridge process could not be synthesized from a Langevin equation. Majumdar

and Orland have shown, that such a bridge process can be written as ¹⁰⁴

$$d\widehat{L}(t) = \left(f(\widehat{L}) + D \frac{\partial}{\partial \widehat{L}} \log \left(P(\widehat{L}_f, t_f | \widehat{L}, t) \right) \right) dt + \sqrt{D} dW_t . \quad (2.18)$$

With $P(L_f, t_f | L, t)$, obtained as the solution of the backward Fokker-Planck equation Eq. (2.4), the guiding term constrains transitions from the current position to the final state. Note that here the time evolves normally. The apparent similarity of the reverse time Langevin equation and the forward bridge will be discussed in chapter 7.

Interestingly, choosing different guiding probabilities, very different stochastic dynamics can be constructed. One example is the meander process, which defines stochastic dynamics, which up to time t_f , have not returned to its initial starting point. The constrained Langevin dynamics formulation of the meander process reads ¹⁰⁴

$$d\widehat{L}(t) = \left(f(\widehat{L}(t)) + D \frac{\partial}{\partial \widehat{L}} \log \left(\int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}, t) \right) \right) dt + \sqrt{D} dW_t , \quad (2.19)$$

where $P(\widehat{L}_f, t_f | \widehat{L}, t)$ is again the solution of the backward FPE, however now with absorbing boundary conditions implied, to prevent sample paths from returning to the starting point at \widehat{L}_0 . The integral inside the logarithm is performed over all possible final states, and thus lifts the conditioning to a specific final position as included in the definition of bridge processes.

In this work, I derive a consistent theory of reverse time and target state aligned ensembles dynamics. While the reverse time part of the formalism strongly relies on Eq. (2.17), is the here derived mathematical construction of target state alignment, inspired by a, to my knowledge, not known construction of meander processes from an ensemble of bridge processes. Adapting ensemble and path inference techniques from section 2.2, I will further demonstrate how the in general non-equilibrium and non-stationary emergent dynamics of developing organisms can be analyzed within the here provided framework.

Reverse-time inference of biological dynamics evolving towards target-states

3.1 Introduction

Accumulating time-dependent, out-of-equilibrium ensembles from single trajectory measurements is an emerging practice in the study of mesoscopic bio-systems^{95,105,133,151}. Examples include force extension measurements of molecules^{34,38,77,100}, direct observations of active transport^{6,17,24,140} and anomalous diffusion in cells¹²². The process under study is then characterized by time dependent ensemble statistics^{22,51,68,69,72,113,116,171}. Today, the development of whole organisms can be recorded in vivo and in single cell resolution¹⁶⁵. Unlike classical in vitro recordings, in vivo experiments face the problem that the onset of consecutive dynamical processes is not defined by the experimental design. Different dynamical processes overlap or change continuously and the concept of well defined initial conditions, core to dynamical theories, blurs.

Instead, and despite lacking clear initial conditions, many biological processes when observed in vivo are well characterized by their convergent dynamics towards a functional target state, i.e. end point. While undetectable at their onset, such dynamics often aggregate into a collective phenomenon, with only a single collective variable necessary to summarize the effective dynamics short before completion. Such target state directed effective dynamics encompass as diverse processes as cytokinetic ring constriction until cell separation^{129,130,179} (Fig. 3.1A), evidence accumulation during decision making^{19,31,66,134,135,141} (Fig. 3.1B), or the dynamical change of protein concentrations within cells (Fig. 3.1C) which trigger a new process upon reaching a threshold^{33,57,63,119}. Due to the intrinsic stochasticity of these examples a single time dependent recording is insufficient. Instead, only an ensemble of independently recorded sample paths is capable of summarizing the respective dynamics and characterize the dynamics in terms of its moments. This raises the question, how such a time dependent ensemble of directional non-equilibrium dynamics with target states has to be constructed. Because the dynamics are most evident in the vicinity of the target state, I propose to analyze such dynamics in their natural frame of reference, i.e. by target state alignment. The dynamics can then be analyzed in reverse time with the target state used as the new initial conditions.

In this chapter, I present a consistent mathematical theory of target state alignment (TSA) and reverse time data inference. I demonstrate the applicability of this theory on sample data from a simple biophysical model of cytokinetic ring constriction. This chapter provides an intuitive approach to TSA dynamics. A detailed derivation of all stated results is provided in chapter 4.

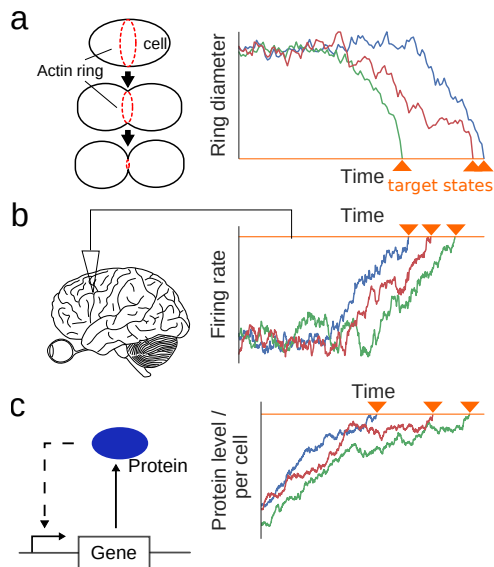


Figure 3.1: Biological processes with target states. (a): Cytokinetic ring constriction until cell separation. **(b):** Neuronal firing rates as a neural correlate of evidence accumulation during decision making. **(c):** Dynamic threshold crossing of protein concentration levels as a mechanism to trigger downstream processes. For each process representative model trajectories (see chapter 4) are shown. The target states (orange line), of full cell separation **(a)**, decision- **(b)** and concentration-threshold **(c)**, and the target state (orange triangle) of each individual sample path are marked.

3.2 Terminal pseudo forces and the mixed nature of TSA ensembles

The intricacies that arise from target-state-alignment become apparent already for the simplest case, i.e. a random (molecular) search process for a target site^{10,60,67,146} (Fig. 3.2a). The TSA ensemble resulting from this undirected motion has intriguing properties. For instance, the proximity of all sample paths to the ensemble mean induced by target state alignment as depicted in Fig. 3.2 suggests that a low noise approximation could be used close to the target state. If the TSA ensemble has an effective stochastic differential equation (SDE) with conventional low noise behavior then the ensemble mean should be the deterministic solution and thus directly reveal the drift term. This however is obviously wrong as the mean, which grows $\sim \tau^{\frac{1}{2}}$, would indicate a diverging force $\propto \frac{1}{L}$ although there is no force at all in the forward dynamics.

How such spurious pseudo forces arise can be understood quite generally for processes $\hat{L}(t)$ satisfying a Langevin equation of the form

$$d\hat{L}(t) = f(\hat{L}) dt + \sqrt{D} dW_t. \quad (3.1)$$

Here $f(\hat{L})$ denotes a deterministic drift term and \sqrt{D} the strength of the fluctuations dW_t . The term dW_t denotes the Wiener process increment with zero mean $\langle dW_t \rangle = 0$ and delta correlated covariance $\langle D dW_t dW_t' \rangle = D\delta(t-t')$. Each observation i consists of a trajectory $\hat{L}_i(t)$ with wall clock time t and length T_i .

The reverse time TSA ensemble is described by a time dependent distribution $R(L, \tau)$ with $L_i(\tau) = \hat{L}_i(T_i - \tau)$ as a function of the reverse time $\tau = T_i - t$. To construct $R(L, \tau)$ two aspects must be taken into account: i.) the underlying dynamics evolve in reverse time; ii.) the lifetime of the trajectories is itself a random variable. Over time less and less trajectories contribute to the ensemble until eventually all trajectories have reached their maximal lifetime at their original point of initiation. $R(L, \tau)$ is thus not normalized and decays with τ . For readability

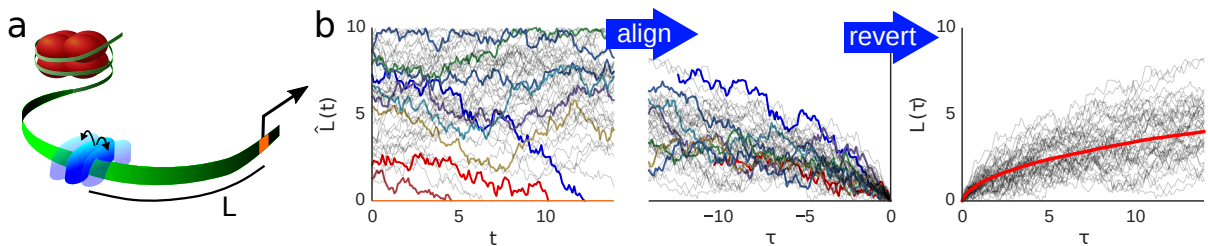


Figure 3.2: Target-state-alignment creates pseudo forces for 1d random target search. (a): Random walk like sliding of transcription factor confined between inaccessible DNA and promoter binding site (orange). (b): (Left): Sample path realizations of a random walk with one reflecting and one absorbing boundary (orange). (Middle): Target state (promoter binding site) aligned sample paths. (Right): Aligned and time reversed ensemble of sample paths. The mean, growing with $\sim \tau^{\frac{1}{2}}$ (red line), indicates the presence of a non-linear alignment force.

we leave out all these dependencies in the notation of this chapter. In chapter 4 a detailed construction of $R(L, \tau)$ is provided and the dependencies stated.

In absence of noise ($D = 0$), time-reversal of Eq. (3.1) is straightforward. All trajectories starting at \hat{L}_0 end at \hat{L}_f at time $t = T$. The reversed process satisfies $L(\tau) = \hat{L}(T - \tau)$. The time-reversed dynamical equation for $L(\tau)$ is $dL(\tau) = -f(L)d\tau$.

In a stochastic system even for a fixed lifetime T , changing the sign in front of the time derivative does not yield the correct time-reversed equation. For instance for a stationary Ornstein-Uhlenbeck processes $f(\hat{L}) = -\hat{L}$ starting from a stationary distribution at $t = 0$, inverting the sign of the drift term $-L \rightarrow L$ results in exponentially diverging trajectories – a completely different behavior as in forward time.

Under the assumption of equal sample path lifetimes T , the correct time-reversal of the SDE Eq. (3.1), is⁵

$$dL(\tau) = (-f(L) + f^G(L)) d\tau + \sqrt{D} dW_\tau, \quad (3.2)$$

where

$$f^G(L) = D \frac{\partial}{\partial L} \log \left(P^{\text{fw}}(L, T - \tau) \right) \quad (3.3)$$

is a guiding force that depends on the solution $P^{\text{fw}}(L, T - \tau)$ of the forward process. For the Ornstein-Uhlenbeck example the correct time-reversed equation can be readily checked. The guiding force $f^G(L)$ ensures that the forward and reverse time dynamics are identical. Hence stating the time-reversed SDE requires knowledge of the solution of the forward process and its initial distribution.

The lifetime T however is not fixed but a random variable denoted as T_i . To assemble the ensemble in reverse time, we partition the trajectories into sub-ensembles $R(L, \tau|T; L_f)$ of fixed lifetime T_i which end at $L_f = \hat{L}_0$ (Fig.3.3). The initial condition of each of the time reversed sub-ensembles is a delta-function $R(L, \tau = 0|T_i; L_f) = \delta(L - L_{\text{ts}})$, where L_{ts} denotes the absorbing target state. Each of these sub-ensembles experiences a different force, because the guiding force Eq. (3.3) depends on the full forward distribution $P^{\text{fw}}(L, T_i - \tau)$ up to time T_i , which is different for every T_i .

To assemble the TSA ensemble, sub-ensembles defined by Eq. (3.2) must be superimposed (Fig.3.3). Three aspects are important. First, each sub-ensemble of lifetime T only contributes to the full ensemble up to this time. Second, the relative weight of trajectories of this lifetime in the full ensemble is given by the hitting time distribution of the forward process $\rho(T|L_f)$, which conditionally depends on the starting points L_f of forward realizations. Third, the relative

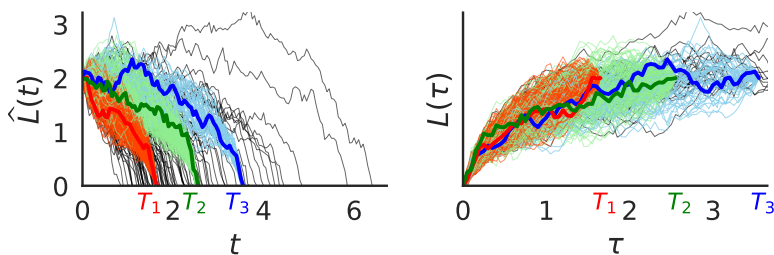


Figure 3.3: Construction of the full aligned time reversed ensemble from sub-ensembles of different lifetimes. (Left): The full forward ensemble is split into sub-ensembles with different completion times T_i . We show three exemplary cases in red (T_1), green (T_2) and blue (T_3). To guide the eye, one sample path per sub-ensemble is highlighted. (Right): After target state alignment and time reversal all sub-ensembles together form the new ensemble $R(L, \tau)$.

weight of each sub-ensemble additionally depends on the initial conditions of the forward process, i.e. on $P^{\text{in}}(L_f)$. Hence, $R(L, \tau)$ for the aligned time-reversed ensemble is

$$R(L, \tau) = \int_{L_{\text{ts}}}^{\infty} dL_f P^{\text{in}}(L_f) \int_{\tau}^{\infty} dT R(L, \tau|T; L_f) \rho(T|L_f) . \quad (3.4)$$

The lower integration limit of the inner integral accounts for the fact that only sub-ensembles that have at least a length of τ contribute to the full ensemble. The outer integral accounts for the distribution of initial values of the forward process. A more in depth discussion of this construction is provided in chapter 4.

We her pause to summarize the implications of our considerations upto this point. First of all, for fixed lifetime sub-ensembles, we can state the general form of the guiding force close to the target state. With $P^{\text{fw}}(L, T - \tau)$ vanishing close to L_{ts} , we can approximate it to leading order by a powerlaw L^δ with $\delta > 0$. The guiding force close to the target state then evaluates to $f^G(L) \propto \frac{1}{L}$, explaining i.a. the above discussed behavior of the mean of the TSA random walk (Fig. 3.2). Second, the construction of the TSA ensemble $R(L, \tau)$ from sub-ensembles of varying lifetimes T_i (and thus varying guiding forces) suggests that in general there is no unique SDE which describes the full reverse time dynamics. Different force laws seem to be active at the same state (L, τ) . The explicit construction of $R(L, \tau)$ as an integral over T Eq. (3.4) however demonstrates that the only dependence of the TSA ensemble on T is given as the constraint, that contributing sample paths must still be alive. We resolve this puzzling observation in the next section.

3.3 Reverse time Fokker-Planck and SDE for the TSA ensemble

Starting from Eq. (3.4), we show in chapter 4 that the dynamics underlying the evolution of $R(L, \tau)$ can be cast into the form of a Fokker-Planck equation (FPE)

$$\frac{\partial}{\partial \tau} R(L, \tau) = - \frac{\partial}{\partial L} ((f(L) + f^{\mathcal{F}}(L)) R(L, \tau)) + \frac{D}{2} \frac{\partial^2}{\partial L^2} R(L, \tau) - P^{\text{in}}(L) \rho(\tau|L) \quad (3.5)$$

with time-dependent sink $-P^{\text{in}}(L) \rho(\tau|L)$, where $P^{\text{in}}(L)$ denotes the distribution of initial positions of the forward dynamics. The sink ensures that the distribution of sample path lifetimes in the reverse time ensemble is the same as in the forward dynamics. Note, that unlike for

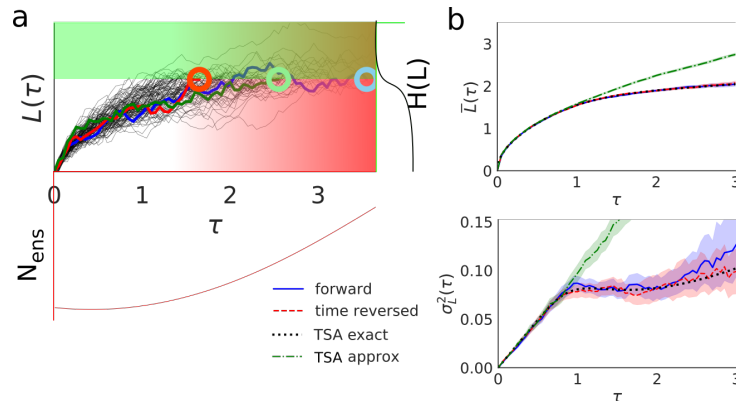


Figure 3.4: TSA FPE and SDE exactly describe the target state aligned ensemble. (a): Schematic depiction of the dependency of the TSA dynamics Eq. (3.5) and Eq. (3.7) on the forward initial condition $H(L)$ and the killing measure $k(L, \tau)$. In the green region, i.e. above the bulk of the initial distribution measured by the sigmoidal $H(L)$, we find $f^{\mathcal{F}}(L) \approx 0$. Below $f^{\mathcal{F}}(L)$ contributes in full strength. In τ direction, an increase in the red gradient indicates that more and more trajectories are killed. The residual white $L - \tau$ plane close to the target state defines the region where $k(L, \tau) \approx 0$ and $H(L) = 1$ holds, and the approximation of well separated initial and final states holds. Circles mark the killing of example trajectories. Superimposed lines and colors are only visualizations of the general TSA FPE and SDE. **(b):** Comparison of the forward (blue), sub-ensemble based (red), exact (black) and approximate (green) reverse time dynamics for $f(L) = -\gamma/L$. Shown are the mean (Top) and variance (Bottom) of all four cases. 95% bootstrap confidence intervals are shown for the cases involving sampling. To exclude numerical inaccuracies due to rarely visited tails of the distribution of completion times $\rho(T|L_f)$, we directly sampled T_i from the numerically obtained hitting time distribution of the forward process. Results were obtained using each 1000 sample path realizations with parameter settings $\gamma = 1$, $D = 0.2$, $\hat{L}_{\text{init}} = 2$.

the sub-ensemble dynamics Eq. (3.2), $f(L)$ is not (!) sign inverted compared to the forward dynamics. The “free energy force”

$$f^{\mathcal{F}}(L) = D \frac{\partial}{\partial L} \log \left(\int_{L_{\text{ts}}}^L dL' e^{-\frac{2\Phi(L')}{D}} H(L') \right) \quad (3.6)$$

captures the combined effect of the total entropy production of all sample paths and the potential difference necessary to overcome for time-reversion of the dynamics up to a position L . The potential $\Phi(L) = \int^L f(L') dL'$ is defined with respect to the sign inverted drift term. $H(L) = 1 - \int_{L_{\text{ts}}}^L P^{\text{in}}(L') dL'$ is a sigmoidal like function which continuously switches from one to zero depending on the forward initial distribution $P^{\text{in}}(L)$. Above the bulk of $P^{\text{in}}(L)$ the free energy force $f^{\mathcal{F}}(L)$ therefore evaluates to zero, and forward and TSA dynamics are indistinguishable in their dynamical law. Below however the free energy force not only reverses the forward dynamics but also adds additional terms, such as the for the guiding force discussed term $\propto \frac{1}{L}$.

We here face the problem that in a target state aligned ensemble our trajectories only have a finite lifetime. For this reason we have to find a mathematical way to include the truncation of stochastic trajectories into our framework. From the theory of reaction diffusion systems¹⁷⁷ we know that a Fokker-Planck equation with sink proportional to the density can be cast into a SDE with killing measure $k(L, \tau)$ proportional to the rate of degradation^{45,70,147}. Adapted to the reverse time FPE for TSA ensembles the corresponding SDE then reads

$$dL(\tau) = (f(L) + f^{\mathcal{F}}(L)) d\tau + \sqrt{D} dW_\tau, \quad (3.7)$$

equipped with a killing measure $k(L, \tau) d\tau = \frac{\rho(\tau|L) P^{\text{in}}(L)}{R(L, \tau)} d\tau$. For TSA ensembles originally comprised of n_{ens} sample paths this implies that the killing measure in each timestep $d\tau$ terminates

$n_{\text{ens}}\rho(\tau|L)d\tau$ of the then still present sample paths. Where this happens is determined by the weight $\frac{P^{\text{in}}(L)}{R(L,\tau)}$. The explicit lifetime T dependence of the sub-ensemble based construction of the TSA-dynamics Eq. (3.4) can thus completely be moved to a time and ensemble dependent boundary condition in the form of a killing measure.

3.4 The TSA ensemble close to the target state

Expressing the dynamics of $R(L, \tau)$ in terms of a single SDE allows to separate the contribution of the forward initial conditions from the pure TSA dynamics close to the target state. This claim is easily verified by inspecting $f^{\mathcal{F}}(L)$ defined in Eq. (3.6). For L sufficiently below the bulk of the forward initial distribution $P^{\text{in}}(L)$ the term $H(L)$, which describes the influence of the forward initial conditions on the reverse time TSA dynamics, evaluates to approximately one. If most of the sample paths have not yet reached the sink at $P^{\text{in}}(L)$, this further implies that almost no sample path has terminated in reverse time. The killing measure $k(L, \tau)$ evaluates to zero and $H(L) \approx 0$. With $k(L, \tau) = 0$ and $H(L) = 1$ the remaining SDE – up to exponentially small corrections – describes TSA dynamics temporally and spatially close to target states and irrespective of any possibly unknown initial conditions. In Fig. 3.4 one example is shown where the ensemble mean and variance of this approximation agree excellent with the exact solution. In chapter 4 we discuss further examples where the approximation holds equally well.

Separating genuine and alignment induced forces

Motivated by the quality of this approximation we searched for a general mapping between forward forces and aligned reverse-time dynamics valid close to its target state. Using Eq.(3.6) with $H(L) = 1$ we find a mapping for power law forces

$$f(\widehat{L}) = -\gamma\widehat{L}^\alpha, \quad (3.8)$$

with $\gamma > 0$, $\alpha \in \mathbb{R}$, target state at $\widehat{L}_{\text{tp}} = 0$ and $\widehat{L}_{\text{in}} \rightarrow \infty$. Under these conditions the free energy force is

$$f^{\mathcal{F}}(L) = \frac{(D\alpha + D) \left(-\frac{2\gamma}{D\alpha+D}\right)^{\frac{1}{\alpha+1}} e^{\frac{2\gamma L^{\alpha+1}}{D\alpha+D}}}{\Theta(\alpha + 1)\Gamma\left(\frac{1}{\alpha+1}\right) - \Gamma\left(\frac{1}{\alpha+1}, -\frac{2L^{\alpha+1}\gamma}{\alpha D+D}\right)}. \quad (3.9)$$

It is valid for $\alpha \neq -1$, with the exceptional connecting case $f^{\mathcal{F}}(L) = \frac{2\gamma+D}{L}$ at $\alpha = -1$. $\Gamma(L)$ is the gamma-function, $\Gamma(n, L)$ the upper incomplete gamma-function and $\Theta(n)$ the Heaviside-function. The derivation is presented in chapter 4.

Noise induced vs. force induced target state arrival

For power law forces, reverse time dynamics fall into two generic classes. For $\alpha \geq 0$ the forward force vanishes at the target state and the system typically terminates *noise induced* (see Fig. 3.5a). Aligned, in reverse time and close to the target state $\widehat{L}_{\text{ts}} = 0$, such an ensemble is indistinguishable from the dynamics of a random walk ($\alpha = 0, \gamma = 0$) with $f^{\mathcal{F}}(L) = \frac{D}{L}$ (see chapter 4 for details). We tested this behavior in Fig. 3.5b for three different noise strengths D . The analytic expression for the mean

$$\bar{L}(\tau) = \sqrt{\frac{8D}{\pi}}\tau, \quad (3.10)$$

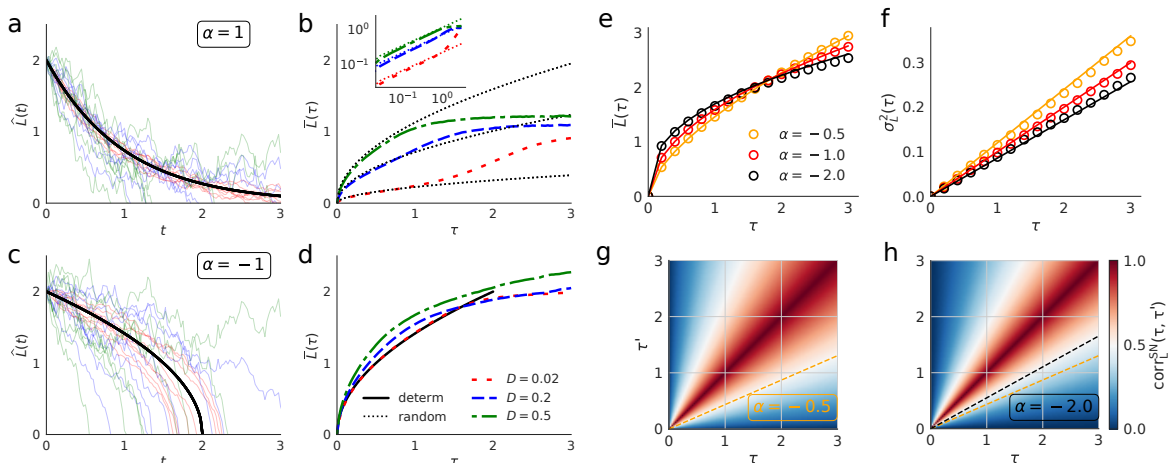


Figure 3.5: TSA ensembles are noise and parameter sensitive, and qualitatively differ for diffusion or drift dominated target-state approaches. (a): For $\alpha > 0$ only noise induced irreversible transitions occur in finite time. (b): Close to their target state, reverse time statistics of noise induced transitions ($\alpha > 0$) are indistinguishable from random fluctuations and thus α -independent. The mean grows $\sim \tau^{\frac{1}{2}}$ (random-walk-like, Eq. (3.10), dotted) independent of the noise strength D as shown in both the log-log inset and main plot. (c): The hitting time of force driven irreversible transitions ($\alpha < 0$) scatters around the deterministic solution (black line). (d): The mean of the aligned reverse time ensemble (dashed line) for $\alpha < 0$ deviates positively from the deterministic solution with increasing noise strength D . Together (c) and (d) motivate a small noise moments expansion. (e),(f): The small noise expansions of mean Eq. (3.12) and variance Eq. (3.13) (lines) agree with the numeric evaluation of the reverse-time TSA dynamics with $f^{\mathcal{F}}(L)$ Eq. (3.9) (circles). (g),(h): The small noise two-time correlation Eq. (3.14) solely depends on the power law exponent α . Linear equi-correlation lines (dashed) allow for an easy distinction. For all plots we chose $\gamma = 1$ and $D = 0.2$ if not stated differently.

of the aligned reverse-time random walk almost perfectly matches the mean of the aligned reverse-time data for the case $\alpha = 1$ and close to $\widehat{L}_{ts} = 0$. It starts to deviate for $Lf(L) > D$ where the approximation is expected to break down (see chapter 4 for further details). Note that Eq. (3.10) also explains the result found for the random target search (Fig. 3.2).

For $\alpha < 0$ the termination of the forward dynamics is *force induced* as $f(\widehat{L})$ diverges at the absorbing boundary. The corresponding reverse time dynamics exhibit the genuine (sign inverted) force law dependence plus corrections proportional to the noise strength D . Fig. 3.5c,d shows exemplary sample paths for the case with $\alpha = -1$. It demonstrates the gradual deviation of the mean from the deterministic solution with increasing noise strength D .

Further insights can be obtained by examining the small L weak noise regime of the free energy force Eq. (3.9) derived in chapter 4. In this regime the reverse time SDE for $\alpha < 0$ simplifies to

$$dL(\tau) = \left(\gamma L^\alpha - \frac{\alpha D}{L} \right) d\tau + \mathcal{O} \left(\frac{D^2}{L^{2+\alpha}} \right) + \sqrt{D} dW_\tau. \quad (3.11)$$

This shows that for all $\alpha < 0$ the leading order correction to the time reversed genuine force is random-walk like ($\sim D/L$), modulated in its strength by the power law exponent α . Despite its simplicity Eq. (3.11) can not be solved analytically for general $\alpha < 0$. Adopting the low noise expansion for moments of ordinary SDEs (Gardiner chapter 6.2⁵⁵), we expand Eq. (3.11)

around its deterministic solution ($D \rightarrow 0$) for small D and obtain

$$\bar{L}(\tau) = ((1 - \alpha)\gamma\tau)^{\frac{1}{1-\alpha}} + D \frac{(7\alpha - 3)((1 - \alpha)\gamma\tau)^{\frac{\alpha}{\alpha-1}}}{4(3\alpha - 1)\gamma} \quad (3.12)$$

$$\sigma_L^2(\tau) = D \frac{1 - \alpha}{1 - 3\alpha} \tau \quad (3.13)$$

$$\text{corr}_L(\tau, \tau') = \left(\frac{\min[\tau, \tau']}{\max[\tau, \tau']} \right)^{\frac{3}{2} + \frac{1}{\alpha-1}} \quad (3.14)$$

for mean, variance and two-time correlation function up to order D (see chapter 4 for details). In Fig. 3.5e,f we compare this approximation to simulations and find excellent agreement.

Motivated by their high accuracy we propose Eqs. (3.12)-(3.14) as basis for fast inference schemes: Given sufficiently good statistics, α can be read off directly from a 2d correlation plot using Eq. (3.14) (see Fig. 3.5g,h). Once α is identified, D can be extracted from the variance Eq. (3.13). The force strength γ follows directly from a α and D constrained fit of the mean Eq. (3.12). Notably this is quite distinct to a normal low noise approximation in which the mean would be the deterministic solution and thus only depend on γ .

3.5 Inferring actomyosin turnover from the terminal dynamics of cytokinesis

To demonstrate the inference of biological dynamics from TSA ensembles we set up a biophysical model of cytokinetic ring constriction which has all ingredients that one expects to complicate the correct identification of directional biological dynamics. In biological systems forces exertion is mediated by myosin motors. The contractile behavior of the actomyosin ring can exhibit very different types of concentration-tension relationships depending on the regime of molecular turnover. Three qualitatively different mechanisms are conceivable. First, the effective constriction force is proportional to the myosin concentration along the cable of length L . The shorter the cable gets, the more myosin per ring perimeter accumulates. The constriction force increases $\propto 1/L$ ($\alpha = -1$), assuming a low myosin turnover rate. A similar reasoning holds for the second scenario. Here, the otherwise individually occurring force exerting molecules, e.g. actin, need a partner to exert forces. The effectively generated constriction force is $\propto 1/L^2$ ($\alpha = -2$). The third mechanism is favored for high turnover rates of the force generating molecules. Independent of L , force molecules are under this scenario present in constant concentration. The constriction force is constant ($\alpha = 0$).

To test our inference scheme and to generate artificial cytokinetic sample paths we fall back on the Zumdieck model¹⁸⁴, which covers all three scenarios as limiting cases. Both the details of the model and the analytic limits are covered in chapter 5. The here used generalization describes the full process of constriction, starting from its equilibrium perimeter via a transition regime towards the final regime which leads to the separation. Internal fluctuations of the ring perimeter are modeled as white noise. After target state alignment we obtain a prototypical TSA ensemble with contributions from both biological and guiding forces. Its target and equilibrium state are well separated and we can hope for the dynamics to assume a pure force law short before cell separation.

For the quantitative inference of the most likely constriction scenario hidden in the forward simulation of the Zumdieck model, we evaluated the path-ensemble likelihood for the TSA ensemble forces Eq. (3.6) with $H(L) = 1$ and for each of the three scenarios with $\alpha = 0, -1, -2$. The details of this approach are stated in chapter 5. The (true) case $\alpha = -1$ was clearly singled out with the highest maximum likelihood value (Fig. 3.6). Further baseline scenarios are tested and unambiguously identified in chapter 5.

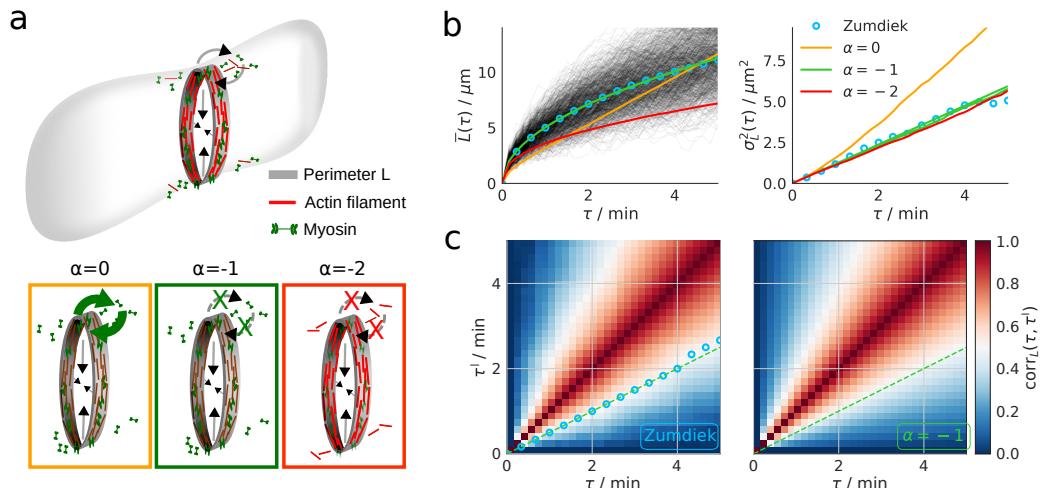


Figure 3.6: Accurate inference of actomyosin turnover from dynamics in the noise extended Zumdieck model¹⁸⁴ of cytokinetic ring constriction. (a): Schematic representation of three different cytokinesis scenarios assuming constriction-forces proportional to molecule concentrations along the ring: Constant shrinkage due to high myosin turnover ($\alpha = 0$, orange), $1/L$ -shrinkage with stalled myosin turnover ($\alpha = -1$, green) and $1/L^2$ -shrinkage for actin-pair driven constriction with stalled turnover ($\alpha = -2$, red). (b): Using path-ensemble maximum likelihood inference (chapter 5) the dynamical law leading to constriction can unambiguously be inferred from realizations of the Zumdieck model (grey lines, blue circles). We visually confirm the inferred underlying force law ($\alpha = -1$) by numerically evaluating mean and variance for the reverse-time TSA dynamics for each of the three inferred maximum likelihood parameter sets ($\gamma_{\text{ML}}^\alpha, D_{\text{ML}}^\alpha$). (c): The two-time correlation of the Zumdieck model and the correlations of the inferred dynamics for $\alpha = -1$ are in perfect agreement. To facilitate the comparison we show one equi-correlation line for both (Zumdieck: blue circles, $\alpha = -1$: green dashes).

Without considering the peculiarities of TSA ensembles several mistakes in interpreting the sample paths might have occurred. First, a scenario of constant force could erroneously be classified as one driven by a divergent force which we now understand as the guiding force correction of the form D/L . Second, the strength of the force for the myosin dominated case could be overestimated as both genuine and guiding force are proportional to $1/L$. Third, for the actin dominated case an erroneous crossover from $1/L$ to $1/L^2$ might be inferred. In contrast, the TSA methods give the genuine force law in all three cases.

3.6 Discussion

In this chapter, we develop the first consistent statistical theory that allows us to infer directed biological processes by target-state-alignment. We show whether and when such a dynamic can be represented by a single SDE and how spurious forces, which inevitably arise due to target-state-alignment can be separated from genuine biological forces. To provide intuition for the effect of target-state-alignment on data we derive the universal low-noise and short-term behavior of TSA ensembles. The biophysical applicability is demonstrated for a model of cytokinetic ring constriction where many complicating factors are present which might hamper the correct identification of directional biological dynamics.

In the literature target-state-alignment is typically avoided in favor of well established but less informative approaches based on hitting time distributions⁵⁷. For this analysis only the last data point from a full trajectory is used and thus most of the information in the data is discarded. Its explicit dependence on (possibly ill defined) initial conditions makes this approach further error prone. Nevertheless, target-state-alignment has been used as means of data analysis

in various fields. It has been employed to infer entropic force differences of DNA in partially blocked spaces¹⁶⁸, or to determine the dynamics that lead up to an irrevocable decision e.g. represented by a saccadic eye movement⁶⁵. The inevitable occurrence of spurious forces in TSA ensembles however has never been noted.

With the present TSA framework one dimensional powerlaw like biological forces in an uncorrelated mean free fluctuating environment can be identified. Erroneous assignments due to inevitable alignment forces can be avoided. It turns out that TSA ensembles fall into two broad classes. For power law exponents $\alpha > 0$, the biological forces vanish at the boundary and directed dynamics become indistinguishable from the dynamics of a target-state aligned random search process with its analytically determinable moments. For power law exponents $\alpha < 0$, directed dynamics are force induced and reverse-time statistics can be treated as random-walk like corrections to the deterministic solution. Expanding on these results, we present small noise expressions for reverse-time mean, variance and two-time correlation function valid for force induced target state convergence. The generic analysis of noise and force induced transitions can be used to guide the eye when searching for the driving phenomenological mechanism underlying target-state aligned data. For direct high precision inference of the true biological forces, we use a path ensemble maximum likelihood scheme adopted to reverse time dynamics. We demonstrate its applicability for different scenarios of cytokinetic ring constriction and point out possible force miss-assignments when ignoring the peculiarities of TSA ensembles.

The TSA approach allows for a new perspective on the design of future experiments. It renders the search for suitable (artificially induced) initial conditions superfluous and allows to study directed dynamics under as natural circumstances as possible using a biologically motivated alignment point. Nevertheless, the identification of transition times, when dynamics change e.g. from an equilibrium state into target state directed dynamics such as during cytokinesis, remains a biologically highly relevant question. The here sketched path ensemble framework provides a possible route to this inference problem. A further generalization of the formalism at hand might include different types of fluctuating environments, the inclusion of external variables and an extension of the framework to more than one degree of freedom. In principle all these generalizations are mathematically straight forward. At least for the case of power law like multiplicative noise analytical extensions of the TSA-theory are possible.

Concluding, the here presented theory provides the mathematical foundations for the inference of target-state-aligned directed dynamics. It offers an intuitive understanding for the ensemble changing effects of target-state-alignment and provides a new tool for the study of non-stationary biological dynamics.

The mathematical theory of target state alignment

4.1 Stochastic dynamics setting

Markov processes

A time dependent stochastic process is defined by its probability to be at its current position \widehat{L}_n at time t_n and by the transitions it takes to get there starting at \widehat{L}_0, t_0 . In this chapter \widehat{L} denotes processes in forward time t . Formally such a process can be written as the transition probability $P(\widehat{L}_n, t_n | \widehat{L}_{n-1}, t_{n-1}, \dots, \widehat{L}_0, t_0)$. For simplicity we assume that the current state of a stochastic process does only depend on its last step and not on its full history, i.e. that the process is memory free and obeys the Markov property

$$P(\widehat{L}_n, t_n | \widehat{L}_{n-1}, t_{n-1}, \dots, \widehat{L}_0, t_0) = P(\widehat{L}_n, t_n | \widehat{L}_{n-1}, t_{n-1}) . \quad (4.1)$$

The Markov property implies that the Chapman Kolmogorov equation

$$P(\widehat{L}_n, t_n | \widehat{L}_{n-2}, t_{n-2}) = \int d\widehat{L}_{n-1} P(\widehat{L}_n, t_n | \widehat{L}_{n-1}, t_{n-1}) P(\widehat{L}_{n-1}, t_{n-1} | \widehat{L}_{n-2}, t_{n-2}) \quad (4.2)$$

is fulfilled which we will use repeatedly in the subsequent derivations.

SDE and Fokker Planck representations

The evolution of a dynamical system under the influence of random forcing can often be studied using a Langevin equation

$$d\widehat{L}(t) = f(\widehat{L}) dt + \sqrt{D} dW_t . \quad (4.3)$$

The first term defines the deterministic drift $f(\widehat{L})$ the second the strength D of random fluctuations. Here the term dW_t denotes the Wiener process increment with zero mean $\langle dW_t \rangle = 0$ and delta correlated covariance $\langle D dW_t dW_{t'} \rangle = D \delta(t - t')$. Throughout this text we adopt the Ito-interpretation of Eq. (4.3) which assures that the random forcing is always applied at the beginning of a discrete time interval dt .

Using Ito's Lemma⁵⁵

$$dg(\widehat{L}(t)) = \left(f(\widehat{L}) \frac{\partial g(\widehat{L})}{\partial \widehat{L}} + \frac{D}{2} \frac{\partial^2 g(\widehat{L})}{\partial \widehat{L}^2} \right) dt + \sqrt{D} \frac{\partial g(\widehat{L})}{\partial \widehat{L}} dW_t , \quad (4.4)$$

which shows how to apply a change of variables for quantities that are governed by Eq. (4.3), we can derive an evolution equation for its probability distribution $P(\widehat{L}, t | \widehat{L}_0, t_0)$. This evolution equation is called the (forward) Fokker-Planck equation (FPE) and is an equivalent description of the stochastic dynamics described in Eq. (4.3). A derivation can be found in Gardiner⁵⁵. Averaging Eq. (4.4) with respect to $P(\widehat{L}, t | \widehat{L}_0, t_0)$, twice integrating by parts and identifying $g(\widehat{L})$ with a delta-function then leads to the sought evolution equation⁵⁵. The resulting partial differential equation

$$\frac{\partial}{\partial t} P(\widehat{L}, t | \widehat{L}_0, t_0) = -\frac{\partial}{\partial \widehat{L}} f(\widehat{L}) P(\widehat{L}, t | \widehat{L}_0, t_0) + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t | \widehat{L}_0, t_0) \quad (4.5)$$

is the sought (forward) Fokker-Planck equation. The adjoint equation is called the backward Fokker-Planck equation

$$\frac{\partial}{\partial t} P(\widehat{L}_f, t_f | \widehat{L}, t) = -f(\widehat{L}) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) - \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}_f, t_f | \widehat{L}, t) . \quad (4.6)$$

and defines the evolution from an initial condition \widehat{L} to a final state \widehat{L}_f . It is important to note that after relabeling both equations yield the same transition probabilities and that it is only a different perspective to solve the same problem. The term 'backward'-Fokker-Planck equation is somewhat ill-posed and can not be exchanged for 'time-reversed'. In analogy to quantum mechanics, the forward Fokker-Planck equation describes the time evolution of the probability distribution and an average of any dynamical property is taken at a specific time point t like in the Schrödinger picture. Conversely, the backward Fokker-Planck equation corresponds to the Heisenberg picture where the focus lies on the time evolution of a dynamical observable and the averages are taken over initial conditions. Like in quantum mechanics both approaches are thus describing exactly the same physics. They are merely offering different perspectives and can be interchanged by a simple transformation. Ultimately both approaches hence refer to the same transition probability. In the next section I show how the concept of time-reversal is to be introduced into a stochastic framework.

4.2 Time reversal and alignment

The mathematical construction of the aligned and time reversed ensemble can be split into three parts. (i) The general definition of time reversal for SDEs, (ii) its adaptation to sample paths $\widehat{L}_i(t)$ that hit a target state \widehat{L}_{ts} after time T_i , and (iii) the alignment and time reversal of a complete ensemble of sample paths $L_i(\tau)$ into one ensemble depending on their "time to completion" $\tau = T_i - t$.

In the first part I present the SDE for the time reversal of sample paths $\widehat{L}_i(t)$ of fixed lifetime T . The ensemble $P(\widehat{L}, t)$ of sample paths $\widehat{L}_i(t)$, evolves freely according to a general stochastic law of motion. The reverse time ensemble then consists of sample paths $L_i(\tau) = \widehat{L}_i(T - \tau)$. All sample paths $\widehat{L}_i(t)$ start at the same time t_0 and evolve for the same time interval T . The time reversed SDE describes precisely this ensemble, but in reverse time. It starts from the final distribution of the forward process $P(L, \tau_0) = P(\widehat{L}, T)$ and evolves into the initial distribution $P(L, T) = P(\widehat{L}, t_0)$. One illustrative example realization of both a forward ensemble $P(\widehat{L}, t)$ and time reversed ensemble $P(L, \tau)$ is shown in Fig. 4.1.

In the second part, we consider the time reversal of sample paths $\widehat{L}_i(t)$ that hit a target state \widehat{L}_{ts} after a fixed time $T_i = T$. In this step, we pick precisely those sample paths L_i of the full forward ensemble $P(\widehat{L}, t)$ that arrive at a target state after a particular time to completion T_i . The time reversed SDE for this sub-ensemble is the same as derived in the first part, with the initial condition of the reverse time process specialized as a delta-function $P(L, \tau_0) = \delta(\widehat{L} - \widehat{L}_{ts})$

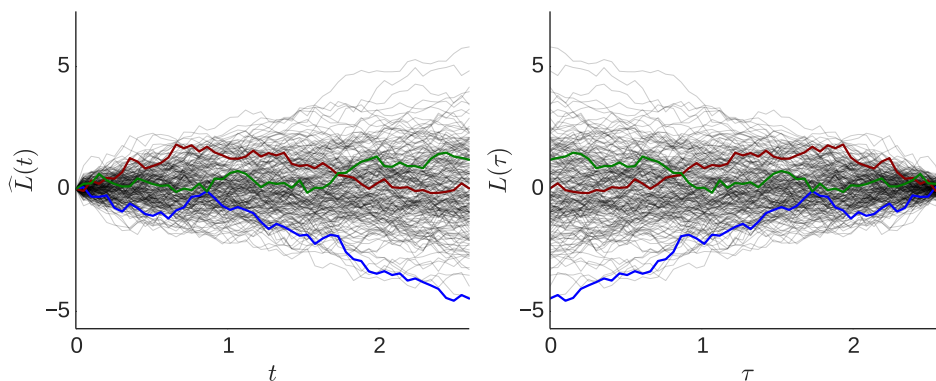


Figure 4.1: Time reversal of an ensemble with natural boundary conditions. (Left): Evolution of a set of sample paths $\hat{L}_i(t)$ up to a time T . (Right): The set of time reverted sample paths $L_i(\tau)$. Natural boundary conditions are defined as $P(\hat{L} \rightarrow \pm\infty, t) = 0$.

localized at the target state. All sample paths of this sub-ensemble start at \hat{L}_{ts} and progress towards the initial conditions of the forward process. They are all of the same lifetime T_i . A realization of this time reversed sub-ensemble together with the full forward ensemble up to time T_i is shown in Fig. 4.2.

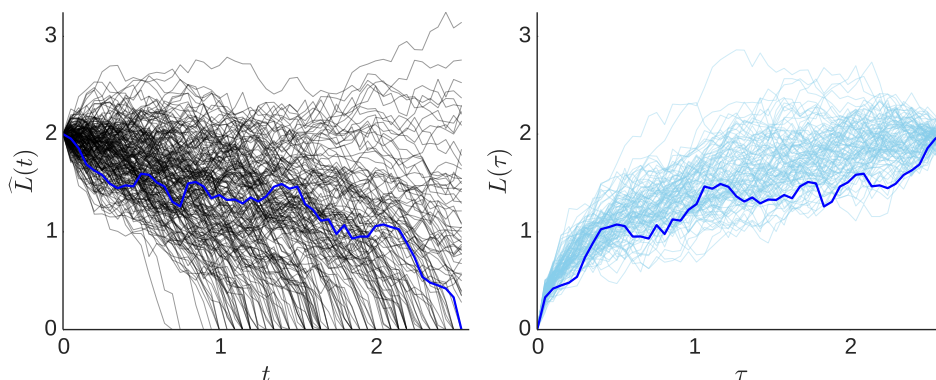


Figure 4.2: Time reversal of a terminating sub-ensemble. (Left): Evolution of a set of sample paths $\hat{L}_i(t)$ up to an irreversible target point. Different sample paths $\hat{L}_i(t)$ hit \hat{L}_{ts} at different times T_i . (Right): Time reversal of a representative sample path $\hat{L}_i(t)$ that has hit \hat{L}_{ts} after time T_i (dark blue). A realization of the sub-ensemble that has terminated after T_i (light blue).

Finally, in the last step, we construct the complete ensemble of time reversed TSA sample paths $L_i(\tau)$, of different lifetimes T_i , in a common time to completion reference frame. Depending on its noise and underlying stochastic law of motion each sample path $\hat{L}_i(t)$ hits \hat{L}_{ts} after a different time T_i , which by itself is a random variable. The distribution $\rho_{\hat{L}_{ts}}(T|\hat{L}_0)$ of lifetimes T is defined by the forward SDE of Eq. (4.3) and its initial conditions. To obtain the complete time reversed ensemble, we first construct a time reversed sub-ensemble for each lifetime T_i . $\rho_{\hat{L}_{ts}}(T|\hat{L}_0)$. The complete time reversed ensemble is then obtained by averaging over these sub-ensembles with respect to the distribution of lifetimes $\rho_{\hat{L}_{ts}}(T|\hat{L}_0)$. This construction of the complete time reversed ensemble out of sub-ensembles of different lifetimes is demonstrated in Fig. 4.3. Target state alignment is achieved by constructing each time reversed sub-ensemble starting from \hat{L}_{ts} and $\tau = 0$. The full time reversed ensemble then also starts from this initial state $(\hat{L}_{ts}, \tau = 0)$.

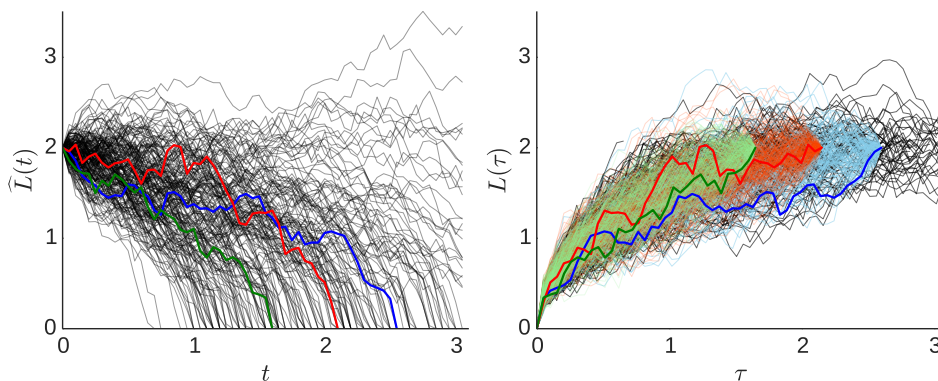


Figure 4.3: Time reversal and alignment of the complete terminating ensemble. (Left): A set of realizations of the complete forward ensemble. Three exemplary sample paths of different lifetimes T_i are highlighted. (Right): Time reversal of terminating sub-ensembles according to their lifetime T_i . The full ensemble is formed of sub-ensembles where three are exemplary highlighted.

The following sections give a mathematically detailed account of the three steps leading to the construction of a target state aligned and time reversed ensemble. We start with the derivation of a general time reversed SDE.

The reverse time Fokker-Planck equation

Time reversal of a stochastic process is a non-trivial operation. While a deterministic process is described by an ODE and a single final condition, time reversal of a stochastic process does not simply turn a forward SDE into a reverse time SDE. This finding is rooted in the observation that transition probabilities of the reverse process must depend both on the initial and final conditions of the forward process. We start by noting that a stochastic process constrained by initial and final conditions can be constructed from the product of two transition probabilities

$$P(\widehat{L}_f, t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0) . \quad (4.7)$$

To ensure that sample paths start at (\widehat{L}_0, t_0) and end at (\widehat{L}_f, t_f) with probability one, this process is normalized using the Chapman Kolmogorov Eq. (4.2). The normalized process

$$P(\widehat{L}, t | \widehat{L}_f, t_f, \widehat{L}_0, t_0) = \frac{P(\widehat{L}_f, t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0)}{P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)} \quad (4.8)$$

is called a bridge process of duration $T = t_f - t_0$ and defines the evolution of stochastic processes conditioned on both the initial (\widehat{L}_0, t_0) and final conditions (\widehat{L}_f, t_f) ¹⁰⁴. We note that $P(\widehat{L}, t | \widehat{L}_f, t_f, \widehat{L}_0, t_0)$ is symmetric under time reversal i.e. relabeling of indices. The initial conditions of the forward process (\widehat{L}_0, t_0) are then the final conditions of the reverse time process (L_f, τ_f) and vice versa. The bridge process defined in reverse time $\tau = t_f - t$ and with $\tau_0 = 0$ then reads $P(L, \tau | L_0, \tau_0, L_f, \tau_f)$. The SDE that describes this process is known⁵ and assumes the form

$$dL(\tau) = \left(-f(L) + D \frac{\partial}{\partial L} \log \left(P^{\text{fw}}(L, T - \tau | L_f, 0) \right) \right) d\tau + \sqrt{D} dW_\tau \quad (4.9)$$

where $P^{\text{fw}}(L, T - \tau | L_f, 0)$ denotes the solution of the forward Fokker Planck equation evaluated in reverse time τ . The initial conditions of this process are given by the final distribution

$P^{\text{fw}}(\widehat{L}_f, t_f | \widehat{L}_0, t_0)$ of the forward process. The guiding force $D \frac{\partial}{\partial L} \log P^{\text{fw}}(L, T - \tau | L_f, 0)$ ensures that every $L_i(\tau)$ synthesized in reverse time τ is statistically indistinguishable from a forward sample path \widehat{L}_i of lifetime T . Close to $\tau \approx T$, $P^{\text{fw}}(L, T - \tau | L_f, 0)$ becomes peaked and deviations from $P(\widehat{L}_0) = \delta(\widehat{L} - \widehat{L}_0)$ are suppressed by strong gradients of the guiding potential. The system is driven towards the initial conditions of the forward process. Note that the guiding force takes the form of an entropic force¹⁵⁹ which reverses the single trajectory entropy production of the forward process¹⁴⁸ until the entropy of the forward initial conditions is reinstalled. An illustration of both an example forward ensemble and its time reversed equivalent is shown in Fig. 4.1.

The SDE Eq. (4.9) captures the dynamics of a time reversed process of fixed duration T and is the first key equation to our approach. Its original derivation⁵ builds on the equivalence of Langevin and Fokker-Planck equations in which Eq. (4.9) is obtained from the reverse time Fokker-Planck Equation

$$\begin{aligned} \frac{\partial P(L, \tau | L_f, \tau_f, L_0, \tau_0)}{\partial \tau} &= \frac{\partial}{\partial L} \left[\left(f(L) - D \frac{\partial}{\partial L} \log P^{\text{fw}}(L, T - \tau | L_f, 0) \right) P(L, \tau | L_f, \tau_f, L_0, \tau_0) \right] \\ &+ \frac{D}{2} \frac{\partial^2}{\partial L^2} P(L, \tau | L_f, \tau_f, L_0, \tau_0) . \end{aligned} \quad (4.10)$$

The probability distribution $P(L, \tau | L_f, \tau_f, L_0, \tau_0)$ describes the evolution of reverse sample paths $L(\tau)$ of fixed lifetime T . It depends on both, the initial and final conditions \widehat{L}_0 and \widehat{L}_f of the forward process. The subsequent derivation of Eq. (4.10) follows Anderson's⁵. I here use the notation employed in chapter 3 and add intermediate steps omitted in Anderson⁵. A mathematically more rigorous derivation can be found in Föllmer⁵².

The derivation proceeds in three steps: First, I will derive an evolution equation of the joint probability distribution $P(\widehat{L}, t, \widehat{L}_f, t_f | \widehat{L}_0, t_0)$ using both the backward and forward Fokker-Planck equation Eq. (4.6) and Eq. (4.5) for the forward sample paths $\widehat{L}(t)$ that evolve with time t . From this intermediate result the evolution equation for the conditional probability $P(\widehat{L}, t | \widehat{L}_f, t_f, \widehat{L}_0, t_0)$ follows directly. It describes sample paths $\widehat{L}_i(t)$ that are constrained to both initial and final conditions. Finally I transform the evolution equation of $P(\widehat{L}, t | \widehat{L}_f, t_f, \widehat{L}_0, t_0)$ into the reverse time Fokker-Planck Equation for $P(L, \tau | L_f, \tau_f, L_0, \tau_0)$ by time reversal.

Note, that I will omit writing out the explicit dependence on the initial conditions (\widehat{L}_0, t_0) during this derivation for notational clarity. Even if not stated, every probability distribution used in this derivation carries (\widehat{L}_0, t_0) as a condition.

I start with writing the joint probability distribution from Eq. (4.7)

$$P(\widehat{L}, t, \widehat{L}_f, t_f) = P(\widehat{L}_f, t_f | \widehat{L}, t) P(\widehat{L}, t) \quad (4.11)$$

as a product of a conditional probability and a probability density assuming $t_f \geq t$. Taking the derivative with respect to t on both sides and multiplying both sides with (-1) , we obtain

$$- \frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} = -P(\widehat{L}, t) \frac{\partial}{\partial t} P(\widehat{L}_f, t_f | \widehat{L}, t) - P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial t} P(\widehat{L}, t) . \quad (4.12)$$

Substituting $\frac{\partial}{\partial t} P(\widehat{L}_f, t_f | \widehat{L}, t)$ using the backward Fokker-Planck Eq. (4.6) and $\frac{\partial}{\partial t} P(\widehat{L}, t)$ using the forward Fokker-Planck Eq. (4.5) yields

$$\begin{aligned} - \frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} &= P(\widehat{L}, t) \left[f(\widehat{L}) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}_f, t_f | \widehat{L}, t) \right] \\ &+ P(\widehat{L}_f, t_f | \widehat{L}, t) \left[\frac{\partial}{\partial \widehat{L}} f(\widehat{L}) P(\widehat{L}, t) - \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t) \right] . \end{aligned} \quad (4.13)$$

In the next step I consider the diffusion terms yielding an effective diffusion like term of the form $\frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}_f, t_f, \widehat{L}, t)$. Application of the product rule to both terms leads to

$$\begin{aligned} & \frac{D}{2} \left[P(\widehat{L}, t) \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}_f, t_f | \widehat{L}, t) \right] - \frac{D}{2} \left[P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t) \right] = \\ & = \frac{D}{2} \left[\frac{\partial}{\partial \widehat{L}} \left(P(\widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) \right) - \left(\frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t) \right) \left(\frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) \right) \right] \\ & - \frac{D}{2} \left[\frac{\partial}{\partial \widehat{L}} \left(P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t) \right) - \left(\frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) \right) \left(\frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t) \right) \right]. \end{aligned} \quad (4.14)$$

Here the product term cancels

$$= \frac{D}{2} \frac{\partial}{\partial \widehat{L}} \left(P(\widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) - P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t) \right), \quad (4.15)$$

using the product rule a second time, but now only for the first term

$$= \frac{D}{2} \frac{\partial}{\partial \widehat{L}} \left(\frac{\partial}{\partial \widehat{L}} \left(P(\widehat{L}, t) P(\widehat{L}_f, t_f | \widehat{L}, t) \right) - P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t) - P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t) \right), \quad (4.16)$$

and summing identical terms, yields

$$= \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} \left(P(\widehat{L}, t) P(\widehat{L}_f, t_f | \widehat{L}, t) \right) - D \frac{\partial}{\partial \widehat{L}} \left(P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t) \right). \quad (4.17)$$

Next I consider all $f(\widehat{L})$ containing terms in Eq. (4.13). Application of the product rule here yields

$$\begin{aligned} P(\widehat{L}, t) f(\widehat{L}) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) + P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} f(\widehat{L}) P(\widehat{L}, t) = \\ = \frac{\partial}{\partial \widehat{L}} \left(f(\widehat{L}) P(\widehat{L}_f, t_f | \widehat{L}, t) P(\widehat{L}, t) \right). \end{aligned} \quad (4.18)$$

Now we can collect terms and apply Eq. (4.11) to join the probability distributions $P(\widehat{L}_f, t_f | \widehat{L}, t)$ and $P(\widehat{L}, t)$, which allows to bring Eq. (4.13) into the form

$$\begin{aligned} -\frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} = \frac{\partial}{\partial \widehat{L}} \left[f(\widehat{L}) P(\widehat{L}, t, \widehat{L}_f, t_f) - D P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t) \right] \\ + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t, \widehat{L}_f, t_f). \end{aligned} \quad (4.19)$$

Note that the term $-D P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t)$, although depending on the diffusion constant D , is now assigned to the drift part. Rewriting this term, using again Eq. (4.11) in the form $P(\widehat{L}_f, t_f | \widehat{L}, t) = P(\widehat{L}_f, t_f, \widehat{L}, t) / P(\widehat{L}, t)$, yields

$$\begin{aligned} -\frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} = \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) - \frac{D}{P(\widehat{L}, t)} \frac{\partial P(\widehat{L}, t)}{\partial \widehat{L}} \right) P(\widehat{L}, t, \widehat{L}_f, t_f) \right] \\ + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t, \widehat{L}_f, t_f). \end{aligned} \quad (4.20)$$

This is our first intermediate result. The evolution equation for the joint probability distribution $P(\widehat{L}, t, \widehat{L}_f, t_f)$.

Dividing this equation by $P(\widehat{L}_f, t_f)$, the distribution of the final state, results in an evolution equation of the forward probability distribution

$$-\frac{\partial P(\widehat{L}, t | \widehat{L}_f, t_f)}{\partial t} = \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) - \frac{D}{P(\widehat{L}, t)} \frac{\partial P(\widehat{L}, t)}{\partial \widehat{L}} \right) P(\widehat{L}, t | \widehat{L}_f, t_f) \right] + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t | \widehat{L}_f, t_f), \quad (4.21)$$

which in its force term can be simplified further

$$-\frac{\partial P(\widehat{L}, t | \widehat{L}_f, t_f)}{\partial t} = \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) - D \frac{\partial}{\partial \widehat{L}} \log P(\widehat{L}, t) \right) P(\widehat{L}, t | \widehat{L}_f, t_f) \right] + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t | \widehat{L}_f, t_f). \quad (4.22)$$

To make the interpretation of this equation and especially of the term $P(\widehat{L}, t)$ explicit, the dependence of every probability term on the initial conditions (\widehat{L}_0, t_0) , suppressed throughout this derivation, is now restored to yield

$$-\frac{\partial P(\widehat{L}, t | \widehat{L}_0, t_0, \widehat{L}_f, t_f)}{\partial t} = \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) - D \frac{\partial}{\partial \widehat{L}} \log P(\widehat{L}, t | \widehat{L}_0, t_0) \right) P(\widehat{L}, t | \widehat{L}_0, t_0, \widehat{L}_f, t_f) \right] + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t | \widehat{L}_0, t_0, \widehat{L}_f, t_f). \quad (4.23)$$

The probability distribution $P(\widehat{L}, t | \widehat{L}_0, t_0, \widehat{L}_f, t_f)$ describes the evolution of sample paths $\widehat{L}(t)$ in wall time t connecting an initial state \widehat{L}_0, t_0 with a final state \widehat{L}_f, t_f . This is our second intermediate result.

The time reversed Fokker-Planck Equation can be obtained from Eq. (4.23) by replacing the wall time t with the reverse time τ . Setting $\tau = t_f - t$, and the duration of the forward process after which it is reverted to $T = t_f - t_0$, the resulting equation reads

$$\frac{\partial P(L, \tau | \widehat{L}_0, t_0, \widehat{L}_f, t_f)}{\partial \tau} = \frac{\partial}{\partial L} \left[\left(f(L) - D \frac{\partial}{\partial L} \log P^{\text{fw}}(L, T - \tau | \widehat{L}_f, 0) \right) P(L, \tau | \widehat{L}_0, t_0, \widehat{L}_f, t_f) \right] + \frac{D}{2} \frac{\partial^2}{\partial L^2} P(L, \tau | \widehat{L}_0, t_0, \widehat{L}_f, t_f). \quad (4.24)$$

The superscript (fw) was added for clarity. In the final step we replace all forward coordinates \widehat{L} by their time reversed equivalent L and obtain the final result already stated in Eq. (4.10). The initial conditions of Eq. (4.10) can be drawn from the final distribution of the forward process $P^{\text{fw}}(\widehat{L}_f, t_f | \widehat{L}_0, t_0)$. Solving or approximating the time reversed Fokker-Planck Equation therefore crucially depends on knowledge about the forward process.

The time reversal of a terminating sub-ensemble

Time reversal of a sub-ensemble which ends at a target state \widehat{L}_{ts} after time $T = t_f - t_0$ is a special case of the general formulation of the reverse time Fokker-Planck Eq. (4.10). From a constructive perspective two specifications are necessary. First, to invoke an irreversible target state, an absorbing boundary condition

$$P(\widehat{L}, t | \widehat{L}_0, t_0) = 0 \quad \text{for} \quad \widehat{L} = \widehat{L}_{\text{ts}} \quad (4.25)$$

must be implemented. Second, with absorbing boundary conditions the product of the two transition probabilities which define the not yet normalized bridge Eq. (4.7), is zero for $\widehat{L}_f \rightarrow \widehat{L}_{ts}$. The second transition $P(\widehat{L}_f, t_t | \widehat{L}, t)$ in Eq. (4.7) must therefore be replaced by the probability to reach the target state \widehat{L}_{ts} starting from \widehat{L} and after a time $t_f - t$. This probability flux into the boundary is called a hitting time distribution and is defined as

$$\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) = -\frac{\partial}{\partial t_f} \int_{\widehat{L}_{ts}}^{\infty} dL_f P(\widehat{L}_f, t_f | \widehat{L}, t). \quad (4.26)$$

The un-normalized bridge, which starts at \widehat{L}_0 , ends at $\widehat{L}_f = \widehat{L}_{ts}$ and only includes sample paths which reach the target state exactly after a time $T = t_f - t_0$, is then defined as

$$\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0). \quad (4.27)$$

To ensure that sample paths starting at \widehat{L}_0 end at $\widehat{L}_f = \widehat{L}_{ts}$ with probability one, Eq. (4.27) is normalized with respect to the hitting time distribution $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}_0, t_0)$. The full bridge ending at a target state then reads

$$R_{\widehat{L}_{ts}}(\widehat{L}, t | \widehat{L}_0, t_0; t_f) := \frac{\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0)}{\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}_0, t_0)}. \quad (4.28)$$

We will show below, that $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t)$ is a solution of the backward Fokker Planck and that the normalization of this bridge process is again a hitting time distribution. The dynamics of Eq. (4.28) written in reverse time τ are therefore formally identical to Eq. (4.10) and read

$$\begin{aligned} \frac{\partial R_{L_{ts}}(L, \tau | L_f, \tau_f; \tau_0)}{\partial \tau} &= \frac{\partial}{\partial L} \left[\left(f(L) - D \frac{\partial}{\partial L} \log P^{\text{fw}}(L, T - \tau | L_f, 0) \right) R_{L_{ts}}(L, \tau | L_f, \tau_f; \tau_0) \right] \\ &+ \frac{D}{2} \frac{\partial^2}{\partial L^2} R_{L_{ts}}(L, \tau | L_f, \tau_f; \tau_0). \end{aligned} \quad (4.29)$$

We again added the upper index (fw) for clarity. The corresponding reverse time SDE is thus identical to Eq. (4.9) with $P^{\text{fw}}(L, T - \tau | \widehat{L}_0, 0)$ assumed with an absorbing boundary at L_{ts} . Assuming absorbing boundary conditions for $P^{\text{fw}}(\widehat{L}, t | \widehat{L}_0, t_0)$, the guiding force in Eq. (4.9) automatically prevents time reversed sample paths $L_i(\tau)$ from returning to \widehat{L}_{ts} for $\tau > 0$. This can be seen by evaluating the guiding force $D \frac{\partial}{\partial L} \log P^{\text{fw}}(L, T - \tau | \widehat{L}_0, 0)$ with respect to the absorbing boundary condition. Lets assume the irreversible target state \widehat{L}_{ts} is a lower boundary of the one dimensional dynamic. As $P^{\text{fw}}(L, T - \tau | \widehat{L}_0, 0) \xrightarrow[L \rightarrow \widehat{L}_{ts}]{} 0$ approaches zero from above close to the absorbing boundary, $\log P^{\text{fw}}(L, T - \tau | \widehat{L}_0, 0) \xrightarrow[L \rightarrow \widehat{L}_{ts}]{} -\infty$ diverges. The guiding force $D \frac{\partial}{\partial L} \log P^{\text{fw}}(L, T - \tau | \widehat{L}_0, 0) \xrightarrow[L \rightarrow \widehat{L}_{ts}]{} \infty$ therefore diverges as well and thus behaves as an infinitely large barrier that prevents all $L_i(\tau)$ from returning to \widehat{L}_{ts} .

The target state bridge normalization

In this sub section we show that $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}_0, t_0)$ is the normalization for the target-state bridge process. One step in this calculation is based on the assumption that the Chapman-Kolmogorov equation Eq. (4.2) also holds, when not the full probability space is considered but only those transitions which do not end at a target state, i.e. $P(\widehat{L}', t' | \widehat{L}, t)$. This assumption is implicit to all calculations of Fokker-Planck equations with absorbing boundary conditions imposed.

Nevertheless, as I am not aware of an explicit demonstration of this claim, I will briefly sketch a derivation below. This subsection first shows that a Chapman Kolmogorov also holds for this subspace of transitions. In a second step this result is used to determine the normalization of a bridge process with target state at an absorbing boundary.

I first demonstrate that a Chapman-Kolmogorov like equation also holds for $P(\widehat{L}, t|\widehat{L}_0, t_0)$ and $P(\widehat{L}_f, t_f|\widehat{L}, t)$ which are not normalized, as transitions to the absorbing boundary are split off as an extra term. Splitting the expression into those two terms yields

$$P^{\text{full}}(\widehat{L}, t|\widehat{L}_0, t_0) = \delta(\widehat{L} - \widehat{L}_{\text{ts}}) P^{\text{ts}}(t|\widehat{L}_0, t_0) + P(\widehat{L}, t|\widehat{L}_0, t_0), \quad (4.30)$$

where $P^{\text{ts}}(t|\widehat{L}_0, t_0)$ denotes the over time $t - t_0$ at the target state accumulated probability mass. For this normalized expression Chapman-Kolmogorov holds

$$P^{\text{full}}(\widehat{L}_f, t_f|\widehat{L}_0, t_0) = \int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L} P^{\text{full}}(\widehat{L}_f, t_f|\widehat{L}, t) P^{\text{full}}(\widehat{L}, t|\widehat{L}_0, t_0) \quad (4.31)$$

and can be rewritten using the definition Eq. (4.30) to yield

$$= \int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L} \left(\delta(\widehat{L}_f - \widehat{L}_{\text{ts}}) P^{\text{ts}}(t_f|\widehat{L}, t) + P(\widehat{L}_f, t_f|\widehat{L}, t) \right) P^{\text{full}}(\widehat{L}, t|\widehat{L}_0, t_0). \quad (4.32)$$

Evaluating the first term of the product we find

$$= \delta(\widehat{L}_f - \widehat{L}_{\text{ts}}) P^{\text{ts}}(t_f|\widehat{L}_0, t_0) + \int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L} P(\widehat{L}_f, t_f|\widehat{L}, t) P^{\text{full}}(\widehat{L}, t|\widehat{L}_0, t_0). \quad (4.33)$$

Due to the absorbing boundary condition, transitions from $\widehat{L}_0 \rightarrow \widehat{L} \rightarrow \widehat{L}_f$, with $\widehat{L}_f \neq \widehat{L}_{\text{ts}}$, are only possible if the intermediate step does not end at the absorbing boundary. We can therefore simplify the remaining integral to include only transitions that do not end at the target state. The final equation reads

$$= \delta(\widehat{L}_f - \widehat{L}_{\text{ts}}) P^{\text{ts}}(t_f|\widehat{L}_0, t_0) + \int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L} P(\widehat{L}_f, t_f|\widehat{L}, t) P(\widehat{L}, t|\widehat{L}_0, t_0). \quad (4.34)$$

Comparing this result with the definition of $P^{\text{full}}(\widehat{L}_f, t_f|\widehat{L}, t)$ in Eq. (4.30), this implies, that the Chapman Kolmogorov like identity

$$P(\widehat{L}_f, t_f|\widehat{L}_0, t_0) = \int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L} P(\widehat{L}_f, t_f|\widehat{L}, t) P(\widehat{L}, t|\widehat{L}_0, t_0) \quad (4.35)$$

holds. We next use this identity to obtain the normalization for the target-state bridge Eq. (4.27) by integration over \widehat{L} . Using the definition of the hitting time distribution

$$\int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L} \rho_{\widehat{L}_{\text{ts}}}(t_f|\widehat{L}, t) P(\widehat{L}, t|\widehat{L}_0, t_0) = - \int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L} P(\widehat{L}, t|\widehat{L}_0, t_0) \frac{\partial}{\partial t_f} \int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L}_f P(\widehat{L}_f, t_f|\widehat{L}, t), \quad (4.36)$$

taking both the derivative with respect to t_f and the integral with respect to \widehat{L}_f to the outside, we arrive at

$$= - \frac{\partial}{\partial t_f} \int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L}_f \int_{\widehat{L}_{\text{ts}}}^{\infty} d\widehat{L} P(\widehat{L}, t|\widehat{L}_0, t_0) P(\widehat{L}_f, t_f|\widehat{L}, t). \quad (4.37)$$

We replace the inner integral by the above derived version of the Chapman Kolmogorov equation Eq. (4.35)

$$= -\frac{\partial}{\partial t_f} \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}_0, t_0) \quad (4.38)$$

and find that the hitting time distribution

$$= \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}_0, t_0) \quad (4.39)$$

for processes starting at (\widehat{L}_0, t_0) and ending at (\widehat{L}_{ts}, t_f) is the normalization of the bridge process Eq. (4.28).

Backward Fokker-Planck for hitting time distribution

To conclude the above section on bridge processes with target states we recall that the hitting time distribution $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t)$ fulfills the backward Fokker-Planck equation Eq. (4.6). This can be easily seen by taking the derivative of Eq. (4.26) with respect to t

$$\frac{\partial}{\partial t} \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) = -\frac{\partial}{\partial t_f} \int_{\widehat{L}_{ts}}^{\infty} dL_f \frac{\partial}{\partial t} P(\widehat{L}_f, t_f | \widehat{L}, t). \quad (4.40)$$

Substituting the backward Fokker-Planck equation Eq. (4.6)

$$= -\frac{\partial}{\partial t_f} \int_{\widehat{L}_{ts}}^{\infty} dL_f \left(-f(\widehat{L}) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) - \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}_f, t_f | \widehat{L}, t) \right)$$

and replacing all $P(\widehat{L}_f, t_f | \widehat{L}, t)$ dependent terms by the definition of the hitting time distribution Eq. (4.26), we obtain the backward Fokker Planck for hitting time distributions

$$= -f(\widehat{L}) \frac{\partial}{\partial \widehat{L}} \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) - \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t). \quad (4.41)$$

The time reversal and alignment of the full terminating ensemble

So far we considered the time reversal of sample paths $\widehat{L}_i(t)$ of identical lifetime $T_i = T$. The lifetime T_i , however, is a stochastic variable. Different sample paths $\widehat{L}_i(t)$ reach the target state after different lifetimes T_i . The distribution of lifetimes is the hitting time distribution $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}_0, t_0)$. The full forward ensemble of all sample paths starting at (\widehat{L}_0, t_0) and ending at \widehat{L}_{ts} after a random time t_f can therefore be understood as comprised of sub-ensembles $R_{L_{ts}}(\widehat{L}, t | \widehat{L}_0, t_0; t_f)$ each of a different lifetime T_i . Each of these sub-ensembles then contributes to the full ensemble with a weight given by the hitting time distribution $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}_0, t_0)$

In this perspective, target state alignment consists of detaching all these sub-ensembles from their initiation point (\widehat{L}_0, t_0) and shifting all sample paths to reach the target state at a common reference time t_f . The resulting target state aligned ensemble in forward time

$$R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f | \widehat{L}_0) = \int_{-\infty}^t dt_0 \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}_0, t_0) R_{\widehat{L}_{ts}}(\widehat{L}, t | \widehat{L}_0, t_0; t_f) \quad (4.42)$$

is then given as the hitting time weighted integral of target-state aligned sub-ensembles. By target state alignment we choose one t_f for all sub-ensembles. Keeping t_f fixed, $T_i = t_f - t_0$

or respectively t_0 becomes the random variable over which we integrate. Notably, not all sub-ensembles contribute to the full TSA ensemble at all times. The longer the lifetime $t_f - t$, the fewer sub-ensembles have a lifetime T_i long enough to still contribute. The integral bounds therefore enforce that at time t only those sub-ensembles with a minimal lifetime $t_f - t$ contribute.

In reverse time τ the expression for the TSA ensemble is

$$R_{L_{\text{ts}}}(L, \tau; \tau_0 | L_f) = \int_{\tau}^{\infty} d\tau_f \rho_{L_{\text{ts}}}(\tau_0 | L_f, \tau_f) R_{L_{\text{ts}}}(L, \tau | L_f, \tau_f; \tau_0). \quad (4.43)$$

With τ_0 identically fixed for all sub-ensembles only those sub-ensembles contribute at time τ with a lifetime $T_i \geq \tau - \tau_0$. The reverse time TSA ensemble $R_{L_{\text{ts}}}(L, \tau; \tau_0 | L_f)$ is therefore not normalized. It decays with the hitting time distribution $\rho_{L_{\text{ts}}}(\tau_0 | L_f, \tau_f)$.

In the next step we generalize this expression to include sample-paths with their initial position \hat{L}_0 drawn from an input distributions $P^{\text{in}}(\hat{L}_0)$. For the reverse time TSA ensemble this generalization simply amounts to integrating over Eq. (4.43) with respect to $P^{\text{in}}(L_f)$ and we arrive at

$$R_{L_{\text{ts}}}(L, \tau; \tau_0) = \int_{L_{\text{ts}}}^{\infty} dL_f P^{\text{in}}(L_f) \int_{\tau}^{\infty} d\tau_f \rho_{L_{\text{ts}}}(\tau_0 | L_f, \tau_f) R_{L_{\text{ts}}}(L, \tau | L_f, \tau_f; \tau_0). \quad (4.44)$$

This expression is exactly the same as stated in chapter 3.4 Eq. (3), where the concept of TSA ensembles is introduced. There, we set $R(L, \tau) := R_{L_{\text{ts}}}(L, \tau; \tau_0)$ and neglected the for the intuition irrelevant dependency on L_{ts} and τ_0 .

Independent of its notation, $R_{L_{\text{ts}}}(L, \tau; \tau_0)$ satisfies an easy to interpret reverse time Fokker-Planck equation which reads

$$\begin{aligned} \frac{\partial}{\partial \tau} R_{L_{\text{ts}}}(L, \tau; \tau_0) &= -P^{\text{in}}(L) \rho_{L_{\text{ts}}}(\tau | L) \\ &- \frac{\partial}{\partial L} \left(\left[f(L) + D \frac{\partial}{\partial L} \log \left(\int_{L_{\text{ts}}}^L dL' e^{-\int^{L'} \frac{2f(L'')}{D} dL''} \left(1 - \int_{L_{\text{ts}}}^{L'} P^{\text{in}}(L'') dL'' \right) \right] \right) R_{L_{\text{ts}}}(L, \tau; \tau_0) \right) \\ &+ \frac{D}{2} \frac{\partial^2}{\partial L^2} R_{L_{\text{ts}}}(L, \tau; \tau_0). \end{aligned} \quad (4.45)$$

The exact derivation is presented further below in this section. The first term constitutes a sink where probability mass proportional to the hitting time distribution is lost. For delta initial conditions $P^{\text{in}}(L) = \delta(L - L_f)$ the sink is only active for $L = L_f$. For arbitrary input distributions $P^{\text{in}}(L_f)$ the rate of probability loss is still proportional to the hitting time distribution but now additionally to the number of sample paths which originated in an interval $P^{\text{in}}(L) dL$. This is exactly in accordance with the construction of $R_{L_{\text{ts}}}(L, \tau; \tau_0)$ in Eq. (4.44). The second line represents the drift and the third a simple diffusion term.

This form of the Fokker Planck equation for the reverse time TSA ensemble Eq. (4.45) corresponds to the Langevin equation

$$dL(\tau) = \left(f(L) + D \frac{\partial}{\partial L} \log \left(\int_{L_{\text{ts}}}^L dL' e^{-\int^{L'} \frac{2f(L'')}{D} dL''} \left(1 - \int_{L_{\text{ts}}}^{L'} P^{\text{in}}(L'') dL'' \right) \right) \right) d\tau + \sqrt{D} dW_{\tau} \quad (4.46)$$

under Ito interpretation.

For further use, we split Eq. (4.46) into conceptually meaningful components. We call the total deterministic contributions to the force, the TSA force $f^{\text{TSA}}(L)$. This allows rewriting Eq. (4.46) as

$$dL(\tau) = f^{\text{TSA}}(L) d\tau + \sqrt{D} dW_{\tau}. \quad (4.47)$$

The TSA force

$$f^{\text{TSA}}(L) = f(L) + f^{\mathcal{F}}(L) \quad (4.48)$$

is comprised of two parts. A term which is exactly identical to the drift term of the forward Fokker Planck Eq. (4.5) and points in the same direction. Second, a “free energy force”

$$f^{\mathcal{F}}(L) = D \frac{\partial}{\partial L} \log \left(\int_{L_{\text{ts}}}^L dL' e^{-\int^{L'} \frac{2f(L'')}{D} dL''} H(L') \right) \quad (4.49)$$

$$=: D \frac{\partial}{\partial L} \log Z(L) \quad (4.50)$$

which reverses the combined entropy production of all sample paths of the force driven system. The partition function $Z(L)$ is here defined as the weighted integral over the Boltzmann factor for the sign inverted forward force $f(L)$ times the sigmoidal function

$$H(L) = 1 - \int_{L_{\text{ts}}}^L P^{\text{in}}(L') dL' . \quad (4.51)$$

For forward initial conditions $P^{\text{in}}(L) = \delta(L - L_f)$, the term $H(L)$ evaluates to a Heaviside step function. Above L_f only the term $f(L)$, identical to the forward dynamics, contributes. Below L_f the free energy force contributes fully.

For stochastic dynamics, the probability of the system to be found at a certain position does not only depend on the potential, but also on the possible number of states it can assume, i.e. on the entropy. When studied over all possible sample path lifetimes as implied by the construction of the TSA ensemble, the potential to be reverted is not only the integral over the force but also the entropic contribution summarized over all times, i.e. (and in analogy to stochastic thermodynamics¹⁴⁸) the single particle free energy. To revert the forward force $f(L)$ and the effect of the stochastic forcing, the free energy force must be proportional to twice the force $f(L)$. Once to compensate for the forward dynamics and once to redirect the dynamics on the free energy surface back to the forward initial conditions. The factor 2 can either be read off from the Boltzmann factor in Eq. (4.49), or more intuitively, by substituting the Einstein relation $D = \frac{2k_B T}{\gamma}$ ⁵⁵. k_B would be the Boltzmann constant, T the temperature and γ the friction coefficient. With this substitution it becomes apparent that the contribution due to the change in free energy comes with a factor of 2.

For an arbitrary input distribution $P^{\text{in}}(L)$, the contribution of the free energy force to the drift vanishes gradually along a sigmoidal-like curve with larger values of L . This implies that above the bulk of initial positions forward and TSA dynamics are indistinguishable in their dynamical law. Below however the free energy force reverses the forward dynamics and adds additional terms which I will study below.

While the drift and diffusion terms are straightforward to transfer, the sink is not represented by the SDE Eq. (4.46) and needs to be handled with care. For diffusion limited reaction diffusion systems¹⁷⁷ it is known that a Fokker-Planck equation of the form

$$\frac{\partial}{\partial \tau} P(L, \tau) = -k(L, \tau) P(L, \tau) - \frac{\partial}{\partial L} a(L) P(L, \tau) + \frac{D}{2} \frac{\partial^2}{\partial L^2} P(L, \tau) \quad (4.52)$$

transforms to a Langevin equation with Ito interpretation and killing measure $k(L, \tau)$ ^{70,147}. The killing measure here defines the probability $p_k = k(L, \tau) d\tau$ of a sample path to be terminated at position L at time τ . It must be evaluated in each timestep and for each sample path. Comparing terms, Eq. (4.52) and Eq. (4.45) tell us, that the killing measure for the reverse time TSA ensemble is given by

$$k(L, \tau) = \frac{P^{\text{in}}(L) \rho_{L_{\text{ts}}}(\tau|L)}{R_{L_{\text{ts}}}(L, \tau; \tau_0)} . \quad (4.53)$$

It defines the fraction of at time τ and position L still present sample paths that must be terminated. In general, it is surprisingly hard to correctly execute the killing measure on a single trajectory level when doing simulations. For example in the case of a delta-sink, i.e. a Robin or radiation boundary condition, it yields the wrong result to simply terminate all sample paths that cross the sink within a finite simulation step $\Delta\tau$ with probability $p_k = k(L, \tau)\Delta\tau$. Instead, the sink crossing must be modeled as a diffusion process of duration $\Delta\tau$ and the case of not seen crossings within the simulation step $\Delta\tau$ must be included. For details we refer to Erban and Chapman⁴⁵.

For the killing measure of reverse time TSA ensembles these problems are in a sense less severe but come with a prize. Notably, the killing measure in Eq. (4.53) can not be executed on a trajectory by trajectory level as it depends on the full distribution of still present sample paths $R_{L_{ts}}(L, \tau; \tau_0)$. This however suggests to simply do counting. We determine the fraction of still present sample paths N_τ , that must be killed in the space-time interval $[L, L + \Delta L] \times [\tau, \tau + \Delta\tau]$, i.e. $\frac{P^{\text{in}}(L)\rho_{L_{ts}}(\tau|L)dLd\tau}{N_\tau}$ and divide it by the fraction of sample paths $\frac{n_{\tau, \Delta L}}{N_\tau}$, which are actually in the interval. $n_{\tau, \Delta L}$ denotes the bin count. For ratios larger than one we terminate all sample paths in the interval. For ratios smaller than one we draw uniform random numbers and compare to the ratio. Formally this is called a metropolis criterion and we obtain the empirical killing measure

$$k(L, \tau)\Delta\tau = \min\left(1, \frac{P^{\text{in}}(L)\Delta L\rho_{L_{ts}}(\tau|L)\Delta\tau}{n_{\tau, \Delta L}\Delta L}\right). \quad (4.54)$$

Comparing the value of the killing measure for each trajectory at time τ to an iid random variable, this criterion allows to determine for each trajectory whether it must be terminated in the current step or not.

Derivation of the reverse time TSA Fokker Planck equation

Our derivation starts from the definition of the TSA ensemble for an arbitrary initial distribution $R_{L_{ts}}(L, \tau; \tau_0)$ Eq. (4.44), here defined in forward time t

$$R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f) = \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}_0, t_0) R_{\widehat{L}_{ts}}(\widehat{L}, t|\widehat{L}_0, t_0; t_f). \quad (4.55)$$

Using the definition of TSA sub-ensembles Eq. (4.28) this expression simplifies to

$$= \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) P(\widehat{L}, t|\widehat{L}_0, t_0). \quad (4.56)$$

To derive the corresponding Fokker Planck equation we take the time derivative on both sides of Eq. (4.55) and Eq. (4.56) and arrive at

$$\frac{\partial}{\partial t} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f) = \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \frac{\partial}{\partial t} \int_{-\infty}^t dt_0 \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) P(\widehat{L}, t|\widehat{L}_0, t_0). \quad (4.57)$$

Applying the Leibniz rule leads to

$$= \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \left(\rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) P(\widehat{L}, t|\widehat{L}_0, t) + \int_{-\infty}^t dt_0 \frac{\partial}{\partial t} \left(\rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) P(\widehat{L}, t|\widehat{L}_0, t_0) \right) \right) \quad (4.58)$$

or equivalently

$$= \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \left(\rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) \delta(\widehat{L} - \widehat{L}_0) + \int_{-\infty}^t dt_0 \frac{\partial}{\partial t} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f|\widehat{L}_0, t_0) \right). \quad (4.59)$$

The first term evaluates to the hitting time distribution times the probability to be at the initial position \widehat{L}_0 . The second term defines a double integral over the dynamics of a single contributing bridge process /sub-ensemble. We obtain

$$= \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) P^{\text{in}}(\widehat{L}) + \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \left(\int_{-\infty}^t dt_0 \frac{\partial}{\partial t} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f|\widehat{L}_0, t_0) \right). \quad (4.60)$$

In the next step we substitute $\frac{\partial}{\partial t} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f|\widehat{L}_0, t_0)$ with the Fokker Planck equation of the non-normalized bridge process as derived above in this section. We substitute the dynamics in the form given by the intermediate result Eq. (4.19), which allows us to handle the integral over t_0 and L_0 more easily. Note, that the transition probability to reach the final state $P(\widehat{L}_f, t_f|\widehat{L}, t)$ in Eq. (4.19) was replaced by the hitting time probability $\rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t)$, and the joint probability $P(\widehat{L}, t, \widehat{L}_f, t_f|\widehat{L}_0, t_0)$ of the two transitions by the respective joint probability $R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f|\widehat{L}_0, t_0)$, to map the general bridge process to a bridge process which ends at a target state. The resulting evolution equation reads

$$\begin{aligned} \frac{\partial}{\partial t} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f) &= \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) P^{\text{in}}(\widehat{L}) \\ &- \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \frac{\partial}{\partial \widehat{L}} \left[f(\widehat{L}) R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f|\widehat{L}_0, t_0) - D \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}, t|\widehat{L}_0, t_0) \right] \\ &- \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f|\widehat{L}_0, t_0). \end{aligned} \quad (4.61)$$

We recall that $R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f|\widehat{L}_0, t_0)$ is defined as $\rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) P(\widehat{L}, t|\widehat{L}_0, t_0)$ according to Eq. (4.27). This allows us to use the defining equation of $R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f)$ Eq. (4.56), and we arrive at the simplified evolution equation

$$\begin{aligned} &= \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) P^{\text{in}}(\widehat{L}) \\ &- \frac{\partial}{\partial \widehat{L}} \left[f(\widehat{L}) R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f) - D \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) \frac{\partial}{\partial \widehat{L}} \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 P(\widehat{L}, t|\widehat{L}_0, t_0) \right] \\ &- \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f). \end{aligned} \quad (4.62)$$

To obtain the guiding force and recover $R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f)$ in the second drift term as well, we multiply $\rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t)$ by $\int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 P(\widehat{L}, t|\widehat{L}_0, t_0)$ both in numerator and denominator

$$\begin{aligned} &= \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t) \delta(\widehat{L} - \widehat{L}_0) \\ &- \frac{\partial}{\partial \widehat{L}} \left[f(\widehat{L}) R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f) \right] \end{aligned} \quad (4.63)$$

$$\begin{aligned} &- \frac{D \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 P(\widehat{L}, t|\widehat{L}_0, t_0) \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t)}{\int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 P(\widehat{L}, t|\widehat{L}_0, t_0)} \frac{\partial}{\partial \widehat{L}} \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 P(\widehat{L}, t|\widehat{L}_0, t_0) \\ &- \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f). \end{aligned} \quad (4.64)$$

The numerator then evaluates to $DR_{\widehat{L}_{ts}}(\widehat{L}, t; t_f)$, using again Eq. (4.56). The denominator together with the rest of the expression can then be simplified to the form of a guiding force with L_0 and t_0 integrated out from the guiding probability. We arrive at our first intermediate result

$$\begin{aligned}
 &= \rho_{\widehat{L}_{ts}}(t_f|\widehat{L}, t)P^{\text{in}}(\widehat{L}) \\
 &- \frac{\partial}{\partial \widehat{L}} \left[f(\widehat{L}) - D \frac{\partial}{\partial \widehat{L}} \log \left(\int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 P(\widehat{L}, t|\widehat{L}_0, t_0) \right) \right] R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f) \\
 &- \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f)
 \end{aligned} \tag{4.65}$$

in the derivation of the full reverse time TSA Fokker-Planck equation Eq. (4.45). Both this intermediate result and the dynamics for the free reverse time bridge Eq. (4.10) depend on the solution of the forward Fokker Planck $P(\widehat{L}, t|\widehat{L}_0, t_0)$. Unfortunately, explicit solutions for $P(\widehat{L}, t|\widehat{L}_0, t_0)$ are only known for very few cases. Interestingly, the above transformations have shown that it is not $P(\widehat{L}, t|\widehat{L}_0, t_0)$ but

$$Q(\widehat{L}) := \lambda \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 P(\widehat{L}, t|\widehat{L}_0, t_0) \tag{4.66}$$

we need an expression for. We introduce λ as a normalization factor to treat $Q(\widehat{L})$ as a normalized distribution. The normalization is necessary as $P(\widehat{L}, t|\widehat{L}_0, t_0)$ is implied with absorbing boundary conditions and decays with the hitting time distribution, as discussed above.

To derive an explicit expression for this integral we start with the forward Fokker-Planck equation

$$\frac{\partial}{\partial t} P(\widehat{L}, t|\widehat{L}_0, t_0) = -\frac{\partial}{\partial \widehat{L}} \left(f(\widehat{L}) P(\widehat{L}, t|\widehat{L}_0, t_0) \right) + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t|\widehat{L}_0, t_0) \tag{4.67}$$

and integrate out the dependence on the initial time t_0

$$\int_{-\infty}^t dt_0 \frac{\partial}{\partial t} P(\widehat{L}, t|\widehat{L}_0, t_0) = \int_{-\infty}^t dt_0 \left(-\frac{\partial}{\partial \widehat{L}} \left(f(\widehat{L}) P(\widehat{L}, t|\widehat{L}_0, t_0) \right) + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}, t|\widehat{L}_0, t_0) \right). \tag{4.68}$$

To evaluate the left hand side we change to an integral over the distance $s = t - t_0$ and adapt the derivative accordingly. The resulting integral

$$\int_{-\infty}^t dt_0 \frac{\partial}{\partial t} P(\widehat{L}, t|\widehat{L}_0, t_0) = -\int_{\infty}^0 ds \frac{\partial}{\partial s} P(\widehat{L}, s|\widehat{L}_0, 0) = -P(\widehat{L}, 0|\widehat{L}_0, 0) = -\delta(\widehat{L} - \widehat{L}_0) \tag{4.69}$$

evaluates to minus the delta initial conditions of the forward process. The $+\infty$ term vanishes due to the natural boundary condition we impose. We define the non-equilibrium stationary probability density with delta insertion probability at \widehat{L}_0

$$Q(\widehat{L}|\widehat{L}_0) := \lambda \int_{-\infty}^t dt_0 P(\widehat{L}, t|\widehat{L}_0, t_0) = -\lambda \int_{\infty}^0 ds P(\widehat{L}, s|\widehat{L}_0, 0) = Q(\widehat{L}, \infty|\widehat{L}_0, 0). \tag{4.70}$$

The generalization $Q(\widehat{L})$ Eq. (4.66) to arbitrary insertion probabilities is obtained by multiplication with $P^{\text{in}}(\widehat{L}_0)$ and subsequent integration over \widehat{L}_0 . The resulting ordinary differential equation then reads

$$\begin{aligned}
 -\lambda P^{\text{in}}(\widehat{L}) &= -\frac{\partial}{\partial \widehat{L}} \left(f(\widehat{L}) Q(\widehat{L}) \right) + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} Q(\widehat{L}) \\
 Q(\widehat{L}_{ts}) &= 0.
 \end{aligned} \tag{4.71}$$

Integrating Eq. (4.71) over the full interval $\widehat{L} \in [\widehat{L}_{ts}, \infty]$ and with $\int_{\widehat{L}_{ts}}^{\infty} P^{\text{in}}(\widehat{L}') d\widehat{L}' = 1$ we find

$$\lambda = \frac{D}{2} \frac{\partial}{\partial \widehat{L}} Q(\widehat{L}_{ts}) \quad (4.72)$$

For completeness note, that λ , as already discussed in Zhang et. al. ¹⁸², defines a steady state probability flux at \widehat{L}_{ts} . Importantly equation Eq. (4.71) can be solved exactly. For completeness we here reproduce the result given in Zhang et. al. ¹⁸².

To solve the ordinary differential equation Eq. (4.71) for $Q(\widehat{L})$ we first integrate over \widehat{L}

$$-f(\widehat{L}')Q(\widehat{L}') \Big|_{\widehat{L}'_{ts}}^{\widehat{L}} + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} Q(\widehat{L}') \Big|_{\widehat{L}'_{ts}}^{\widehat{L}} = -\lambda \int_{\widehat{L}'_{ts}}^{\widehat{L}} P^{\text{in}}(\widehat{L}') d\widehat{L}' \quad (4.73)$$

and obtain

$$-f(\widehat{L})Q(\widehat{L}) + \frac{D}{2} \frac{\partial}{\partial \widehat{L}} Q(\widehat{L}) = \lambda \left(1 - \int_{\widehat{L}'_{ts}}^{\widehat{L}} P^{\text{in}}(\widehat{L}') d\widehat{L}' \right). \quad (4.74)$$

Combining the terms on the left hand side into a single expression yields

$$\frac{D}{2} e^{\int_{\widehat{L}}^{\widehat{L}} \frac{2f(\widehat{L}')}{D} d\widehat{L}'} \frac{\partial}{\partial \widehat{L}} \left(Q(\widehat{L}) e^{-\int_{\widehat{L}}^{\widehat{L}} \frac{2f(\widehat{L}')}{D} d\widehat{L}'} \right) = \lambda \left(1 - \int_{\widehat{L}'_{ts}}^{\widehat{L}} P^{\text{in}}(\widehat{L}') d\widehat{L}' \right). \quad (4.75)$$

In the last step we again integrate from \widehat{L}_{ts} to \widehat{L} and solve for $Q(\widehat{L})$ to obtain

$$Q(\widehat{L}) = \frac{2\lambda}{D} e^{\int_{\widehat{L}}^{\widehat{L}} \frac{2f(\widehat{L}')}{D} d\widehat{L}'} \int_{\widehat{L}'_{ts}}^{\widehat{L}} d\widehat{L}' e^{-\int_{\widehat{L}'}^{\widehat{L}} \frac{2f(\widehat{L}'')}{D} d\widehat{L}''} \left(1 - \int_{\widehat{L}'_{ts}}^{\widehat{L}'} P^{\text{in}}(\widehat{L}'') d\widehat{L}'' \right), \quad (4.76)$$

which is the second intermediate result.

In the final step of the derivation of the full reverse time TSA dynamics we plug the obtained result for $Q(\widehat{L})$ into Eq. (4.65). Using the obtained result for $Q(\widehat{L})$, and $P^{\text{in}}(\widehat{L})$ as normalized distribution, we find for the guiding force

$$D \frac{\partial}{\partial \widehat{L}} \log Q(\widehat{L}) = 2f(\widehat{L}) + D \log \left(\int_{\widehat{L}'_{ts}}^{\widehat{L}} d\widehat{L}' e^{-\int_{\widehat{L}'}^{\widehat{L}} \frac{2f(\widehat{L}'')}{D} d\widehat{L}''} \left(1 - \int_{\widehat{L}'_{ts}}^{\widehat{L}'} P^{\text{in}}(\widehat{L}'') d\widehat{L}'' \right) \right). \quad (4.77)$$

With this, we obtain the final reverse time Fokker Planck written in forward time

$$\begin{aligned} \frac{\partial}{\partial t} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f; \widehat{L}_0) &= P^{\text{in}}(\widehat{L}) \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) \\ &+ \frac{\partial}{\partial \widehat{L}} \left(\left[f(\widehat{L}) + D \frac{\partial}{\partial \widehat{L}} \log \left(\int_{\widehat{L}'_{ts}}^{\widehat{L}} d\widehat{L}' e^{-\int_{\widehat{L}'}^{\widehat{L}} \frac{2f(\widehat{L}'')}{D} d\widehat{L}''} \left(1 - \int_{\widehat{L}'_{ts}}^{\widehat{L}'} P^{\text{in}}(\widehat{L}'') d\widehat{L}'' \right) \right] \right) R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f; \widehat{L}_0) \\ &- \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f; \widehat{L}_0). \end{aligned} \quad (4.78)$$

To revert the time axis we set $\tau = t_f - t$, $L_f = \widehat{L}_0$ and obtain the full reverse time TSA Fokker-Planck equation as stated in Eq. (4.45).

4.3 Normalized TSA dynamics

The TSA ensemble $R_{L_{ts}}(L, \tau; \tau_0)$, derived in section 4.2, is the exact mathematical description of an ensemble of target state aligned trajectories. In reverse time more and more trajectories reach the end of their lifetime and stop to contribute to the ensemble. At $\tau \rightarrow \infty$ the ensemble vanishes. To calculate the moments of this ensemble, averages however must be taken with respect to a for each time τ normalized distribution. This normalized TSA ensemble then reads

$$R_{L_{ts}}^N(L, \tau; \tau_0) = \frac{R_{L_{ts}}(L, \tau; \tau_0)}{\int_{L_{ts}}^{\infty} dL R_{L_{ts}}(L, \tau; \tau_0)} \quad (4.79)$$

In this section we show that Eq. (4.79) can not only be obtained subsequently by normalization, but also directly from a constitutive Fokker-Planck equation. We first observe that, using the definition of $R_{L_{ts}}(L, \tau; \tau_0)$ as stated in Eq. (4.44) and integrating over L , the normalized TSA ensemble can be written as

$$R_{L_{ts}}^N(L, \tau; \tau_0) = \frac{R_{L_{ts}}(L, \tau; \tau_0)}{\int_{L_{ts}}^{\infty} P^{\text{in}}(L_f) \int_{\tau}^{\infty} d\tau_f \rho_{L_{ts}}(\tau_0|L_f, \tau_f)} \quad (4.80)$$

Using this definition of the normalized TSA ensemble we can derive a constitutive Fokker-Planck equation of the form

$$\begin{aligned} \frac{\partial}{\partial \tau} R_{L_{ts}}^N(L, \tau; \tau_0) = & - \frac{\rho_{L_{ts}}(\tau_0|L, \tau) P^{\text{in}}(L) - R_{L_{ts}}^N(L, \tau; \tau_0) \int_{L_{ts}}^{\infty} dL_f P^{\text{in}}(L_f) \rho_{L_{ts}}(\tau_0|L_f, \tau)}{\int_{L_{ts}}^{\infty} dL_f P^{\text{in}}(L_f) \int_{\tau}^{\infty} d\tau_f \rho_{L_{ts}}(\tau_0|L_f, \tau_f)} \\ & - \frac{\partial}{\partial L} \left(\left[f(L) + D \frac{\partial}{\partial L} \log \left(\int_{L_{ts}}^L dL' e^{-\int^{L'} \frac{2f(L'')}{D} dL''} \left(1 - \int_{L_{ts}}^{L'} P^{\text{in}}(L'') dL'' \right) \right) \right] R_{L_{ts}}^N(L, \tau; \tau_0) \right) \\ & + \frac{D}{2} \frac{\partial^2}{\partial L^2} R_{L_{ts}}^N(L, \tau; \tau_0) \quad (4.81) \end{aligned}$$

The detailed derivation is given below. Importantly, the dynamics of the with τ -decaying TSA ensemble Eq. (4.45) and the normalized ensemble Eq. (4.81) are identical. They differ however in the sink. While the un-normalized TSA dynamics loses probability mass according to their forward initial and lifetime statistics, is the normalized ensemble probability mass preserving. The same amount of trajectories that are terminated are simultaneously created. Sink and source however differ in their position. While the sink is positioned according to the forward initial distribution is the source proportional to the normalized TSA ensemble at time τ and the instantaneous change in the normalization. The whole expression is then evaluated with respect to the TSA ensemble normalization.

Under these considerations, the Langevin equations of both the not- and normalized ensemble are identical. While the killing measures are identical in their location, the normalized ensemble features an additional source term that compensates for all terminated sample paths. Reverse time trajectories in this ensemble thus not only start at $\tau = \tau_0$, but over the whole time course of the reverse time dynamics. In analogy to the discussion in section 4.2, the killing measure reads

$$k(L, \tau) = \frac{P^{\text{in}}(L) \rho_{L_{ts}}(\tau_0|L, \tau)}{R_{L_{ts}}(L, \tau; \tau_0)} \quad (4.82)$$

Note, that the normalization of the sink and of the normalized ensemble cancel and only the un-normalized distribution $R_{L_{ts}}(L, \tau; \tau_0)$ remains. In terms of the killing measure, normalized

and un-normalized TSA Langevin dynamics are thus virtually identical. For the normalized TSA ensemble additionally a source of the form

$$s(L, \tau) = \frac{\int_{L_{ts}}^{\infty} dL_f P^{\text{in}}(L_f) \rho_{L_{ts}}(\tau_0 | L_f, \tau)}{\int_{L_{ts}}^{\infty} dL_f P^{\text{in}}(L_f) \int_{\tau}^{\infty} d\tau_f \rho_{L_{ts}}(\tau_0 | L_f, \tau_f)} . \quad (4.83)$$

occurs. We note that the source is independent of the position L and equals the instantaneous change of the normalization divided by the normalization.

The normalized TSA ensemble is clearly useful to calculate moments. From a conceptual perspective however, an ensemble which creates sample paths “on the fly” is clearly not the ensemble which matches the experimental protocol of target state alignment. It thus depends on the question we ask, whether the normalized ensemble could be profitably used. With the exception of the following derivation, in the rest of this chapter we always refer to the un-normalized case and indicate if a normalized version is used.

Derivation of the normalized TSA dynamics

In this section we derive the normalized TSA dynamics. We start from the definition of the normalized ensemble Eq. (4.79) written in forward time

$$R_{L_{ts}}^N(\widehat{L}, t; t_f) = \frac{R_{L_{ts}}(\widehat{L}, t; t_f)}{\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{L_{ts}}(t_f | \widehat{L}_0, t_0)} .$$

We then substitute the definition Eq.(4.55) of the un-normalized ensemble

$$= \int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \frac{\rho_{L_{ts}}(t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0)}{\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{L_{ts}}(t_f | \widehat{L}_0, t_0)} , \quad (4.84)$$

to obtain our starting point expression. In forward time the time-derivative of the normalized ensemble reads

$$\frac{\partial}{\partial t} R_{L_{ts}}^N(\widehat{L}, t; t_f) = \frac{\partial}{\partial t} \int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \frac{\rho_{L_{ts}}(t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0)}{\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{L_{ts}}(t_f | \widehat{L}_0, t_0)} . \quad (4.85)$$

Applying the Leibniz rule we find

$$\begin{aligned} &= \frac{\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \rho_{L_{ts}}(t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t)}{\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{L_{ts}}(t_f | \widehat{L}_0, t_0)} \\ &+ \int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \frac{\partial}{\partial t} \frac{\rho_{L_{ts}}(t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0)}{\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{L_{ts}}(t_f | \widehat{L}_0, t_0)} . \end{aligned} \quad (4.86)$$

Recognizing $P(\widehat{L}, t | \widehat{L}_0, t)$ in the first term as a delta function and evaluating the time derivative in the second term then leads to

$$\begin{aligned} &= \frac{\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \rho_{L_{ts}}(t_f | \widehat{L}, t) \delta(\widehat{L} - \widehat{L}_0)}{\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{L_{ts}}(t_f | \widehat{L}_0, t_0)} \\ &- \int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \frac{\rho_{L_{ts}}(t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0) \int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \rho_{L_{ts}}(t_f | \widehat{L}_0, t)}{\left(\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{L_{ts}}(t_f | \widehat{L}_0, t_0) \right)^2} \\ &+ \int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \frac{\frac{\partial}{\partial t} \rho_{L_{ts}}(t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0)}{\int_{L_{ts}}^{\infty} d\widehat{L}_0 P^{\text{in}}(\widehat{L}_0) \int_{-\infty}^t dt_0 \rho_{L_{ts}}(t_f | \widehat{L}_0, t_0)} . \end{aligned} \quad (4.87)$$

We next evaluate the delta-function in the first term. The resulting term can be identified with the sink in the non-normalized reverse time TSA, here divided by the TSA-normalization. The second term can be associated as a source that ensures that the reverse time TSA stays normalized. We substitute $R_{\hat{L}_{ts}}^N(\hat{L}, t; t_f)$ as defined in Eq. (4.80) where possible. The third term describes the dynamics of the TSA ensemble as in the non-normalized ensemble Eq. (4.58), divided by the normalization. The normalized reverse time SDE written in forward time t then reads

$$\begin{aligned}
 &= \frac{\rho_{\hat{L}_{ts}}(t_f|\hat{L}, t)P^{\text{in}}(\hat{L}) - R_{\hat{L}_{ts}}^N(\hat{L}, t; t_f) \int_{\hat{L}_{ts}}^{\infty} d\hat{L}_0 P^{\text{in}}(\hat{L}_0) \rho_{\hat{L}_{ts}}(t_f|\hat{L}_0, t)}{\int_{\hat{L}_{ts}}^{\infty} d\hat{L}_0 P^{\text{in}}(\hat{L}_0) \int_{-\infty}^t dt_0 \rho_{\hat{L}_{ts}}(t_f|\hat{L}_0, t_0)} \\
 &+ \frac{\partial}{\partial \hat{L}} \left(\left[f(\hat{L}) + D \frac{\partial}{\partial \hat{L}} \log \left(\int_{\hat{L}_{ts}}^{\hat{L}} d\hat{L}' e^{-\int^{\hat{L}'} \frac{2f(\hat{L}'')}{D} d\hat{L}''} \left(1 - \int_{\hat{L}_{ts}}^{\hat{L}'} P^{\text{in}}(\hat{L}'') d\hat{L}'' \right) \right] \right) R_{\hat{L}_{ts}}^N(\hat{L}, t; t_f) \right) \\
 &- \frac{D}{2} \frac{\partial^2}{\partial \hat{L}^2} R_{\hat{L}_{ts}}^N(\hat{L}, t; t_f). \tag{4.88}
 \end{aligned}$$

Switching from forward time to the reverse time description we arrive at the final expression stated in Eq. (4.81).

4.4 Exactly solvable reverse time dynamics

Our construction of the reverse time TSA ensemble in section 4.2 has revealed three conceptually different approaches to obtain information on its distribution and dynamics. First, the TSA ensemble can be obtained from its definition Eq. (4.55). Given both the transition probability and the hitting time distribution of the forward problem are known we can construct the non-normalized bridge Eq. (4.27). Integration over all possible initial times and switching to the reverse time notation then leads to the reverse time TSA ensemble ending at L_f . If an initial distribution is known, it is easy to generalize the TSA ensemble to include a distribution of forward initial conditions. Below we discuss one example of this type.

The second method to obtain the reverse time TSA ensemble is based on sampling the reverse time dynamics. We derive an analytic expression for the reverse time sub-ensemble SDE Eq. (4.9) and synthesize sample paths with duration T drawn from the forward lifetime time distribution. Three analytically tractable cases are shown below.

The third approach is based on directly solving the full reverse time TSA Fokker Planck equation Eq. (4.45). We discuss one example.

Constructing the TSA ensemble from its definition – the random walk

TSA ensembles can be constructed from their defining equation Eq. (4.55), which demands to integrate over t_0 with respect to the non-normalized bridge process. We demonstrate this construction for the simple case of a target state aligned Gaussian random walk, or equivalently the diffusion process with $f(\hat{L}) = 0$. The transition probability of a diffusion process with absorbing boundary conditions at $\hat{L}_{ts} = 0$ is

$$P(\hat{L}, t|\hat{L}_0, t_0) = \frac{e^{-\frac{(\hat{L}-\hat{L}_0)^2}{2D(t-t_0)}} - e^{-\frac{(\hat{L}+\hat{L}_0)^2}{2D(t-t_0)}}}{\sqrt{2\pi D(t-t_0)}} \tag{4.89}$$

and can easily be obtained by using the technique of mirror charges. The lifetime (or hitting) time distribution is

$$\rho_0(t_f|\widehat{L}, t) = \frac{\widehat{L} e^{-\frac{\widehat{L}^2}{2D(t_f-t)}}}{\sqrt{2\pi D(t_f-t)^3}} \quad (4.90)$$

and follows directly from Eq. (4.26). Combining these two terms into the non-normalized bridge Eq. (4.27), and integrating over t_0 according to Eq. (4.55), we obtain

$$R_0(L, \tau; 0|L_f) = \frac{L(L + L_f - |L - L_f|)e^{-\frac{L^2}{2D\tau}}}{\sqrt{2\pi D\tau^3}}, \quad (4.91)$$

here written in reverse time τ and $L_f = \widehat{L}_0$. The initial time of the reverse time process is set to zero. It is the solution of the target state aligned diffusion process with a delta distributed forward initial probability at $P^{\text{in}}(L) = \delta(L - L_f)$. Integrating over L we obtain the cumulative of the hitting time distribution of the random walk. The probability mass of the TSA ensemble changes proportional to the hitting time distribution Eq. (4.90) and thus with the rate enforced by the sink in the full reverse time TSA Fokker-Planck equation Eq. (4.45).

The Fokker-Planck equation for the reverse time TSA ensemble Eq. (4.45) can be used to directly verify this result by insertion. As the process under consideration is a random walk, for the forward force holds $f(L) = 0$. We briefly discuss this substitution for the equivalent and easier to read TSA Langevin equation Eq. (4.46) and therein the free energy force Eq. (4.49). Substituting $f(L) = 0$ into Eq. (4.49), the exponential simplifies to one. With delta distributed forward initial conditions, $H(L)$ becomes a step function with jump at $L = L_f$. Below L_f , the step function $H(L)$ is one and the integral inside the logarithm evaluates to L . For $L > L_f$ we find $H(L) = 0$ and thus no contribution of the free energy force to the TSA dynamics. For $L \neq L_f$, the resulting Langevin equation is

$$dL(\tau) = \begin{cases} \frac{D}{L} d\tau + \sqrt{D} dW_\tau & \text{for } L < L_f \\ \sqrt{D} dW_\tau & \text{for } L > L_f . \end{cases} \quad (4.92)$$

The drift term of the Fokker Planck equation is thus strongly simplified and Eq. (4.91) can be checked by insertion. Note, that the TSA force law assumes the important form $\frac{D}{L}$ for $L < L_f$ that we will meet time and again in the following considerations. For $L = L_f$ the reverse time TSA Fokker-Planck Eq. (4.45) assumes a more complicated form as we must evaluate the delta initial conditions of the forward process $P^{\text{in}} = \delta(L - L_f)$ inside the free energy force, which yields a non-analytical expression. To circumvent this problem, and to demonstrate the validity of Eq. (4.45) for non-delta distributed forward initial conditions, we mix the TSA ensemble solutions with delta insertion Eq. (4.91), using a truncated Gaussian insertion probability

$$P^{\text{in}}(L_f) = \frac{\sqrt{\frac{2}{\pi}} e^{-\frac{(L_{\text{in}} - L_f)^2}{2\sigma^2}}}{\sigma \left(\text{erf}\left(\frac{L_{\text{in}}}{\sqrt{2}\sigma}\right) + 1 \right)} \quad (4.93)$$

normalized to the interval $L_f \in [0, \infty]$, and marginalize over L_f . We obtain the TSA reverse

time ensemble

$$\begin{aligned}
 R_0(L, \tau; 0) &= e^{-\frac{L^2 \left(\frac{\sigma^2}{D\tau} + 1 \right) + 2L_{\text{in}}^2}{2\sigma^2}} \\
 & \times \frac{L e^{\frac{L^2 + L_{\text{in}}^2}{2\sigma^2}} \left(\sqrt{2\pi} e^{\frac{L_{\text{in}}^2}{2\sigma^2}} \left((L_{\text{in}} - L) \operatorname{erf} \left(\frac{L - L_{\text{in}}}{\sqrt{2}\sigma} \right) + L_{\text{in}} \operatorname{erf} \left(\frac{L_{\text{in}}}{\sqrt{2}\sigma} \right) + L \right) + 2\sigma \right) - 2\sigma L e^{\frac{L_{\text{in}}(2L + L_{\text{in}})}{2\sigma^2}}}{\pi(D\tau)^{3/2} \left(\operatorname{erf} \left(\frac{L_{\text{in}}}{\sqrt{2}\sigma} \right) + 1 \right)}
 \end{aligned} \tag{4.94}$$

for diffusive target search starting with Gaussian insertion probability. This expression satisfies Eq. (4.45) with $f(L) = 0$.

Sampling TSA ensembles via reverse time random bridges

In this section, we discuss three different example cases for which the full TSA ensemble is explicitly constructed from fixed lifetime sub-ensembles. This can either be done by numerical integration of analytical obtainable sub-ensembles, or directly via sampling of reverse time bridge SDEs, where the lifetime is drawn from the hitting time distribution for each realization. We here discuss only the latter case for the three example cases of the Bessel, advected random walk and Ornstein-Uhlenbeck process.

Bessel process

The dynamical law of the Bessel process reads¹⁸

$$d\widehat{L}(t) = -\frac{\gamma}{\widehat{L}} dt + \sqrt{D} dW_t. \tag{4.95}$$

Its solution with $\gamma > 0$, absorbing boundary condition at $\widehat{L}_{\text{ts}} = 0$, and natural boundary conditions at $\widehat{L} \rightarrow \infty$ is known as Bessel process and reads¹⁸

$$P^{\text{fw}}(\widehat{L}, t | \widehat{L}_0, t_0 = 0) = \frac{\widehat{L}_0 e^{-\frac{\widehat{L}^2 + \widehat{L}_0^2}{2Dt}} \left(\frac{\widehat{L}}{\widehat{L}_0} \right)^{\frac{1}{2} - \frac{\gamma}{D}} I_{\frac{\gamma}{D} + \frac{1}{2}} \left(\frac{\widehat{L}\widehat{L}_0}{Dt} \right)}{Dt}. \tag{4.96}$$

$I_\nu(z)$ is the modified Bessel function of the first kind, with series expansion

$$I_\nu(z) = \left(\frac{z}{2} \right)^\nu \sum_{k=0}^{\infty} \frac{\left(\frac{z^2}{2} \right)^k}{k! \Gamma(\nu + k + 1)} \tag{4.97}$$

for real valued ν from Abramovitz & Stegun Eq. (AS 9.6.10)¹. A derivation and the lifetime distribution

$$\rho_0(T | \widehat{L}_0) = \frac{2^{-\frac{\gamma}{D} - \frac{1}{2}} e^{-\frac{\widehat{L}_0^2}{2DT}} \left(\frac{\widehat{L}_0^2}{DT} \right)^{\frac{\gamma}{D} + \frac{1}{2}}}{T \Gamma \left(\frac{\gamma}{D} + \frac{1}{2} \right)} \tag{4.98}$$

can be found in^{18,107}. Knowing the forward probability distribution Eq. (4.96), the time reversed sub-ensemble SDE of this process can directly be obtained by evaluating the derivative of the logarithm of Eq. (4.96) using the identity

$$\frac{d}{dz} \frac{I_{\nu + \frac{1}{2}}(z)}{\sqrt{z}} = \frac{I_{\nu + \frac{3}{2}}(z)}{\sqrt{z}} + \frac{\nu I_{\nu + \frac{1}{2}}(z)}{z^{\frac{3}{2}}} \tag{4.99}$$

from (AS 10.2.21)¹. We arrive at

$$dL(\tau) = \left(\frac{D}{L} - \frac{L}{(T-\tau)} + \frac{\gamma}{L} + \frac{L_f I_{\frac{\gamma}{D} + \frac{3}{2}} \left(\frac{LL_f}{D(T-\tau)} \right)}{(T-\tau) I_{\frac{\gamma}{D} + \frac{1}{2}} \left(\frac{LL_f}{D(T-\tau)} \right)} \right) d\tau + \sqrt{D} dW_\tau, \quad (4.100)$$

where the first term results from the power law term in Eq. (4.96), the second from the exponential and the last two from $I_\nu(z)$ using the stated identity. For the interpretation, recall that the initial conditions of the forward process are the final conditions of the reverse time process and thus $L_f = \widehat{L}_0$. To obtain the full result for the time reversed dynamics, we first need to solve Eq. (4.100) for fixed T and subsequently average over all T according to the distribution of lifetimes Eq. (4.98). In Fig. 4.4, we compare two ways to synthesize this ensemble. One generated from Eq. (4.95), subsequently aligned and time reverted, the other directly generated using Eq. (4.100) and Eq. (4.98). Comparing statistics of both the mean and variance, as shown in Fig. 4.4, confirms our calculations.

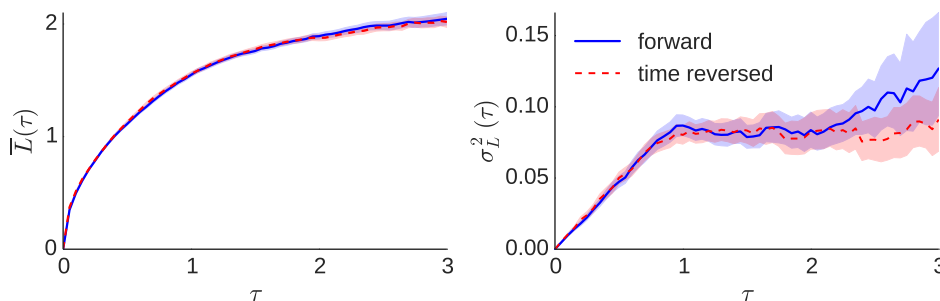


Figure 4.4: Forward (blue) and time reversed dynamics (red) show excellent agreement with respect to the Bessel process. Shown are the mean (Left) and variance (Right) of both processes with 95% bootstrap confidence intervals. To exclude numerical inaccuracies due to rarely visited tails of Eq. (4.98), we directly sampled from the numerically obtained hitting time distribution of the forward process. Results were obtained using each 1000 realizations of the respective ensemble with parameter settings $\gamma = 1$, $D = 0.2$ and $\widehat{L}_0 = 2$. The larger τ , the broader the confidence intervals become, as less and less sample paths contribute to the averaging due to differences in their lengths T_i .

Advection random walk

The advected random walk changes with constant drift velocity ($\gamma > 0$) until the target state \widehat{L}_{ts} is reached. Its forward SDE is

$$d\widehat{L}(t) = -\gamma dt + \sqrt{D} dW_t. \quad (4.101)$$

We set the absorbing boundary to $\widehat{L}_{\text{ts}} = 0$ and use natural boundary conditions at $\widehat{L} \rightarrow \infty$. The conditional probability for the forward process is

$$P^{\text{fw}}(\widehat{L}, t | \widehat{L}_0, t_0 = 0) = \frac{1}{\sqrt{2\pi Dt}} \left(e^{-\frac{(\widehat{L} + \gamma t - \widehat{L}_0)^2}{2Dt}} - e^{-\frac{(\widehat{L} + \gamma t + \widehat{L}_0)^2}{2Dt} + \frac{2\gamma \widehat{L}_0}{D}} \right), \quad (4.102)$$

and the associated distribution of lifetimes is

$$\rho_0(T | \widehat{L}_0) = \frac{\widehat{L}_0}{\sqrt{2\pi DT^3}} e^{-\frac{(\widehat{L}_0 - \gamma T)^2}{2DT}}, \quad (4.103)$$

as given in Redner¹³⁸.

Substituting Eq. (4.102) into the reverse time SDE for a sub-ensemble of fixed lifetime T Eq. (4.9), yields

$$dL(\tau) = -\frac{L - L_f \coth\left(\frac{LL_f}{D(T-\tau)}\right)}{T - \tau} d\tau + \sqrt{D} dW_\tau . \quad (4.104)$$

Recall that the initial conditions of the forward process are the final conditions of the reverse time process and thus $L_f = \hat{L}_0$. Interestingly this time reversed SDE is independent of the drift velocity and thus identical to the random walk bridge. This puzzling finding is resolved when considering not only the sub-ensemble of fixed lifetime T , but the complete aligned and time reversed ensemble according to Eq. (4.42). Averaging all sub-ensembles defined by Eq. (4.104) over T with respect to the distribution of lifetimes Eq. (4.103), recovers the expected γ -dependence. The γ -dependence is thus fully mediated by the lifetime distribution. A numerical comparison of the forward ensemble (after time reversion and alignment), and the time reversed ensemble, defined by Eq. (4.104) and Eq. (4.103), is displayed in Fig. 4.5.

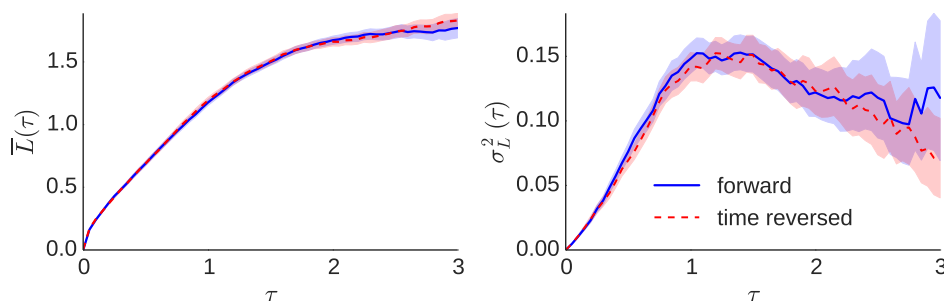


Figure 4.5: Forward (blue) and time reversed dynamics (red) of the advected random walk show excellent agreement. Shown are the mean (Left) and variance (Right) of both processes with 95% bootstrap confidence intervals. To exclude numerical inaccuracies due to rarely visited tails of Eq. (4.103), we directly sampled from the numerically obtained hitting time distribution of the forward process. Results were obtained using 1000 realizations of the respective ensemble with parameter settings $\gamma = 1$, $D = 0.2$ and $\hat{L}_0 = 2$. The larger τ , the broader the confidence intervals become, as less and less sample paths contribute to the averaging due to differences in their lengths T_i .

Ornstein-Uhlenbeck type process

The stochastic relaxation process describes dynamical processes with a linear restoring force. Its forward SDE is

$$d\hat{L}(t) = -\gamma\hat{L} dt + \sqrt{D} dW_t , \quad (4.105)$$

where $\gamma > 0$. The conditional probability density of this process with an absorbing boundary at $\hat{L}_{\text{ts}} = 0$, can be constructed using the method of mirror charges. We obtain

$$P^{\text{fw}}(\hat{L}, t | \hat{L}_0, t_0 = 0) = \sqrt{\frac{\gamma}{\pi D(1 - e^{-2\gamma t})}} \left(e^{-\frac{\gamma(\hat{L} - \hat{L}_0 e^{-\gamma t})^2}{D(1 - e^{-2\gamma t})}} - e^{-\frac{\gamma(\hat{L} + \hat{L}_0 e^{-\gamma t})^2}{D(1 - e^{-2\gamma t})}} \right) \quad (4.106)$$

and the associated distribution of lifetimes T

$$\rho_0(T | \hat{L}_0) = \frac{2D\hat{L}_0 \left(\frac{\gamma}{D}\right)^{3/2} e^{2\gamma T - \frac{\gamma\hat{L}_0^2}{D(e^{2\gamma T} - 1)}}}{\sqrt{\pi} (e^{2\gamma T} - 1)^{3/2}} , \quad (4.107)$$

given in Ricciardi and Sato¹³⁹. Using the forward probability distribution from Eq. (4.106), the time reversed SDE describing sub-ensembles of length T reads

$$\begin{aligned} dL(\tau) &= \left(\gamma L + D \frac{\partial}{\partial L} \left(e^{-\frac{\gamma(L-L_f e^{-\gamma(T-\tau)})^2}{D(1-e^{-2\gamma(T-\tau)})}} - e^{-\frac{\gamma(L+L_f e^{-\gamma(T-\tau)})^2}{D(1-e^{-2\gamma(T-\tau)})}} \right) \right) d\tau + \sqrt{D} dW_\tau \\ &= \left(-\frac{\gamma L}{\tanh((T-\tau)\gamma)} + \frac{L_f \gamma \coth(\frac{LL_f \gamma}{D \sinh((T-\tau)\gamma)})}{\sinh((T-\tau)\gamma)} \right) d\tau + \sqrt{D} dW_\tau . \end{aligned} \quad (4.108)$$

Note, that the initial conditions of the forward process are now the final conditions of the reverse time process with $L_f = \widehat{L}_0$. The full aligned and time reversed ensemble is then obtained by averaging over all possible sub-ensembles, defined by Eq. (4.108), with respect to the distribution of lifetimes given in Eq. (4.107). In Fig. 4.6, we compare mean and variance obtained by terminal alignment of Eq. (4.105), and directly from Eq. (4.108) sampling lifetimes from Eq. (4.107). We find excellent agreement.

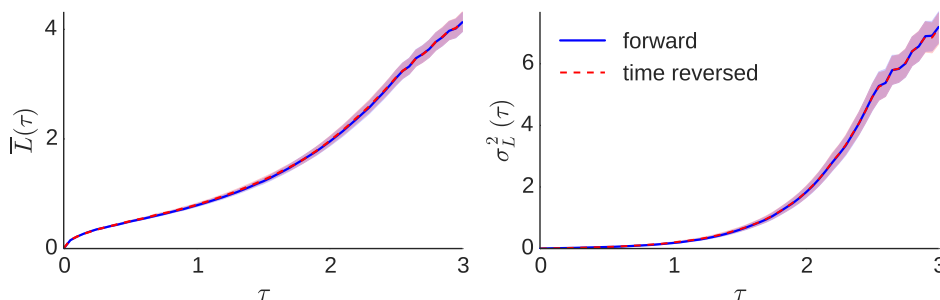


Figure 4.6: Forward (blue) and time reversed dynamics (red) of Ornstein-Uhlenbeck type processes agree excellent. Shown are the mean (Left) and variance (Right) of both processes with 95% bootstrap confidence intervals. To exclude numerical inaccuracies due to rarely visited tails of Eq. (4.107), we directly sampled from the numerically obtained hitting time distribution of the forward process. Results were obtained using 1000 realizations of sample paths with parameter settings $\gamma = 1$, $D = 0.2$ and $\widehat{L}_0 = 10$. Unlike for the previous examples, we used a larger value for \widehat{L}_0 to reveal the interesting concave convex shape of the mean, which, for too small \widehat{L}_0 values, is not detectable. The larger τ , the broader the confidence intervals become, as less and less sample paths contribute to the averaging due to differences in their lifetimes T_i .

Constructing TSA ensembles from their reverse time TSA Fokker-Planck equation – the Bessel process

The third approach to obtain the density of complete reverse time TSA ensembles, directly solves the defining Fokker-Planck Eq. (4.45) or SDE Eq. (4.46). We here consider the case of the Bessel process for which this approach is analytically tractable. The forward process is defined by Eq. (4.95) with initial condition $L_f = \widehat{L}_0$. The probability density which describes the forward process is given by Eq. (4.96). Defined as a TSA ensemble the SDE for the Bessel process reads

$$dL(\tau) = \left(-\frac{\gamma}{L} + D \frac{\partial}{\partial L} \log \left(\int_0^L dL' L'^{\frac{2\gamma}{D}} \Theta(\widehat{L}_0 - L') \right) \right) d\tau + \sqrt{D} dW_\tau , \quad (4.109)$$

where we substituted $f(L) = -\frac{\gamma}{L}$ into Eq. (4.46). We assume $L_{\text{ts}} = 0$ and $L_f = \widehat{L}_0$. For $L \neq L_f$ this SDE simplifies to

$$dL(\tau) = \begin{cases} \frac{\gamma+D}{L} d\tau + \sqrt{D} dW_\tau & \text{for } L < L_f \\ -\frac{\gamma}{L} d\tau + \sqrt{D} dW_\tau & \text{for } L > L_f, \end{cases} \quad (4.110)$$

which both have known solutions in terms of their transition probability densities when defined on the complete interval $L \in [0, \infty]$ ¹⁸. These solutions can be used to piecewise solve the full SDE. We start with the solution of the process defined in Eq. (4.110) for $L < L_f$, but assume for the moment that we are interested in the solution on the complete interval $L \in [0, \infty]$. The transition probability, for arbitrary L_0 , which describes this process, is then given as ¹⁸

$$P^<(L, \tau | L_0, \tau_0 = 0) = \frac{L e^{-\frac{L^2+L_0^2}{2D\tau}} \left(\frac{L}{L_0}\right)^{\frac{\gamma}{D}+\frac{1}{2}} I_{\frac{\gamma}{D}+\frac{1}{2}}\left(\frac{LL_0}{D\tau}\right)}{D\tau}. \quad (4.111)$$

Using the series expansion of $I_\nu(z)$ Eq. (4.97) from Abramovitz & Stegun (AS 9.9.10)¹, we can evaluate this process for $L_0 \rightarrow 0$ and obtain

$$P^<(L, \tau | L_0 = 0, \tau_0 = 0) = \frac{2^{-\frac{\gamma}{D}-\frac{1}{2}} (D\tau)^{-\frac{\gamma}{D}-\frac{3}{2}} e^{-\frac{L^2}{2D\tau}} L^{\frac{2\gamma}{D}+2}}{\Gamma\left(\frac{\gamma}{D} + \frac{3}{2}\right)}. \quad (4.112)$$

The needed contribution of the sub-process defined in Eq. (4.110) ($L > L_f$) is more complicated to obtain, as its transition probability Eq. (4.96) goes to zero for $L_0 \rightarrow L_{\text{ts}}$. We first guess that the L dependent part of the solution of Eq. (4.110) ($L > L_f$), now defined on $L \in [0, \infty]$, can be obtained by first normalizing the known solution of the Bessel process Eq. (4.96). The resulting ensemble consists of sample paths which at time t have not been absorbed at the target state but are generated by the process defined in Eq. (4.110) ($L > L_f$). We then take the limit $L_0 \rightarrow L_{\text{ts}}$. The normalization is simply done by the cumulative of the hitting time distribution stated in Eq. (4.98). The seeked expression for $L_0 \rightarrow 0$ then reads

$$P^>(L, \tau | L_0 = 0, \tau_0 = 0) \propto \lim_{L_0 \rightarrow 0} \frac{L_0 e^{-\frac{L^2+L_0^2}{2D\tau}} \left(\frac{L}{L_0}\right)^{\frac{1}{2}-\frac{\gamma}{D}} I_{\frac{\gamma}{D}+\frac{1}{2}}\left(\frac{LL_0}{D\tau}\right)}{D\tau \left(1 - \frac{\Gamma\left(\frac{\gamma}{D}+\frac{1}{2}, \frac{L_0^2}{2D\tau}\right)}{\Gamma\left(\frac{\gamma}{D}+\frac{1}{2}\right)}\right)} = \frac{L e^{-\frac{L^2}{2D\tau}}}{D\tau}, \quad (4.113)$$

where ($>$) denotes the forward force law $-\frac{\gamma}{L}$. To find the limit $L_0 \rightarrow 0$ we again use the zero-th order of the series expansion for $I_\nu(z)$ stated in Eq. (4.97). With the help of Eq.(AS 6.5.3), Eq.(AS 6.5.4) and Eq.(AS 6.5.29) ¹, we replace the upper incomplete gamma function in the denominator of Eq. (4.113) by the series

$$\Gamma(\nu, z) = \Gamma(\nu) - z^\nu \sum_{n=0}^{\infty} \frac{(-z)^n}{(\nu+n)n!} \quad \text{for } |z| < \infty. \quad (4.114)$$

The zero-th order cancels the one in the denominator. The next order cancels all remaining L_0 -terms of the nominator. The resulting term is proportional to the probability density of those sample paths which have survived until τ and started at $L_{\text{ts}} = 0$.

In the next step, we combine Eq. (4.111) and Eq. (4.113) into one expression which solves Eq. (4.109). Assuming continuity but not differentiability at $L = L_f$, we can easily connect the

two solutions at L_f , as Eq. (4.113) is fully included in Eq. (4.111). We obtain

$$R_0(L, \tau; 0|L_f) = \begin{cases} \frac{2(2D\tau)^{-\frac{\gamma}{D}-\frac{3}{2}} e^{-\frac{L_f^2}{2D\tau}} L^{\frac{2\gamma}{D}+2}}{\Gamma(\frac{\gamma}{D}+\frac{3}{2})} & L < L_f \\ \frac{2(2D\tau)^{-\frac{\gamma}{D}-\frac{3}{2}} e^{-\frac{L_f^2}{2D\tau}} L_f^{\frac{2\gamma}{D}+1} L}{\Gamma(\frac{\gamma}{D}+\frac{3}{2})} & L \geq L_f . \end{cases} \quad (4.115)$$

For $L \neq L_f$, we confirm this solution by insertion into the Fokker-Planck equation corresponding to Eq. (4.110). Furthermore, integrating $R_0(L, \tau; 0|L_f)$ over L yields one minus the cumulative of the forward hitting time Eq. (4.98) and shows that $R_0(L, \tau; 0|L_f)$ from Eq. (4.115) decays with the rate given by the hitting time distribution as demanded by the reverse time TSA Fokker Planck equation Eq. (4.45). As L_{ts} , according to the Feller boundary classification¹⁰⁷ is an entrance boundary and the other boundary is a natural boundary all lost probability mass must be due to the kink of $R_0(L, \tau; 0|L_f)$ at L_f , i.e. due to the sink at $\delta(L - L_f)$. This confirms that Eq. (4.115) is the full solution of the reverse time TSA Fokker-Planck associated with Eq. (4.109).

To calculate mean and variance with respect to all not yet killed sample paths of the reverse-time ensemble, we normalize Eq. (4.115) to unity with respect to L for all times τ . We obtain for the mean

$$\bar{L}(\tau) = \frac{A1 + A2}{B} \quad (4.116)$$

with

$$\begin{aligned} A1 &= e^{\frac{L_f^2}{2D\tau}} \left(\sqrt{2\pi D\tau} L_f^{\frac{2\gamma}{D}+1} \operatorname{erfc} \left(\frac{L_f}{\sqrt{2D\tau}} \right) + 2(2D\tau)^{\frac{\gamma}{D}+1} \left(\Gamma \left(\frac{\gamma}{D} + 2 \right) - \Gamma \left(\frac{\gamma}{D} + 2, \frac{L_f^2}{2D\tau} \right) \right) \right) \\ A2 &= 2L_f^{\frac{2(\gamma+D)}{D}} \\ B &= 2L_f^{\frac{2\gamma}{D}+1} + 2(2D\tau)^{\frac{\gamma}{D}+\frac{1}{2}} e^{\frac{L_f^2}{2D\tau}} \left(\Gamma \left(\frac{\gamma}{D} + \frac{3}{2} \right) - \Gamma \left(\frac{\gamma}{D} + \frac{3}{2}, \frac{L_f^2}{2D\tau} \right) \right) \end{aligned}$$

and for the variance

$$\sigma_L^2(\tau) = \frac{(2D\tau)^{\frac{\gamma}{D}+\frac{3}{2}} e^{\frac{L_f^2}{2D\tau}} \left(\Gamma \left(\frac{\gamma}{D} + \frac{5}{2} \right) - \Gamma \left(\frac{\gamma}{D} + \frac{5}{2}, \frac{L_f^2}{2D\tau} \right) \right) + L_f^{\frac{2\gamma}{D}+1} (2D\tau + L_f^2)}{L_f^{\frac{2\gamma}{D}+1} + 2(2D\tau)^{\frac{\gamma}{D}+\frac{1}{2}} e^{\frac{L_f^2}{2D\tau}} \left(\Gamma \left(\frac{\gamma}{D} + \frac{3}{2} \right) - \Gamma \left(\frac{\gamma}{D} + \frac{3}{2}, \frac{L_f^2}{2D\tau} \right) \right)} - \bar{L}(\tau)^2 . \quad (4.117)$$

Both mean and variance are depicted in Fig. 3.4b.

4.5 TSA dynamics close to the target state

The dynamics valid for final stages of convergence to the target state can be characterized by a single SDE valid for all forward initial conditions. This claim is easily verified by inspecting Eq. (4.46). For L sufficiently below the bulk of the forward initial distribution $P^{\text{in}}(L)$, the term $H(L) = 1 - \int_{L_{ts}}^L P^{\text{in}}(L') dL'$, which describes the influence of the forward initial conditions on the reverse time TSA dynamics, evaluates to approximately one. If most of the sample paths have not yet reached the sink at $P^{\text{in}}(L)$, this further implies that almost no sample path has terminated in reverse time. The killing measure $k(L, \tau)$ (Eq. (4.53)), i.e. the product of hitting

time and initial distribution normalized to the current number of still existing sample paths, therefore defines a sector of the $L - \tau$ plane, where, for $k(L, \tau) \approx 0$, TSA ensembles are – up to exponentially small corrections – governed by

$$dL(\tau) = \left(f(L) + D \frac{\partial}{\partial L} \log \left(\int_{L_{ts}}^L dL' e^{-\int^{L'} \frac{2f(L'')}{D} dL''} \right) \right) d\tau + \sqrt{D} dW_\tau . \quad (4.118)$$

This form of the reverse time TSA ensemble, valid for final stages of convergence to the target state, provides a powerful tool to infer forward dynamics irrespective of unknown initial conditions.

The reverse time SDE for TSA ensembles close to completion Eq. (4.118) can be analytically expressed for a broad class of forward forces. Exemplary, we study the properties of power law like forward dynamics with force term

$$f(\widehat{L}) = -\gamma \widehat{L}^\alpha , \quad (4.119)$$

with $\gamma > 0$, $\alpha \in \mathbb{R}$ and constant additive noise D . We derive analytical expressions for the free energy force using forward forces of the form Eq. (4.119). We check their validity numerically and use the derived expressions to characterize general time-reversed ensembles near the target state.

The TSA reverse time dynamics for forward dynamics Eq. (4.119) are

$$dL(\tau) = \left(-\gamma L^\alpha + \frac{(D\alpha + D) \left(-\frac{2\gamma}{D\alpha + D} \right)^{\frac{1}{\alpha+1}} e^{\frac{2\gamma L^{\alpha+1}}{D\alpha + D}}}{\Theta(\alpha + 1)\Gamma\left(\frac{1}{\alpha+1}\right) - \Gamma\left(\frac{1}{\alpha+1}, -\frac{2L^{\alpha+1}\gamma}{\alpha D + D}\right)} \right) d\tau + \sqrt{D} dW_\tau \quad \alpha \neq -1 , \quad (4.120)$$

where $\Gamma(z)$ is the gamma-function, $\Gamma(\nu, z)$ the upper incomplete gamma function and $\Theta(\nu)$ the Heaviside step function. The case for $\alpha = -1$ connecting $\alpha > 1$ and $\alpha < 1$ is the Bessel process with reverse time TSA dynamics

$$dL(\tau) = \frac{\gamma + D}{L} d\tau + \sqrt{D} dW_\tau \quad \alpha = -1 . \quad (4.121)$$

A detailed derivation and an exemplary comparison to the exact dynamics, already discussed in section 4.4, is provided below.

TSA dynamics independent from forward initial conditions

In this section we derive reverse time TSA dynamics under the assumption that the forward initial conditions are spatially and temporally well separated from the target state. We derive Eq. (4.120) for a forward force of the form $f(\widehat{L}) = -\gamma \widehat{L}^\alpha$, as stated in Eq. (4.119).

We start from the definition of the free energy force as stated in Eq. (4.49) for general reverse time TSA dynamics with

$$f^{\mathcal{F}}(L) = D \frac{\partial}{\partial L} \log \widetilde{I}(L) , \quad (4.122)$$

where the integral

$$\widetilde{I}(L) = \int_0^L e^{\frac{2\Phi(L')}{D}} H(L') dL' \quad (4.123)$$

takes the form of a canonical partition function for the potential

$$\Phi(L) = - \int^L f(L') dL' , \quad (4.124)$$

associated with the sign inverted forward force. The influence of the forward initial conditions on the TSA dynamics is captured by $H(L)$ Eq. (4.51). It is a sigmoidal like function, which is one at the target state and decreases the closer L is to the bulk of the forward initial condition. It approaches zero beyond. For well separated target states and forward initial conditions we can therefore safely set $H(L) = 1$, close to L_{ts} , and define $I(L) = \tilde{I}(L)|_{H(L)=1}$.

The derivation of Eq. (4.120) then starts with evaluating the potential

$$\Phi(L) = \int^L \gamma L^\alpha dL' = \frac{\gamma L^{\alpha+1}}{\alpha+1} \quad \alpha \neq -1 . \quad (4.125)$$

Substituting the resulting $\Phi(L)$ into $I(L)$ and with help of Eq. (AS 6.5.25) from Abramowitz & Stegun¹, yields

$$I(L) = \tilde{C}(\alpha) \Gamma \left(\frac{1}{\alpha+1}, -\frac{2\gamma L^{\alpha+1}}{D(\alpha+1)} \right) \Big|_0^L \quad \alpha \neq -1 , \quad (4.126)$$

where $\Gamma(n, z)$ is the upper incomplete gamma function and $\tilde{C}(\alpha) := -\frac{1}{\alpha+1} \left(-\frac{2\gamma}{D(\alpha+1)} \right)^{-\frac{1}{\alpha+1}}$ is a prefactor depending on α and independent of L . Depending on α , the lower boundary assumes two different values. For $\alpha > -1$, the argument $z := -\frac{2\gamma L^{\alpha+1}}{D(\alpha+1)}$ of the upper incomplete gamma function $\Gamma(\nu, z)$, goes to zero and $\Gamma(\nu, z)$ reduces to $\Gamma(\nu)$. For $\alpha < -1$, the argument z diverges ($z \rightarrow \infty$) as $L \rightarrow 0$ and $\Gamma(\nu, z) \rightarrow 0$, as can be read off from Eq. (AS 6.5.32) in Abramowitz & Stegun¹. Combined into a single expression using the Heaviside step-function $\Theta(\nu)$ we arrive at

$$I(L) = \tilde{C}(\alpha) \left(\Gamma \left(\frac{1}{\alpha+1}, -\frac{2\gamma L^{\alpha+1}}{D(\alpha+1)} \right) - \Theta(\alpha+1) \Gamma \left(\frac{1}{\alpha+1} \right) \right) \quad \alpha \neq -1 . \quad (4.127)$$

To obtain the final expression for the free energy force $f^{\mathcal{F}}$ we rewrite

$$f^{\mathcal{F}}(L) = D \frac{\partial}{\partial L} \log I(L) = \frac{D}{I(L)} \frac{\partial}{\partial L} I(L) , \quad (4.128)$$

which allows us to directly read off the sought expression, as $\frac{\partial}{\partial L} I(L)$ is simply the integrand of $I(L)$, and obtain

$$f^{\mathcal{F}}(L) = \frac{(D\alpha + D) \left(-\frac{2\gamma}{D\alpha + D} \right)^{\frac{1}{\alpha+1}} e^{\frac{2\gamma L^{\alpha+1}}{D\alpha + D}}}{\Theta(\alpha+1) \Gamma \left(\frac{1}{\alpha+1} \right) - \Gamma \left(\frac{1}{\alpha+1}, -\frac{2\gamma L^{\alpha+1}}{D\alpha + D} \right)} \quad \alpha \neq -1 . \quad (4.129)$$

Substituting the resulting $f^{\mathcal{F}}(L)$ into the equation for the time reversed SDE Eq. (4.118) yields the final result as stated in Eq. (4.120).

Comparing TSA dynamics with and without forward initial conditions

To demonstrate the excellent agreement of the TSA approximation stated in Eq. (4.120) and Eq. (4.121) with the corresponding aligned forward simulations close to completion we use the three, from section 4.4 known, analytically tractable example cases of constriction ($\alpha = -1$), advection ($\alpha = 0$) and relaxation ($\alpha = 1$). From above, we further know that the TSA-approximation of well separated initial and final conditions breaks down when a relevant fraction of time reversed sample paths reaches the level of the initial state of the forward process. To observe this phenomenon, we state the initial conditions of the forward process explicitly and adapt the plot range accordingly.

Bessel process

The Bessel process within the TSA close to completion approximation is given by Eq. (4.121). Fig. 4.7 shows the comparison between this approximation and the corresponding exact forward

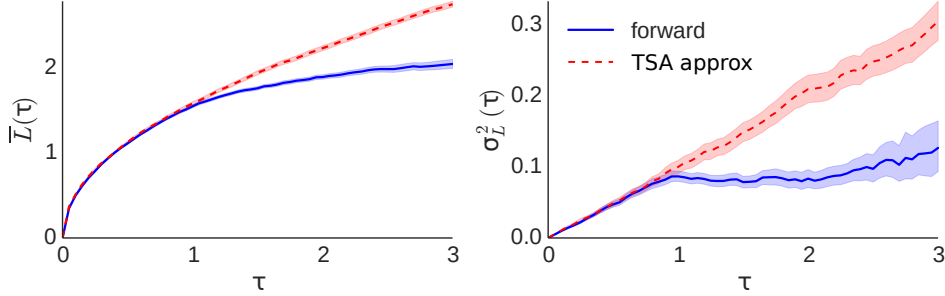


Figure 4.7: Forward dynamics (blue) and the approximate time reversed TSA dynamics (red) of the Bessel process agree excellently close to the target state. Shown are the mean (Left) and variance (Right) of both processes with 95% bootstrap confidence intervals. Results were obtained using 1000 realizations for each of the respective ensembles with parameter settings $\gamma = 1$, $D = 0.2$ and $\widehat{L}_{\text{in}} = 2$.

process starting at $\widehat{L}_{\text{in}} = 2$. Close to completion clearly both mean and variance are very well reproduced by Eq. (4.121).

Important for benchmark tests in section 4.6, Eq. (4.121) can be solved analytically and results in the transition probability stated in Eq. (4.111).

Specifying Eq. (4.111) with respect to the initial conditions $(L_0, \tau_0) = (0, 0)$ yields Eq. (4.112), which can be used to obtain expressions for mean

$$\bar{L}(\tau) = \frac{\sqrt{2}\sqrt{D}\tau\Gamma\left(\frac{\gamma}{D} + 2\right)}{\Gamma\left(\frac{\gamma}{D} + \frac{3}{2}\right)} \quad (4.130)$$

and variance

$$\sigma_L^2(\tau) = D\tau \left(\frac{2\gamma}{D} - \frac{2\Gamma\left(\frac{\gamma}{D} + 2\right)^2}{\Gamma\left(\frac{\gamma}{D} + \frac{3}{2}\right)^2} + 3 \right). \quad (4.131)$$

Forming the joint probability distribution $P(L, \tau, L', \tau')$ from Eq. (4.111) and Eq. (4.112) we obtain the two-time covariance function

$$C_L(\tau, \tau') = \frac{2D\Gamma\left(\frac{\gamma}{D} + 2\right)^2 \left(\sqrt{\tau'\tau}^{-\frac{\gamma}{D}-2} (\tau - \tau')^{\frac{\gamma}{D} + \frac{5}{2}} {}_2F_1\left(\frac{\gamma}{D} + 2, \frac{\gamma}{D} + 2; \frac{\gamma}{D} + \frac{3}{2}; \frac{\tau'}{\tau}\right) - \sqrt{\tau'\tau} \right)}{\Gamma\left(\frac{\gamma}{D} + \frac{3}{2}\right)^2} \quad (4.132)$$

stated for $\tau > \tau'$, where ${}_2F_1(a, b; c; z)$ is the Gaussian hypergeometric function. The obtained mean Eq. (4.130), variance Eq. (4.131) and two-time covariance Eq. (4.132) are later in this text used as a baseline reference for further derivations.

Advected random walk

Assuming advection and thus $\alpha = 0$, the TSA approximation close to completion stated in Eq. (4.120) simplifies to

$$dL = \gamma \left(\frac{2}{1 - e^{-\frac{2\gamma L}{D}}} - 1 \right) d\tau + \sqrt{D} dW_\tau. \quad (4.133)$$

Close to completion, the dynamics match both in mean and variance with the exact aligned forward simulation as displayed in Fig. 4.8. With help of Eq. (4.133), we can now explain the general

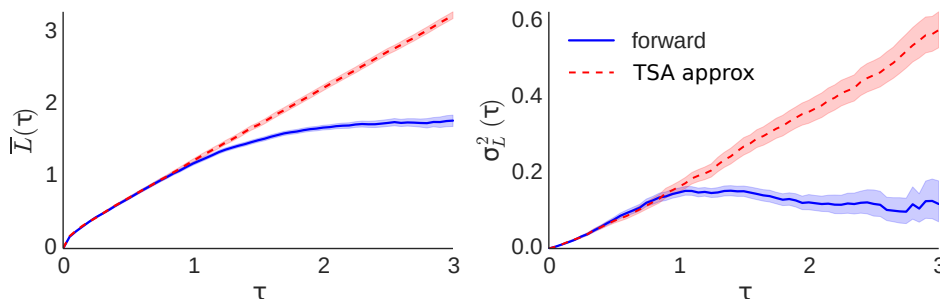


Figure 4.8: Forward dynamics (blue) and the approximate time reversed TSA dynamics (red) of the advected random walk agree excellently close to the target state. Shown are the mean (Left) and variance (Right) of both processes with 95% bootstrap confidence intervals. Results were obtained using each 1000 realizations of the respective ensemble with parameter settings $\gamma = 1$, $D = 0.2$ and $\hat{L}_{\text{in}} = 2$.

behavior of aligned advection processes close to completion. For small L , Eq. (4.133) reduces to $\frac{D}{L}$ and induces a square root like behavior of the mean (see chapter 3 or Eq. (4.130)($\gamma = 0$)). Further away from the absorbing boundary, the time reversed force term simplifies to γ , yielding a linear behavior of the mean. Then, eventually, the influence of the initial conditions ($\hat{L}_{\text{in}} = 2$) of the forward process can no longer be neglected and the approximated and exact mean start to deviate.

Ornstein-Uhlenbeck type process

The TSA-approximation of well separated initial and final conditions Eq. (4.120) reduces for the case of relaxation ($\alpha = 1$) to the form

$$dL = \left(-\gamma L + \frac{\sqrt{\gamma D}}{F\left(L\sqrt{\frac{\gamma}{D}}\right)} \right) d\tau + \sqrt{D} dW_\tau, \quad (4.134)$$

where $F(x) = e^{-x^2} \int_0^x e^{y^2} dy$ is the Dawson integral¹. In Fig. 4.9, we compare this approximation

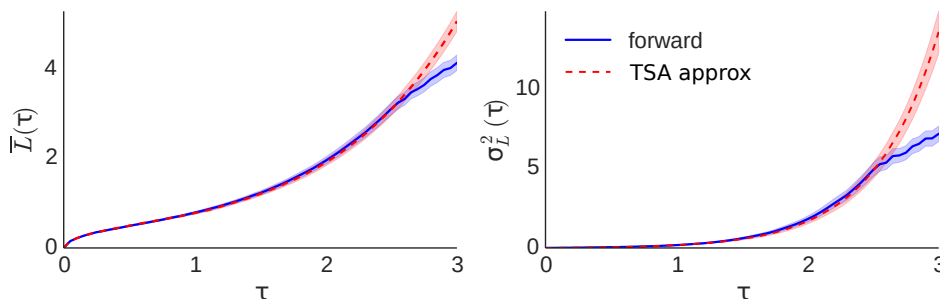


Figure 4.9: Forward dynamics (blue) and the approximate time reversed TSA dynamics (red) of an Ornstein-Uhlenbeck type process agree excellently close to the target state. Shown are the mean (Left) and variance (Right) of both processes with 95% bootstrap confidence intervals. Results were obtained using each 1000 realizations of the respective ensemble with parameter settings $\gamma = 1$, $D = 0.2$ and $\hat{L}_{\text{in}} = 10$.

with the forward simulation. In particular noteworthy here is the behavior of the mean, changing from a root like to a convex shape. We further observe that the approximate variance is a convex function, which seems to be a preserved feature of all processes with $\alpha > 0$.

4.6 Small L and weak noise approximation – Force and Noise driven TSA dynamics

Depending on the form of the force law $f(\widehat{L})$, dynamics close to the target state are either dominated by the deterministic drift or by random fluctuations which eventually determine when a trajectory reaches the target state. To quantitatively characterize this effect we assume that sufficiently close to the target state, most dynamics are well approximated by a powerlaw. Evaluating a small L -expansion of Eq. (4.120) then yields separable contributions due to the forward force law and due to the noise. Evaluating the limit $L \rightarrow L_{\text{ts}}$ allows to characterize which contribution dominates.

We start our derivation with the expansion of Eq. (4.120) for small L . The behavior of the expansion depends on the argument of both the upper incomplete gamma function and the exponential $\left| \frac{2\gamma L^{\alpha+1}}{D(\alpha+1)} \right|$ and whether it diverges or vanishes for $L \rightarrow 0$. If we only consider L then its behavior is clear and we observe divergence for $\alpha < -1$ and convergence for $\alpha > -1$. If we, however, additionally demand that our results should not deteriorate in the limit of small noise, i.e. $L \rightarrow 0$ and $D \rightarrow 0$, then the case $-1 < \alpha < 0$ switches its class and starts to diverge for $L \rightarrow 0$. A detailed argument and derivation is given below in this section. We here only state the results for the two regimes $\alpha \geq 0$ and $\alpha < 0$.

For $\alpha \geq 0$ the reverse time SDE Eq. (4.120) reduces to

$$dL(\tau) = \left(\frac{D}{L} - \gamma \frac{\alpha}{\alpha+2} L^\alpha + \mathcal{O}\left(\frac{L^{2\alpha+1}}{D}\right) \right) d\tau + \sqrt{D} dW_\tau \quad \alpha \geq 0. \quad (4.135)$$

Independent of the exact value of α , the force that drives the time reverse ensemble away from \widehat{L}_{ts} is thus always the pure free energy force $\frac{D}{L}$ as exhibited by the random walk case Eq. (4.92). This finding has important implications for the reconstruction of $f(L)$ from TSA data. Close to \widehat{L}_{ts} all time reversed processes with $\alpha \geq 0$ mimic pure diffusion. The identification of genuine forces has thus to rely on dynamics sufficiently far away from \widehat{L}_{ts} . For larger values of L , however, higher order terms can no longer be neglected. Their contributions mix and no clear universal approximation is achievable. We conclude that a reliable reconstruction of the genuine force $f(L)$ for $\alpha > 0$ is only possible using the full time reversed expression as stated in Eq. (4.120).

For the domain $\alpha < 0$, and for $-1 < \alpha < 0$ only under the assumption of small noise ($D \rightarrow 0$), Eq. (4.120) reduces to

$$dL(\tau) = \left(\gamma L^\alpha - \frac{\alpha D}{L} + \mathcal{O}\left(\frac{D^2}{L^{2+\alpha}}\right) \right) d\tau + \sqrt{D} dW_\tau \quad \alpha < 0. \quad (4.136)$$

For $\alpha < -1$, and for $0 < \alpha < -1$ given sufficiently small noise, the TSA dynamics close to the target state are fully governed by the sign inverted forward force law plus a contribution due to the noise similar to Eq. (4.135), however, here reweighted by a factor α . We note that the case $\alpha = -1$, i.e. the reverse time Bessel process Eq. (4.121), is exactly included in Eq. (4.136). Dynamics with $\alpha < 0$ should thus be distinguishable by inspecting their behavior near the target state.

In summary, Eq. (4.135) and Eq. (4.136) together allow to distinguish between noise driven Eq. (4.135) and force driven Eq. (4.136) TSA dynamics close to the target state. The noise driven case ($\alpha \geq 0$) close to the target state is indistinguishable from a pure TSA random

walk with free energy force $\frac{D}{L}$. This process is a Bessel process and fully characterized by the transition probability Eq. (4.111) ($\gamma = 0$) and density Eq. (4.112) ($\gamma = 0$). To form expectations about TSA ensemble data, we calculate then mean, variance and two-time correlation function below in this section.

The force driven case ($\alpha < 0$) has a non negligible contribution due to the sign inverted forward force close to the target state. This sensitivity to the parameter of the forward dynamics close to L_{ts} motivates to seek for a small noise expansion around the sign inverted forward law. Further below, we demonstrate that expressions for mean, variance and two-time correlation function can be obtained in the limit of small noise.

The derivation of the small L expansion

The expansion of $f^{\mathcal{F}}(L)$ Eq. (4.129) depends on the limit properties of the argument $z = -\frac{2L^{\alpha+1}\gamma}{\alpha D + D}$ of the rewritten free energy force

$$f^{\mathcal{F}}(L) = \frac{C(D)e^{-z}}{\Theta(\frac{1}{\nu})\Gamma(\nu) - \Gamma(\nu, z)} \quad \alpha \neq -1 \quad (4.137)$$

and whether the limit of $|z|$, with respect to D and L , rather approaches zero or infinity. For readability, we collect all non z -dependent terms into

$$C(D) = (D\alpha + D) \left(-\frac{2\gamma}{D\alpha + D} \right)^{\frac{1}{\alpha+1}} \quad (4.138)$$

and set $\nu = \frac{1}{1+\alpha}$. In the following three subsections we discuss the expansion of $f^{\mathcal{F}}(z)$ with respect to z , given small L , for the regimes $\alpha \in (-\infty, -1), (-1, \infty)$. We then reassess the reverse time forces in the interval $\alpha \in (-1, 0)$ and with respect to its dependency on the size of the diffusion constant D .

The case $|z| \rightarrow 0$

For $\alpha > -1$ ($\nu > 0$) the expansion variable $|z|$ decreases with decreasing L . For these cases the denominator of Eq. (4.137) can be written as

$$\Gamma(\nu) - \Gamma(\nu, z) = \Gamma(\nu)z^\nu e^{-z} \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\nu + n + 1)} \quad \text{for } |z| < \infty, \quad (4.139)$$

using Eq. (AS 6.5.3), Eq. (AS 6.5.4) and Eq. (AS 6.5.29)¹. The free energy force $f^{\mathcal{F}}(L)$ is

$$f^{\mathcal{F}}(L) = \frac{C(D)}{\Gamma(\nu)z^\nu \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\nu+n+1)}} \quad \alpha > -1, \quad (4.140)$$

where we canceled the exponential function in numerator and denominator. Truncating the series in the denominator to include only the terms up to first order, we arrive at

$$f^{\mathcal{F}}(L) = \frac{\nu C(D)z^{-\nu}}{1 + \frac{z}{\nu+1} + \mathcal{O}(z^2)} \quad \alpha > -1. \quad (4.141)$$

We here used the recurrence relation $\Gamma(\nu + 1) = \nu\Gamma(\nu)$ from Eq. (AS 6.1.15) Abramowitz & Stegun¹ to simplify the expression. In the last step we approximate the denominator to lowest order in z and obtain

$$f^{\mathcal{F}}(L) = \nu C(D)z^{-\nu} \left(1 - \frac{z}{\nu+1} + \mathcal{O}(z^2) \right) \quad \alpha > -1. \quad (4.142)$$

Replacing the effective parameters and variable z with the original parameters ν, γ, D and variable L , we arrive at the small L expansion

$$f^{\mathcal{F}}(L) = \frac{D}{L} + \frac{2\gamma L^\alpha}{\alpha + 2} + \mathcal{O}\left(\frac{L^{1+2\alpha}}{D}\right) \quad \alpha > -1. \quad (4.143)$$

Substituting Eq. (4.143) into Eq. (4.118) we arrive at the small L -expansion stated in Eq. (4.135).

The case $|z| \rightarrow \infty$

For $\alpha < -1$ ($\nu < 0$) the expansion variable z goes to infinity for small L . With $\nu < 0$ the Heaviside step function evaluates to zero and the denominator simplifies to $\Gamma(\nu, z)$ which we expand for large z , using the known asymptotic expansion Eq. (AS 6.5.32)¹

$$\Gamma(\nu, z) = z^{\nu-1} e^{-z} \left(1 + \frac{\nu-1}{z} + \mathcal{O}(z^{-2}) \right) \quad \text{for } |z| \rightarrow \infty. \quad (4.144)$$

We substitute this expansion into Eq. (4.137) and arrive at

$$f^{\mathcal{F}}(L) = -\frac{C(D)z^{1-\nu}}{1 + \frac{\nu-1}{z} + \mathcal{O}(z^{-2})} \quad \alpha < -1 \quad (4.145)$$

after canceling the exponential terms. In a final step we approximate the denominator for $z \rightarrow \infty$, i.e. Taylor expand for small orders of $1/z$, and obtain

$$f^{\mathcal{F}}(L) = -C(D)z^{1-\nu} \left(1 - \frac{\nu-1}{z} + \mathcal{O}(z^{-2}) \right) \quad \alpha < -1. \quad (4.146)$$

Re-substituting the original parameters and variable L , the final free energy force reads

$$f^{\mathcal{F}}(L) = 2\gamma L^\alpha - \frac{\alpha D}{L} + \mathcal{O}\left(\frac{D^2}{L^{2+\alpha}}\right) \quad \alpha < -1. \quad (4.147)$$

The L and D dependent range $-1 < \alpha < 0$

For α in the range $-1 < \alpha < 0$, the value of the expansion variable z depends on the choice of the diffusion constant D . For large D the limit behavior of z is fully defined by L and we find

$$\lim_{L \rightarrow 0} |z| = \lim_{L \rightarrow 0} \left| -\frac{2L^{\alpha+1}\gamma}{\alpha D + D} \right| = 0 \quad \text{for } -1 < \alpha < 0. \quad (4.148)$$

If we additionally assume, that the expansion holds in the limit of small noise, the limiting behavior changes and we find

$$\lim_{L \rightarrow 0, D \rightarrow 0} |z| = \lim_{L \rightarrow 0, D \rightarrow 0} \left| -\frac{2L^{\alpha+1}\gamma}{\alpha D + D} \right| = \infty \quad \text{for } -1 < \alpha < 0. \quad (4.149)$$

This implies that for a given L , we can always find a small enough D so that the expansion variable z becomes large. Thus, we also consider $|z| \rightarrow \infty$ in the regime $-1 < \alpha < 0$. Under this condition, the $\Theta(\alpha + 1)$ dependent term in the denominator of Eq. (4.137) does no longer vanish automatically. Eq. (4.145) then reads

$$f^{\mathcal{F}}(L) = -\frac{C(D)z^{1-\nu}}{m(z) + 1 + \frac{\nu-1}{z} + \mathcal{O}(z^{-2})} \quad \alpha < -1 \quad (4.150)$$

with

$$m(z) = -e^z z^{1-\nu} \Gamma(\nu) . \quad (4.151)$$

In the following we argue, that $m(z)$ becomes negligible for large $|z|$. We show that the limes of $m(z)$ with respect to $D \rightarrow 0$ is zero and conclude that we can always find a D , which is small enough to neglect $m(z)$ in the expansion of $f^{\mathcal{F}}(L)$. For small D , Eq. (4.150) therefore follows the same expansion as the case $\alpha < -1$ discussed before.

To proof these statements, we first recall, that $z = -\frac{2L^{1+\alpha}\gamma}{(1+\alpha)D}$ is negative for $-1 < \alpha < 0$. The exponential thus decays with increasing $|z|$. With $\nu = \frac{1}{1+\alpha}$, the exponent $1 - \nu$ is negative for $-1 < \alpha < 0$ and the power law $|z|^{1-\nu}$ decays with increasing $|z|$. For $|z| \rightarrow \infty$ we therefore find

$$\lim_{|z| \rightarrow \infty} |m(|z|)| = 0 \quad \text{for} \quad -1 < \alpha < 0 . \quad (4.152)$$

We next show, that $|m(z)|$ also approaches zero for $D \rightarrow 0$. We start by rewriting $m(z)$ in terms of α, γ, D and L and obtain

$$m(L) = \Gamma\left(\frac{1}{\alpha+1}\right) L^\alpha \left(-\frac{2\gamma}{D(1+\alpha)}\right)^{\frac{\alpha}{\alpha+1}} e^{-\frac{2\gamma L^{\alpha+1}}{D(1+\alpha)}} . \quad (4.153)$$

To clarify the behavior of the first D -dependent term we note, that its absolute value can be written as $\left|\left(-\frac{2\gamma}{D(1+\alpha)}\right)^{\frac{\alpha}{\alpha+1}}\right| = \left|\frac{D(1-|\alpha|)}{2\gamma}\right|^{\frac{1-|\alpha|}{|\alpha|}}$. With $|\alpha| \in (0, 1)$, the exponent is greater than zero and the full term goes to zero for $D \rightarrow 0$. The second D dependent term in Eq. (4.153) is the exponential. For $D \rightarrow 0$ the exponent goes to $-\infty$ and the exponential vanishes. We can therefore write

$$\lim_{D \rightarrow 0} |m(L)| = 0 \quad \text{for} \quad -1 < \alpha < 0 . \quad (4.154)$$

In the final step we show that $|m(L)|$ also vanishes in the simultaneous limit of $D \rightarrow 0$ and $L \rightarrow 0$. To perform this limit we set w.l.o.g. $D = L$ and rewrite Eq. (4.153) accordingly. In the resulting equation

$$m(L) = \Gamma\left(\frac{1}{\alpha+1}\right) L^{\frac{\alpha^2}{1+\alpha}} \left(-\frac{2\gamma}{1+\alpha}\right)^{\frac{\alpha}{\alpha+1}} e^{-\frac{2\gamma L^\alpha}{1+\alpha}} , \quad (4.155)$$

we can directly read off that the powerlaw $L^{\frac{\alpha^2}{1+\alpha}}$ decays to zero for decreasing L , as does the exponential. We conclude, that

$$\lim_{\substack{L \rightarrow 0, D \rightarrow 0 \\ D \sim L}} |m(L)| = 0 \quad \text{for} \quad -1 < \alpha < 0 \quad (4.156)$$

holds. It is thus always possible to find a small enough D , where the large z expansion of $f^{\mathcal{F}}(L)$, discussed above, holds also in the regime $-1 < \alpha < 0$.

We conclude this section by a remark. Reinspecting Eq. (4.153) we observe that all D 's are multiplied by $1 + \alpha$. For α decreasing from zero to minus one we therefore expect $m(L)$ to become smaller and smaller until it approaches zero for $\alpha \rightarrow -1$ using the very same argument as for D . The large $|z|$ expansion will therefore hold for the larger D values, the closer α is to -1 .

Moments and universal correlation function for noise driven TSA ensembles

For $\alpha > 0$, the SDE Eq. (4.135) with respect to the free energy force $f^{\mathcal{F}}(L)$ expanded in small L tells us that the dynamics close to the target state are purely governed by the force term $\frac{D}{L}$. From section 4.4, or simply by inserting $f(L) = 0$ into Eq. (4.118), we know, that the SDE

$$dL(\tau) = \frac{D}{L}d\tau + \sqrt{D} dW_\tau \quad (4.157)$$

describes the dynamics of a TSA random walk close to the target state and well separated from its forward initial conditions. For completeness and easy use we here state the mean

$$\bar{L}(\tau) = \sqrt{\frac{8D\tau}{\pi}}, \quad (4.158)$$

variance

$$\sigma_L^2(\tau) = \frac{(3\pi - 8)D\tau}{\pi} \quad (4.159)$$

and two-time covariance

$$C_L(\tau, \tau') = \begin{cases} \frac{2D(-4\sqrt{\tau\tau'} + 3\sqrt{\tau(\tau'-\tau)} + (2\tau + \tau') \sin^{-1}(\sqrt{\frac{\tau}{\tau'}}))}{\pi} & \text{for } \tau < \tau' \\ \sigma_L^2(\tau) & \text{for } \tau = \tau' \\ \frac{2D(-4\sqrt{\tau'\tau} + 3\sqrt{\tau'(\tau-\tau')} + (2\tau' + \tau) \sin^{-1}(\sqrt{\frac{\tau'}{\tau}}))}{\pi} & \text{for } \tau > \tau'. \end{cases} \quad (4.160)$$

They can be obtained either from the known transition probability Eq. (4.111) and density Eq. (4.112) of the Bessel process with an entrance boundary at $L_{\text{ts}} = 0$ and in the limit $\gamma \rightarrow 0$, or from the moments stated in section 4.4. Using the obtained variance and covariance, the correlation function $\text{corr}_L(\tau, \tau') = C_L(\tau, \tau')/(\sigma_L(\tau)\sigma_L(\tau'))$ reads

$$\text{corr}_L(\tau, \tau') = \begin{cases} \frac{2(-4\sqrt{\tau\tau'} + 3\sqrt{\tau(\tau'-\tau)} + (2\tau + \tau') \sin^{-1}(\sqrt{\frac{\tau}{\tau'}}))}{(3\pi - 8)\sqrt{\tau\tau'}} & \text{for } \tau < \tau' \\ 1 & \text{for } \tau = \tau' \\ \frac{2(-4\sqrt{\tau'\tau} + 3\sqrt{\tau'(\tau-\tau')} + (2\tau' + \tau) \sin^{-1}(\sqrt{\frac{\tau'}{\tau}}))}{(3\pi - 8)\sqrt{\tau'\tau}} & \text{for } \tau > \tau'. \end{cases} \quad (4.161)$$

Let us briefly discuss the obtained reverse time moments. For the mean, we observe that it grows with $\sqrt{\tau}$. It is thus similar to a forward random walk which starts at the boundary but never returns to it. Such a random walk can be constructed from Eq. (4.89) (random walk with mirror charges), after normalization to one (for every τ), and in the limit $\tilde{L}_0 \rightarrow 0$. The result is the same we obtained as an intermediate step in the derivation of the exact TSA reverse time Bessel process Eq. (4.113)

$$P(L, \tau) = \frac{Le^{-\frac{L^2}{2D\tau}}}{D\tau} \quad (4.162)$$

and is known as a Rayleigh distribution. For the mean of this restricted random walk we find

$$\bar{L}(\tau) = \sqrt{\frac{\pi D\tau}{2}}, \quad (4.163)$$

which shows the same τ dependence as the mean of the TSA random walk Eq. (4.158) but with a different pre-factor. This observations motivates to also study the variance of the restricted random walk

$$\sigma_L^2(\tau) = \frac{1}{2}(4 - \pi)D\tau . \quad (4.164)$$

We again find the same dependence on τ as in the variance of the TSA random walk but again a different pre-factor. This shows that fitting mean or variance independently can lead to wrong conclusions. Comparing for example the coefficient of variation $CV = \frac{\sigma_L(\tau)}{\bar{L}(\tau)}$ we find the two different constants $CV_{\text{TSA-rw}} = \sqrt{\frac{3\pi}{8} - 1} \approx 0.42$ for the TSA random walk and $CV_{\text{rest-rw}} = \sqrt{\frac{4}{\pi} - 1} \approx 0.52$ for the restricted random walk. This shows the TSA random walk, as one might have assumed, is not the same as a restricted random walk.

While sample paths in both cases are generated by the same dynamics, they differ in the selection of sample paths which contribute to the ensemble. In the TSA ensemble, sample paths in principle have had the time to explore the full half plane $L \in [0, \infty]$ before they reach the target state. The fraction of sample paths which predominantly “walk” in one direction close to the target state is therefore enriched compared to a restricted random walk, which only neglects all sample paths which ever return to the initial conditions. Formally, its the difference between a TSA ensemble and a meander process¹⁰⁴, which never returns to its initial position. For mathematical details on meander processes we refer to Majumdar and Orland¹⁰⁴, for a constructive perspective that helps understanding this effect to section 7.4.

Moments for force driven TSA ensembles

For $\alpha < -1$ and for $-1 < \alpha < 0$ under the additional constraint of small D , TSA dynamics are force driven close to the target state. This motivates to try a small noise expansion around the deterministic time reversed dynamics to obtain expression for the moments of the SDE Eq. (4.136). We adopt the approach defined for a standard SDE (see Gardiner chapter 6.2⁵⁵) to include order D drift terms and find for the mean

$$\bar{L}(\tau) = ((1 - \alpha)\gamma\tau)^{\frac{1}{1-\alpha}} + D \frac{(7\alpha - 3)((1 - \alpha)\gamma\tau)^{\frac{\alpha}{\alpha-1}}}{4(3\alpha - 1)\gamma} + \mathcal{O}(D^{3/2}) , \quad (4.165)$$

variance

$$\sigma_L^2(\tau) = D \frac{1 - \alpha}{1 - 3\alpha} \tau + \mathcal{O}(D^2) \quad (4.166)$$

and covariance

$$C_L(\tau, \tau') = \begin{cases} \left(\frac{\tau}{\tau'}\right)^{\frac{\alpha}{\alpha-1}} \sigma_L^2(\tau) + \mathcal{O}(D^2) & \text{for } \tau < \tau' \\ \left(\frac{\tau'}{\tau}\right)^{\frac{\alpha}{\alpha-1}} \sigma_L^2(\tau') + \mathcal{O}(D^2) & \text{for } \tau > \tau' \\ \sigma_L^2(\tau) + \mathcal{O}(D^2) & \text{for } \tau = \tau' \end{cases} \quad (4.167)$$

by a perturbation theory calculation around the deterministic solution of Eq. (4.136) in the Diffusion constant up to order D . The correlation function $\text{corr}_L(\tau, \tau') = C_L(\tau, \tau')/(\sigma_L(\tau)\sigma_L(\tau'))$ reads

$$\text{corr}_L(\tau, \tau') = \begin{cases} \left(\frac{\tau}{\tau'}\right)^{\frac{3}{2} + \frac{1}{\alpha-1}} & \text{for } \tau < \tau' \\ \left(\frac{\tau'}{\tau}\right)^{\frac{3}{2} + \frac{1}{\alpha-1}} & \text{for } \tau > \tau' \\ 1 & \text{for } \tau = \tau' . \end{cases} \quad (4.168)$$

Note, that for all expressions (Eq. (4.165)-Eq. (4.168)) boundary conditions are already implicitly included due to the contribution of the free energy force. We further recall that the expansion stated in Eq. (4.136) was truncated at order D . Only terms up to this order are therefore considered in Eq. (4.165), Eq. (4.166) and Eq. (4.167). Given this approximation, we have shown above in this section that only the mean Eq. (4.165) carries a relevant contribution from the approximate free energy force $-\frac{\alpha D}{L}$.

The obtained expression can be used to infer power law forward forces $f(L) = -\gamma L^\alpha$ from TSA ensembles in the limit of small noise. For example, starting from the two-time correlation function, we read off the power law exponent α . With α we can fit the variance and determine the diffusion constant D . The force strength γ is obtained from the mean.

For the small noise approximation to hold well, and as a rule of thumb, the coefficient of variation $CV = \sigma_L(\tau)/\bar{L}(\tau)$, approximated as the lowest order small noise terms, should be well below one. For example for $\alpha = -1$ this implies $CV = \sqrt{D/(4\gamma)} < 1$. For all other α -values we find $CV \sim \tau^{\frac{\alpha+1}{2(\alpha-1)}}$. The approximation thus becomes better the larger τ given $-1 < \alpha < 0$, and for $\alpha < -1$ the smaller τ . In the following, I will briefly comment on the expressions for $\bar{L}(\tau)$, $\sigma_L^2(\tau)$ and $C_L(\tau, \tau')$, and present a more detailed discussion and a step by step derivation below.

The mean $\bar{L}(\tau)$ (Eq. (4.165)) is given as the deterministic solution to Eq. (4.136) ($D \rightarrow 0$) plus a (for all $\alpha < 0$) solely positive contribution of order D . The case $\alpha < 0$ thus shows that a purely deterministic approach to extract the mean from any experimentally obtained and terminally aligned data is bound to fail. As the mean, given by Eq. (4.165), adds up a deterministic term and a deterministic positive contribution due to the noise, a pure deterministic approach overestimates the underlying forward force by the contribution of the noise. The difference between the deterministic approximation (order D^0 of Eq. (4.165)) and the correct analytic obtainable solution Eq. (4.130) is exemplary shown in Fig. 4.10 (Left) for the case $\alpha = -1$. To

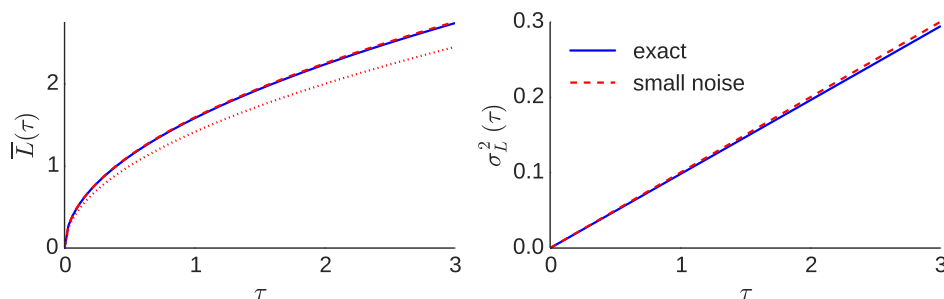


Figure 4.10: The small noise approximation up to order D is an excellent approximation of the exact dynamics. Shown are small noise expressions (red) for the mean (Left) Eq. (4.165) and variance (Right) Eq. (4.166) in comparison to the exact analytic expressions (blue), as stated in Eq. (4.130) Eq. (4.131) for the Bessel process. For the small noise expression of the mean we additionally show the curve up to order D^0 (dotted). Curves are obtained for the parameters $\alpha = -1$, $\gamma = 1$, $D = 0.2$.

be precise, in Fig. 4.10 we compare the exact solution of the TSA Bessel process with forward initial conditions shifted to infinity, and the small noise approximation stated in Eq. (4.165) and Eq. (4.166). We find excellent agreement.

The variance $\sigma_L^2(\tau)$, up to order D (Eq. (4.166)) is a γ -independent function linearly growing with $D\tau$. It thus shows the same dependence on $D\tau$ as the exact variance result Eq. (4.131) for the Bessel process SDE Eq. (4.136) with $\alpha = -1$. Furthermore, if we assume $D/\gamma \rightarrow 0$, the variance for $\alpha = -1$, as stated in Eq. (4.131) approaches the small noise solution from below until both assume the form $\frac{D\tau}{2}$. Taken together this indicates that in the limit of small noise, characterized by the CV , all relevant features of $\sigma_L^2(\tau)$ and $C_L(\tau, \tau')$, are capture within an

expansion up to order D . To further test this for general $\alpha < 0$ we compare numerical solutions of Eq. (4.136) with our analytic small noise expressions for mean Eq. (4.165) and variance Eq. (4.166). Results for $\alpha = -0.5$ and $\alpha = -2$ are displayed in Fig. 4.11 and compared to the case $\alpha = -1$. All considered, we conclude that $\bar{L}(\tau)$ and $\sigma_L^2(\tau)$, as stated in Eq. (4.165) and

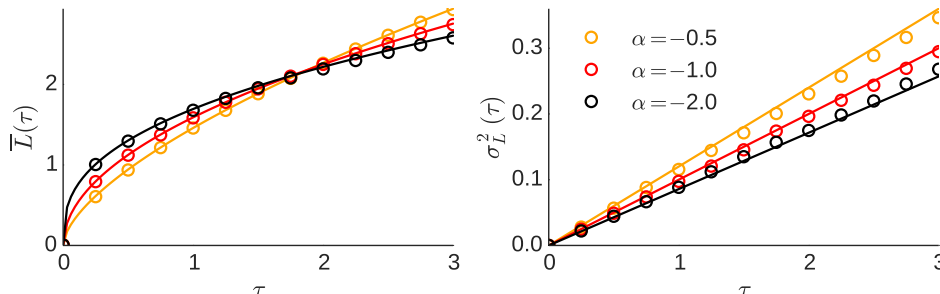


Figure 4.11: The small noise expansion is a very good approximation of its underlying TSA-SDE. Comparison between a numerical evaluation of Eq. (4.136) (circles) and its approximation as analytic small noise expansion up to order D^1 (lines) as given in Eq. (4.165) and Eq. (4.166). Shown are mean (Left) and variance (Right) for the cases $\alpha = -1/2$ (orange), $\alpha = -1$ (red) and $\alpha = -2$ (black). All other parameters are chosen identical with $\gamma = 1$ and $D = 0.2$.

Eq. (4.166), are sufficient to reconstruct $f(L)$ and D for general $\alpha < 0$ and in the limit of small noise.

The two-time covariance function $C_L(\tau, \tau')$ is highly informative for force reconstruction. If enough sample paths are available ($\mathcal{O}(100) - \mathcal{O}(1000)$), the covariance stated in Eq. (4.167) is suitable to uniquely infer both the general form of the forward force $f(L)$ and the noise level.

The fraction $C_L(\tau, \tau')/(\sigma_L(\tau)\sigma_L(\tau'))$ represents the two time correlation function $\text{corr}_L(\tau, \tau')$. With its sole dependence on α , $\text{corr}_L(\tau, \tau')$ uniquely fixes the general form of the force law, irrespective of the reconstructed γ and the noise level D . To check for the quality of Eq. (4.168), we compare in Fig. 4.12 the case $\alpha = -1$ to its analytical counterpart obtained from Eq. (4.131) and Eq. (4.132). Using the same parameters as already in Fig. 4.10 we observe in Fig. 4.12

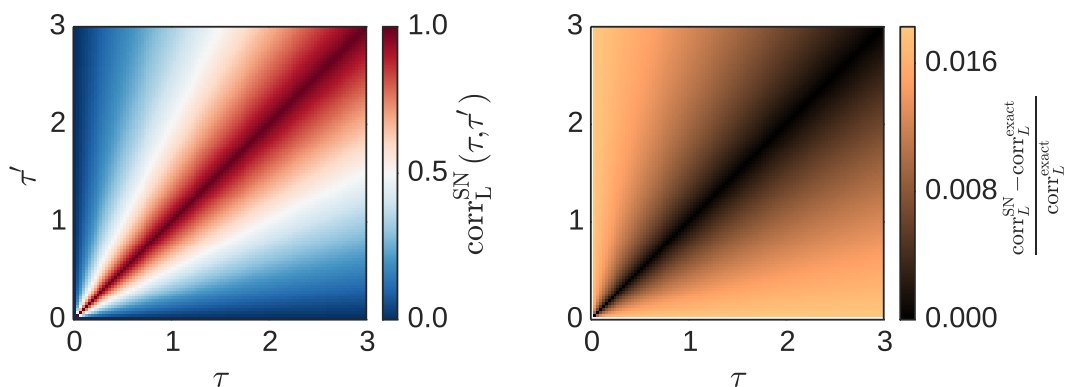


Figure 4.12: The small noise approximation of the correlation function is an excellent approximation of the exact dynamics. Shown is the small noise correlation function as given by Eq. (4.168) for the case $\alpha = -1$ (Left) and its deviations from the known exact correlation function of the Bessel process (Right) directly obtainable from Eq. (4.131) and Eq. (4.132).

(Right) a maximal deviation of 2% for large correlation times, and even smaller deviations for shorter correlation times.

Interestingly, the 2d representation of $\text{corr}_L(\tau, \tau')$ as stated in Eq. (4.168) and displayed for the cases $\alpha = -0.5$ and $\alpha = -2$ in Fig. 4.13, shows an easy to identify signature. Due to the sole

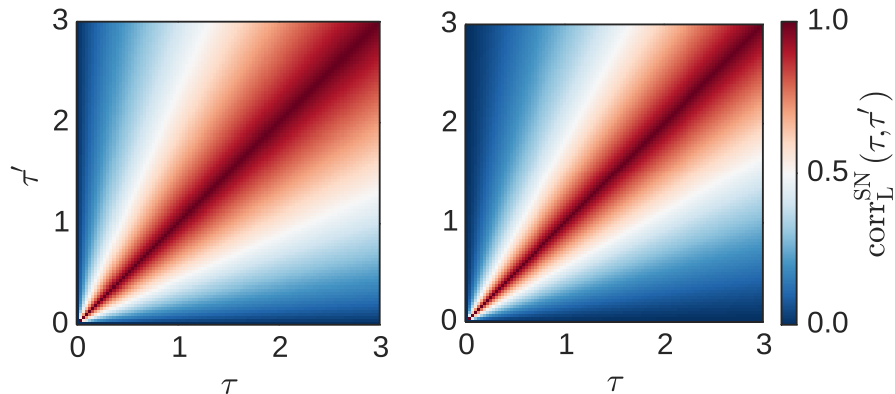


Figure 4.13: Exemplary correlation functions obtained from the analytic small noise expression as stated in Eq. (4.168). Shown are the cases $\alpha = -1/2$ (Left) and $\alpha = -2$ (Right). The two remaining parameters are chosen identical with $\gamma = 1$ and $D = 0.2$. For both we observe characteristic straight equi-value lines as demanded by Eq. (4.168).

dependence of Eq. (4.168) on the ratio of τ and τ' the 2d correlation function exhibits straight level lines ($\tau' = \text{const} \cdot \tau$). Given we are in the regime of small noise and do not observe this straight line structure in our experimental data, we can thus directly exclude all pure forward force laws of the form stated in Eq. (4.119) with $\alpha < 0$. We further know from above that processes with $\alpha \geq 0$ follow a force law D/L close to completion. The respective correlation function can be calculated and is stated in Eq. (4.161). Connecting the two cases $\alpha < 0$ and $\alpha > 0$ we now have a theoretical understanding of the mean, variance and two-time correlation close to target states.

Weak noise approximation of moments for force driven TSA ensembles

Starting from Eq. (4.136), we here derive the expressions Eq. (4.165)-Eq. (4.168) for mean $\bar{L}(\tau)$, variance $\sigma_L^2(\tau)$, covariance $C_L(\tau, \tau')$ and the correlation function $\text{corr}_L(\tau, \tau')$. The calculation is valid in the domain $\alpha < 0$, as discussed above, and in the limit of small noise. With some minor modifications we follow the weak noise approximation approach discussed in Gardiner (chapter 6.2)⁵⁵.

The derivation starts with rewriting Eq. (4.136) in the general form

$$dL = a(L)d\tau + \epsilon^2 b(L)d\tau + \epsilon dW_\tau \quad (4.169)$$

with \sqrt{D} substituted by the expansion parameter ϵ . For small ϵ we expand

$$L(\tau) = L_0(\tau) + \epsilon L_1(\tau) + \epsilon^2 L_2(\tau) + \dots \quad (4.170)$$

around the deterministic solution $L_0(\tau)$. We further assume that we can write

$$a(L) = a(L_0 + \epsilon L_1 + \epsilon^2 L_2 + \dots) = a_0(L_0) + \epsilon a_1(L_0, L_1) + \epsilon^2 a_2(L_0, L_1, L_2) + \dots \quad (4.171)$$

The general expression for $a(L)$ then reads

$$a(L) = a\left(L_0 + \sum_{m=1}^{\infty} \epsilon^m L_m\right) = \sum_{p=0}^{\infty} \frac{1}{p!} \frac{d^p a(L_0)}{dL_0} \left(\sum_{p=0}^{\infty} \epsilon^m L_m\right)^p \quad (4.172)$$

and after sorting terms we find for the first three terms

$$a_0(L_0) = a(L_0) \quad (4.173)$$

$$a_1(L_0, L_1) = L_1 \frac{da(L_0)}{dL_0} \quad (4.174)$$

$$a_2(L_0, L_1, L_2) = L_2 \frac{da(L_0)}{dL_0} + \frac{1}{2} L_1^2 \frac{d^2a(L_0)}{dL_0^2} . \quad (4.175)$$

The same procedure can be applied to $b(L)$. Taken together the equations to lowest order in ϵ are

$$dL_0 = a(L_0)d\tau \quad (4.176)$$

$$dL_1 = a_1(L_1, L_0)d\tau + dW_\tau \quad (4.177)$$

$$dL_2 = a_2(L_2, L_1, L_0)d\tau + b(L_0)d\tau \quad (4.178)$$

The first of these equations is the deterministic equation

$$dL_0 = \gamma L_0^\alpha d\tau . \quad (4.179)$$

As the process starts at the boundary at $L_0(0) = 0$, its solution is simply given as

$$L_0 = ((1 - \alpha)\gamma\tau)^{\frac{1}{1-\alpha}} \quad \text{with } \alpha < 0 . \quad (4.180)$$

The second equation Eq. (4.177) describes a generalized Ornstein-Uhlenbeck process

$$dL_1 = k(L_0(\tau))L_1 d\tau + dW_\tau \quad (4.181)$$

with the time-dependent drift coefficient

$$k(L_0) = \frac{da(L_0)}{dL_0} = \gamma \frac{dL_0^\alpha}{dL_0} = \gamma\alpha L_0^{\alpha-1} . \quad (4.182)$$

Inserting the solution for $L_0(\tau)$ we find

$$k(L_0) = \gamma\alpha ((1 - \alpha)\gamma\tau)^{-1} . \quad (4.183)$$

Employing the initial conditions $L_1(0) = 0$, the solution of equation Eq. (4.181) reads

$$L_1(\tau) = \int_0^\tau e^{\int_{\tau'}^\tau k(L_0(s))ds} dW_{\tau'} = \int_0^\tau \left(\frac{\tau}{\tau'}\right)^{-\frac{\alpha}{\alpha-1}} dW_{\tau'} . \quad (4.184)$$

Together, Eq. (4.180) for $L_0(\tau)$ and Eq. (4.184) for $L_1(\tau)$ form the basis of all further calculations. Inserting these solutions into the expansion stated in Eq. (4.170) and taking the ensemble average we obtain expressions for mean

$$\langle L(\tau) \rangle = \langle L_0(\tau) \rangle + \epsilon \langle L_1(\tau) \rangle + \dots , \quad (4.185)$$

variance

$$\sigma_L^2(\tau) = \langle (L(\tau) - \langle L(\tau) \rangle)^2 \rangle = \epsilon^2 \langle L_1^2(\tau) \rangle + \dots \quad (4.186)$$

and two-time covariance

$$C(\tau, \tau') = \langle (L(\tau) - \langle L(\tau) \rangle) (L(\tau') - \langle L(\tau') \rangle) \rangle = \epsilon^2 \langle L_1(\tau) L_1(s) \rangle + \dots \quad (4.187)$$

in lowest order of ϵ and valid for $\alpha < 0$. We first evaluate these expressions to lowest order, which for $\sigma_L^2(\tau)$ and $C_L(\tau, \tau')$ is already at order $\mathcal{O}(D)$. Corrections for the mean up to order $\mathcal{O}(D)$ are evaluated subsequently. Note, that the average over a single Wiener integral, such as $\langle L_1(\tau) \rangle$ always evaluates to zero and therefore corrections to the mean will be of order ϵ^2 .

To 0th order, the solution for the mean is equivalent to the deterministic solution Eq. (4.180). For the variance, we have to calculate the second moment of equation Eq. (4.184), i.e.

$$\begin{aligned} \sigma_L^2(\tau) &= D \langle L_1^2(\tau) \rangle = D \int_0^\tau \int_0^\tau \left(\frac{\tau}{\tau'}\right)^{-\frac{\alpha}{\alpha-1}} \left(\frac{\tau}{\tau''}\right)^{-\frac{\alpha}{\alpha-1}} dW_{\tau'} dW_{\tau''} \\ &= D \int_0^\tau \left(\frac{\tau}{\tau'}\right)^{-\frac{2\alpha}{\alpha-1}} d\tau' = D \frac{1-\alpha}{1-3\alpha} \tau \quad \text{for } \alpha < 0, \end{aligned} \quad (4.188)$$

where we replace ϵ^2 by the original D . Note that $\langle L_1 \rangle^2 = 0$ and thus drops out of the calculation. Similarly, we obtain the two-time covariance function

$$\begin{aligned} C(\tau, \tau') &= D \int_0^\tau \int_0^{\tau'} \left(\frac{\tau}{s}\right)^{-\frac{\alpha}{\alpha-1}} \left(\frac{\tau'}{s'}\right)^{-\frac{\alpha}{\alpha-1}} dW_s dW_{s'} \\ &= D \int_0^{\min(\tau, \tau')} \left(\frac{\tau}{s}\right)^{-\frac{\alpha}{\alpha-1}} \left(\frac{\tau'}{s}\right)^{-\frac{\alpha}{\alpha-1}} ds \\ &= \begin{cases} D \frac{(\alpha-1)\tau^{\frac{1}{\alpha-1}+2}\tau'^{-\frac{\alpha}{\alpha-1}}}{3\alpha-1} & \text{for } \tau < \tau' \\ D \frac{(\alpha-1)\tau'^{\frac{1}{\alpha-1}+2}\tau^{-\frac{\alpha}{\alpha-1}}}{3\alpha-1} & \text{for } \tau > \tau' \\ D \langle L_1^2(\tau) \rangle & \text{for } \tau = \tau' \end{cases} \quad \text{and } \alpha < 0 \end{aligned} \quad (4.189)$$

again in terms of D . Using the variance Eq. (4.188), we can rewrite this result in the final form stated in Eq. (4.166). The two-time correlation function also directly follows from Eq. (4.189). Using the definition

$$\text{Corr}(\tau, \tau') := \frac{C(\tau, s)}{\sigma_L(\tau)\sigma_L(\tau')} \quad (4.190)$$

we obtain the final form stated in Eq. (4.168).

In a last step we calculate the first D -dependent contribution to the mean, characterizing deviations from the deterministic solution. As contributions due to $L_1(\tau)$ average out, terms of order ϵ^2 become important and thus the contribution $-\alpha D/L$ of the free energy force as stated in Eq. (4.136). Using the series expansion of $L(\tau)$ in orders of ϵ as stated in Eq. (4.170), we next evaluate the contribution of $L_2(\tau)$ to the mean. This is most simply achieved by directly evaluating an averaged Eq. (4.178). Using the specifications due to Eq. (4.175) we then obtain a linear differential equation

$$\frac{d\langle L_2 \rangle}{d\tau} = \frac{da(L_0)}{dL_0} \langle L_2 \rangle + \frac{1}{2} \frac{d^2a(L_0)}{dL_0^2} \langle L_1^2 \rangle - \frac{\alpha}{L_0} \quad (4.191)$$

with the solution

$$\langle L_2(\tau) \rangle = \frac{(7\alpha - 3)((1 - \alpha)\gamma\tau)^{\frac{\alpha}{\alpha-1}}}{4(3\alpha - 1)\gamma} \quad \text{for } \alpha < 0. \quad (4.192)$$

The final result Eq. (4.165) is obtained from

$$\bar{L}(\tau) = L_0(\tau) + D \langle L_2(\tau) \rangle. \quad (4.193)$$

Testing TSA model inference and distinction on a simple model of cytokinesis

5.1 A model of cytokinetic ring constriction with distinct effective force laws

To assess whether TSA ensemble analysis can distinguish different regimes of directed dynamics we here study the exemplary process of Cytokinesis. Cytokinesis is the process of separation of a mother cell into two daughter cells at the end of the cell cycle. For animal cells this separation is driven by a contractile ring formed from actin filaments, myosin and associated proteins. We used a model that describes the contraction and molecular turnover of force generating actomyosin bundles, i.e. an assembly of filaments and motor complexes¹⁸⁴. Two coupled differential equations define the model. The first describes how the balance of viscous and elastic forces determines the ring diameter and leads to ring constriction or dilation. The viscous force is written as $-\xi^{-1}\dot{R}$, where \dot{R} denotes the time derivative of the ring radius and ξ an effective friction coefficient. The elastic response of the cell upon cell deformation derives from a harmonic energy term

$$E(R) = \frac{K}{2}(R - R_0)^2 \quad (5.1)$$

with elastic modulus $K > 0$ and equilibrium radius R_0 . Combined with a term Σ for the contractile stress, we obtain a force balance equation of the form

$$\dot{R} = -\xi \left(\frac{\partial E}{\partial R} + 2\pi\Sigma \right). \quad (5.2)$$

The dynamics of the filament concentration $c = N/(2\pi R)$ with N filaments is defined along a ring of length $2\pi R$. Without turnover the ring perimeter therefore changes with $R \propto 1/c$. Including turnover we obtain the material balance equation

$$\dot{c} = k_p - k_d c - \frac{\dot{R}}{R} c \quad (5.3)$$

with filament polymerization and depolymerization rates k_p and k_d . The contractile stress $\Sigma = AN_b c^2$ models the active force generation of actin filaments and myosin motors. A is the (positively defined) effective material coefficient, N_b the number of distinct filament bundles and c , the number density of filaments per unit length along a bundle. The contractile force is

proportional to c^2 to account for the fact that actin filaments need a partner to exert forces. Together Eq. (5.2) and Eq. (5.3) define an elementary model of cytokinetic ring constriction.

Zumdieck et. al. used parameters from independent measurements curated from the literature¹⁸⁴. Comparing numerical solutions of the model, using these parameters with time laps recordings of the first-division constriction dynamics of *C. elegans* oocytes, Zumdieck et. al. found good agreement between model and measurement.

To allow for different stoichiometrics of the potential force generating molecules such as actin and myosin, and to include a stochastic component, we slightly generalized this model dynamics

$$d\widehat{L}(t) = - \left(\gamma_1^\lambda (\widehat{L}(t) - \widehat{L}_m) + \gamma_2^\lambda \widehat{c}(t)^n \right) dt + \sqrt{D} dW_t \quad (5.4)$$

$$d\widehat{c}(t) = \left(k_{\text{on}}^\lambda - k_{\text{off}}^\lambda \widehat{c}(t) \right) dt + \frac{d\widehat{L}(t)}{\widehat{L}(t)} \widehat{c}(t). \quad (5.5)$$

We define $\widehat{L} := R$, $\gamma_1^\lambda := K\xi$, $\widehat{L}_m := R_0$, $\gamma_2^\lambda := 2\pi AN_b$, $k_{\text{on}}^\lambda := k_p$ and $k_{\text{off}}^\lambda := k_d$. To allow for single and pairwise interactions of the force generating molecules we replace the exponent of the concentration by a parameter $n := \{1, 2\}$. In this version, the index $\lambda := \{m, c\}$ further accounts for the possibility to switch between two different regimes. A maintenance (m) phase where a fixed ring perimeter is stabilized, and a constriction (c) phase where the ring diameter shrinks until cell separation. For simplicity the switching between the two regimes $m \rightarrow c$ is modeled as a discrete shift in parameter space $(\gamma_1^m, \gamma_2^m, k_{\text{on}}^m, k_{\text{off}}^m) \rightarrow (\gamma_1^c, \gamma_2^c, k_{\text{on}}^c, k_{\text{off}}^c)$. Both generalizations, stoichiometrics of myosin and stochasticity, were already suggested by the original authors¹⁸⁴.

Below we show that this generalized model exhibits three qualitatively different biophysically plausible regimes, which can be approximated with a single effective SDE of the form

$$d\widehat{L}(t) = -\gamma_{\text{eff}} \widehat{L}^\alpha + \sqrt{D} dW_t \quad (5.6)$$

with $\alpha = 0, -1, -2$. These three effective dynamics are recovered from Eq. (5.4) and Eq. (5.5), for the biological meaningful limits of fast turnover, stalled myosin turnover, and stalled actin turnover.

The fast turnover limit

In the limit of fast turnover ($k_{\text{on}}^\lambda \gg 1$, $k_{\text{off}}^\lambda \gg 1$) and $k_{\text{on}}^\lambda/k_{\text{off}}^\lambda = \text{const.}$, the above model reduces to 1d dynamics of the ring perimeter \widehat{L} which, during the contraction phase, is driven by a constant force γ_{eff} . In this limit $\widehat{c}(t)$ relaxes very fast to its equilibrium concentration $k_{\text{on}}^\lambda/k_{\text{off}}^\lambda$. The remaining SDE for the ring perimeter \widehat{L} is

$$d\widehat{L}(t) = - \left(\gamma_1^\lambda (\widehat{L} - \widehat{L}_m) + \gamma_2^\lambda \left(\frac{k_{\text{on}}^\lambda}{k_{\text{off}}^\lambda} \right)^n \right) dt + \sqrt{D} dW_t. \quad (5.7)$$

Assuming a very “soft” cell shape and a strong constriction force we set $\gamma_1^c = 0$ and find an equation of the form of Eq. (5.6). This is the force to deform the entire cell. An analytic expression for the reverse time force close to completion can be obtained using Eq. (4.118). It turns out that the effect of $\gamma_1 > 0$ on the moments of the aligned reverse time ensemble close to completion is manure. Only the mean changes marginally and rises slightly slower than in the case with $\gamma_1 = 0$.

Stalled turnover limits

In the limit of stalled turnover ($k_{\text{on}}^\lambda \approx 0$, $k_{\text{off}}^\lambda \approx 0$) all force generating molecules are trapped in the ring over the time scale of the constriction phase. We thus set k_{on}^c and k_{off}^c to zero. The

concentration dynamics from Eq. (5.5) and during constriction then reduces to

$$\frac{d(\widehat{c}(t)\widehat{L}(t))}{dt} = \dot{\widehat{c}}(t)\widehat{L}(t) + \widehat{c}(t)\dot{\widehat{L}}(t) = 0, \quad (5.8)$$

implying a fixed number of force generating molecules on the ring. In this limit

$$\widehat{c}(t)\widehat{L}(t) = \text{const.} \quad (5.9)$$

As this is maintained during constriction, the constant value is inherited from the value of this product at the end of the maintenance phase, the switching time t_{sw} . We can therefore replace $\widehat{c}(t)$ in the perimeter dynamics Eq. (5.4), and obtain

$$d\widehat{L}(t) = - \left(\gamma_1^c (\widehat{L} - \widehat{L}_m) + \gamma_2^c \left(\frac{\overline{\widehat{c}(t_{\text{sw}})\widehat{L}(t_{\text{sw}})}}{\widehat{L}(t)} \right)^n \right) dt + \sqrt{D} dW_t \quad (5.10)$$

with $\overline{\widehat{c}(t_{\text{sw}})}$ approximated by the ratio of $k_{\text{on}}^m/k_{\text{off}}^m$ and

$$\overline{\widehat{L}(t_{\text{sw}})} = \widehat{L}_m - \frac{\gamma_2^m}{\gamma_1^m} \overline{\widehat{c}(t_{\text{sw}})}, \quad (5.11)$$

where $\overline{\widehat{c}(t_{\text{sw}})}$ and $\overline{\widehat{L}(t_{\text{sw}})}$ are the average values during the maintenance phase. For the example displayed in chapter 3, we again assume dominance of the constriction force ($\gamma_1^c = 0$) during the constriction phase. Note that, as the concentration of force generating molecules diverges, this is a particular mild assumption here. The effective dynamics then terminates either $\propto 1/\widehat{L}$ or $\propto 1/\widehat{L}^2$ depending on the model variants ($n = 1$: myosin; $n = 2$: actin) for the force generating molecule.

Simulations

Sample paths for the three different constriction scenarios with constriction forces constant ($\alpha = 0$), proportional to $\propto 1/\widehat{L}$ ($\alpha = -1$) or $\propto 1/\widehat{L}^2$ ($\alpha = -2$) were generated from Eq. (5.4) and Eq. (5.5). All simulations were started from the equilibrium distribution of the maintenance phase. After a time t_{sw} (chosen larger than the window used for inference), the dynamics for each sample path was switched to the constriction regime. The case corresponding to $\alpha = -1$ is provided in Fig. 3.6 of chapter 3. Here we additionally display the two cases with $\alpha = 0, -2$ (see Fig. 5.1). Tab.5.1 lists the parameter settings for all simulations. For integration we used

Parameters	Simulation 1 ($\alpha = 0$)		Simulation 2 ($\alpha = -1$)		Simulation 3 ($\alpha = -2$)	
	maintenance	constriction	maintenance	constriction	maintenance	constriction
n	1	1	1	1	2	2
$k_{\text{on}}/\mu\text{m s}^{-1}$	400	2000	0.5	0	0.7	0
k_{off}/s^{-1}	10	20	0.01	0	0.1	0
γ_1/s^{-1}	0.075	0	0.075	0	0.075	0
$\gamma_2/\mu\text{m}^{1+n} s^{-1}$	$2\pi \cdot 3.6 \cdot 10^{-5}$	$2\pi \cdot 3.6 \cdot 10^{-5}$	$2\pi \cdot 3.6 \cdot 10^{-5}$	$2\pi \cdot 3.6 \cdot 10^{-5}$	$2\pi \cdot 3.6 \cdot 10^{-5}$	$2\pi \cdot 3.6 \cdot 10^{-5}$
$\widehat{L}_m/\mu\text{m}$	14.5	14.5	14.5	14.5	14.5	14.5
$D/\mu\text{m}^2 s^{-1}$	0.04	0.04	0.04	0.04	0.04	0.04

Table 5.1: Parameter settings used for 3 different sample path ensembles created from Eq. (5.4) and Eq. (5.5). Simulation 2 is discussed in chapter 3.

a standard Euler scheme with time step $\Delta t = 0.05s$.

Maximum likelihood path ensemble inference

To infer the most likely constriction scenario hidden in each of the forward simulations as defined in Table 5.1, we maximize the path likelihood of the respective ensemble after alignment and time reversion. We implement the approach discussed in section 5.2. We infer the reverse time force from the TSA ensemble and reconstruct the underlying forward force. We assumed the forward force law is of power law form $f(\widehat{L}) = -\gamma\widehat{L}^\alpha$. We can therefore directly read off the forward force $f(L)$ from the inferred reverse time force $f^{\text{TSA}}(L)$.

For each of the three simulations defined in Table 5.1 the inference procedure is identical. We here discuss Simulation 2 (with true effective $\alpha = -1$) shown in the chapter 3. The idea is to treat the simulation data as black box, as experimental data would be, and determine which forward force law $f(\widehat{L})$ is most suitable to describe the constriction dynamics. In principle the maximal included lifetime of the reverse time trajectories $N\Delta\tau$ should be optimized as well. Starting from very short trajectories for which contributions of the maintenance phase are negligible, and increasing the length until distinctions from a pure ensemble become detectable, then marks the transition from constriction to maintenance. We here, for simplicity, directly choose N based on visual inspection of the aligned reverse time ensemble.

There are two options to find the best model. First, to simultaneously optimize for $\alpha, \gamma_{\text{eff}}, D$. Second to choose one α from the three above described mechanisms, assuming it contains the correct model, and infer the best fitting γ_{eff}, D . With the first approach we find the best model with its maximum likelihood parameters. With the latter we find the best fitting parameters for each of the candidates ($\alpha = 0, -1, -2$). This then allows to quantitatively compare the three different models either in terms of their maximum likelihood values or, in terms of their moments which we each calculate from simulations of the respective reverse time SDE Eq. (4.120) using the newly inferred maximum likelihood parameter $\gamma_{\text{MAP}}, D_{\text{MAP}}$.

For Simulation 2 (with true effective $\alpha = -1$) the results of this comparison are shown in Fig. 3.6 of chapter 3. For the “true” case $\alpha = -1$, the full terminal aligned simulation and the inferred dynamics perfectly coincide in their moments. For $\alpha = 0, -2$ they strongly deviate. A similar observation holds for Simulation 1 and 3 shown in Fig. 5.1. For both, the correct dynamics are recovered while the deviant models can not be matched in their statistics. In total this shows that our maximum likelihood path ensemble inference approach for reverse time analysis allows to reliably infer the correct underlying mechanism while rejecting deviant models. The theory of TSA ensemble inference thus provides a reliable framework for the inference of effective dynamics. In the following sections we will explore generalizations to this baseline framework.

5.2 Reverse-time ensemble path inference

The reverse time TSA ensemble approach aims at the inference of the underlying forward dynamics from dynamics with identifiable dynamics only close to the target state. We here propose to directly infer the reverse time force $f^{\text{TSA}}(L)$ from TSA sample paths using a path-ensemble-likelihood inference scheme.

The construction idea of this likelihood inference scheme follows classical maximum likelihood and Bayesian path inference approaches^{36,41,43,78,108,109,167}. We first define the transition probability for discrete sample paths starting at L_{ts} at time $\tau = 0$ and ending at τ_f for simplicity. Sample points are taken with equal spacing $\Delta\tau$. We consider sample paths of duration $T = N\Delta\tau$. The transition probability for individual realizations of such a Markovian process

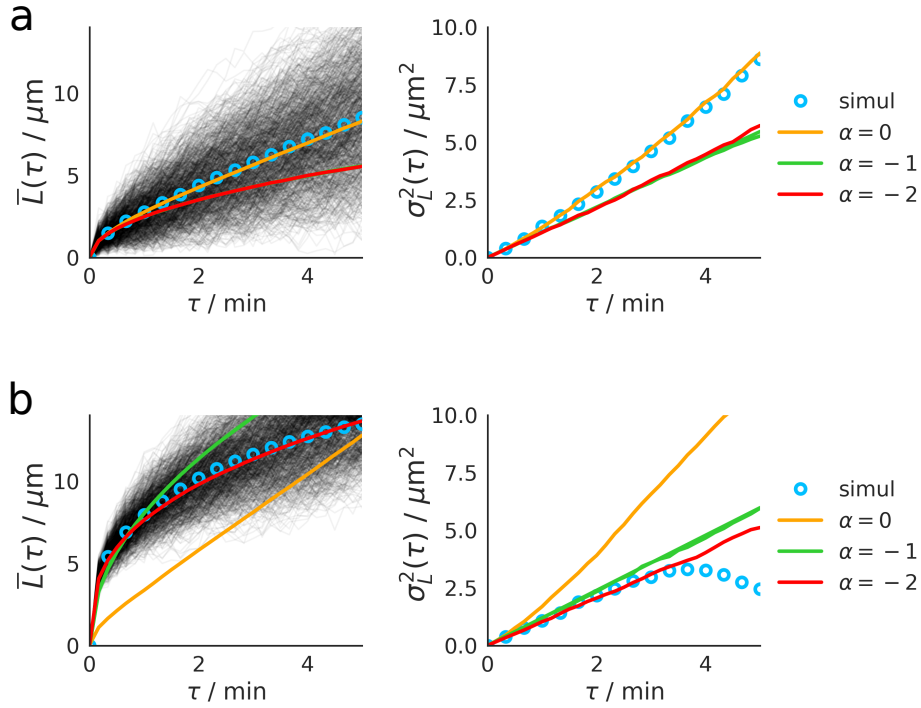


Figure 5.1: The dynamical law leading to constriction can unambiguously be inferred from realizations of the Zumdieck model, using path integral maximum likelihood inference Eq. (5.15). We visually confirm the inferred underlying force law $f(\hat{L})$ by numerically evaluating mean and variance for the reverse-time dynamics under the assumption of well separated initial and final conditions, for each of the three inferred maximum likelihood parameter sets $(\gamma_{\text{ML}}^\alpha, D_{\text{ML}}^\alpha)$ **(a)**: Given the correct underlying force law is $f(\hat{L}) = -\gamma$ ($\alpha = 0$), only the case $\alpha = 0$ (yellow line) predicts both the mean and variance of the Zumdieck model (blue circles). **(b)**: Given the correct underlying force law is $f(\hat{L}) = -\frac{\gamma}{\hat{L}^2}$ ($\alpha = -2$), only the case $\alpha = -2$ (red line) predicts both the mean and variance of the Zumdieck model (blue circles).

reads

$$P(L_f, \tau_f | L_{\text{ts}}, 0) = \prod_{i=0}^{N-1} P(L_{i+1}, \tau_{i+1} | L_i, \tau_i) . \quad (5.12)$$

Assuming individual, statistically independent sample paths, the total likelihood of a finite size sample n_{ens} of sample paths is the product of individual transition probabilities. Taking the logarithm the log likelihood function is

$$\log \mathcal{L} = \sum_{j=1}^{n_{\text{ens}}} \sum_{i=0}^{N-1} \log P(L_{i+1}^{(j)}, \tau_{i+1}^{(j)} | L_i^{(j)}, \tau_i) \quad (5.13)$$

which can be used for maximum likelihood inference of model parameter. The exact transition probability $P(L_{i+1}, \tau_{i+1} | L_i, \tau_i)$ is not available analytically for most stochastic processes. A simple approximation for $P(L_{i+1}, \tau_{i+1} | L_i, \tau_i)$ useful for closely sampled trajectories is of Gaussian form

$$P(L_{i+1}, \tau_{i+1} | L_i, \tau_i) = \frac{1}{\sqrt{2\pi D \Delta\tau}} e^{-\frac{(L_{i+1} - L_i - f^{\text{TSA}}(L_i) \Delta\tau)^2}{2D \Delta\tau}} . \quad (5.14)$$

This is equivalent to a first order stochastic simulation of an SDE in Ito interpretation, where the change dL with respect to the deterministic force times $\Delta\tau$ plus a random step with zero mean and variance $D\Delta\tau$ is evaluated. The log-likelihood under the approximation of Gaussian transition probabilities thus reads

$$\log \mathcal{L} = \sum_{j=1}^{n_{\text{ens}}} \sum_{i=0}^{N-1} \left(-\frac{\left(L_{i+1}^{(j)} - L_i^{(j)} - f^{\text{TSA}}(L_i^{(j)})\Delta\tau \right)^2}{2D\Delta\tau} - \frac{1}{2} \log(2\pi D\Delta\tau) \right). \quad (5.15)$$

This approximation becomes exact for arbitrary small time steps, and deviations occur as the time-steps become bigger. It is therefore recommended to check after optimization, if the residual increments after subtraction of the deterministic part are actually Gaussian. If not, more sophisticated path inference techniques can be used^{44,88,124}, which can cope with “almost” Gaussian transition probabilities^{2,3} or even arbitrary step size¹³.

For forward power law forces of the form $f(L) = -\gamma L^\alpha$, and close to the target state, the corresponding reverse time forces $f^{\text{TSA}}(L) = f(L) + f^{\mathcal{F}}(L)$ can be stated exactly using Eq. (4.120) for the free energy force $f^{\mathcal{F}}(L)$. Using $f^{\mathcal{F}}(L)$, in the form as defined in Eq. (4.120), the log-likelihood is not easily evaluated numerically for small L . As both $\Gamma(\nu)$, $\Gamma(\nu, z)$ and $\exp(z)$ need to be evaluated numerically and separately, slight inconsistencies can lead to numerical instabilities. We therefore transformed $f^{\mathcal{F}}(L)$ to forms in which only one special function is evaluated numerically. This is feasible when treating $\alpha < -1$ and $\alpha > -1$ separately. The full free energy force then splits into three cases

$$f^{\mathcal{F}}(L) = \begin{cases} \frac{D}{-\frac{2L^{\alpha+1}\gamma}{\alpha D+D} {}_1F_1(1; 1+\frac{1}{1+\alpha}; -\frac{2L^{\alpha+1}\gamma}{\alpha D+D})} & \text{for } \alpha > -1 + \epsilon \\ \frac{2\gamma}{L} - \frac{\alpha D}{L} & \text{for } -1 + \epsilon < \alpha < -1 - \epsilon \\ \frac{-D(1+\alpha)}{-\frac{2L^{\alpha+1}\gamma}{\alpha D+D} U(1; 1+\frac{1}{1+\alpha}; -\frac{2L^{\alpha+1}\gamma}{\alpha D+D})} & \text{for } -1 - \epsilon < \alpha < -1 - \epsilon. \end{cases} \quad (5.16)$$

In our experience this representation with $\epsilon = 0.05$ worked robustly. Here ${}_1F_1(a; b; z)$ is the Kummer confluent hypergeometric function, $U(a; b; z)$ is Tricomi’s confluent hypergeometric function. To obtain this representation we used the substitutions

$$\Gamma(a) - \Gamma(a, z) = a^{-1} z^a e_1^{-z} F_1(1; 1+a; z) \quad (5.17)$$

which can be derived from from Abramovitz & Stegun (AS 6.5.3) and (AS 6.5.12)¹, and

$$\Gamma(a, z) = z^a U(1; 1+a; z) e^{-z} \quad (5.18)$$

which follows from Abramovitz & Stegun (AS 13.1.29) and (AS 13.6.28)¹. We point out that $f^{\text{TSA}}(L(\tau_0))$ diverges for $L(\tau_0) = L_{\text{ts}}$. We therefore leave the target state data point out and start the inference with the next or second next state.

The strength of the path ensemble inference formalism compared to optimizing predicted moments such as mean and variance is that the optimization is performed with respect to the full likelihood. No considerations about sufficient statistics or how to weigh the different moments relative to each other in a loss function are required. By directly inferring the underlying reverse time SDE, mean, variance etc. become predictions. The comparison between observed and predicted moments can then be used to test the inference and accept or reject the model.

Understanding cell contact constriction in epithelial morphogenesis through data driven reverse-time inference

6.1 Introduction

How an unstructured agglomerate of individual cells develops into a complex organism is one of the fundamental biological questions of our time. Research dating back as far as to the beginning of the last century¹⁷² has uncovered conserved mechanisms that drive development and seem to prevail in all metazoan embryos⁸⁵. One of the recurrent themes is convergent extension of tissue, which describes the simultaneous narrowing and lengthening of tissue during development (Fig. 6.1a)⁸⁶. It is probably the main mechanism of tissue rearrangement in embryogenesis⁸⁶. Convergently expanding tissues are found and studied in organisms ranging from nematodes and arthropods to vertebrates⁷³. It is found in organ development such as the frog (*Xenopus*) gut²⁷ and kidney⁹⁹ and mouse palate development¹⁶⁶. Convergent extension is of particular importance to understand body axis elongation, which describes the separation of tissue parts which at a later stage develop into the head and tail region. It occurs exemplarily in axis elongation during *Drosophila* germband extension, but is also preserved in all examined chordates⁸⁵.

The fundamental mechanism that drives convergent extension is (oriented) cell intercalation. Two different modes of cell intercalation are observed (Fig. 6.1b). The first is based on cell crawling. Here cells squeeze in between two previously neighboring cells using actin rich protrusion that facilitate crawling supposedly combined with a weakening of cell-cell adhesion¹⁵³. The tissue thereby elongates perpendicular to the direction of cell migration. This mode of cell intercalation has originally been discovered in *Xenopus* notochord formation^{84,86} and was later also found in mouse neural plate elongation^{153,178}.

The second mode of cell intercalation is mediated by acto-myosin driven junction shrinkage between two neighboring cells and subsequent junction expansion between the two new neighboring cells and perpendicular to the shrunken contact^{39,73,153}. This process occurs either as a neighbor exchange of four cells called a T1 transition, or with more cells involved as rosette formation with subsequent directional resolution. (Fig. 6.1c)

While originally thought to be two distinct modes of cell rearrangement, recent findings suggest that even in the prototype process for the crawling mode (i.e. *Xenopus* notochord formation) junction shrinkage might play a role¹⁵³. The same holds for the prototype of junction shrinkage and subsequent expansion (i.e. *Drosophila* germband extension), where some of the rosette forming processes seem to include as a first step a crawling mode at the basal side of the

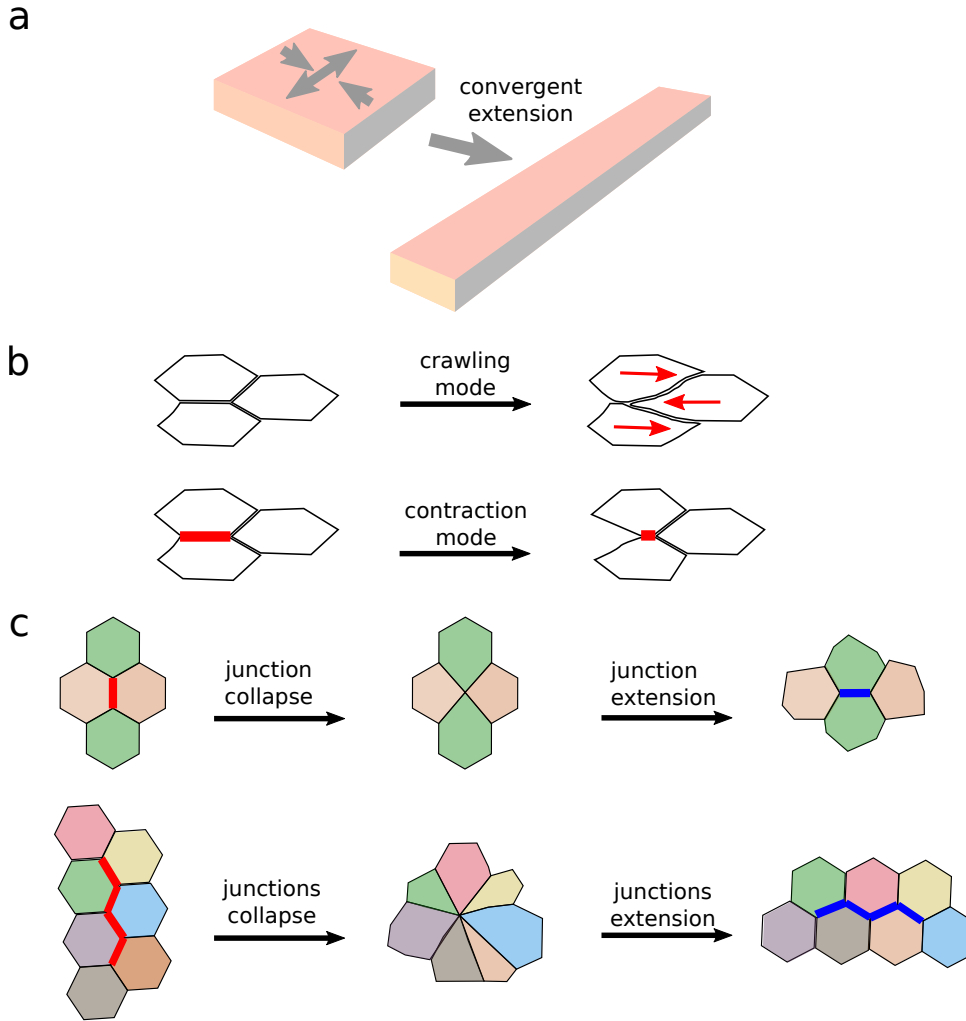


Figure 6.1: Modes of convergent extension. (a): Schematic depiction of convergent extending tissue. (b): Two modes of cell intercalation that drive convergent extension. (Top): Intercalation by cell crawling. Cells squeeze in between two previously neighboring cells thereby extending the tissue. (Bottom): Cell intercalation by junction remodeling. (c): Neighbor exchange by junction remodeling. (Top): The topological T1 transition defines the standard mode of cell intercalation by junction remodeling. It is comprised of the three parts, junction shrinkage, resolution in a 4-vertex and subsequent junction expansion. (Bottom): Cell intercalation by rosette formation and subsequent resolution. Shrinking junctions are depicted in red, extending in blue. Scheme concepts are taken from Keller et.al.⁸⁶, Shindo¹⁵³ and Kong et.al.⁹¹.

cells^{73,153}. Whether these findings will lead to a unified theory or are just rare events is unclear today⁵⁹.

In this work, we only focus on one mode of cell intercalation, that is convergent extension by junction shrinkage (and subsequent expansion) in *Drosophila melanogaster*. More specifically we study the contraction phase. Contact constriction is driven by the coordinated interplay of molecules that drive junction remodeling and molecules that ensure tissue integrity. It involves a complicated cascade of feedback mechanisms and allows for extracellular clues to influence modes of operation. So far we only possess a qualitative understanding of what matters for junction shrinkage and phenomenological models that describe these processes. None of these modeling approaches however allows to distinguish between models of seemingly similar qualitative behavior but completely different underlying dynamics. Unlike previous approaches that

constructed effective biological models mainly based on biological intuition, we here aim to *infer* a model that satisfies all observed statistics and allows us to make *quantitative* predictions. Our goal is thus to infer the full time dependent probability of the stochastic dynamics of junction shrinkage. Such an effective model would provide a reference point to which all bio-physically motivated models must compare to. Instead of building a model from intuition we constrain possible model classes by the data.

We will start our discussion with a review on the current biological knowledge about junction shrinkage during cell intercalation in *Drosophila* morphogenesis. In section 6.2 we review the two most frequently used biophysical modeling approaches of cell rearrangement, i.e. vertex and visco-elastic modeling. The rest of this chapter introduces the data (section 6.3), our modeling approach and the final model we find to predict the full quantitative phenotype of junction shrinkage (section 6.4).

Early *Drosophila* morphogenesis

In *Drosophila* morphogenesis, the force driven rearrangement of epithelial tissue sets in right after cellularization and about 3h after fertilization^{91,97}. At this stage, the onset of gastrulation, the nucleus of the fertilized egg has undergone 13 rounds of division leading to approximately 6000 nuclei which after cellularization by plasma membrane invagination from the surfaces of the egg⁹⁷ have formed a one layered epithelium spanning the whole football shaped embryo²³. Triggered already at fertilization by maternally positioned RNA bicoid, an ongoing cascade of morphogenes has by then established the body plan with its anterior-posterior (AP) and dorso-ventral (DV) axis. The at the onset of gastrulation fairly homogeneous epithelium is ready to drastically change its shape^{59,80,97}.

Beginning with mesoderm invagination and a few minutes later germband extension, force driven tissue dynamics then start to drastically reshape the embryo^{23,91}. At the same time the tissue is further shaped by the formation of the cephalic furrow^{91,161} and the invagination of the posterior midgut^{91,97}. We here focus on the process of germband extension, which is prototypical for convergent extension (Fig. 6.2).

During germband extension a about 18 cells wide and 60 cells long stripe of cells at the ventral side of the embryo elongates by more than a factor of 2 in anterior-posterior direction (AP) while shrinking in dorso-ventral direction (DV)^{96,163}. This process of convergent extension lasts about 2.5 h with a rapid initial extension phase of 45 min^{23,91}.

Germband extension is driven by the two genetically separable mechanisms of cell intercalation and cell shape changes⁹¹. The currently emerging picture describes cell intercalation as a local phenomenon driven by local forces^{79,180,181}. Additionally anisotropic tissue-wide forces induced by posterior midgut invagination are present³⁵. In wildtype embryos these forces seem to be dissipated by cell intercalation within the first 15 min of germband extension and thus effectively increase the total number of AP to DV intercalation events (reviewed by Kong⁹¹).

Directed cell intercalation can be realized by topological T1 transitions^{12,137} or rosette formation¹⁶ and subsequent directional resolution (Fig. 6.2(Middle)). During a T1 transitions four cells exchange neighbors by junction remodeling¹² (Fig. 6.2(Bottom)). This process consists of two parts. i) First, the junction between two neighboring cells constricts until four cells are connected in a four vertex. ii) In a second independent process a new cell contact between the previously not contacting cells is established by expansion⁹¹. During germband extension the process of constriction is (almost always) observed in DV direction, while contact expansion happens in AP direction^{12,137}. Rosette formation and subsequent directional resolution is a conceptually identical mechanism with the only difference that more than four cells are connected into a common vertex⁶². We here focus on T1 transitions which predominantly occur at the very beginning of germband extension (stage 6). Rosette formation seems to be setting in slightly

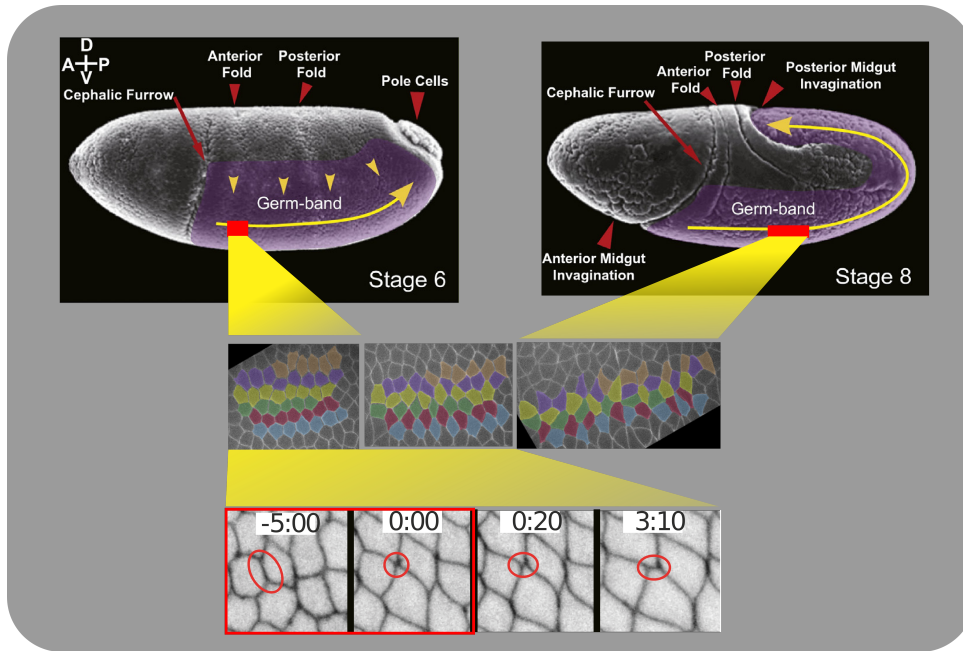


Figure 6.2: Convergent extension of the germband via cell intercalation through junction remodeling. (Top): 2.5 fold elongation of germband. Exemplary region of cell intercalation is schematically depicted on top of a scanning electron microscopy photomicrograph¹⁴² (modified from Kong et.al.⁹¹). (Middle): Three exemplary stages of tissue rearrangements via cell intercalation (provided by Zhang et.al.¹⁸³). (Bottom): Cell intercalation is mediated by T1-transitions, i.e. neighbor-exchange by “vertical”-junction shrinkage and “horizontal” junction expansion.

later (\sim stage 7), probably linked to a feedback loop of ectopic tension and associated myosin cable formation to create linked AP junctions which subsequently jointly constrict into rosettes^{16,47}.

To facilitate oriented cell intercalation planar cell polarity (PCP) must be established^{12,16,180,181}. This implies that relevant components of the cell architecture are oriented with respect to the AP and DV axis¹⁸⁰. In *Drosophila* PCP is tightly linked to the AP patterning system which originates from the maternally positioned RNA bicoid and, via the hierarchical order of gap, pair rule and segment polarity genes, defines the positional information in the embryo^{59,80}. It is important to note however that PCP is not established via long range interactions but by interactions of neighboring cells^{180,181}. Recently one possible link between the AP patterning system and PCP has been established in *Drosophila*¹²⁶. Knockout mutants of the gap genes runt and eve almost lose their planar polarity and show strongly reduced germband extension^{79,181}. These gap genes (among many other targets) regulate the Toll receptor family which in turn is associated with PCP⁹¹. Without Toll-2,6 and 8 the enrichment of myosin II along vertical junctions and the enrichment of Bazooka/Par along horizontal junctions is lost, which are both known to be essential for oriented cell intercalation^{91,126}. The observed effects of Toll-2,6 and 8 combined knockouts is however not strong enough to convincingly demonstrate that this is the only used pathway to ensure PCP⁹¹. While essential in other contexts of convergent extension (cochlea extension, kidney tubule elongation), the PCP pathway does not seem to play an essential role^{73,153}. In summary, with the discovery of the importance of 3 Toll family receptors in germband elongation, a link between AP patterning, PCP and oriented cell intercalation has been established. This pathway is however not sufficient to explain the full extent of PCP, demanding for further parallel mechanisms to be resolved⁹¹.

Cell contact constriction is driven by non-muscle myosin II motors^{12,91,181}. These myosin

II motors form complexes in a network together with F-actin fibers, which are linked to the cell junctions via the cell adhesion molecule E-cadherin and adapter proteins such as β -catenin, α -catenin and vinculin⁹³ (Fig. 6.3). In an activated state (discussed below), these motors form

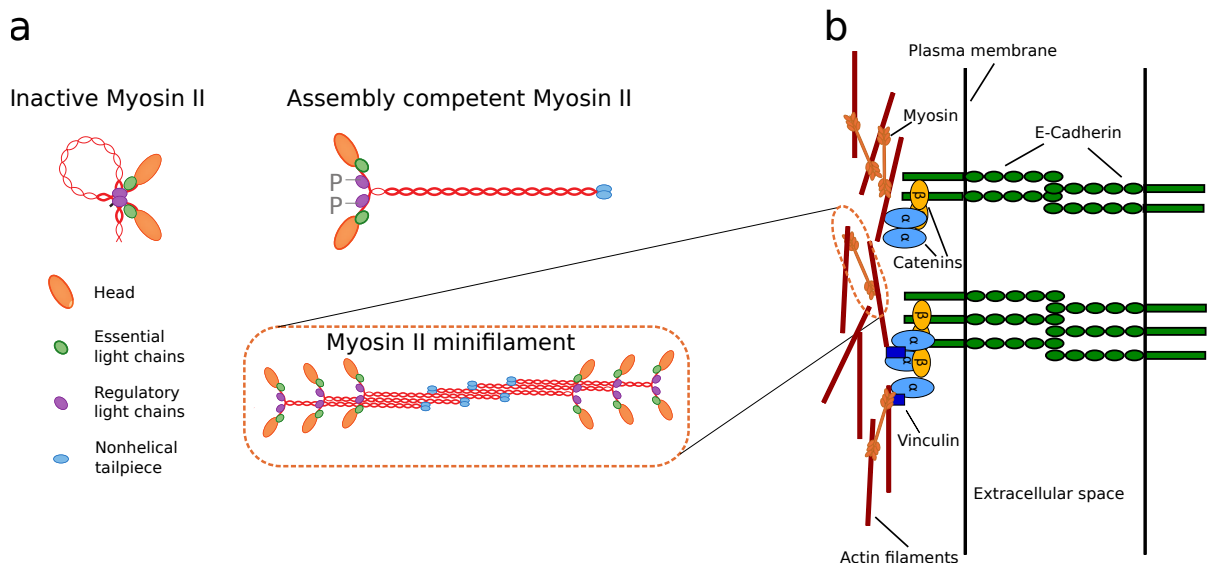


Figure 6.3: Adhesion complexes link neighboring cells and are coupled to the dynamics of the acto-myosin network. (a): Schematic depiction of inactive, through phosphorylation of the regulatory light chain unfolded, and into minifilaments assembled myosin II. The details and dynamics of minifilament assembly, and especially of various phosphorylation mechanisms, are still under debate⁹³ (Figure adapted from Lecuit et. al.⁹³). (b): E-Cadherin clusters stabilize adherens junctions. These clusters are coupled to the inner cell acto-myosin network via α and β -catenin and vinculin. The details of the coupling mechanisms are still an active field of research²¹.

myosin mini-filaments with several head domains on both ends to grab on closely actin fibres⁹³ (Fig. 6.3a). Under ATP hydrolysis these motors then move along the actin filaments and cause the network, and via its couplings to the cell junctions also the cell contact to constrict⁹¹. It has been shown that planar polarized myosin II (i.e. enriched along vertical junctions) is necessary for junction shrinkage^{12,16}. This observation might be linked to the general observation that myosin accumulates in regions of increased tension (leading e.g. to myosin cable formation and subsequent rosette formation) and dissociates rapidly when the tension is released, establishing a feedback loop for myosin recruitment⁴⁷. A more recent study suggest that asymmetric pulsatile flows of medial myosin towards these junctions promote junction shrinkage, and postulates junctional myosin only as a stabilizing agent¹³⁶. The proposed mechanism is ratchet like with phases of pulse driven constriction and subsequent stabilization of a new equilibrium cell contact length¹³⁶. With Rab 35 there has been a recent report of an oscillatory ratchet that prevents relaxation to the original pre-constriction length⁸¹.

While the exact mechanism of junction shrinkage is not yet fully understood, a main regulatory pathway that regulates the dynamics of acto-myosin networks in the context of junction shrinkage has been established^{37,87}. To function, i.e. to form myosin mini-filaments, myosin II must first be activated. This activation is mediated by the kinase Rok which phosphorylates the non-muscle myosin II regulatory light chain – a prerequisite for mini-filament formation and thus force exertion⁹¹ (Fig. 6.4). Rok itself is activated by the GTPase Rho1 in its activated (GTP) state^{110,118}. Rho1 in turn is regulated, depending on whether it regulates medio-apical or junctional myosin by two different guanine nucleotide exchange factors (GEFs). While RhoGEF2 regulates medio-apical myosin, p114 RhoGEF/Wireless regulates junctional myosin³⁷. Both GEFs are regulated by a G protein coupled receptor (GPCR) and their hetero trimeric

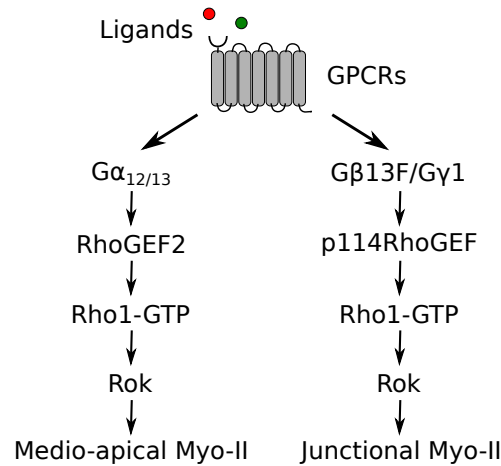


Figure 6.4: In *Drosophila* embryonic ectoderm, medio-apical and junctional myosin are regulated by two distinct pathways³⁷. Following Ligand stimulation, GPCRs release active $G\alpha_{12/13}$ and $G\beta_{13F}/G\gamma_1$ which promote RhoGEF2 and p114RhoGEF signaling, respectively. The precise mechanism of the following activation of Rho1 is unknown. Rho1-GTP then activates the kinase Rok, which phosphorylates the myosin regulatory light chain and thus allows for myosin minifilament formation. Myosin minifilaments, which are a prerequisite for force-exertion, are thus recruited medio-apically and junctional by two distinct pathways (Scheme adapted from De Las Bayonas³⁷).

G proteins^{37,87}. From these G proteins $G\alpha_{12/13}/Cta$ -GTP regulates RhoGEF2 and ultimately activates the apico-medial myosin pool, while $G\beta_{13F}/G\gamma_1$ (among other regulators) regulates junctional myosin via p114RhoGEF/Wireless and the Rho/Rok pathway^{37,87}. Together these components establish an unbroken chain from the internal planar polarized acto-myosin dynamics to an external triggering and modification by extracellular ligands. How and if this pathway relates to the planar polarization of myosin 2 by the toll receptors (Toll 2,6 and 8) and thus to the AP-patterning system via *eve* and *runt* is currently an open question³⁷.

While the relevance of planar polarized acto-myosin dynamics for junction shrinkage is well established, are the molecular mechanisms at the junction still poorly understood. From a conceptual perspective it is clear that junctions must be remodeled to fully constrict. As E-cadherins act as anchors that link cell interfaces to form adherens junctions, E-cadherin must simultaneously protect the epithelium against rupture under stress while still allowing the junctions to be remodeled. Interestingly, these two mechanisms seem to occur on different time scales²⁸. As an immediate response to stress the complex of E-cadherin α -catenin-actin or E-cadherin-vinculin-actin act as a catch bond, i.e. their lifetime increases under load²⁸. On longer time scales, the response of E-cadherin to stress depends on the circumstances⁹⁴. Both E-cadherin enrichment and endocytosis upon stress have been reported⁸². For vertical junctions during germband elongation it seems however to be true that E-cadherin endocytosis is upregulated mediated by Rho1⁹⁸, and E-cadherin clusters are downregulated by the kinase Rock^{94,111}. Interestingly, recent evidence points towards a combined picture where E-cadherin fulfills a double role. It protects against tensile (vertical to junction) stress but allows shear stress (parallel to the junction) to dynamically remodel the tissue while dynamically preserving adhesion⁸².

In summary, contact remodeling during germband elongation is a highly complicated local mechanism, based on local time changing dynamics. It incorporates junction stabilizing forces and forces that dynamically drive contact remodeling. It incorporates feedback mechanisms, and regulatory mechanisms on all scales. Based on planar polarized concentrations of molecules, junction remodeling is further linked to the very onset of its morphogen gradient based patterning system. How to disentangle all these effects and quantify their relative importance is still an

open question. With a more and more detailed molecular understanding however it becomes also increasingly difficult to say what is relevant and what could be a minimal model that explains the core components. In the next section I will review how modelers have incorporated the at the time of modeling available biological knowledge into phenomenological dynamics. Subsequently, I will introduce a different approach where I show how starting from an ensemble of ~ 1000 contact constriction events different model classes can be excluded or respectively selected based on data driven inference.

6.2 Biophysical models

To answer the question how cell contact constriction in epithelial tissues is possible researchers have resorted to biophysical modeling. Typically the goal is to find a sufficient model that explains convergent extension driven by contact constriction and expansion. In this review section, I will mostly focus on efforts to model the constriction process.

Three different model classes have been proposed. The first uses 2d vertex modeling, where in the default model each cell in a tissue is replaced by 6 vertices with position r_i . These vertices are linked by line elements L_{ij} (Fig. 6.5). To calculate the dynamics of such tissue an energy function^{46,74}

$$E(r_i) = \sum_{\alpha} \frac{K_{\alpha}}{2} (A_{\alpha} - A_{\alpha}^0)^2 + \sum_{\langle i,j \rangle} \gamma_{ij} L_{ij} + \sum_{\alpha} \frac{\Gamma_{\alpha}}{2} P_{\alpha}^2 \quad (6.1)$$

is minimized. The first term accounts for the elasticity K_{α} of the cell area A_{α} and its deviation from its equilibrium size A_{α}^0 . The area is calculated from a functional of the vertex positions. The sum runs over all cells α . The second term captures the line tension with strength γ_{ij} . The

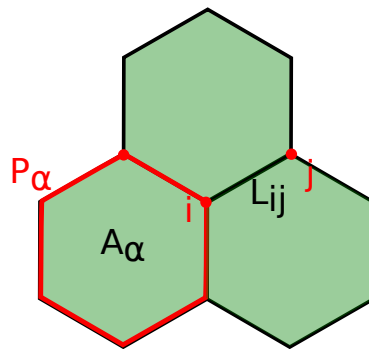


Figure 6.5: Schematic depiction of the typical components in vertex models. L_{ij} denotes the junction length, P_{α} the perimeter of a cell α , and A_{α} the cell area.

sum is over all junctions $\langle i, j \rangle$. This term can be generalized to include planar polarized tension¹³⁷. The third term describes the perimeter P_{α} and its contractility Γ_{α} as can be induced by an acto-myosin ring. The vertex dynamics are then given as

$$\eta \frac{dr_i}{dt} = - \frac{\partial}{\partial r_i} E(r_i), \quad (6.2)$$

where η is the drag coefficient.

Minimizing the energy expression Eq. (6.1) with respect to vertex position shows that within these models the tension typically establishes an equilibrium cell contact length. To break this static pictures and introduce T1 transitions (or Rosette formation with subsequent directional

resolution) researchers have therefore either introduced orientation dependent line tension coefficients¹³⁷ or introduced further terms that model medio-apical and junctional myosin concentrations and their assumed regulators¹⁵⁴. In summary, vertex models have granted insights into possible mechanisms of cell intercalation and convergent extension. Without an explosion of parameter dimensionality, they do however not offer an easily accessible way to study the complicated, heavily regulated mechanisms that actually occur during junction shrinkage (and subsequent contact expansion). Even in their most complicated forms, with dozens of tunable parameters, these models are not designed to capture very general properties of real tissue, i.e. contact fluctuations and the heterogeneity of equilibrium cell contact lengths. To understand what details matter for junction shrinkage, a more local approach is suitable and reviewed below.

The second model class describes the tissue as a visco-elastic medium and models the constricting cell contact as independent collective coordinate L . These effective models start from the assumption that the change of the cell contact length with time

$$\mu_l \frac{dL}{dt} = T - T(m) \quad (6.3)$$

is proportional to the tension T exerted by the neighboring cells and the increased tension T_m due to myosin accumulation along the cell contact¹⁵. μ_l denotes the drag coefficient. Depending on the model assumptions researchers have included a further term proportional to the change in the tension $\frac{dT(m)}{dt}$ to incorporate myosin pulsatility³², or a term proportional to the cell contact length $T(L)$ to model elastic restraints of the contact itself. The full equation reads

$$\mu_l \frac{dL}{dt} = -T(L) - T(m) - \mu_m \frac{dT(m)}{dt}, \quad (6.4)$$

where we included the static terms into $T(L)$. Let's consider the different contributions to the model separately. Ignoring the myosin (m) dependent terms for the moment the resulting terms are known as the Kelvin-Voigt model¹²⁵. In its simplest form $T(L)$ can be approximated as a spring with $T(L) = k_T(L - L_0)$ and equilibrium length L_0 . Upon sudden external stress application and later release such a model relaxes back to its equilibrium length. On the other hand, only studying the myosin dependent terms, we end up with the Maxwell dash-pot model¹²⁵. For simplicity and as an example, we could assume a linear relation between tension and myosin. In this model stress application then leads to an irreversible change in the contact length.

The third model class describes the tissue as an active material^{112,120,131}. These models allow to infer global and local shear stresses and include cell rearrangements and deformation. However, in their coarse grained view they do not provide the here needed means to study junction remodeling as a locally driven and regulated process.

Depending on whether the elastic or viscous property of tissue should be emphasized, all three models have been used to model cellular dynamics^{32,40,112}. Especially Clement et. al.³² have shown that a model based on only the myosin dependent terms in Eq. (6.4) is sufficient to fit trajectory data based on myosin concentration and its derivative. With several free fitting parameters per trajectory, such a model however only demonstrates that certain components of the model should not be neglected. I here aim for a different question. I want to show which components of an effective model are necessary and sufficient to explain the observed cell contact constriction dynamics. The next section therefore establishes the framework to rule out certain model classes and demonstrate, how to find essential components to such an effective model of contact constriction.

6.3 From modeling to inference – a paradigm shift

The exact biophysical mechanisms that drive PCP contact constriction are extremely complicated. They involve a plethora of proteins that have been shown to be essential in the process (see section 6.1). It involves tension exerted by the surrounding tissue, contact remodeling and mechano-chemical feedback loops. Using biophysical modeling and especially vertex simulations, progress has been made to understand how in principle tissue can be rearranged to yield convergent extension. To study the constricting cell contact as an independent process, visco-elastic modeling, using the cell contact length as a collective coordinate, has been established successfully. Both approaches however only allow to give plausibility arguments that certain cellular mechanisms might evolve according to the principles assumed in the model. It does not allow to distinguish between effective models with similar qualitative behavior, but different underlying dynamics. To distinguish between different models it is clearly not sufficient to only study the typical or mean behavior of the observed process – here contact constriction. Instead, I claim in this chapter, it is necessary to infer effective models from an ensemble of time dependent trajectories including their noisy fluctuations. A such inferred effective model would allow to predict the time dependent probability distribution of constricting cell contacts. It would allow to provide hard bounds on any biophysical model. If a complicated biophysical model cannot be coarse grained to the inferred model it fails. With the rise of machine learning techniques it is now possible to acquire and analyze such data sets and perform model prediction.

The ensemble of constricting cell contacts

Data driven model inference is based on high throughput data analysis and the acquisition of a suitable ensemble that describes time dependent dynamics. In line with visco-elastic modeling, I here propose to study the dynamics of constricting cell contacts as the stochastic dynamics of a one dimensional, accessible collective variable, i.e. the junction length L . The dynamic changes of this length represent the full process of junction shrinkage during T1 transitions in germband elongation. To infer an effective model of junction shrinkage we start with the experimental acquisition and definition of the ensemble of constricting cell contacts. Previous work has mostly been based on observation of a few dozen selected cell intercalation events. To infer the full probability distribution, or at least the time dependent mean, variance and two-time covariance of such events during germband extensions, we need a data set of at least one or two orders of magnitude larger. To acquire such a data set, our group has developed a fully automatic pipeline to segment germband extension movies and extract contact lengths L of T1 transitions (Lars Reichl, unpublished). From 65 fly embryos we obtained 958 T1 transitions which we merge into a single time dependent ensemble.

When analyzing ensemble statistics of such a time-dependent data set of non-stationary dynamics, it is not a priori clear what a common reference point for the independently acquired cell-contact constriction trajectories could be. Inspecting proto typical trajectories of junction shrinkage (Fig. 6.6), we observe that all seem to descend from very different positions of first recordance by the algorithm. They all however have in common that they end with $L = 0\mu\text{m}$. We here suggest to align all sample paths to this target state, i.e. the T1 point after which the contact ceases to exist (Fig. 6.7). The thus formed ensemble can be analyzed working in reverse time τ . All trajectories start from the same point $L = 0\mu\text{m}$ at the same time $\tau = 0\text{min}$. For example a point at $\tau = 3\text{min}$ would be 3 min away from full constriction. The resulting ensemble is shown in Fig. 6.8a. With well defined initial conditions at $L = 0\mu\text{m}$ and $\tau = 0\text{min}$, the thus constructed ensemble can now be characterized with its reverse-time τ -dependent moments such as mean and variance, and tested against different models (Fig. 6.8b,c).

Starting from the assumption, that the contact length L is indeed a meaningful collective

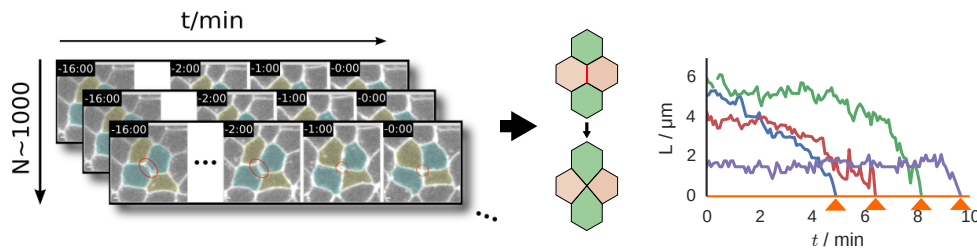


Figure 6.6: Fully automated contact length reconstruction of “vertical” junction shrinkage during T1 transitions. (Left): Schematic depiction of the acquired time dependent ensemble of ~ 1000 T1 transitions. (Middle): “Vertical” junction shrinkage of stylized cells starting from equilibrium length to resolution in 4-vertex target state. (Right): Four selected sample paths that represent the variety of observed dynamics in their maintenance length, their transition dynamics, and their constriction dynamics towards the 4-vertex target state.

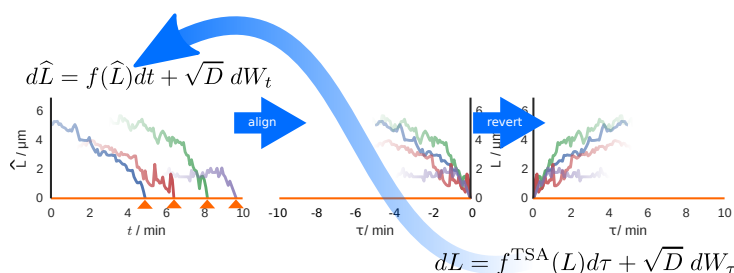


Figure 6.7: Target state alignment and time reversion allows to reconstruct forward dynamics of unknown onset from TSA-dynamics with well defined new “initial” conditions. Shown are four selected sample paths in the process of alignment and time reversion. The shading of sample paths suggests that a common dynamical constriction law irrespective of any precursor dynamics will only be present close to the target state. The superimposed graphic indicates the process of inferring the reverse time dynamics $f^{\text{TSA}}(L)$ and the subsequent reconstruction of the true forward force law $f(\hat{L})$.

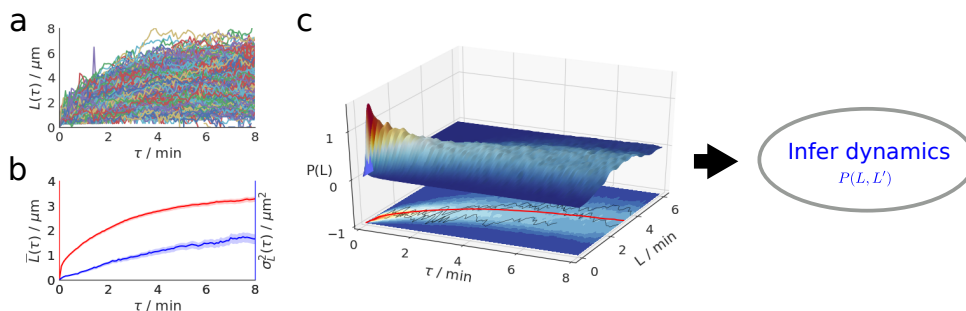


Figure 6.8: Data driven model inference from an ensemble of target state aligned constricting cell contacts. (a): The full TSA ensemble. (b): TSA-ensemble mean (red) and variance (blue) with 95% bootstrap confidence intervals. (c): Kernel density estimation of the TSA-ensemble probability distribution $P(L)$, its projection onto the $\tau - L$ plane and the ensemble mean (red). The goal is to infer a stochastic dynamical model and reconstruct the joint probability distribution $P(L(\tau), L(\tau'))$.

variable, it should be possible to define stochastic dynamics that are capable to reproduce all properties of the experimentally acquired ensemble of contact constriction. In forward time t , and at least close to the alignment point, the ensemble should be in a mode of constriction and a Langevin equation of the form

$$d\hat{L} = f(\hat{L})dt + \sqrt{D} dW_t \quad (6.5)$$

seems a plausible description (Fig. 6.7). $f(\hat{L})$ describes a time independent effective force that

models constriction. The second term approximates the experimentally observed stochastic fluctuations as uncorrelated Gaussian white noise with variance D . We denote forward in time sample paths as \widehat{L} . With the inclusion of stochastic fluctuations, Eq. (6.5) goes beyond the classic deterministic visco-elastic modeling discussed in section 6.2 and opens the path for an ensemble based inference scheme. Written in forward time, these dynamics are however ill defined as the onset of constriction can not easily be read of from the experimental data (cf. Fig. 6.7).

To relate these stochastic dynamics to the acquired TSA ensemble of constricting cell contacts, we rewrite Eq. (6.5) as a reverse time TSA-SDE

$$dL = f^{\text{TSA}}(L)d\tau + \sqrt{D} dW_\tau \quad (6.6)$$

whose original derivation, stated in chapter 4, was driven by precisely this inference problem. Under the approximation of well separated initial and final junction length, the reverse time force is defined as

$$f^{\text{TSA}}(L) = f(L) + D \frac{\partial}{\partial L} \log \left(\int_0^L dL' e^{-\int^{L'} \frac{2f(L'')}{D} dL''} \right) \quad (6.7)$$

(see chapter 3 and 4 for a detailed derivation). To distinguish between stochastic forward and reverse time dynamics, reverse time sample paths are written as L . With this analytic expression for the reverse time force, we are now in a position to analyze the experimentally obtained TSA junction collapse ensemble in reverse time and extract the effective force $f(\widehat{L})$ of the forward process, irrespective of the unknown onset of the dynamics.

Quantifying measurement noise in TSA ensembles

The aggregation of an ensemble from individually recorded junction collapse events allows to separate the measurement noise from the intrinsic fluctuations of the system. I apply three different techniques to separate the measurement noise from intrinsic fluctuations. For all three we assume that the observed junction length L_k at time $\tau_k = k\Delta\tau$ with $k \in [0, N]$ is given by

$$L_k = L_k^{\text{true}} + \epsilon_k . \quad (6.8)$$

N is the total number of reverse time steps and ϵ_k is an additive measurement noise. For simplicity, we assume the measurement noise ϵ_k to be a zero mean $\langle \epsilon_k \rangle = 0$ random gaussian variable with $\langle \epsilon_j \epsilon_k \rangle = (\sigma_N^2)_j \delta_{jk}$. Here $(\sigma_N^2)_j$ denotes the noise variance at time τ_j .

The first method to separate the measurement noise from the true dynamics of $L(\tau)$ is based on the observation that both mean and variance of the true TSA ensemble are zero at $\tau = 0$ min. With the variance of L_k separable into

$$\text{Var}[L_k] = \text{Var}[L_k^{\text{true}} + \epsilon_k] = \text{Var}[L_k^{\text{true}}] + \text{Var}[\epsilon_k] + 2\text{Cov}[L_k^{\text{true}}, \epsilon_k] \quad (6.9)$$

and under the assumption of independence of measurement noise and junction length ($\text{Cov}[L_k^{\text{true}}, \epsilon_k] = 0$), we can directly read off $(\sigma_N^2)_0$ from the extrapolated intersection of the observed variance with the y-axis at time $\tau = 0$ min. At the intersection and by construction of the TSA ensemble $\sigma_L^{\text{true}}(0) = 0\mu\text{m}^2$. We find $(\sigma_N^2)_0 = 0.049\mu\text{m}^2$.

The second independent method to determine the noise variance $\text{Var}[\epsilon_k]$, is based on the observation that the off-diagonal elements of the two-time ensemble covariance

$$\text{Cov}[L_j, L_k] = \text{Cov}[L_k^{\text{true}}, L_j^{\text{true}}] \quad \text{for } j \neq k \quad (6.10)$$

are measurement error free under the above assumption of uncorrelated measurement noise. The “true” variance can be determined for each time τ_k extrapolating $\text{Var}[L_k^{\text{true}}] = \text{Cov}[L_k^{\text{true}}, L_k^{\text{true}}]$

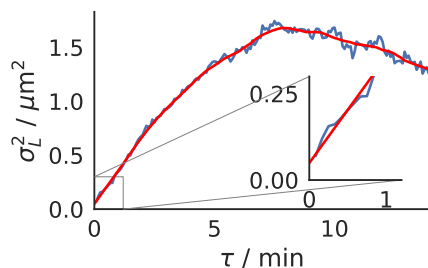


Figure 6.9: The measurement error close to the target state $(\sigma_N^2)_0$ can be read off from the to $\tau = 0$ min extrapolated ensemble variance. We first smooth the calculated experimental variance with the Savitzky Golay filter and subsequently linearly extrapolate the obtained function to $\tau = 0$ min. By construction $\sigma_L^{\text{true}}(0) = 0\mu\text{m}^2$ at the target state. The measurement error can thus be read off from the intersection with the y-axis. We find $(\sigma_N^2)_0 = 0.049\mu\text{m}^2$.

from the non-diagonal covariance terms. By subtraction of the measured variance and the extrapolation, we obtain an estimate for the time τ dependent measurement error (Fig. 6.10). With this method we find $(\sigma_N^2)_0 = 0.043\mu\text{m}^2$ which coincides nicely with the estimate obtained

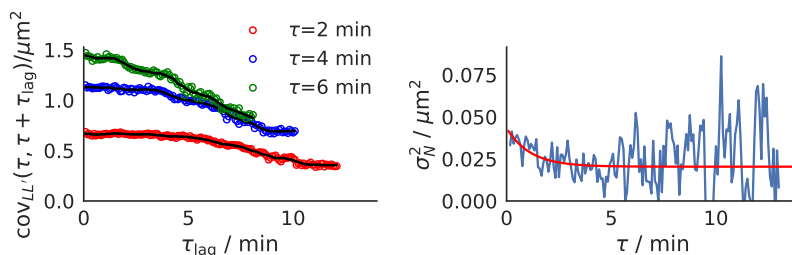


Figure 6.10: The time dependency of the measurement error can be estimated from the difference of the variance and the covariance extrapolated to its diagonal. (Left): Three representative cuts through the covariance with one time τ kept constant. Black lines are obtained using the Savitzky Golay filter and subsequently extrapolated to $\tau_{\text{lag}} = 0$ min. (Right): The difference between $\text{Var}[L_k]$ and the extrapolated covariance value yields an estimate of the measurement error. Fitting the first 7min with $\sigma_N^2(\tau) = a \exp(-b\tau) + c$ we find $a = 0.023\mu\text{m}^2$, $b = 0.882\text{min}^{-1}$, $c = 0.020\mu\text{m}^2$.

from the to $\tau = 0$ min extrapolated variance.

The third method is well known in the field of single particle tracking¹¹. This approach is based on the simple observation, that the first off-diagonal term of the increment covariance of a diffusion process with uncorrelated intrinsic and additive Gaussian measurement noise yields $\text{Cov}[\Delta L_k, \Delta L_{k-1}] = -(\sigma_N^2)_k$. Up to first order this approach is also valid for Langevin dynamics of the form Eq. (6.6). The theoretic increment covariance is

$$\text{Cov}[\Delta L_j, \Delta L_k] = \begin{cases} D\Delta\tau + 2\sigma_N^2 & \text{for } j = k \\ -(\sigma_N^2)_{\max(j,k)} & \text{for } j = k \pm 1 \\ 0 & \text{otherwise} \end{cases} + \mathcal{O}(\Delta\tau\sigma_N^2, \Delta\tau^2, \sigma_N^4) \quad (6.11)$$

Calculating the experimental increment covariance we find relevant non-zero contributions only for the diagonal and first off-diagonal term. All other off-diagonals fluctuate around zero (Fig. 6.11a). We confirm this observation calculating the auto-covariance of the increments, which is approximately zero for lags larger $\Delta\tau$ (Fig. 6.11b,c). With no long-range correlations

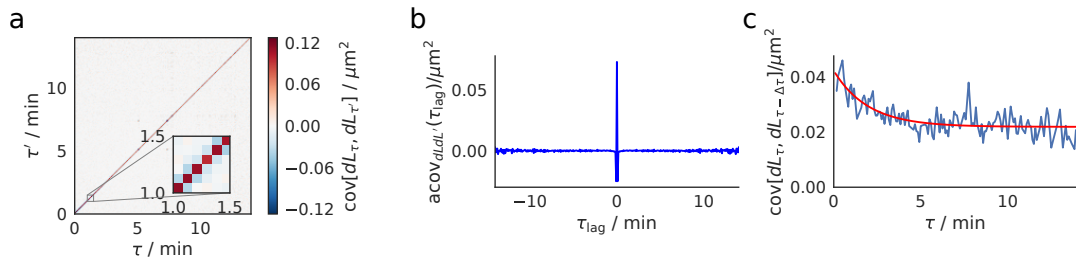


Figure 6.11: The measurement noise is estimated from the first off-diagonal of the increment covariance. (a): The increment covariance with zoom on the diagonal shows that only the diagonal and the first off-diagonal have relevant non-zero contributions. (b): The increment auto-covariance with lag-time τ_{lag} confirms that only the diagonal (positive center peak) and the first off-diagonal (negative peaks) have relevant non-zero contributions. (c): An exponential fit $a \exp(-b\tau) + c$ with $a = 0.020 \mu\text{m}^2$, $b = 0.506 \text{min}^{-1}$, $c = 0.022 \mu\text{m}^2$ yields an estimate for the with time τ changing measurement noise $\sigma_N^2(\tau)$.

detectable in the increment covariance the assumption of uncorrelated intrinsic and measurement noise is justified.

The estimated measurement noise $\sigma_N^2(0) = 0.042 \mu\text{m}^2$ is largest close to the target state, where it matches the one obtained by variance and covariance extrapolation. Within approximately 5 min, it then decays to its equilibrium value $\sigma_N^2(\infty) = 0.022 \mu\text{m}^2$. With known $\sigma_N^2(\tau)$ and with help of Eq. (6.11) one can estimate $D\Delta\tau$. For reference we here state the obtained value at $\tau = 8 \text{min}$. We find $D\Delta\tau = 0.021 \mu\text{m}^2$ and, using $\Delta\tau = 5 \text{sec}$, the estimated intrinsic noise $D = 0.25 \mu\text{m}^2 / \text{min}$. Below we will infer D from the data.

Comparing the intrinsic noise variance and the measurement noise, we find them to be of the same order. A model for the measurement noise must therefore be included into any analysis of the ensemble dynamics. We note that, when not evaluated at a certain time τ or as an average, D increases slightly towards the target state. Whether this is due to a change in the dynamics or a change in the fluctuations, can not be resolved without further knowledge about the underlying dynamics obtained by model inference.

In the following, we base all estimates on the measurement error defined in the third method, i.e. the increment covariance. In the final inferred model we further use the thus obtained measurement error variance $\sigma_N^2(\tau)$ to additively add measurement noise to simulated trajectories $L^{\text{true}}(\tau)$. We draw the measurement error of the junction length from the Rice distribution

$$P(L|L^{\text{true}}) = \frac{L}{\sigma^2} e^{-\frac{(L^2 + L^{\text{true}2})}{2\sigma^2}} I_0\left(\frac{LL^{\text{true}}}{\sigma^2}\right), \quad (6.12)$$

which can be obtained assuming additive Gaussian measurement noise with standard deviation σ for each vertex position. $I_0(z)$ denotes the Bessel function of the first kind with zero order. We note that Eq. (6.12) becomes Gaussian for almost all used junction lengths. Using the Rice distribution to correct simulated sample paths instead of a Gaussian, thus only prevents artificially introduced negative sample path lengths for trajectories close to the target state. It ensures a non-zero y-axis intersection of the simulated and measurement noise corrected variance, and slightly adds to the mean value in the last couple of steps before the target state. We therefore refrain from iteratively matching the variance of Eq. (6.12) with the experimental variance and simply set $\sigma^2 = \sigma_N^2$.

6.4 Models of increasing complexity

Models of increasing complexity allow us to find the most simple model that is capable of reproducing all relevant summary statistics. In the following sections we stepwise increase the complexity of the assumed model. For each model we determine whether it is in principle capable to reproduce the experimental TSA ensemble. We assume measurement noise of the form reconstructed by the increment covariance method, which decays in reverse time following the fit in Fig. 6.11.

Rejecting the null hypothesis of junction collapse by random fluctuations

The most simple conceivable effective model of junction shrinkage are random fluctuations that eventually lead to contact collapse. Although biologically highly implausible, our model selection approach must be able to rule out such a model purely based on the time-dependent statistics of the experimentally obtained ensemble. We start from the observation, that target state aligned random walk trajectories would follow a reverse time SDE of the form

$$dL = \frac{D}{L}d\tau + \sqrt{D} dW_\tau, \quad (6.13)$$

assuming well separated initial and final conditions. Eq. (6.13) can be obtained from Eq. (6.6) with $f(L) = 0$. In the literature this process is known as Bessel process and can be solved exactly (see chapter 4 for details). Surprisingly both the time dependent mean of Eq. (6.13)

$$\bar{L}(\tau) = \sqrt{\frac{8D\tau}{\pi}}, \quad (6.14)$$

with its square root like dependence on τ , and the variance

$$\sigma_L^2(\tau) = \frac{(3\pi - 8)D\tau}{\pi}, \quad (6.15)$$

which increases linearly in τ , seem to be proportional to mean and variance obtained from the experiment. To reject this hypothesis we fit the mean Eq. (6.14) of the TSA random walk to the mean of the data. Comparing only the last 3 min before junction collapse, we find good

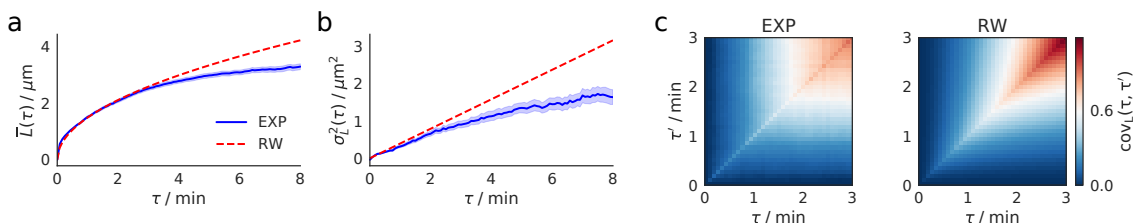


Figure 6.12: The null hypothesis of junction collapse by random fluctuations can only be rejected when statistics up to the two-time covariance are considered. (a): A least square fit of the theoretical TSA-random walk mean Eq. (6.14) (RW) to the first 3 min of the experimental mean (EXP) yields $D_{\text{rw}} = 0.87\mu\text{m}^2/\text{min}$. (b): Using the inferred D_{rw} to predict the ensemble variance, we find good agreement with the measurement noise corrected experimental variance in the first 3 min. (c): Comparing the experimental and TSA-random walk covariance, we find them strongly disagreeing in their lines of equi-covariance. The experimental covariance shows perpendicular equi-covariance lines, the random walk a wedge.

agreement. The inferred diffusion constant $D_{\text{rw}} = 0.87\mu\text{m}^2/\text{min}$ is used to predict the variance of the experimental data using Eq. (6.15). We obtain the puzzling result that the variance of

the data is well predicted by the inferred diffusion constant D_{rw} . A one parameter pure noise model perfectly fits the mean and variance of a biological process which is known to be strongly directional.

Comparing the above estimated diffusion constant D , with the here inferred D_{rw} , first doubts are appropriate. To resolve this puzzling predictive power of the random walk model, I compare the two-time covariance of both TSA-random walk and experiment (see Fig. 6.12c) and find strong disagreement. While lines of constant covariance are perfectly perpendicular in the experimental data, indicating a constant non decaying covariance when evaluated in reverse time, they form a wedge in the analytically obtained covariance expression Eq. (4.160) of the TSA-random walk.

This finding has strong implications for the quantitative analysis of contact constriction. With only a small number of junction collapses at hand it is not possible to reject the null hypothesis of junction remodeling as a random event. It further leads to a wrong estimate of the diffusion constant. Only when enough data is available to calculate the two-time covariance the null hypothesis can be rejected. With this in mind we increase the model complexity to also include an effective deterministic force component $f(L)$.

A model of power law like effective constriction dynamics

In this section we ask whether the full quantitative phenotype of contact constriction, comprised of mean variance and two-time covariance can in principle be explained by a deterministic effective force law $f(L)$ plus random fluctuations and measurement noise. We start with the inference of the most likely reverse time dynamics under the assumption of a power law like effective force law

$$f(L) = -\gamma L^\alpha . \quad (6.16)$$

To first order and close to the target state this approximation should hold for any one dimensional continuous effective force law. The study of power laws will therefore (to first order) also inform about the principal applicability of a single one dimensional potential for the distribution of junction shrinkage. For the inference I use a path maximum likelihood scheme that includes Gaussian measurement noise with variance σ_N^2 ^{43,109,167}. I extend this approach to an ensemble of trajectories of the same number of increments N where each trajectory contributes with the same weight to the log-likelihood function (see chapter 5 for details). The used log-likelihood function reads

$$\log \mathcal{L} = \sum_{j=1}^{n_{\text{ens}}} \sum_{i=0}^{N-1} \left(-\frac{\left(L_{i+1}^{(j)} - L_i^{(j)} - f^{\text{TSA}}(L_i^{(j)})\Delta\tau \right)^2}{2D\Delta\tau + 4\sigma_N^2} - \frac{1}{2} \log (2\pi (D\Delta\tau + 2\sigma_N^2)) \right) . \quad (6.17)$$

The inner sum runs over all N increments of a trajectory, the outer over all trajectories n_{ens} . The used reverse time TSA-force

$$f^{\text{TSA}}(L) = -\gamma L^\alpha + \frac{(D\alpha + D) \left(-\frac{2\gamma}{D\alpha + D} \right)^{\frac{1}{\alpha+1}} e^{\frac{2\gamma L^{\alpha+1}}{D\alpha + D}}}{\Theta(\alpha + 1)\Gamma\left(\frac{1}{\alpha+1}\right) - \Gamma\left(\frac{1}{\alpha+1}, -\frac{2L^{\alpha+1}\gamma}{\alpha D + D}\right)} \quad (6.18)$$

can be obtained for all power law exponents α from Eq. (6.7) using Eq. (6.16) (for details see chapter 3 and 4). $\Gamma(z)$ is the gamma-function, $\Gamma(\nu, z)$ the upper incomplete gamma function and $\Theta(\nu)$ the Heaviside step function. The connecting case for $\alpha = -1$ is the Bessel process with reverse time TSA force $f^{\text{TSA}}(L) = \frac{\gamma + D}{L}$. For numerical stability we use the form stated in

Eq. (5.16). Maximum Likelihood parameters valid for the last 4 min before reaching the target state are obtained using nested sampling^{20,48,49,50,157}. For a brief introduction to the method see section 6.5. We find $\alpha_{\text{ML}} = -0.64$, $\gamma_{\text{ML}} = 0.70\mu\text{m}^{1-\alpha}/\text{min}$ and $D_{\text{ML}} = 0.33\mu\text{m}^2/\text{min}$. The inferred D is expectedly similar to the one obtained from the estimates of the increment variance in section 6.3. Simulating the reverse time dynamics with respect to the inferred parameter and subtracting the estimated measurement noise from the experimental variance, we find experiment and theory to disagree in both mean and variance (Fig. 6.13(Left)).

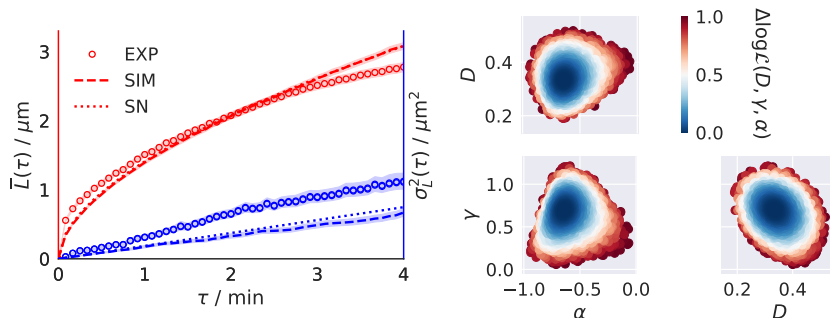


Figure 6.13: Maximum likelihood reverse time ensemble inference shows that power laws of the form Eq. (6.16) are not suitable to describe junction constriction dynamics. (Left): Both mean (red) and variance (blue) of the experiment and reverse time simulations of the maximum likelihood force law with $\alpha_{\text{ML}} = -0.64 \pm 0.02$, $\gamma_{\text{ML}} = 0.70 \pm 0.03\mu\text{m}^{1-\alpha}/\text{min}$ and $D_{\text{ML}} = 0.33 \pm 0.01\mu\text{m}^2/\text{min}$ disagree. The precision of the obtained values is stated in one standard deviation to both sides. All curves based on trajectory averages are shown with 95 % bootstrap intervals. In this plot we additionally show that the small noise theory (dotted) almost perfectly coincides with the reverse time simulation. (Right): The normalized relative change of the log-likelihood indicates that, although positive and negative α -values in the range $[-2, 2]$ were allowed, only negative values are selected. For the inference the last 4min before the target state were used.

To study whether power law models for the constriction dynamics can in general be rejected for all α and irrespective of the inferred maximum likelihood α_{ML} , we resort to an analytical approach. We first expand Eq. (6.18) for small L and under the assumption of weak noise (see chapter 4). A subsequent small noise moment expansion⁵⁵ yields for the mean

$$\bar{L}(\tau) = ((1 - \alpha)\gamma\tau)^{\frac{1}{1-\alpha}} + D \frac{(7\alpha - 3)((1 - \alpha)\gamma\tau)^{\frac{\alpha}{\alpha-1}}}{4(3\alpha - 1)\gamma} + \mathcal{O}(D^{3/2}), \quad (6.19)$$

and for the variance

$$\sigma_L^2(\tau) = D \frac{1 - \alpha}{1 - 3\alpha} \tau + \mathcal{O}(D^2). \quad (6.20)$$

For details see chapter 4. For the inferred parameter values, we find almost perfect agreement with the mean and variance of the exact reverse time simulation (Fig. 6.13(Right)).

The most obvious statistic, that neither the noise driven nor the power law approach was capable to explain, is the experimentally observed two-time covariance with its straight equi-covariance lines $C_L(\tau, \tau') \approx \min(\sigma_L^2(\tau), \sigma_L^2(\tau'))$ (see Fig. 6.14(Left)). For $\alpha = \alpha_{\text{ML}}$, the small noise covariance, valid for $\alpha < 0$,

$$C_L(\tau, \tau') = \begin{cases} \left(\frac{\tau}{\tau'}\right)^{\frac{\alpha}{\alpha-1}} \sigma_L^2(\tau) + \mathcal{O}(D^2) & \text{for } \tau < \tau' \\ \left(\frac{\tau'}{\tau}\right)^{\frac{\alpha}{\alpha-1}} \sigma_L^2(\tau') + \mathcal{O}(D^2) & \text{for } \tau > \tau' \\ \sigma_L^2(\tau) + \mathcal{O}(D^2) & \text{for } \tau = \tau' \end{cases} \quad (6.21)$$

clearly differs from the experimentally observed (see Fig. 6.14). For the derivation see chapter 4. In general, for any $\alpha < 0$, the covariance Eq. (6.21) decays for $\tau > \tau'$ and does not preserve

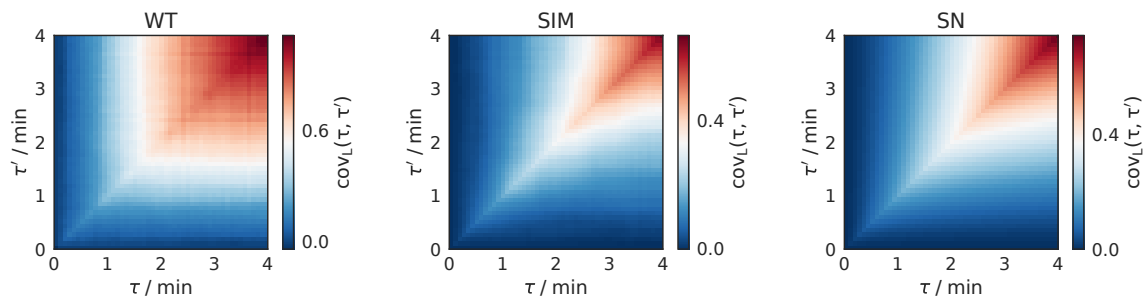


Figure 6.14: The two-time covariance comparison of experiment (WT), reverse time simulation (SIM) and small noise (SN) covariance, using maximum likelihood parameters $\alpha_{\text{ML}} = -0.64$, $\gamma_{\text{ML}} = 0.70\mu\text{m}^{1-\alpha}/\text{min}$ and $D_{\text{ML}} = 0.33\mu\text{m}^2/\text{min}$, exemplary shows, that power laws of the form Eq. (6.16) are genuinely not capable of reproducing the statistics of the experiment. While experiment and power law dynamics strongly differ in their lines of equi-covariance, do the simulation and its small noise approximation agree.

the constant value $\sigma_L^2(\tau')$ observed in the experimental covariance. The limit $\alpha \rightarrow 0$, where the condition of a non decaying covariance seems fulfilled, is not included in the validity of the approximation. For α approaching 0, the stated covariance is only valid further away from the target state (see chapter 4). More important however, for α close to zero, these cases can be ruled out as there the small noise mean Eq. (6.19) becomes almost linear in contradiction to the observed experimental mean (Fig.6.13).

Most cases with $\alpha > 0$ can be ruled out by the convex shape of the deterministic solution and thus of their mean. Only fully fluctuation (i.e. D) dominated cases, which can be approximated by Eq. (6.13), could show a square root like behavior of the mean Eq. (6.14) (see chapter 3). These cases are however ruled out by their wedge-like shape of their covariance shown in Fig. 6.14(Middle,Right). For a more detailed discussion of the two regimes of α see chapter 3 and 4.

In summary, using both evidence from stochastic inference and small noise theory, pure power law constriction dynamics can be ruled out. In the next sections, I will discuss biologically possible scenarios that are in principle capable of generating the observed two-time covariance.

Relaxing assumptions of the power law model

The failure of the power law model to explain the experimental data allows to plausibly reject any constriction mechanism that assumes a single effective potential plus white noise as a sufficient model for contact constriction. In this section, we study whether one of the implicit and simplifying assumptions underlying the Langevin equation Eq. (6.6) leads to the failure of the power law dynamics to reproduce all observed statistics. We study three modifications to Eq. (6.6). (i) Correlated or multiplicative noise, (ii) heterogeneity in the realizations of the force law, (iii) the importance of the forward initial conditions.

Adding correlated or multiplicative noise

The first possibility to induce a non-decaying covariance are correlations in the noise, e.g. induced by elastic interactions with neighboring cells and junctions. From the Mori-Zwanzig projection formalism, it is known that projecting a high dimensional system (such as the complex epithelial tissue) onto a set of effective coordinates (one junction) causes correlations in the noise if the

time scales of the fluctuations are not separated from the time scale of the dynamics^{115,117,185}. From the increment analysis discussed above we know however that no prevailing correlations are present in the increments. We therefore rule out correlations in the noise as cause for the non-decaying covariance.

Projecting a high dimensional system onto a low dimensional representation can further justify to generalize the diffusion constant D as multiplicative noise term $D(L)$. Given the time scales of the low dimensional dynamics (here one junction) are well separated from the residual changes of the system, i.e. occur on a longer time scale, the diffusion with D in L -direction might effectively slow down and (with some technical restrictions) a multiplicative noise terms of the form

$$D(L) \approx \frac{D}{1 + \left(\frac{d}{dL}\sigma(L^\perp)_L\right)^2} \quad (6.22)$$

occurs.⁹ Here $\sigma(L^\perp)_L$ denotes the standard deviation of orthogonal fluctuations in L^\perp -direction averaged at position L . An increasing variance in the orthogonal dynamics thus leads to less effective fluctuations in junction collapse direction. Guided by pure sterical considerations we next assume that with a shorter contact length also the space of possible perpendicular fluctuations decreases. Assuming a roughly linear dependency of $\sigma(L^\perp)_L$ on L , $D(L)$ will therefore either slightly increase or decrease with L . Simplifying again, we approximate the effective diffusion as a power law

$$D(L) = DL^\beta \quad (6.23)$$

with β either slightly positive or negative.

From the theory of TSA ensembles extended to multiplicative noise we know that the expressions for multiplicative noise in forward and reverse time are identical $D(\widehat{L}) = D(L)$ (see chapter 7 for details). Using again a small L weak noise expansion, now extended to multiplicative noise, we can next rule out multiplicative noise of the form Eq. (6.23). We start from the observation, that the variance of the experiment is linear in τ , close to the target state (Fig. 6.12b). In the small L and weak noise limit of TSA dynamics with powerlaw like multiplicative noise, we find however, that noise driven dynamics show a variance

$$\sigma_L^2(\tau) \sim (D\tau)^{\frac{1}{1-\beta/2}} \quad (6.24)$$

for $\alpha \geq \beta$ and $\beta < 2$, and force driven dynamics a variance

$$\sigma_L^2(\tau) = \frac{D \left((1-\alpha)\gamma\tau \right)^{\frac{1-\alpha+\beta}{1-\alpha}}}{\gamma \left(1 - 3\alpha + \beta \right)} \quad (6.25)$$

for $\alpha < \beta$ and $\alpha < 0$. Both expressions become non-linear in τ for $\beta \neq 0$. We therefore rule out multiplicative noise as a relevant contribution to the experimentally observed dynamics (see chapter 7 for a detailed derivation and visualization).

Heterogeneity in the force law realizations

When studying an ensemble of different realizations of the same biological dynamics, differences in the observed dynamics are most likely not only due to intrinsic random fluctuations but also due to different realizations of the machinery at work. For example, different amounts of e.g. actin, myosin or E-cadherin molecules are aggregated into a single machinery that leads to constriction. While the form of the effective force law should be the same for all events, is the constricting strength γ probably different for each realization. Using again the small L weak

noise limit of TSA dynamics, now extended to heterogeneous dynamics, we find, that mainly the covariance changes under the influence of heterogeneity. We find

$$\begin{aligned}
C_L(\tau, \tau') &= \left(\langle \gamma^{\frac{2}{1-\alpha}} \rangle - \langle \gamma^{\frac{1}{1-\alpha}} \rangle^2 \right) (1 - \alpha)^{\frac{2}{1-\alpha}} (\tau \tau')^{\frac{1}{1-\alpha}} \\
&\quad + C_L^{\text{hom}}(\tau, \tau') \\
&\quad + D \left(1 - \langle \gamma^{\frac{1}{1-\alpha}} \rangle_{P(\gamma)} \langle \gamma^{\frac{\alpha}{\alpha-1}-1} \rangle_{P(\gamma)} \right) \frac{(\alpha - 1)(7\alpha - 3)}{4 - 12\alpha} \left(\tau^{\frac{1}{1-\alpha}} \tau'^{\frac{\alpha}{\alpha-1}} + \tau'^{\frac{1}{1-\alpha}} \tau^{\frac{\alpha}{\alpha-1}} \right),
\end{aligned} \tag{6.26}$$

where $C_L^{\text{hom}}(\tau, \tau')$ denotes the homogeneous contribution already stated in Eq. (6.21). $\langle \cdot \rangle_{P(\gamma)}$ denotes the ensemble average with respect to $P(\gamma)$. In chapter 7 we show that the effect of the two additional terms leads to a broadening of the covariance and that, for short durations, effectively perpendicular equi-covariance lines can be achieved. Preserving a perpendicular structure for more than 8 min, and into a regime where constriction has not yet started, as observed in the experimental covariance (Fig. 6.10(Left)), is however not possible with this approach. On these grounds, we rule out heterogeneity of this kind as the sole reason for the experimentally obtained covariance. The sought mechanism must be heterogeneous, but also be preserved both in the constriction and maintenance regime. A sole heterogeneous power law is not sufficient as an explanation.

The influence of forward initial conditions

In this section we want to evaluate how different initial conditions influence the reconstructed ensemble statistics. From the theory of TSA dynamics in chapter 3 and 4 we know that in the very vicinity of the target state the influence of the initial conditions is marginal. Further away the free energy force $f^{\mathcal{F}}(L)$ Eq. (3.6) however changes with changing initial conditions.

Inspecting the ensemble of experimental trajectories in Fig. 6.6 it is clear that no initial conditions can be defined. Instead, moving backwards in time the trajectories eventually seem to approach an equilibrium regime with constant mean and variance. To be precise, each junction seems to assume its own equilibrium length. The full reverse time dynamics are therefore comprised of a constriction part, followed by a transition regime and subsequently a regime of maintained junction length.

In its current form the TSA theory does not provide the means to include such explicit time dependent changes in the effective force law. To nevertheless assess the influence of forward initial conditions on TSA dynamics we compare simulations of the same dynamics, once in forward time including initial conditions and once in reverse time assuming initial conditions at infinity. The forward dynamics are

$$d\hat{L} = \begin{cases} -\gamma_{\text{OU}} \left(\hat{L} - \mu_j \right) dt + \sqrt{D} dW_t & \text{for } t < T_{\text{sw}} \\ -\gamma \hat{L}^\alpha dt + \sqrt{D} dW_t & \text{otherwise,} \end{cases} \tag{6.27}$$

where we chose T_{sw} large enough to ensure that all such synthesized trajectories are at least as long as the maximal sample path duration in the experimental ensemble. As a proxy for the equilibrium cell contact length μ_j , the initial conditions are drawn from the instantaneous position of the experimentally obtained TSA ensemble at time $\tau = 8\text{min}$ (Fig. 6.15). This equals roughly the time where the TSA ensemble mean and variance seem to have reached its equilibrium value.

The transition probability of the initializing equilibrium Ornstein-Uhlenbeck process with Gaussian measurement noise can be calculated analytically⁷¹. One obtains for the corresponding

log-likelihood function⁷¹

$$\log \mathcal{L} = \sum_{j=1}^{n_{\text{ens}}} \sum_{i=0}^{N-1} \left(-\frac{\left(\widehat{L}_{i+1}^{(j)} - \mathbb{E}[\widehat{L}_{i+1}^{(j)} | \widehat{L}_i^{(j)}] \right)^2}{2 \text{Var}[\widehat{L}_{i+1}^{(j)} | \widehat{L}_i^{(j)}]} - \frac{1}{2} \log \left(2\pi \text{Var}[\widehat{L}_{i+1}^{(j)} | \widehat{L}_i^{(j)}] \right) \right), \quad (6.28)$$

with conditional moments

$$\mathbb{E}[\widehat{L}_{i+1}^{(j)} | \widehat{L}_i^{(j)}] = \frac{\widehat{L}_i^{(j)} D + 2\gamma_{\text{OU}} \mu_j \sigma_N^2}{D + 2\gamma_{\text{OU}} \sigma_N^2} e^{-\gamma_{\text{OU}} \Delta t} + \mu_j (1 - e^{-\gamma_{\text{OU}} \Delta t}), \quad (6.29)$$

$$\text{Var}[\widehat{L}_{i+1}^{(j)} | \widehat{L}_i^{(j)}] = \frac{D \sigma_N^2}{D + 2\gamma_{\text{OU}} \sigma_N^2} e^{-2\gamma_{\text{OU}} \Delta t} + \frac{D}{2\gamma_{\text{OU}}} (1 - e^{-2\gamma_{\text{OU}} \Delta t}) + \sigma_N^2. \quad (6.30)$$

For reference, in the limit $\gamma_{\text{OU}} \rightarrow 0$, we recover the diagonal of the increment covariance Eq. (6.11) used to estimate the measurement noise.

Equating the experimental estimated increment variance $\text{var}[dL_\tau]$ (Fig. 6.15(Left)) with the OU variance with measurement error stated in Eq. (6.30) (averaged over all sample paths) a good estimate for γ_{OU} can be obtained. We evaluate the equation at $\tau = 8\text{min}$, and with $\text{var}_{dL}(8\text{min}) = 0.064\mu\text{m}^2$, $\sigma_N^2(8\text{min}) = 0.022\mu\text{m}^2$, $D = D_{\text{ML}} = 0.33\mu\text{m}^2/\text{min}$ and $\Delta\tau = 5\text{sec}$, we find $\gamma_{\text{OU}} = 0.88\text{min}^{-1}$. With this, all parameters are determined and we can evaluate the influence of forward initial conditions on reverse time TSA ensemble statistics.

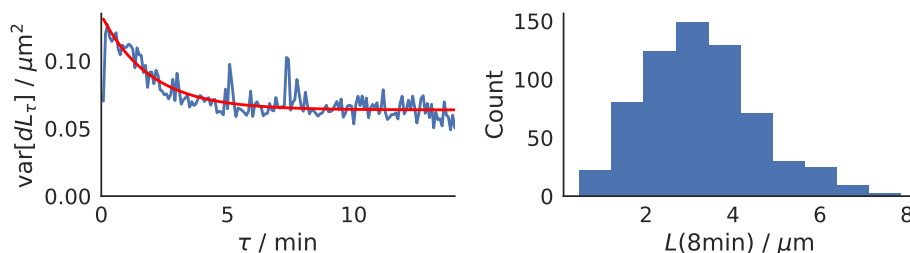


Figure 6.15: A simple estimate of the maintenance dynamics is obtained from the increment variance and the junction length statistics at 8min. The increment variance with exponential fit $\text{var}_{dL}(\tau) = a \exp(-b\tau) + c$ (Left), and the distribution of instantaneous sample path positions at 8min (Right), together with the noise estimation shown in Fig. 6.11, are used to approximate the equilibrium dynamics before constriction as a harmonic potential (Eq. (6.27)). We find $a = 0.070\mu\text{m}^2$, $b = 0.521\text{min}^{-1}$ and $c = 0.064\mu\text{m}^2$ and thus $\text{var}_{dL}(8\text{min}) = 0.064\mu\text{m}^2$.

As expected, the ensemble statistics of the TSA forward and in reverse time simulated dynamics coincide close to the target state (Fig. 6.16). After approximately 4 min, mean and variance, of dynamics with and without specified initial conditions, start to diverge. The same holds true for the two-time covariance. Close to the target state both approaches indistinguishably show the wedge like structure of a covariance which decays with increasing τ and $\tau > \tau'$. For later times τ however, the covariance, which includes fluctuations around an equilibrium contact length, starts to show the signature of the experimental covariance. We will use this observation as a guidance for the following considerations.

Explicitly time dependent and heterogeneous dynamics

Combining the effect of heterogeneous maintenance dynamics with power law constriction dynamics in the previous section, has revealed two properties of such systems. (i) Heterogeneous

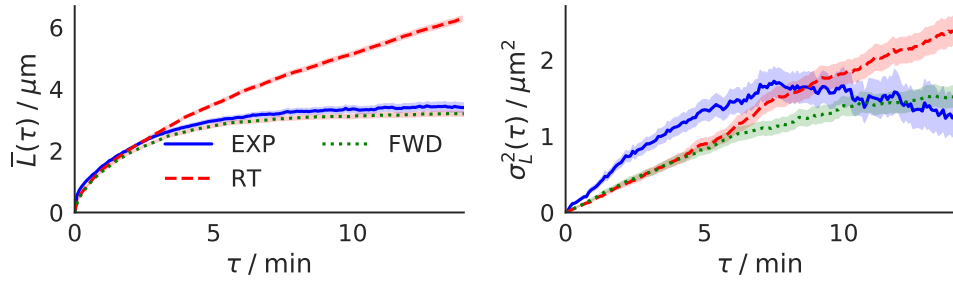


Figure 6.16: Compared to pure reverse time power law dynamics, combining power law constriction dynamics with equilibrium dynamics and a hard switch starts to change the observed ensemble statistics at $\tau \gtrsim 4$ min. (Left): The mean of the reverse time simulation (RT) with forward initial conditions at $+\infty$ starts to deviate from the experimental (EXP) mean at $\tau \approx 4$ min. Parameters used for the power law are again $\alpha_{\text{ML}} = -0.64$, $\gamma_{\text{ML}} = 0.70 \text{min}^{-1}$ and $D_{\text{ML}} = 0.33 \mu\text{m}^2/\text{min}$. Using a forward equilibrium simulation with the estimated $\gamma_{\text{OU}} = 0.88 \text{min}^{-1}$ and a hard switch into the same power law, as it underlies the reverse time dynamics, leads to a very good prediction of the mean. (Right): As already expected from Fig. 6.15, the variance of both the forward and reverse time approach do not coincide with the experimental observation. The deviation in the variance of the forward and reverse time approach at $\tau \approx 4$ min indicate the time where the forward initial conditions start to change the statistics.

dynamics are capable to reproduce the experimentally observed covariance. (ii) In the current construction, the influence of these dynamics is however limited to times $\tau \gtrsim 4$ min.

Combining these two findings we reason that the heterogeneity, induced by the different equilibrium cell contact lengths, must be “transported” into the constriction regime, while preserving its heterogeneity. The most parameter saving approach to realize this idea is therefore to assume, that not a position dependent force, but a to first order purely time dependent force drives constriction. In this picture each cell contact fluctuates around an equilibrium length which gradually changes over time. The assumed effective Langevin equation reads

$$d\hat{L} = -\gamma \left(\hat{L} - \mu_j(t) \right) dt + \sqrt{D} dW_t . \quad (6.31)$$

To mimic the increased speed of junction collapse observed in the ensemble mean, we model the time dependent change as

$$\mu_j(t) = \mu_j \left(1 - e^{-\frac{t-t_\mu^j}{\Delta t_s}} \right) , \quad (6.32)$$

where Δt_s provides the time scale and t_μ^j the time where $\mu_i(t)$ reaches the target state. In reverse time these parameters are denoted as $\Delta \tau_s$ and τ_μ^j . For simulations we set $t_\mu^j = T_{\text{sw}}$. In reverse time, t_μ^j becomes an inference parameter that reveals, whether the target state was reached by force or diffusion.

Inferring explicit time dependent dynamics

The reverse time TSA inference of an ensemble with time dependent force law is a currently unsolved problem. Researchers trying to solve this problem are confronted with the observation that after target state alignment, each sample path has experienced a different force law at the same space time point (L, τ) . A single, well defined TSA-force can not any longer be obtained.

With reverse time inference prevented by the supposedly explicit time dependence of the problem, we resort to the biased forward in time inference of the dynamics. We use the path

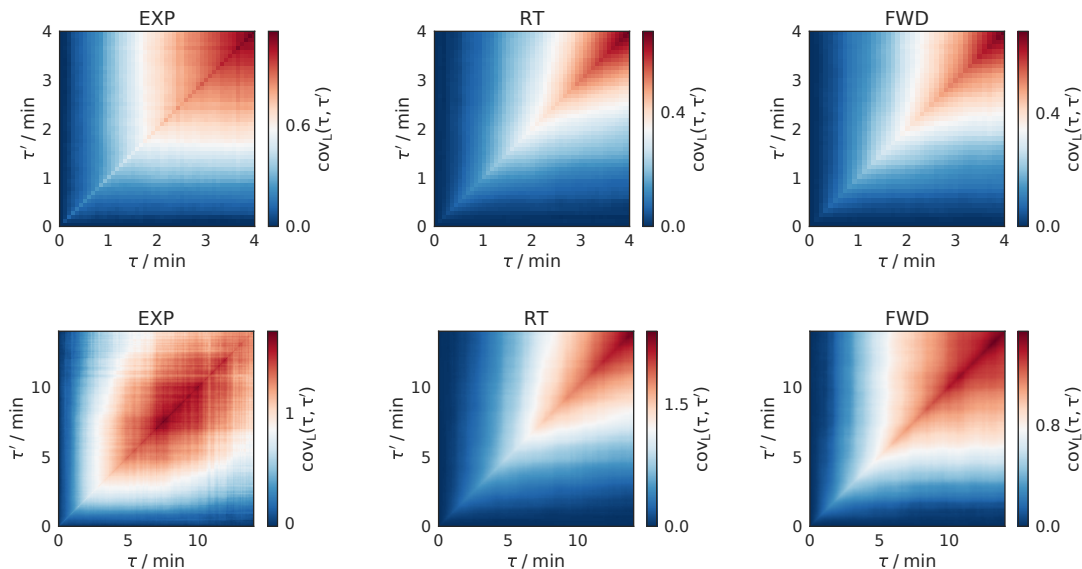


Figure 6.17: The comparison of experimental (Exp), reverse-time (RD) and forward simulated dynamics (FWD), with respect to their two-time covariance at different times, shows, that heterogeneity, due to a distribution of equilibrium cell contact lengths, is capable to reproduce perpendicular equi-covariance lines. The upper row shows the covariance for the first 4 min, the lower for the first 14 min. While the reverse time simulation (Middle) is unchanged over time, we find almost perpendicular equi-covariance lines for the late time forward simulation. In comparison to the experimental covariance (Bottom Left) they do however occur at too late times τ .

ensemble inference algorithm described in section 6.4 Eq. (6.28)-Eq. (6.30), with μ replaced by Eq. (6.32). In general, using forward in time inference for targeted processes, the used parameter inference is biased. Selecting forward trajectories, based on any other criterion than being from the same fixed time interval and with identical lifetimes, introduces pseudo forces into the observed ensemble. Different positions are sampled with statistics unlike the ones given by the forward dynamics. It is this effect that is characterized exactly in the reverse time TSA ensemble and ignored in the forward approach (see chapter 3 and 4).

Lets consider the problem in its most simple form. We assume that an ensemble of trajectories originates at the same position and time (\hat{L}_0, t_0) . Until the fastest trajectory hits the target state all trajectories are represented with the same weight in the ensemble used for inference. As time evolves, the typical lifetime, a not yet terminated trajectory in the ensemble has when it reaches a certain position, increases. The apparent force needed to reach the target state is decreased. In the limit of $t \rightarrow \infty$ it appears as if no force at all is necessary. With all trajectories eventually reaching the target state this effect is strongest in the vicinity of the target state, where now not only the fastest but all trajectories contribute to the inferred apparent force at this position. Using a forward inference scheme therefore adds a convex bias to the true force. For example, a power law force $\sim \hat{L}^\alpha$ would be inferred as $\sim \hat{L}^{\alpha+|c|}$.

The only dynamics where these problems do not occur are equilibrium dynamics. Here, forward and reverse time dynamics are identical and the temporal start and endpoint are irrelevant. Importantly, if the change in the equilibrium position is slow enough, the suggested time dependent model Eq. (6.31) is precisely such an equilibrium process with quasi statically changing equilibrium position. Independent of the initial equilibrium distribution and the direction of time, it is then always the same process with a shift in its reference position.

Studying these dynamics as an ensemble with target state, will nevertheless introduce a small “convexifying” bias, as some trajectories live longer than others. This bias will however

only effect the inferred dynamics of the mean position, which is solely responsible for the shift towards the target state and thus for the bias. We therefore expect only Δt_s , which defines the speed of the change in the reference position, to be slightly biased. Following the above reasoning, the true value will be marginally smaller than the inferred.

Inferring heterogeneous dynamics

The inference of dynamics, with each trajectory starting from its own equilibrium length, adds another level of complexity to the inference scheme. With μ_j and t_μ^j each trajectory of the model Eq. (6.31) has its own “local” parameters to be inferred, additional to the “global” ensemble parameter γ , D and Δt_s . We approximate this high dimensional inference problem by always keeping the local parameter at the maximum likelihood value (μ_j^*, t_μ^{j*}) , with respect to the current sample point in the global parameter space spanned by γ, D and Δt_s . The log-likelihood function from Eq. (6.28) then reads

$$\log \mathcal{L} = \sum_j^{n_{\text{ens}}} \log \mathcal{L}_j(\gamma, D, \Delta t_s | \mu_j^*, t_\mu^{j*}). \quad (6.33)$$

The inference proceeds in two steps. (i) The inference algorithms, here nested sampling^{20,48,49,50,157} (and section 6.5), proposes a data point in the global parameter space and subsequently (ii) a second optimizer, here Nelder-Mead, finds for each local log-likelihood $\log \mathcal{L}_j$ the maximizing local parameters (μ_j^*, t_μ^{j*}) . The obtained summed $\log \mathcal{L}$ value is used for the next update of the global parameters γ , D and Δt_s .

The maximum likelihood model

Constriction, modeled as an elastic process with time changing equilibrium position Eq. (6.31), allows to reproduce all observed summary statistics. We find $\gamma_{\text{ML}} = 0.83 \text{min}^{-1}$, $D_{\text{ML}} = 0.30 \mu\text{m}^2/\text{min}$ and $\Delta t_{\text{sML}} = 2.45 \text{min}$, using the approximated inference scheme laid out above in this section. The ensemble statistics are produced with forward simulations, using the inferred global maximum likelihood parameters and the respective local inferred values μ_j in combination with the measurement noise model defined in Eq. (6.12). The sample paths lifetimes are adapted to the lifetimes of the experimental ensemble.

Both mean and variance coincide with the experimental ones over the full available time span (Fig. 6.18). Only close to the target state we find slight deviations in the mean. With only the single parameter Δt_s available to effectively tune the slope this is not surprising. A more detailed model, although facing the danger of overfitting, would need to either incorporate a second time scale, or a slightly position dependent force. Furthermore, the elastic model with time changing equilibrium position is the first so far discussed model that is capable of generating a non-decaying covariance, i.e. perpendicular lines of equi-covariance close to the target state (Fig. 6.19(Bottom,Left)). This covariance structure is preserved over the full lifetime of the TSA ensemble (Fig. 6.19(Bottom,Middle-Right)). Interestingly, after about 8min, the experimental covariance loses its particular structure, indicating that the clear ordering in equilibrium positions is disrupted and the modeling assumptions break down Fig. 6.19(Top).

The inferred model is further confirmed by the increment auto-covariance of the inferred and simulated increments (Fig. 6.20). In these increment statistics we find perfect agreement of simulation and experiment, indicating that both measurement and intrinsic noise are captured correctly. The increment auto-covariance of the inferred increments is only marginally different from the other two (Fig. 6.20). The small difference for $\tau_{\text{lag}} > \Delta \tau$ however indicates that the inferred forces are not perfectly subtracted from the increments but with a systematic bias.

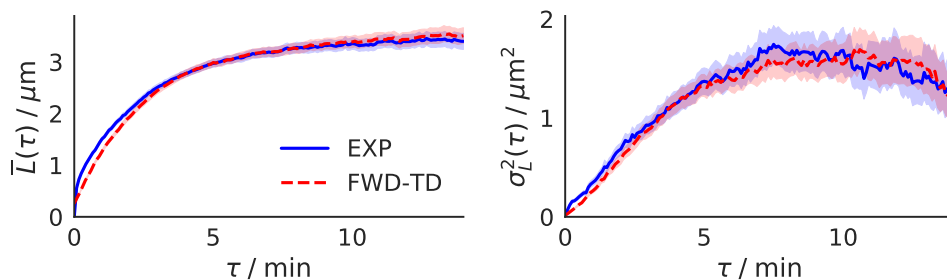


Figure 6.18: The ratchet like model of an elastic process with time changing equilibrium position Eq. (6.31) predicts the mean and variance over a time span of 14 min. The simulation (FWD-TD) is performed in forward time with parameters $\gamma_{ML} = 0.83 \pm 0.01 \text{min}^{-1}$, $D_{ML} = 0.30 \pm 0.01 \mu\text{m}^2/\text{min}$ and $\Delta\tau_{sML} = 2.45 \pm 0.03 \text{min}$. The precision of the obtained values is stated in one standard deviation to both sides. Both ensembles consist of 958 trajectories. After target state alignment τ -dependent measurement noise according to Eq. (6.12) is added. Each simulated trajectory is adjusted in its length to match its partner in the experimental data. (Left): Simulated and experimental mean coincide well over the full evaluated time span. The deviation close to the target state is due to the limited flexibility of the exponential force law to match the exact shape of the observed mean. (Right): Simulated variance and mean coincide over the whole time span. 95 % bootstrap confidence intervals are shown for both mean and variance. Notably only the last 8min until the target state is reached were used in the inference to predict the full 14min.

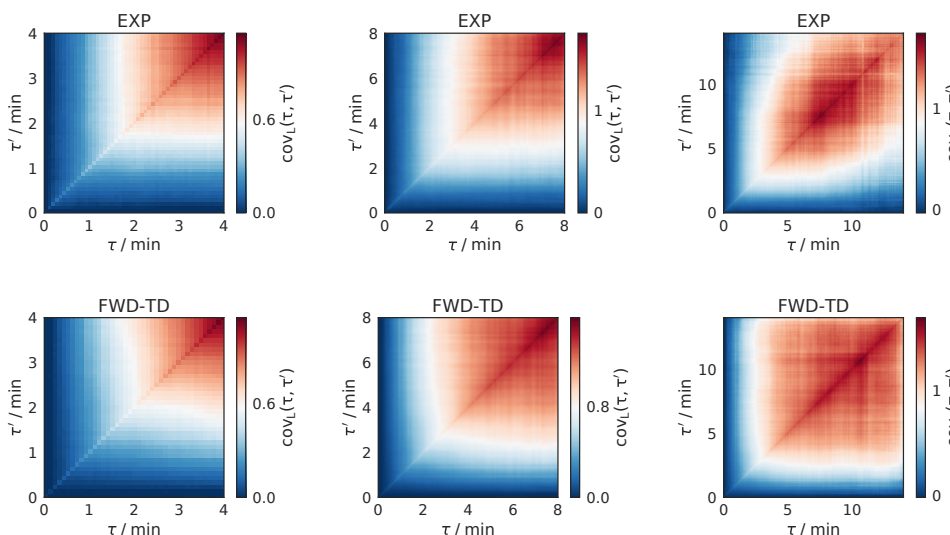


Figure 6.19: The two-time covariance of experiment and the ratchet like time dependent elastic process Eq. (6.31) agree well in their shape up to $\tau = 8 \text{min}$. At later times a new dynamical regime seems to be present which can not be traced to the finite length of the sample paths which is incorporated in the model.

This bias creates a detectable non-decaying auto-covariance of the residual increments. As above already observed for the mean, this bias can be attributed to the lack of flexibility in the assumed model. The almost not detectability of this bias however also indicates, that the model is a very good approximation to the true underlying dynamics.

For completeness, we also display the histogram of inferred equilibrium cell-contact lengths before constriction and the distribution of times τ_μ^i where the equilibrium length equals the target state (Fig. 6.21). We find $P(\mu)$ to be a broadly distributed skewed distribution. Whether this is just a fingerprint of with a certain rate intercalating cells, or even to facilitate intercalation, has to be resolved in the future. The distribution $P(\tau_\mu)$ scatters around zero, where zero is the noise

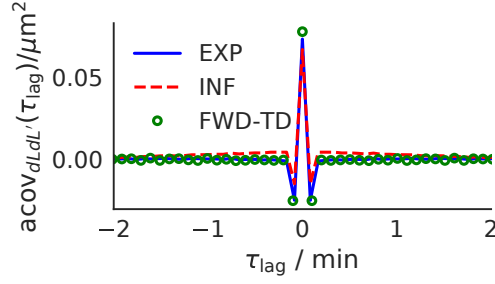


Figure 6.20: Measurement noise (lower peaks) and random fluctuations (included in center peak) are captured exactly by the model. The increment auto-covariance shows that experiment (EXP) and simulation of the ratchet like explicitly time dependent model with measurement noise (FWD-TD) coincide almost perfectly. The additionally shown auto-covariance of the increments slightly deviates, indicating that the strong simplification of a single time scale for the time dependent change of the equilibrium position introduces a slight but systematic bias.

free solution (Fig. 6.21). The joint distribution $P(\tau_\mu, \mu)$ can then be used as a health check for the model (Fig. 6.21). For most sample paths the inferred values are only slightly correlated, indicating that most sample paths are well described by the model. Only some trajectories, with small μ_i and large negative values τ_μ^i , indicate where the assumption of one time scale for the junction shrinkage of all cell contacts is too restrictive.

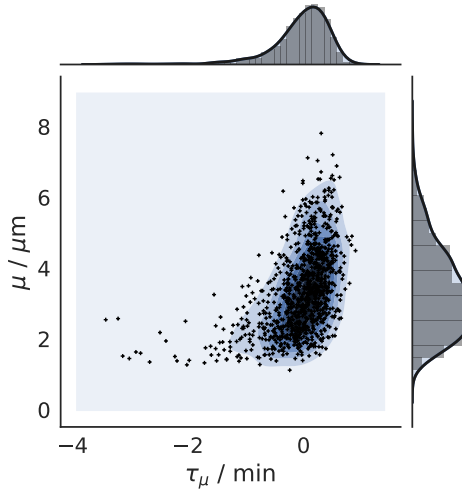


Figure 6.21: The joint scatter plot of the inferred equilibrium cell contact length μ_i and the time τ_μ^i , when $\mu(\tau)$ reaches the target state, confirms the model as a good description of the data. Most τ_μ^i are centered around zero. Sample paths which originate from a small equilibrium cell contact length μ_i sometimes select a clearly negative τ_μ^i . These “outliers” indicate the limitations of the model, which describes the change in the equilibrium position with just one global parameter $\Delta\tau_s$. In this plot we ignore three τ_μ^i values, which the inference algorithm shifts towards its lower negative limit. As already suggested by the ensemble length statistics at $\tau = 8\text{min}$ in Fig. 6.15, the distribution of equilibrium cell contact lengths μ_i is skewed and broadly distributed.

With the here used approach, the dynamical force law that leads to constriction can be inferred for each individually observed shrinking junction. In Fig. 6.22, we show a representation of this effective model for three experimentally obtained sample paths, which coincide in the inferred μ_i and τ_μ^i , but are otherwise completely independent realizations. All three nicely follow the inferred time dependent potential.

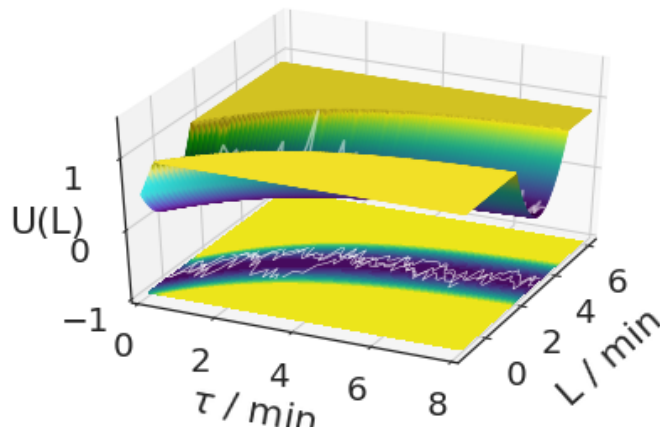


Figure 6.22: The reconstructed potential of trajectories with similar inferred equilibrium cell contact length represents a good effective description of the dynamics. For the plot we chose 3 trajectories with μ_j equal to the ensemble mean at 8 min. For graphical reasons the potential is truncated at higher values. Here $U(L)$ denotes the integrated force term including all parameters and is thus in units $\mu\text{m}^2\text{s}^{-1}$.

The biological interpretation

With a spring-like behavior on short time scales and viscous behavior on longer time scales the inferred model confirms the observed time scale separation in tissue²⁸. In this effective model γ is independent of time τ and realization. The inferred value for $\gamma_{\text{ML}} = 0.83\text{min}^{-1}$ can be interpreted as the inverse time scale of the spring. In the visco-elastic context dissipation thus occurs on a time scale of 1.2min. This value is in line with a similar estimate obtained by junction stretching experiments with optical tweezers^{8,32}.

The viscous contribution is modeled as an explicitly time dependent force that transports the equilibrium cell contact length towards the target state. To fulfill the ensemble statistics and in particular the mean, it is necessary that the equilibrium position $\mu_j(t)$ changes with increasing speed, here heuristically modeled as one minus an exponential. The necessity of a with time increased speed of junction shrinkage is in line with the observed feedback loop of myosin recruitment upon tension⁴⁷.

The time scale $\Delta\tau_{\text{sML}}$ is with 2.45 min larger than the typical duration of myosin pulses during constriction which are estimated to last about 80-100s³². Phases of pulse driven shrinkage and subsequent incomplete relaxation are not incorporated in this effective dynamics approach. Nevertheless, if the effective change of the equilibrium position would be oscillatory, a fingerprint of this behavior should be detectable in the residual increments after subtracting the inferred deterministic force $f(\hat{L}, t, \mu_j, T_i)$ defined by Eq. (6.31) and Eq. (6.32) for each trajectory. Interestingly, neither the increment covariance, nor the residual increment covariance shows signs of oscillatory behavior not captured by the drift term (Fig. 6.20). The not detectability of residual oscillations might have several causes. They could be buried in the measurement noise, which, with our choice of resolution, is probably higher than in the original report¹³⁶. It could however also be a selection bias in the original report¹³⁶ and depend on which contacts are chosen at which time at which position in the intercalating germband. Concluding, it could imply that although present, myosin pulses only matter in their average effect and in combination with the full ratchet like machinery which leads to monotonous shrinkage of the equilibrium cell contact length. It will be interesting to compare the temporal dynamics of molecules involved in the ratchet mechanism⁸¹ with the pulsatile dynamics. One could then check whether a com-

plete suppression of pulses by a ratchet mechanism is possible, or whether the oscillations could dissipate undetected in the noisy fluctuations around the equilibrium length.

The here inferred dynamics are thus truly effective dynamics, which integrate, e.g. myosin driven junction shrinkage and remodeling, e.g. by E-cadherin endocytosis, into a single effective force Eq. (6.31), Eq. (6.32). The inferred, explicitly time dependent force is thus a baseline to which any more detailed junction remodeling model must compare to. Based on our current knowledge of the biological details of junction shrinkage (see section 6.1), one plausible assumption would be to assume that junction remodeling happens with a constant rate independent of τ . The increased speed could then be fully attributed to myosin recruitment upon junction shrinkage. The inferred, with time changing "equilibrium" position $\mu_i(t)$ from Eq. (6.32), should then be a strong predictor of the average time dependent myosin concentration upon junction shrinkage.

Outlook: Myosin dynamics

Acto-myosin dynamics are the basis of force exertion in cells⁹³. Actin fibers provide the scaffold and myosin, bundled in mini filaments and attached to different fibers, exerts pulling forces⁹³. To go beyond the here inferred effective model of junction shrinkage, we need to understand how the inferred time dependence of junction remodeling is linked to the acto-myosin machinery and its regulations.

The understanding of myosin driven junction remodeling has grown steadily. During germband extension, mediated by cell intercalation, myosin is enriched along "vertical" junctions^{12,16}. More recently, pulsed flows of medio-apical myosin towards junctions with subsequent junction shrinkage, were suggested as mechanism of junction collapse. In this model, junction shrinkage is followed by a ratchet like mechanism, which prevents the junctions to relax back to their initial configuration. The details of this mechanism is currently an active and debated field of research^{32,81}. Furthermore, junction shrinkage has been linked to several mechano-chemical feedback loops^{47,94,98,111}.

The here inferred effective model of junction shrinkage allows to make progress in deciphering the relative importance of these three mechanisms, i.e. myosin enrichment and the importance of pulsatility, the ratchet like mechanism, feedback loops and regulation. While the proposed mechanisms for the ratchet must be compatible with the observed dynamical change in the equilibrium position, could the increased speed of the reference junction shrinkage also be a marker for a feedback mechanism. I here only discuss the analysis for myosin but analog considerations hold for bio-chemical reporters that are linked to ratchet like dynamics and feedback loops.

With the TSA path ensemble approach, simultaneous recordings of the constricting contact and the junctional myosin dynamics can easily be incorporated in the inference scheme. In the most simple version of this approach myosin and its changes in local concentration are treated to linearly cause changes in the equilibrium position and the noisy fluctuations. The measured myosin dynamics then replace $\mu_i(t)$ and D in the model Eq. (6.31). This approach conceptually follows Clement et. al. ³², however with an important difference. While Clement et. al. ³² study this problem only on the individual sample path level, is the here introduced TSA ensemble approach capable of testing predictions on the ensemble level. In our approach, if such myosin driven dynamics are not capable of approximating the observed ensemble statistics, it can be rejected.

The second approach assumes the inferred explicit time dependent model Eq. (6.31) as ground truth. It uses the result of the maximum likelihood path ensemble inference scheme to split each trajectory in its deterministic drift and noisy contributions. In its simplest form this is done by subtracting the for each time step inferred force from the respective increment ΔL_i . The residual force and noise increments can then be correlated with myosin and the change in

myosin concentration. From the analysis in section 6.3 we know that the residual fluctuations are uncorrelated. Only the effective deterministic dynamics matter for constriction. Nevertheless, correlating the residual increments with myosin and its derivative provides valuable insights into the question what causes the observed junction length fluctuations. Correlating myosin with the inferred incremental force then provides information about the dynamical mechanism of junction shrinkage.

On an ensemble level the average myosin concentration can further be correlated with the averaged dynamics of the equilibrium position $\mu_i(t)$. A strong correlation would emphasize the importance of average effects such as myosin enrichment along “vertical” junctions.

To test these ideas my collaborator simultaneously recorded both junctional and myosin dynamics. When we started to explore the joint measurement, we realized that the so far for all experiments used junction marker GFP-Spider^{16,90}, in combination with a myosin marker¹⁰⁶, led to an alignment error. The apparent junctions, characterized by myosin enrichment, and the junctions represented by GFP-Spider were offset to each other. We reasoned that this is due to GFP-Spider labeling the whole plasma membrane, while the myosin dynamics occur mostly apically. Switching to E-cad-GFP¹²³ as a marker cured this problem.

Interestingly, this change in the z-axis representation of the junction-length dynamics, i.e. the measurement depth below the tissue surface, lead to a detectable change in the dynamics of the ensemble mean. Compared to the junction dynamics marked with GFP-Spider, junction

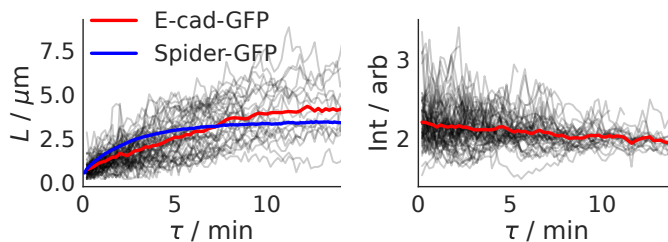


Figure 6.23: Junctional dynamics are dependent on the z-level and on the total junctional myosin concentration. (Left): Comparing junctional dynamics derived from E-cad GFP (apical) and Spider-GFP (wide z-range, comparably more basal) we find different dynamical behavior. While the apical dynamics change with constant rate (red), do the more basal / average dynamics (blue) show a square root like behavior. Interestingly, apical dynamics seem to lead and are only followed by the more basal dynamics. A future study based on z-stack scans, similar to a recent report¹⁷⁰, will hopefully shed light on this question. (Right): Myosin fluorescence intensities in arbitrary units normalized by the junction length to yield concentrations. We find a mild increase in the average concentration with no onset time detectable. Whether this result stands under improved experimental conditions (see section 6.5) remains to be checked. Black lines indicate individual samples of the experiment with junctions marked by E-cad-GFP.

shrinkage seems to occur with almost constant rate. Furthermore, junctions represented by a more apical marker (E-cad-GFP) seem at the same time and on average to be shorter than when represented more basal. This would motivate a model where junctions constrict first apically while more basal parts follow subsequently. The finding is in line with a recent report on different dynamics at different z-levels¹⁷⁰. Interestingly, not all junctions seem to follow the same scheme and only a fraction of maximally 60% of the junctions constrict first apically, with the basal movements lagging behind¹⁷⁰.

In preliminary results we find that the change in junction length anti-correlates with the total myosin intensity (Fig. 6.24(Left)) and its derivative (Fig. 6.24(Right)). Intensities are normalized by the junction length and thus represent the myosin concentration. With measurement artifacts

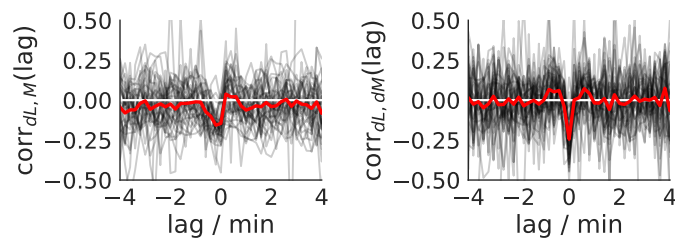


Figure 6.24: The incremental change of the junction length anti-correlates with both the myosin concentration and its derivative. (Left): Junction shrinkage seems to predate the increase in the myosin concentration confirming previous reports¹³⁶. (Right): The incremental change of both junction and myosin are anti-correlated at lag zero. Whether this is a measurement artifact or can be used to explain the intrinsic junction fluctuations with fluctuations in the myosin concentration remains to be answered in a more detailed study. In both plots, the red line indicates the average over all pairwise correlations, the black line individual realizations.

larger and of equal size as the signal, the reported intensity changes have to be interpreted with care (see section 6.5 for the data cleaning).

We find that the change in junctional myosin intensity seems to follow the change in the junction length, confirming previous reports of a delay of 9s¹³⁶. This would support the hypothesis that junctional myosin only has a stabilizing role, while apico-medial pulses drive constriction¹³⁶. With the here used temporal resolution of 12s a more detailed analysis is elusive. We furthermore find the incremental change of the junctional myosin concentration to be anti-correlated with the incremental junction length change. The observed lag is zero. Whether this is an artifact or an indicator that the intrinsic stochasticity observed in the TSA ensemble dynamics is due to changes in the myosin concentration will be answered in future experiments.

Unfortunately, with the currently available experimental setup, we are not able to provide truly reliable values for the myosin dynamics to test the above suggested inference strategies. The found importance of the measurement depth for epithelial tissue dynamics however provides interesting evidence for a recent shift in the perception of the importance of the third dimension of epithelial tissue dynamics. It will be interesting to explore the effective TSA dynamics approach for the full tissue depth, starting in slices from apical to basal. Combining these slices should provide us with a better understanding of the steric movements of intercalating cells and possibly provide evidence for the recently reported importance of basal dynamics.

6.5 Methods

Inference

Throughout this text maximum likelihood inference was performed within the framework of nested sampling^{20,48,49,50,157}. In general, nested sampling allows not only to obtain maximum likelihood values but the full posterior distribution

$$P(\Theta|D, M) = \frac{P(D|\Theta, M)P(\Theta|M)}{P(D|M)}, \quad (6.34)$$

or, assuming a flat prior $P(\Theta|M) = \text{const.}$, the full likelihood $P(D|\Theta, M)$. Here Θ is defined as the parameter vector, D the data vector and M the model. With the full likelihood distribution available, we can characterize the precision of the inference in terms of its mean and variance. As the mean value, within the stated precision, coincides in all here discussed cases with the maximum likelihood value, only the standard deviation is stated.

Designed for Bayesian model comparison, nested sampling calculates the evidence of a model, i.e. the probability $P(D|M)$ of the data D given the model M . Within the Bayesian framework it is obtained from the prior $P(\Theta)$ weighted integral

$$P(D|M) = \int P(D|\Theta, M)P(\Theta)d\Theta \quad (6.35)$$

over the likelihood $P(D|\Theta, M)$. To obtain the true quantity of interest, that is the probability of the model given the data, one applies Bayes law and finds

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)}. \quad (6.36)$$

While this quantity itself is not accessible, the ratio of two model probabilities, the so called Bayes factor

$$K = \frac{P(D|M_1)P(M_1)}{P(D|M_2)P(M_2)} \quad (6.37)$$

can be calculated from the ratio of two evidences times their model prior probability. For practical applications one assumes equal prior probability of the models.

For the nested sampling inference scheme the integral Eq. (6.35) is evaluated numerically via sampling. To avoid problems of convergence to the global maximum, the here used version of nested sampling multinest^{20,48,49} relies on an ensemble of life points, i.e. a vector of points in parameter space sorted by their likelihood²⁰. The prior defines which parameter values can be drawn. In each sampling step the likelihood list is updated and a more likely value added while the least likely value in the list is removed. The algorithm thus scans, beginning from the least likely cases, different layers of the likelihood until a convergence criterion is reached.^{20,48}. The fully evaluated evidence then allows for model comparison.

Unfortunately the above discussed models of power law constriction and visco-elastic shrinkage are designed for different time windows of the data set. The power law model suggest a mechanism close to the target state, the visco-elastic model refers to the full time window including the original equilibrium regime. The obtained evidence values are thus not comparable. Any future modification or proposed extension to the viso-elastic model can however be compared to this approach. No further modifications to the here suggested framework are necessary.

Experiments

All experiments were performed in the lab of Jörg Großhans by Deqing Kong with experimental protocols stated in his PhD thesis⁹⁰. Junction lengths were extracted with a fully automated pipeline written in the group of Fred Wolf by Lars Reichl (currently unpublished) based on ilastik¹⁶⁰ and EDGE⁵⁶. For the experiment with combined myosin and junctional dynamics recordings (section 6.4), Matthias Häring used a combination of machine learning for segmentation⁶⁴ and ImageJ^{144,145} for vertex reconstruction. We then reconstructed the myosin intensity along the junction using a $1.6\mu\text{m}$ wide rectangle up to but not including the vertex position. Based on the obtained raw data, consisting of junction length and junctional myosin intensity time series, I performed correction and analysis steps. The image cleaning is shown in Fig. 6.25.

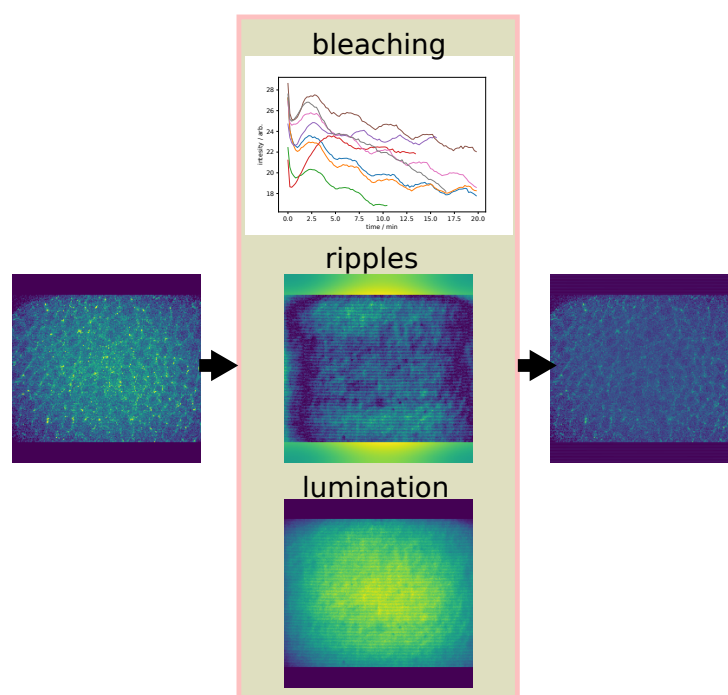


Figure 6.25: Schematic representation of the data cleaning procedure to obtain myosin intensities. (Left): A raw image of the myosin intensity. (Top): Average image intensities are used to correct for bleaching (and oscillations introduced by the setup) by normalization. (Middle): Horizontal ripples in the intensities are inferred in their amplitude and subtracted from the data. The respective frequency and an average over all data is shown. (Bottom): A with time changing illumination halo is fitted by an ellipse and corrected for by normalization. (Right): A fully corrected sample frame.

Generalizations of the theory of target state alignment

7.1 Heterogeneity in the realizations of the strength γ of the force law

To approximate biological processes with a single force law is always an idealization. We here relax this assumption to allow different realizations in the strength of the force law for individual trajectories while keeping the form of the force law fixed. For the above used example of power law forces $f(\widehat{L}) = -\gamma\widehat{L}^\alpha$ this implies that for each sample path realization, γ is drawn from a distribution $P(\gamma)$. To incorporate this form of heterogeneity we include a dependency on γ into the definition of the homogeneous TSA density, multiply by $P(\gamma)$ and integrate over all possible γ . The resulting ensemble captures the full heterogeneity. To infer the most likely realization of γ for each sample path we can use our path ensemble inference scheme, with the only difference that γ must now be optimized for each samplepath and not any longer on the ensemble level.

While difficult to observe on the single trajectory level, heterogeneity is clearly visible on the ensemble level when studying the moments and cumulants of the dynamics. We here denote the normalized ensemble distribution of the heterogeneous ensemble as $R^N(L, \tau; \gamma)$, and the normalized and in γ homogeneous contributions as $R^N(L, \tau|\gamma)$. As all moments for the homogeneous dynamics

$$\langle L^n(\tau) \rangle_{R^N(L, \tau|\gamma)} = \int R^N(L, \tau|\gamma) L^n dL \quad (7.1)$$

are linear with respect to the ensemble distribution $R^N(L, \tau|\gamma)$, the moments for the heterogeneous distribution

$$\begin{aligned} \langle L^n(\tau) \rangle_{R^N(L, \tau; \gamma)} &= \int \int R^N(L, \tau; \gamma) L^n dL d\gamma = \int P(\gamma) \left(\int R^N(L, \tau|\gamma) L^n dL \right) d\gamma \\ &= \langle \langle L^n(\tau) \rangle_{R^N(L, \tau|\gamma)} \rangle_{P(\gamma)} \end{aligned} \quad (7.2)$$

can simply be obtained by integration of $\langle L^n(\tau) \rangle_{R^N(L, \tau|\gamma)}$ with respect to $P(\gamma)$. In the next section we use this approach to generalize the small noise moments of homogeneous dynamics to heterogeneous dynamics.

The small noise expansion for heterogeneous dynamics with $\alpha < 0$

We here discuss the effects of heterogeneity on moments for the exemplary and analytically tractable case of the small noise expansion of power law force $f(L) = -\gamma L^\alpha$ with $\alpha < 0$. This

case of force driven TSA dynamics is particularly simple as we already know the moments for the homogeneous case (section 4.6).

Using Eq. (7.2) the moment expansion stated in section 4.6 above generalizes to

$$\bar{L}(\tau) = \langle\langle L(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) = \langle\langle L_0(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) + D \langle\langle L_2(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) + \dots \quad (7.3)$$

for the mean and

$$\begin{aligned} \sigma_L^2(\tau) &= \langle\langle \left(L(\tau) - \langle\langle L(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \right)^2 \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \\ &= \langle\langle \left(L_0(\tau) - \langle\langle L_0(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \right)^2 \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \\ &\quad + D \langle\langle L_1(\tau)^2 \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \\ &\quad + 2D \left(\langle\langle L_0(\tau) L_2(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) - \langle\langle L_0(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \langle\langle L_2(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \right) \dots, \end{aligned} \quad (7.4)$$

for the variance, where we used the expansion of $L(\tau)$ in small orders of \sqrt{D} stated in section 4.6. While for the mean, all displayed terms occur already in the homogeneous case and are only averaged with respect to γ , are the two not $L_1(\tau)$ dependent terms in the variance genuinely new and due to the introduced heterogeneity in γ . With $L_0(\tau)$ defined as the deterministic solution of the homogeneous case, the average with respect to $R^N(L, \tau|\gamma)$ can be neglected and the variance simplifies to

$$\begin{aligned} &= \langle\langle \left(L_0(\tau) - \langle L_0(\tau) \rangle_{P(\gamma)} \right)^2 \rangle\rangle_{P(\gamma)} \\ &\quad + D \langle\langle L_1(\tau)^2 \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \\ &\quad + 2 \left(\langle\langle L_0(\tau) \rangle\rangle_{P(\gamma)} \langle\langle L_2(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) - \langle L_0(\tau) \rangle_{P(\gamma)} \langle\langle L_2(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \right) \dots \end{aligned} \quad (7.5)$$

For the covariance the same considerations apply and we find

$$\begin{aligned} C(\tau, \tau') &= \langle\langle \left(L(\tau) - \langle\langle L(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \right) \left(L(\tau') - \langle\langle L(\tau') \rangle\rangle_{R^N(L,\tau'|\gamma)} P(\gamma) \right) \rangle\rangle_{R^N(L,\tau;L,\tau'|\gamma)} P(\gamma) \\ &= \langle\langle \left(L_0(\tau) - \langle L_0(\tau) \rangle_{P(\gamma)} \right) \left(L_0(\tau') - \langle L_0(\tau') \rangle_{P(\gamma)} \right) \rangle\rangle_{P(\gamma)} \\ &\quad + D \langle\langle L_1(\tau) L_1(\tau') \rangle\rangle_{R^N(L,\tau;L,\tau'|\gamma)} P(\gamma) \\ &\quad + D \left(\langle\langle L_0(\tau) \rangle\rangle_{P(\gamma)} \langle\langle L_2(\tau') \rangle\rangle_{R^N(L,\tau'|\gamma)} P(\gamma) - \langle L_0(\tau) \rangle_{P(\gamma)} \langle\langle L_2(\tau') \rangle\rangle_{R^N(L,\tau'|\gamma)} P(\gamma) \right) \\ &\quad + D \left(\langle\langle L_0(\tau') \rangle\rangle_{P(\gamma)} \langle\langle L_2(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) - \langle L_0(\tau') \rangle_{P(\gamma)} \langle\langle L_2(\tau) \rangle\rangle_{R^N(L,\tau|\gamma)} P(\gamma) \right) + \dots \end{aligned} \quad (7.6)$$

for the two-time covariance. With $L_0(\tau)$, $\langle L_1^2(\tau) \rangle_{R^N(L,\tau|\gamma)}$, $\langle L_1(\tau) L_1(\tau') \rangle_{R^N(L,\tau|\gamma)}$ and $\langle L_2(\tau) \rangle_{R^N(L,\tau|\gamma)}$ already derived in section 4.6 and stated in Eq. (4.180), Eq. (4.192), Eq. (4.188) and Eq. (4.189), the heterogeneous moments can easily be calculated. We find for the mean

$$\bar{L}(\tau) = \langle\langle \gamma^{\frac{1}{1-\alpha}} \rangle\rangle_{P(\gamma)} ((1-\alpha)\tau)^{\frac{1}{1-\alpha}} + \langle\langle \gamma^{-\frac{1}{1-\alpha}} \rangle\rangle_{P(\gamma)} D \frac{(7\alpha-3)((1-\alpha)\tau)^{\frac{\alpha}{\alpha-1}}}{4(3\alpha-1)}. \quad (7.7)$$

The mean of the homogeneous and of heterogeneous dynamics are thus indistinguishable in their form. The only difference is, that for the heterogeneous case the γ -dependency has to be averaged with respect to the distribution of γ .

Unlike for the mean, the small noise variance not only changes in its parameter values but also in its function. The heterogeneity in the deterministic contribution to the dynamics adds

two extra variance terms

$$\begin{aligned} \sigma_L^2(\tau) &= \left(\langle \gamma^{\frac{2}{1-\alpha}} \rangle_{P(\gamma)} - \langle \gamma^{\frac{1}{1-\alpha}} \rangle_{P(\gamma)}^2 \right) ((1-\alpha)\tau)^{\frac{2}{1-\alpha}} \\ &\quad + D \frac{1-\alpha}{1-3\alpha} \tau \\ &\quad + 2D \left(1 - \langle \gamma^{\frac{1}{1-\alpha}} \rangle_{P(\gamma)} \langle \gamma^{\frac{\alpha}{\alpha-1}-1} \rangle_{P(\gamma)} \right) \frac{(\alpha-1)(7\alpha-3)\tau}{4-12\alpha}, \end{aligned} \quad (7.8)$$

where only the $\mathcal{O}(D)$ term in the second line is already known from the homogeneous case. The new D -independent term in the first line is proportional to the variance of $\gamma^{\frac{1}{1-\alpha}}$ and the squared zero-order mean of the ensemble. The new contribution of $\mathcal{O}(D)$, stated in the third line, is linear in τ and increases with increasing heterogeneity. The combined impact of heterogeneity on mean and variance is shown in Fig. 7.1.

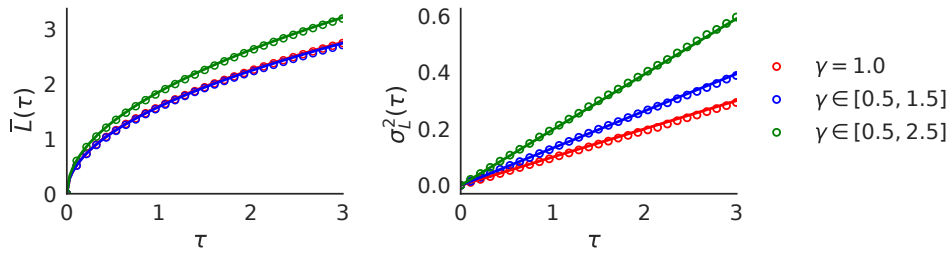


Figure 7.1: Heterogeneity in the sample path realizations increases the variance. Comparison of forward (circles) and small noise (lines) TSA expression for dynamics with $f(L) = -\frac{\gamma}{L}$ and γ randomly drawn from a fixed interval for each sample path realization. Chosen are $\gamma \in [0.5, 1.5]$, $\gamma \in [0.5, 2.5]$ and for comparison $\gamma = 1.0$. Shown are the mean (Left) and variance (Right) for all three cases. The mean changes depending on the mean of the chosen interval. The variance increases with increasing heterogeneity. The statistics of the forward dynamics are based on 20000 trajectories that start at $\widehat{L}_0 = 20$. The analytic expression for the mean and variance are stated in Eq. (7.7) and Eq. (7.8). The diffusion constant is chosen as $D = 0.2$.

The observed covariance however is most telling about the degree of heterogeneity in the dynamics as it allows for visually very different equi-covariance lines which can not be obtained in the homogeneous case by any parameter combination. Evaluating the contribution to the covariance, which is due to the heterogeneity in the force law, we find

$$\begin{aligned} C_L(\tau, \tau') &= \left(\langle \gamma^{\frac{2}{1-\alpha}} \rangle - \langle \gamma^{\frac{1}{1-\alpha}} \rangle^2 \right) (1-\alpha)^{\frac{2}{1-\alpha}} (\tau\tau')^{\frac{1}{1-\alpha}} \\ &\quad + C_L^{\text{hom}}(\tau, \tau') \\ &\quad + D \left(1 - \langle \gamma^{\frac{1}{1-\alpha}} \rangle_{P(\gamma)} \langle \gamma^{\frac{\alpha}{\alpha-1}-1} \rangle_{P(\gamma)} \right) \frac{(\alpha-1)(7\alpha-3)}{4-12\alpha} \left(\tau^{\frac{1}{1-\alpha}} \tau'^{\frac{\alpha}{\alpha-1}} + \tau'^{\frac{1}{1-\alpha}} \tau^{\frac{\alpha}{\alpha-1}} \right), \end{aligned} \quad (7.9)$$

where $C_L^{\text{hom}}(\tau, \tau')$ is defined in Eq. (4.189) and for clarity marked with the flag “hom”. The observed effects conceptually generalize to arbitrary force laws. The broadening effect on the two-time covariance, i.e. the slower decay of the covariance in reverse time, is shown in Fig. 7.2.

7.2 Generalization to multiplicative noise

In the previous sections we have discussed the stochastic dynamics of processes with a fixed diffusion constant D . In this section we generalize our results to also include multiplicative noise.

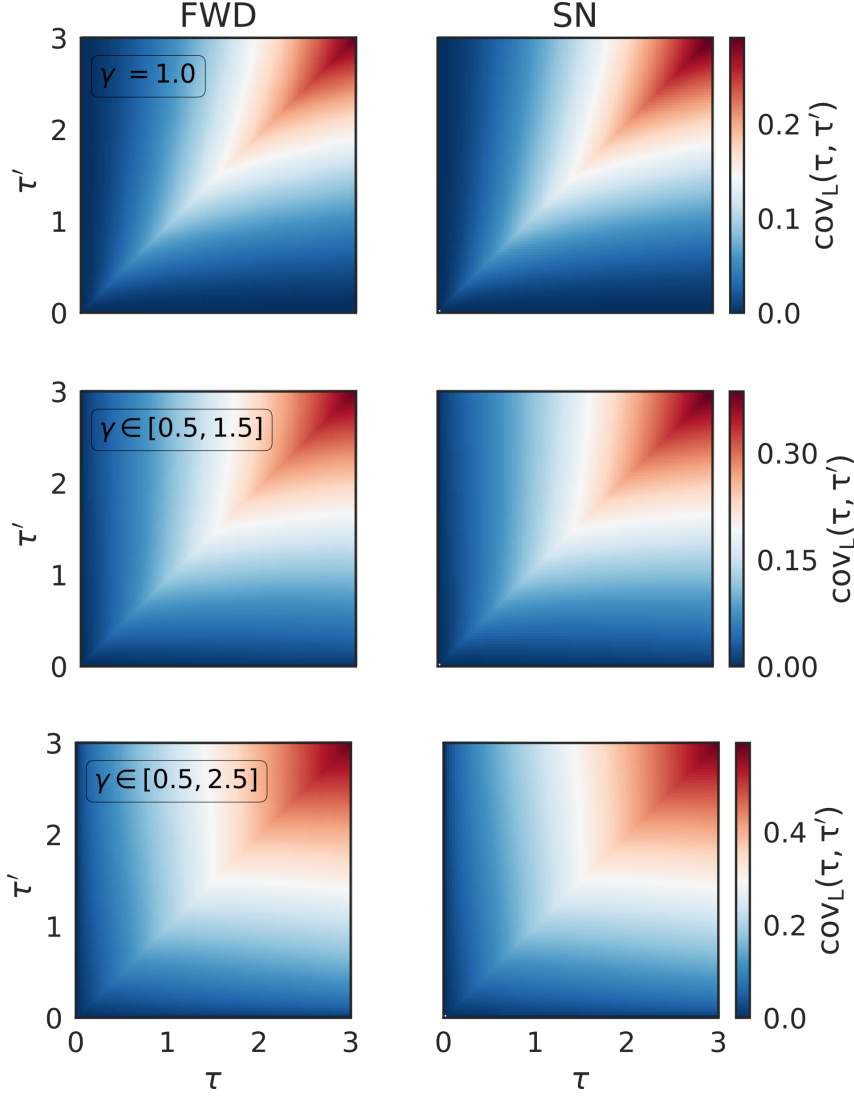


Figure 7.2: Heterogeneity in the sample path realizations leaves a characteristic fingerprint in the two-time covariance due to a change in the lines of equal covariance. Comparison of forward (FWD) and small noise (SN) TSA expression for dynamics with $f(L) = -\frac{\gamma}{L}$ and γ randomly drawn from a fixed interval for each sample path realization. Chosen are $\gamma \in [0.5, 1.5]$, $\gamma \in [0.5, 2.5]$ and for comparison $\gamma = 1.0$. Shown are the two-time covariance for all three cases. The statistics of the forward dynamics are based on 20000 trajectories that start at $\hat{L}_0 = 20$. The analytic expression for the small noise covariance is stated in Eq. (7.9). The used diffusion constant is $D = 0.2$.

For such a process the diffusion “constant” $D(\hat{L})$ is a twice differentiable function dependent on the current position \hat{L} . We start our discussion with an adaptation of both the forward

$$\frac{\partial}{\partial t} P(\hat{L}, t | \hat{L}_0, t_0) = -\frac{\partial}{\partial \hat{L}} f(\hat{L}) P(\hat{L}, t | \hat{L}_0, t_0) + \frac{1}{2} \frac{\partial^2}{\partial \hat{L}^2} D(\hat{L}) P(\hat{L}, t | \hat{L}_0, t_0) \quad (7.10)$$

and backward Fokker-Planck equation

$$\frac{\partial}{\partial t} P(\hat{L}_f, t_f | \hat{L}, t) = -f(\hat{L}) \frac{\partial}{\partial \hat{L}} P(\hat{L}_f, t_f | \hat{L}, t) - \frac{1}{2} D(\hat{L}) \frac{\partial^2}{\partial \hat{L}^2} P(\hat{L}_f, t_f | \hat{L}, t) \quad (7.11)$$

to multiplicative noise. We here chose the Ito interpretation to define these two Fokker-Planck equations. For the forward Fokker-Planck equation this implies that the diffusion “constant”

$D(\widehat{L})$ is taken inside of the second derivative, for the backward Fokker-Planck equation it is evaluated outside of the derivative. For a thorough discussion which interpretation should be chosen under different experimental conditions and how to transform between different interpretations we refer to van Kampen¹⁶⁹ and Sokolov¹⁵⁸. Throughout this text we use the Ito-interpretation where the stochastic forcing is always evaluated at the onset of each time increment⁵⁵. The following sections are organized in accordance with the derivation without multiplicative noise. We start with the derivation of the theory, discuss analytical examples and generalize the small noise theory to include both multiplicative noise and heterogeneity.

The reverse time Fokker-Planck equation with multiplicative noise

The time reversal of stochastic dynamics with multiplicative noise can be derived following the exact same derivation scheme as laid out in section 4.2. We here only highlight how including $D(\widehat{L})$ modifies some of the terms. We start with the evolution equation of the joint probability distribution $P(\widehat{L}, t, \widehat{L}_f, t_f)$ Eq. (4.13) adapted to multiplicative noise

$$-\frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} = P(\widehat{L}, t) \left[f(\widehat{L}) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) + \frac{1}{2} D(\widehat{L}) \frac{\partial^2}{\partial \widehat{L}^2} P(\widehat{L}_f, t_f | \widehat{L}, t) \right] + P(\widehat{L}_f, t_f | \widehat{L}, t) \left[\frac{\partial}{\partial \widehat{L}} f(\widehat{L}) P(\widehat{L}, t) - \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) P(\widehat{L}, t) \right]. \quad (7.12)$$

Evaluating products, we observe that $P(\widehat{L}, t)$ is linked to $D(\widehat{L})$ in the diffusion term. The derivation of the reverse time Fokker-Planck equation can therefore be carried on analog to section 4.2 with the original diffusion constant D set to one and $P(\widehat{L}, t)$ in all formerly D dependent terms replaced by $D(\widehat{L})P(\widehat{L}, t)$. Using this replacement we can directly rewrite Eq. (4.19) as

$$-\frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} = \frac{\partial}{\partial \widehat{L}} \left[f(\widehat{L}) P(\widehat{L}, t, \widehat{L}_f, t_f) - P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} D(\widehat{L}) P(\widehat{L}, t) \right] + \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) P(\widehat{L}, t, \widehat{L}_f, t_f). \quad (7.13)$$

Note that the diffusion “constant” $D(\widehat{L})$ here surfaces inside the second derivative with respect to \widehat{L} , as the respective joint probability distribution $P(\widehat{L}, t, \widehat{L}_f, t_f)$ is obtained from the product of $P(\widehat{L}_f, t_f | \widehat{L}, t)$ and $P(\widehat{L}, t)$. Using this product again to rewrite the guiding term $-P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} D(\widehat{L}) P(\widehat{L}, t)$, we arrive at

$$-\frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} = \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) - \frac{1}{P(\widehat{L}, t)} \frac{\partial D(\widehat{L}) P(\widehat{L}, t)}{\partial \widehat{L}} \right) P(\widehat{L}, t, \widehat{L}_f, t_f) \right] + \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) P(\widehat{L}, t, \widehat{L}_f, t_f). \quad (7.14)$$

Extending the guiding term both in numerator and denominator with $D(\widehat{L})$ we arrive at a guiding force of the familiar logarithmic form $-D(\widehat{L}) \frac{\partial}{\partial \widehat{L}} \log \left(D(\widehat{L}) P(\widehat{L}, t) \right)$. We observe that $D(\widehat{L})$ is now additionally included inside of the logarithm.

From here on all further derivation steps do not involve $D(\widehat{L})$ and we can follow the trodden path of section 4.2 to arrive at the final reverse time Fokker-Planck equation with multiplicative

noise

$$\begin{aligned} \frac{\partial}{\partial \tau} P(L, \tau | L_f, \tau_f, L_0, \tau_0) &= \\ &= \frac{\partial}{\partial L} \left[\left(f(L) - D(L) \frac{\partial}{\partial L} \log \left(D(L) P^{\text{fw}}(L, T - \tau | L_f, 0) \right) \right) P(L, \tau | L_f, \tau_f, L_0, \tau_0) \right] \\ &\quad + \frac{1}{2} \frac{\partial^2}{\partial L^2} D(L) P(L, \tau | L_f, \tau_f, L_0, \tau_0) , \end{aligned} \quad (7.15)$$

where we again used $T = t_f - t_0$ and $\tau = t_f - t$ with τ_0 set to zero. We note that, although the direction of time switched from t to reverse time τ , the resulting Fokker-Planck equation must be interpreted in the Ito sense.

The TSA dynamics with multiplicative noise

Starting from the reverse time Fokker-Planck equation Eq. (4.10), the derivation of the full TSA ensemble dynamics in section 4.2 involved three crucial steps. First, averaging sub-ensembles – defined by the reverse time Fokker-Planck adapted to target states analog to section 7.2 – with respect to the hitting time distribution. Second, taking the time derivative with respect to t over exactly this ensemble. Third, averaging with respect to the forward initial conditions. Clearly all these operations do not affect the multiplicative noise term $D(\hat{L})$. Analog to the extension of the reverse time Fokker-Planck equation to multiplicative noise in section 7.2, we can therefore keep $D(\hat{L})$ directly attached to $P(\hat{L}, t | \hat{L}_0, t_0)$. The term $D(\hat{L})$ thus shows up on the very same positions as in the reverse time Fokker-Planck equation, that is inside of the second derivative of the diffusion term, and both inside and outside of the logarithm of the guiding force. Our first intermediate result in the derivation of the TSA dynamics in section 4.2 Eq. (4.65) can therefore be generalized to

$$\begin{aligned} \frac{\partial}{\partial t} R_{\hat{L}_{\text{ts}}}(\hat{L}, t; t_f) &= \rho_{\hat{L}_{\text{ts}}}(t_f | \hat{L}, t) P^{\text{in}}(\hat{L}) \\ &\quad - \frac{\partial}{\partial \hat{L}} \left[f(\hat{L}) - D(\hat{L}) \frac{\partial}{\partial \hat{L}} \log \left(D(\hat{L}) \int_{\hat{L}_{\text{ts}}}^{\infty} d\hat{L}_0 P^{\text{in}}(\hat{L}_0) \int_{-\infty}^t dt_0 P(\hat{L}, t | \hat{L}_0, t_0) \right) \right] R_{\hat{L}_{\text{ts}}}(\hat{L}, t; t_f) \\ &\quad - \frac{1}{2} \frac{\partial^2}{\partial \hat{L}^2} D(\hat{L}) R_{\hat{L}_{\text{ts}}}(\hat{L}, t; t_f) . \end{aligned} \quad (7.16)$$

As for the above derivation of the TSA dynamics with constant D , we next seek a more accessible expression for the double integral over the solution of the forward FPE with absorbing boundary conditions, inside of the logarithm in Eq. (7.16). Already in section 4.2 Eq. (4.66), we termed this expression times a normalization factor λ , $Q(\hat{L})$ and showed that it follows the forward Fokker-Planck equation with the time evolution replaced by the source term $-\lambda P^{\text{in}}(\hat{L})$ Eq. (4.71). Analog to the above consideration that neither the integral over t_0 or \hat{L}_0 act on $D(\hat{L})$, the defining ODE for $Q(\hat{L})$, given multiplicative noise, then reads

$$\begin{aligned} -\lambda P^{\text{in}}(\hat{L}) &= -\frac{\partial}{\partial \hat{L}} \left(f(\hat{L}) Q(\hat{L}) \right) + \frac{1}{2} \frac{\partial^2}{\partial \hat{L}^2} D(\hat{L}) Q(\hat{L}) \\ Q(\hat{L}_{\text{ts}}) &= 0. \end{aligned} \quad (7.17)$$

The normalization factor $\lambda = \frac{1}{2} \frac{\partial}{\partial \hat{L}} D(\hat{L}_{\text{ts}}) Q(\hat{L}_{\text{ts}})$ is obtained by integration over the full interval $[\hat{L}_{\text{ts}}, \infty]$ with $\int_{\hat{L}_{\text{ts}}}^{\infty} P^{\text{in}}(\hat{L}') d\hat{L}' = 1$. Solved in a different context, the solution of Eq. (7.17) is known¹⁸². Integrating over the interval \hat{L}_{ts} to \hat{L} analogous to Eq. (4.74) one arrives at

$$-f(\hat{L}) Q(\hat{L}) + \frac{1}{2} \frac{\partial}{\partial \hat{L}} D(\hat{L}) Q(\hat{L}) = \lambda \left(1 - \int_{\hat{L}_{\text{ts}}}^{\hat{L}} P^{\text{in}}(\hat{L}') d\hat{L}' \right) . \quad (7.18)$$

Combining the terms on the left hand side into one expression then yields

$$\frac{1}{2} e^{\int^{\widehat{L}} \frac{2f(\widehat{L}')}{D(\widehat{L}')} d\widehat{L}'} \frac{\partial}{\partial \widehat{L}} \left(D(\widehat{L}) Q(\widehat{L}) e^{-\int^{\widehat{L}} \frac{2f(\widehat{L}')}{D(\widehat{L}')} d\widehat{L}'} \right) = \lambda \left(1 - \int_{\widehat{L}_{\text{ts}}}^{\widehat{L}} P^{\text{in}}(\widehat{L}') d\widehat{L}' \right). \quad (7.19)$$

Again integrating from \widehat{L}_{ts} to \widehat{L} and solving for $Q(\widehat{L})$ yields the final seeked expression

$$Q(\widehat{L}) = \frac{2\lambda}{D(\widehat{L})} e^{\int^{\widehat{L}} \frac{2f(\widehat{L}')}{D(\widehat{L}')} d\widehat{L}'} \int_{\widehat{L}_{\text{ts}}}^{\widehat{L}} d\widehat{L}' e^{-\int^{\widehat{L}'} \frac{2f(\widehat{L}'')}{D(\widehat{L}'')} d\widehat{L}''} \left(1 - \int_{\widehat{L}_{\text{ts}}}^{\widehat{L}'} P^{\text{in}}(\widehat{L}'') d\widehat{L}'' \right). \quad (7.20)$$

The free energy force then reads

$$D(\widehat{L}) \frac{\partial}{\partial \widehat{L}} \log D(\widehat{L}) Q(\widehat{L}) = 2f(\widehat{L}) + D(\widehat{L}) \log \left(\int_{\widehat{L}_{\text{ts}}}^{\widehat{L}} d\widehat{L}' e^{-\int^{\widehat{L}'} \frac{2f(\widehat{L}'')}{D(\widehat{L}'')} d\widehat{L}''} \left(1 - \int_{\widehat{L}_{\text{ts}}}^{\widehat{L}'} P^{\text{in}}(\widehat{L}'') d\widehat{L}'' \right) \right), \quad (7.21)$$

where, importantly, the $D(\widehat{L})$ terms inside the logarithm have canceled.

Substituting this free energy force in Eq. (7.16) directly leads to the reverse time TSA Fokker-Planck equation with multiplicative noise

$$\begin{aligned} \frac{\partial}{\partial \tau} R_{L_{\text{ts}}}(L, \tau; \tau_0) &= -P^{\text{in}}(L) \rho_{L_{\text{ts}}}(\tau|L) \\ &- \frac{\partial}{\partial L} \left(\left[f(L) + D(L) \frac{\partial}{\partial L} \log \left(\int_{L_{\text{ts}}}^L dL' e^{-\int^{L'} \frac{2f(L'')}{D(L'')} dL''} \left(1 - \int_{L_{\text{ts}}}^{L'} P^{\text{in}}(L'') dL'' \right) \right] \right) R_{L_{\text{ts}}}(L, \tau; \tau_0) \\ &+ \frac{1}{2} \frac{\partial^2}{\partial L^2} D(L) R_{L_{\text{ts}}}(L, \tau; \tau_0) \end{aligned} \quad (7.22)$$

rewritten in reverse time. We observe that the resulting equation is identical to Eq. (4.45) when D is replaced by $D(L)$. Note, that the resulting ensemble is not normalized but decays with a rate proportional to the hitting time $\rho_{L_{\text{ts}}}(\tau|L)$, in line with the construction idea of TSA ensembles. The extension of the normalized ensemble for multiplicative, which is useful to e.g. determine moments from simulations, follows analogously using the steps laid out in section 4.3.

In the next step we associate the reverse time TSA Fokker-Planck equation with a Langevin equation. With the diffusion “constant” $D(L)$ inside of the second derivative and the forces inside of the first derivative, Eq. (7.22) is clearly of Ito-Form. The corresponding Langevin equation thus reads

$$\begin{aligned} dL(\tau) &= \left(f(L) + D(L) \frac{\partial}{\partial L} \log \left(\int_{L_{\text{ts}}}^L dL' e^{-\int^{L'} \frac{2f(L'')}{D(L'')} dL''} \left(1 - \int_{L_{\text{ts}}}^{L'} P^{\text{in}}(L'') dL'' \right) \right) \right) d\tau \\ &+ \sqrt{D(L)} dW_\tau \end{aligned} \quad (7.23)$$

The killing measure is unaffected by the change to multiplicative noise and thus stays the same as in Eq. (4.53).

Exactly solvable reverse time TSA dynamics with multiplicative noise

The exact solution of the reverse time TSA Fokker-Planck equation with multiplicative noise is obtainable only for a few cases. We her exemplary study the case

$$d\widehat{L}(t) = -\gamma dt + \sqrt{D\widehat{L}} dW_t. \quad (7.24)$$

where the SDE is interpreted in Ito sense. For simplicity we choose δ -initial conditions at $\widehat{L}_0 = L_f$. The TSA reverse time SDE to be solved then reads

$$dL(\tau) = \begin{cases} (\gamma + D) d\tau + \sqrt{DL} dW_\tau & \text{for } L < L_f \\ -\gamma d\tau + \sqrt{DL} dW_\tau & \text{for } L > L_f, \end{cases} \quad (7.25)$$

and is directly obtained from the definition of the reverse time TSA SDE with multiplicative noise Eq. (7.23).

To construct a solution for Eq. (7.25) we can either follow the receipts described in section 4.4, or exploit that Eq. (7.24) is equivalent to a Bessel process Eq. (4.95), as can be seen using Ito's lemma Eq. (4.4). Under this lemma one can show that a SDE of the form of the Bessel-process

$$dx(t) = -\frac{\gamma}{x} dt + \sqrt{D} dW_t \quad (7.26)$$

transforms to

$$dy(t) = (-2\gamma + D) dt + \sqrt{4Dy} dW_t \quad (7.27)$$

using a variable transformation of the form $y = x^2$. Substituting $D = \frac{\widetilde{D}}{4}$ and $\gamma = \frac{\widetilde{\gamma}}{2} + \frac{\widetilde{D}}{8}$, we arrive at Eq. (7.24) with the “ \sim ” removed from display for notational unity. Knowing the correct variable transformation and parameter substitutions we can directly transform the solution of the TSA Bessel process as stated in Eq. (4.115) into the solution of Eq. (7.25) and find

$$R_0(L, \tau; L_f) = \begin{cases} \frac{4^{\frac{\gamma+D}{D}} (D\tau)^{-\frac{2(\gamma+D)}{D}} e^{-\frac{2L}{D\tau}} L^{\frac{2\gamma}{D}+1}}{\Gamma\left(\frac{2(D+\gamma)}{D}\right)} & L \leq L_f \\ \frac{4^{\frac{\gamma+D}{D}} (D\tau)^{-\frac{2(\gamma+D)}{D}} e^{-\frac{2L_f}{D\tau}} L_f^{\frac{2\gamma}{D}+1}}{\Gamma\left(\frac{2(D+\gamma)}{D}\right)} & L \geq L_f. \end{cases} \quad (7.28)$$

Note, that we additionally transformed L_f using the transformation from x to y . We validate Eq. (7.28) by substitution into the to Eq. (7.25) corresponding reverse time Fokker-Planck for $L > L_f$ and $L < L_f$ respectively. To check whether Eq. (7.28) decays with a rate equal to the forward hitting time we calculated it from the cumulative of Eq. (7.28) and compare it to the hitting time distribution

$$\rho_0(\tau|L_f) = \frac{L_f 4^{\frac{\gamma}{D}} e^{-\frac{2L_f}{D\tau}} \left(\frac{D\tau}{L_f}\right)^{-\frac{2\gamma}{D}}}{\gamma \tau^2 \Gamma\left(\frac{2\gamma}{D}\right)} \quad (7.29)$$

of the forward Bessel process. The latter is obtained from Eq. (4.96) using the same transformations and substitutions as for Eq. (7.28), and the definition of the hitting time distribution Eq. (4.26).

To confirm the validity of both the reverse time TSA-SDE with multiplicative noise and the solution of above's TSA-dynamics, we calculate the mean

$$\bar{L}(\tau) = \frac{D^2 \tau^2 \left(\Gamma\left(\frac{2\gamma}{D} + 3\right) - \Gamma\left(\frac{2\gamma}{D} + 3, \frac{2L_f}{D\tau}\right) \right) + L_f 2^{\frac{2\gamma}{D}+1} e^{-\frac{2L_f}{D\tau}} (D\tau + 2L_f) \left(\frac{D\tau}{L_f}\right)^{-\frac{2\gamma}{D}}}{2\tau(2\gamma + D) \left(\Gamma\left(\frac{2\gamma}{D} + 1\right) - \Gamma\left(\frac{2\gamma}{D} + 1, \frac{2L_f}{D\tau}\right) \right)} \quad (7.30)$$

and variance

$$\sigma_L^2(\tau) = \frac{D^3\tau^3 \left(\Gamma\left(\frac{2\gamma}{D} + 4\right) - \Gamma\left(\frac{2\gamma}{D} + 4, \frac{2L_f}{D\tau}\right) \right) + L_f 4^{\frac{\gamma+D}{D}} e^{-\frac{2L_f}{D\tau}} \left(D^2\tau^2 + 2D\tau L_f + 2L_f^2 \right) \left(\frac{D\tau}{L_f} \right)^{-\frac{2\gamma}{D}}}{4\tau(2\gamma + D) \left(\Gamma\left(\frac{2\gamma}{D} + 1\right) - \Gamma\left(\frac{2\gamma}{D} + 1, \frac{2L_f}{D\tau}\right) \right)} - \bar{L}(\tau)^2 \quad (7.31)$$

with respect to Eq. (7.28) after normalization. In Fig. 7.3, we compare this mean and variance to the ones of the forward dynamics after target state alignment, and a numerical evaluation of the Fokker-Planck associated with Eq. (7.25). Given sufficiently many forward trajectories for

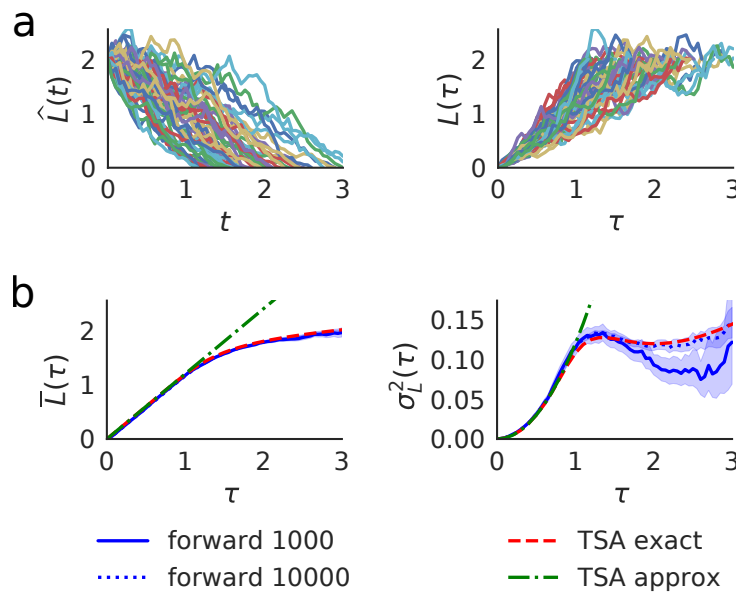


Figure 7.3: The TSA FPE and SDE also with multiplicative noise exactly describe the target state aligned ensemble. (a): Exemplary forward and reverse time trajectories. **(b):** Comparison of the forward (blue), analytic (red) and approximate reverse time dynamics (green) for $f(L) = -\gamma$ and $D(L) = DL$. Shown are the mean and variance of all three cases. The statistics of the forward dynamics are shown twice based on 1000 and 10000 trajectories to demonstrate sampling errors due to finite numbers. 95% bootstrap confidence intervals are shown for the cases involving sampling. Results were obtained with $\gamma = 1, D = 0.2$ and $\hat{L}_0 = L_f = 2$.

convergence in the mean and variance we find perfect agreement.

In general, care must be taken when using Ito's lemma to construct equivalent reverse time TSA dynamics from known expressions. First, the hitting time distribution of the forward problem must be recalculated with respect to the transformed forward transition probability $P^{\text{fw}}(\hat{L}, t)$. And second, using arbitrary initial distributions $P^{\text{in}}(L)$, it is important to clarify in which variable space the initial distribution is defined and whether we need to transform it to the new coordinates.

The TSA-dynamics close to the target state for multiplicative noise

The dynamics both spatially and temporally close to the target state can be deduced from Eq. (7.23) with the same arguments as given in section 4.5. Assuming well separated initial and final conditions, the dependency on the initial conditions in Eq. 7.23 can be neglected. The initial condition dependent term $H(L)$ defined in Eq. (4.51) can be set to one and the killing

measure $k(L, \tau)$ to zero. We arrive at

$$dL(\tau) = \left(f(L) + D(L) \frac{\partial}{\partial L} \log \left(\int_{L_{ts}}^L dL' e^{-\int^{L'} \frac{2f(L'')}{D(L'')} dL''} \right) \right) d\tau + \sqrt{D(L)} dW_\tau, \quad (7.32)$$

which is formally equivalent to Eq. (4.118) with all diffusion constants D substituted by the L -dependent term $D(L)$. Analog to section 4.5, we call

$$f^{\mathcal{F}}(L) = D(L) \frac{\partial}{\partial L} \log \left(\int_{L_{ts}}^L dL' e^{-\int^{L'} \frac{2f(L'')}{D(L'')} dL''} \right) \quad (7.33)$$

a free energy force. To form a better understanding of how these TSA-dynamics differ from their forward counterpart, we study the exemplary case of power law like dynamics both in the force

$$f(L) = -\gamma L^\alpha \quad (7.34)$$

and diffusion term

$$D(L) = DL^\beta \quad (7.35)$$

with $\gamma, D > 0$ and $\alpha, \beta \in \mathbb{R}$. We observe, that, when evaluating the fraction $\frac{2f(L)}{D(L)} = \frac{2L^{\alpha-\beta}}{D}$ in the free energy force of Eq. (7.32), we end up with a formally equivalent ($\alpha \rightarrow \alpha - \beta$) logarithmic term as the one evaluated in section 4.5 for constant D . The free energy force with respect to power law like multiplicative noise can therefore be deduced from the constant D case Eq. (4.129), which, for distinguishability, we here call $f_0^{\mathcal{F}}(L, \alpha)$. The mapping is then defined as

$$f^{\mathcal{F}}(L) = L^\beta f_0^{\mathcal{F}}(L, \alpha - \beta). \quad (7.36)$$

The term L^β corrects for the now explicitly L dependent diffusion constant $D(L)$ in front of the logarithm of the guiding force. Stated explicitly we find

$$f^{\mathcal{F}}(L) = L^\beta \frac{D(\alpha - \beta + 1) \left(-\frac{2\gamma}{D(\alpha - \beta + 1)} \right)^{\frac{1}{\alpha - \beta + 1}} e^{\frac{2\gamma L^{\alpha - \beta + 1}}{D(\alpha - \beta + 1)}}}{\Theta(\alpha - \beta + 1) \Gamma\left(\frac{1}{\alpha - \beta + 1}\right) - \Gamma\left(\frac{1}{\alpha - \beta + 1}, -\frac{2L^{\alpha - \beta + 1}\gamma}{D(\alpha - \beta + 1)}\right)} \quad \alpha - \beta \neq -1. \quad (7.37)$$

The case $\alpha - \beta = 1$ can be obtained using the same mapping as above, or simply by evaluating Eq. (7.32). We find

$$dL(\tau) = (\gamma + D)L^\alpha d\tau + \sqrt{DL^{\alpha+1}} dW_\tau \quad \alpha - \beta = -1. \quad (7.38)$$

for the corresponding TSA SDE.

Concluding, the simplicity of the transformation, from a constant D to a multiplicative noise term $D(L)$ for power law like dynamics, allows us to also generalize the ensemble likelihood inference scheme discussed in section 5.2. The diffusion constant in the Gaussians is simply replaced by its multiplicative version and the used versions of the free energy forces are simply multiplied by L^β with all α replaced by $\alpha - \beta$ according to the above substitution. We further note that Eq. (7.37) also allows us to calculate the exact analytical free energy force for one-component chemical Langevin equations ($\alpha = \beta$)⁵⁸, which are often used when the noise is proportional to the reaction under study. For more complicated reaction dynamics the free energy force can be obtained numerically or in some cases, such as simple production and degradation of one species, also analytically.

Comparing the completion dynamics close to the target state and extended to multiplicative noise to the full TSA-dynamics

To study the quality of the TSA approximation for multiplicative noise we refer to the most simple case, i.e. Eq. (7.38) with $\alpha < 1$. For these values of α we can again exploit Ito's lemma Eq. (4.4) and transform the reverse time TSA version of the Bessel process with well separated initial and final conditions Eq. (4.121) into Eq. (7.38). We use the substitution $L_{\text{new}} = L_{\text{old}}^{\frac{2}{1-\alpha}}$, and replace $D_{\text{old}} = \frac{(1-\alpha)^2}{4} D_{\text{new}}$ and $\gamma_{\text{old}} = \frac{1-\alpha}{2} \gamma_{\text{new}} + \frac{D_{\text{new}}}{8} (1-\alpha^2)$. Note that for $\alpha > 1$ the used target state at L_{ts} would change its position from zero to infinity under the given transformation, leaving results unintuitive to interpret.

In the next step we use this transformation starting from Eq. (4.121) to solve Eq. (7.38) for $\alpha < 1$ and with target state $L_{\text{ts}} = 0$. For the density we find

$$R(L, \tau) = - \frac{(\alpha - 1) 2^{\frac{2\gamma - \alpha D + 2D}{D - \alpha D}} ((\alpha - 1)^2 D \tau)^{\frac{2\gamma + D}{(\alpha - 1)D} - 1} L^{-\alpha + \frac{2\gamma}{D} + 1} e^{-\frac{2L^{1-\alpha}}{(\alpha - 1)^2 D \tau}}}{\Gamma\left(\frac{-\alpha D + 2D + 2\gamma}{D - D\alpha}\right)} \quad \alpha < 1, \quad (7.39)$$

leaving us with well defined mean

$$\bar{L}(\tau) = \frac{2^{\frac{1}{\alpha-1}} (1 - \alpha)^{-\frac{2}{\alpha-1}} \Gamma\left(\frac{-\alpha D + 3D + 2\gamma}{D - D\alpha}\right)}{\Gamma\left(\frac{-\alpha D + 2D + 2\gamma}{D - D\alpha}\right)} (D\tau)^{\frac{1}{1-\alpha}} \quad \alpha < 1, \quad (7.40)$$

and variance

$$\sigma_L^2(\tau) = \frac{4^{\frac{1}{\alpha-1}} (1 - \alpha)^{-\frac{4}{\alpha-1}} \left(\Gamma\left(\frac{-\alpha D + 2D + 2\gamma}{D - D\alpha}\right) \Gamma\left(\frac{-\alpha D + 4D + 2\gamma}{D - D\alpha}\right) - \Gamma\left(\frac{-\alpha D + 3D + 2\gamma}{D - D\alpha}\right)^2 \right)}{\Gamma\left(\frac{-\alpha D + 2D + 2\gamma}{D - D\alpha}\right)^2} (D\tau)^{-\frac{2}{\alpha-1}} \quad (7.41)$$

for $\alpha < 1$. Despite the complicated prefactors, the dependency on τ is simple. In Fig. 7.3 we compare the case $\alpha = 0$ to the aligned and time reverse forward dynamics and the exact solution for mean and variance stated in Eq. (7.40) and Eq. (7.41).

The small L expansion for multiplicative noise TSA dynamics close to the target state

With the substitution formula for free energy forces Eq. (7.36) valid for power law noise and forces, the generalization of the small L reverse time TSA dynamics Eq. (4.135) and Eq. (4.136) to multiplicative noise is straightforward. Starting from the respective approximate small L free energy forces Eq. (4.143) and Eq. (4.147) we find

$$dL(\tau) = \left(DL^{\beta-1} - \gamma \frac{\alpha - \beta}{\alpha - \beta + 2} L^\alpha \right) d\tau + \sqrt{D} L^\beta dW_\tau \quad \text{for} \quad \alpha \geq \beta, \quad (7.42)$$

and

$$dL(\tau) = \left(\gamma L^\alpha - D(\alpha - \beta) L^{\beta-1} \right) d\tau + \sqrt{D} L^\beta dW_\tau \quad \text{for} \quad \alpha < \beta. \quad (7.43)$$

As for the case of constant D , the dynamics of Eq. (7.42) with $\alpha \geq \beta$ are dominated by the noise dependent term $DL^{\beta-1}$, which for constant D simplifies to the from above well known $\frac{D}{L}$. The exact form of the force $f(L)$ and the assumed exponent α are thus irrelevant. The

dynamics close to the target state behave as if only noisy fluctuations lead to absorption at the target. Formally neglecting the next order term, i.e. setting $\gamma = 0$ in Eq. (7.42), we end up with a special case of the dynamics stated in Eq. (7.38) ($\gamma = 0, \alpha = \beta - 1$). Noise driven TSA dynamics are therefore well described by the analytic solution of Eq. (7.38) ($\gamma = 0, \alpha = \beta - 1$). We can use the obtained expressions for mean Eq. (7.40) and variance Eq. (7.41) given $\beta < 2$, and with $\gamma = 0$ and $\alpha = \beta - 1$.

For $\alpha < \beta - 1$, and for $0 < \alpha - \beta < 1$ given sufficiently small noise strength D , the TSA dynamics Eq. (7.43) follow the sign inverted forward force law $f(L)$ plus a strictly positive contribution due to the noise. While in the noise driven case Eq. (7.42), the residual force due to target state alignment is given as $DL^{\beta-1}$, this contribution is modified in the force driven case by a factor $\alpha - \beta$. Below, we calculate the small noise mean, variance and covariance for the force driven TSA dynamics. We then generalize the results to also include heterogeneity in γ .

Moments for the noise driven TSA dynamics with multiplicative noise

For $\alpha \geq \beta$, the dynamics close to the target state are noise dominated, as we read off from Eq. (7.42). From above, or simply by inserting $f(L) = 0$ into the defining equation of the free energy force Eq. (7.33), we know that the respective TSA SDE therefore evaluates to

$$dL(\tau) = DL^{\beta-1}d\tau + \sqrt{DL^\beta} dW_\tau \quad \text{for} \quad \alpha \geq \beta. \quad (7.44)$$

The solution of Eq. (7.44) with $\beta < 2$ is a special case of Eq. (7.38). With $\gamma = 0$ and $\alpha = \beta - 1$, we find for mean

$$\bar{L}(\tau) = \frac{2^{\frac{\beta-1}{\beta-2}}(2-\beta)^{-\frac{2}{\beta-2}}\Gamma\left(-\frac{2}{\beta-2}\right)}{\Gamma\left(\frac{1}{2-\beta}\right)}(D\tau)^{\frac{1}{2-\beta}} \quad (7.45)$$

and variance

$$\sigma_L^2(\tau) = \frac{(2-\beta)^{-\frac{4}{\beta-2}}\left(3 \cdot 4^{\frac{1}{\beta-2}}\Gamma\left(\frac{1}{2-\beta}\right)\Gamma\left(-\frac{3}{\beta-2}\right) - 4^{\frac{\beta-1}{\beta-2}}\Gamma\left(-\frac{2}{\beta-2}\right)^2\right)}{\Gamma\left(\frac{1}{2-\beta}\right)^2}(D\tau)^{\frac{1}{1-\beta/2}}. \quad (7.46)$$

Despite their lengthy prefactors both mean and variance show a very simple dependency on τ . For $\beta = 0$, both terms simplify to the case of constant D , with a square-root dependence of the mean Eq. (4.158) and a linear dependence on τ for the variance Eq. (4.159). Three exemplary cases are shown in Fig. 7.4. For values $\beta > 0$, i.e. with vanishing noise at the target state and an increasing contribution further away, the variance becomes a convex function. For $\beta < 0$ the noise starts to diverge close to the target and the variance becomes concave. The very same trends can be observed for the mean with the turning point from concave to convex at $\beta = 1$.

Moments for the force driven TSA dynamics with multiplicative noise

The analysis of the force driven dynamics close to the target state including multiplicative noise is closely related to the discussion with constant noise D . With the constant noise dynamics simply scaled by L^β all previous conclusions can be generalized. For $\alpha < \beta - 1$ and for $-1 < \alpha - \beta < 0$ under the additional constraint of small D the dynamics close to the target state are therefore force driven. Analog to section 4.6 this motivates to try a small noise expansion around the deterministic solution to obtain analytic expressions for mean, variance and covariance. For the

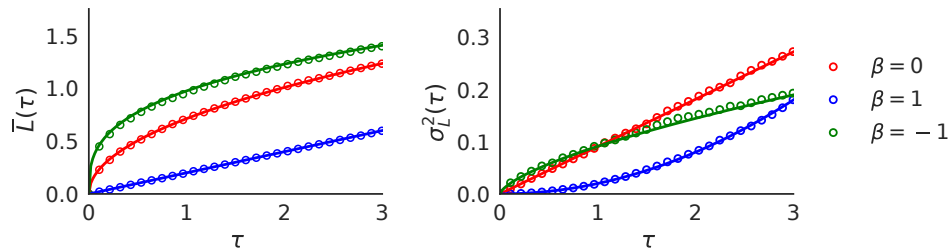


Figure 7.4: TSA ensembles with dominant multiplicative noise are distinguishable in their mean and variance. Comparison of forward (circles) and exact (lines) multiplicative noise driven TSA dynamics with $f(L) = 0$ and $D(L) = DL^\beta$ with $\beta = -1, 0, 1$. Shown are the mean (Left) and variance (Right) for all three cases. The statistics of the forward dynamics are based on 20000 trajectories that start at $\hat{L}_0 = 20$, to effectively get rid of the initial conditions. The exact analytic expression of reverse time TSA mean and variance are taken from Eq. (7.45) and Eq. (7.46). Curves were obtained with $\gamma = 1$ and $D = 0.2$.

following stated results, which we derive in this section below, we only considered the case $\alpha < 0$, analog to the discussion in section 4.6 for constant D . This is partially motivated by consistency, partial by the insight that for $\alpha > 0$ no general expression can be derived. Considering for example the case $\alpha = 3$ – as already discussed in Gardiner⁵⁵ – we do not expect the OU process ($\alpha = 1$) which is foundational to the small noise expansion to be a good approximation for a cubic potential. For a more detailed discussion on the problem of $\alpha > 0$ see Gardiner⁵⁵.

Comparing the small noise moments with and without multiplicative noise, we find the mean

$$\bar{L}(\tau) = ((1 - \alpha)\gamma\tau)^{\frac{1}{1-\alpha}} + \frac{D}{\gamma} \frac{(7\alpha^2 - \alpha(8\beta + 3) + 2\beta(\beta + 1)) ((1 - \alpha)\gamma\tau)^{\frac{\alpha-\beta}{\alpha-1}}}{2(2\alpha - \beta)(3\alpha - \beta - 1)} \quad (7.47)$$

expectedly unchanged in its zero-order deterministic contribution but with a generalized version of its order- D contribution. Keeping in mind that $\alpha < 0$ holds, the slope of the variance

$$\sigma_L^2(\tau) = \frac{D}{\gamma} \frac{((1 - \alpha)\gamma\tau)^{\frac{1-\alpha+\beta}{1-\alpha}}}{1 - 3\alpha + \beta} \quad (7.48)$$

is – as in the noise driven case – determined by β . For $\beta > 0$ the variance becomes convex, for $\beta < 0$ it is concave and for $\beta = 0$ it is linear. Three cases exemplary depicting this behavior are shown in Fig. 7.5. Together with the finding from section 7.2, that the variance in the noise driven case increases with $\tau^{\frac{1}{1-\beta/2}}$, this is an important result for practical applications. Experimental TSA ensembles with a priori unknown homogeneous dynamical laws seem to be characterizable by their variance close to the target state. Given we find a linear variance, we can safely assume $\beta = 0$. For a convex variance β should be positive and negative for a concave variance.

The two-time covariance with multiplicative noise is just a straightforward generalization of the constant D case discussed in section 4.6. We find

$$C(\tau, \tau') = \begin{cases} D \frac{((1-\alpha)\gamma\tau')^{-\frac{\alpha}{\alpha-1}} ((1-\alpha)\gamma\tau)^{\frac{-2\alpha+\beta+1}{1-\alpha}}}{\gamma(1-3\alpha+\beta)} & \text{for } \tau < \tau' \\ D \frac{((1-\alpha)\gamma\tau)^{-\frac{\alpha}{\alpha-1}} ((1-\alpha)\gamma\tau')^{\frac{-2\alpha+\beta+1}{1-\alpha}}}{\gamma(1-3\alpha+\beta)} & \text{for } \tau > \tau' \\ \sigma_L^2(\tau) & \text{for } \tau = \tau' . \end{cases} \quad (7.49)$$

For three different β values the covariance is shown in Fig. 7.6. Importantly, while the covariance

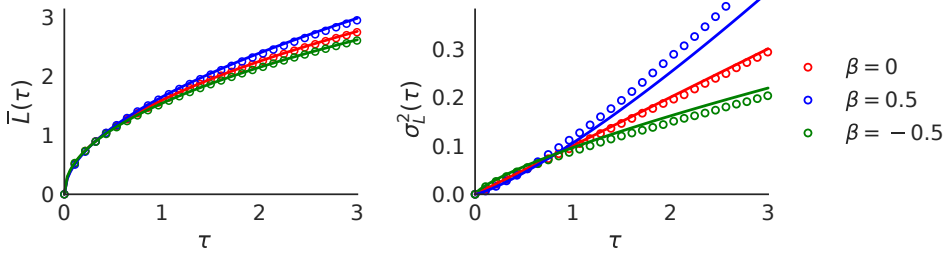


Figure 7.5: Force dominated TSA dynamics with multiplicative noise are distinguishable in their variance. Comparison of forward (circles) and small noise (lines) TSA expression for dynamics with $f(L) = -\frac{\gamma}{L}$ and multiplicative noise $D(L) = DL^\beta$ with $\beta = 0, 0.5, -0.5$. Shown are the mean (Left) and variance (Right) for all three cases. In the mean the influence of β is marginal. In the variance $\beta > 0$ leads to an up sloping, $\beta < 0$ to a down sloping, and $\beta = 0$ to a linear variance. The statistics of the forward dynamics are based on 20000 trajectories that start at $\widehat{L}_0 = 20$. The analytic expressions for the mean and variance are stated in Eq. (7.47) and Eq. (7.48). The used parameters are $\gamma = 1$ and $D = 0.2$.

depends on its scaling parameters D and γ , the two time correlation function

$$\text{Corr}(\tau, \tau') = \begin{cases} \left(\frac{\tau}{\tau'}\right)^{\frac{-3\alpha+\beta+1}{2-2\alpha}} & \text{for } \tau < \tau' \\ \left(\frac{\tau'}{\tau}\right)^{\frac{-3\alpha+\beta+1}{2-2\alpha}} & \text{for } \tau > \tau' \\ 1 & \text{for } \tau = \tau' \end{cases} \quad (7.50)$$

is scale free and solely depends on the power law exponents α and β , and the ratio of τ and τ' . For the here valid parameter regimes $\alpha < 0$ and $\alpha < \beta$ we therefore find the multiplicative noise case to be indistinguishable from the case with constant D i.e. $\beta = 0$, assuming we only consider the correlation function. While at first sight seemingly disadvantageous for an inference scheme, this finding allows to distinguish heterogeneous dynamics from homogeneous dynamics, also in the case of multiplicative noise. With the same argument as for the constant D case homogeneous dynamics with $\alpha < 0$ can always be rejected when the straight equi-correlation lines in a 2d correlation plot are not preserved.

The derivation of small noise moments for force driven TSA dynamics with multiplicative noise

The derivation of the small noise moments given multiplicative noise closely follows the special case with constant D discussed in section 4.6. We again follow the approach discussed in Gardiner⁵⁵.

We start our derivation with a rewritten version of Eq. (7.43)

$$dL = a(L)d\tau + \epsilon^2 b(L)d\tau + \epsilon c(L)dW_\tau, \quad (7.51)$$

with \sqrt{D} substituted by the order parameter ϵ . We then use the already above stated ϵ expansion of $L(\tau)$ and $a(L)$ (Eq. (4.170)-Eq. (4.175)). Following the same expansion scheme for $b(L)$ and $c(L)$ we arrive at an ordered set of SDEs

$$dL_0 = a(L_0)d\tau \quad (7.52)$$

$$dL_1 = a_1(L_1, L_0)d\tau + c(L_0)dW_\tau \quad (7.53)$$

$$dL_2 = a_2(L_2, L_1, L_0)d\tau + b(L_0)d\tau + c(L_1, L_0)dW_\tau, \quad (7.54)$$

which we truncate after the second contributing order to the full solution $L(\tau)$. The zeroth order term is clearly unchanged by adding multiplicative noise. Its solution, the zeroth order contribution to the mean, is thus identical to Eq. (4.180).

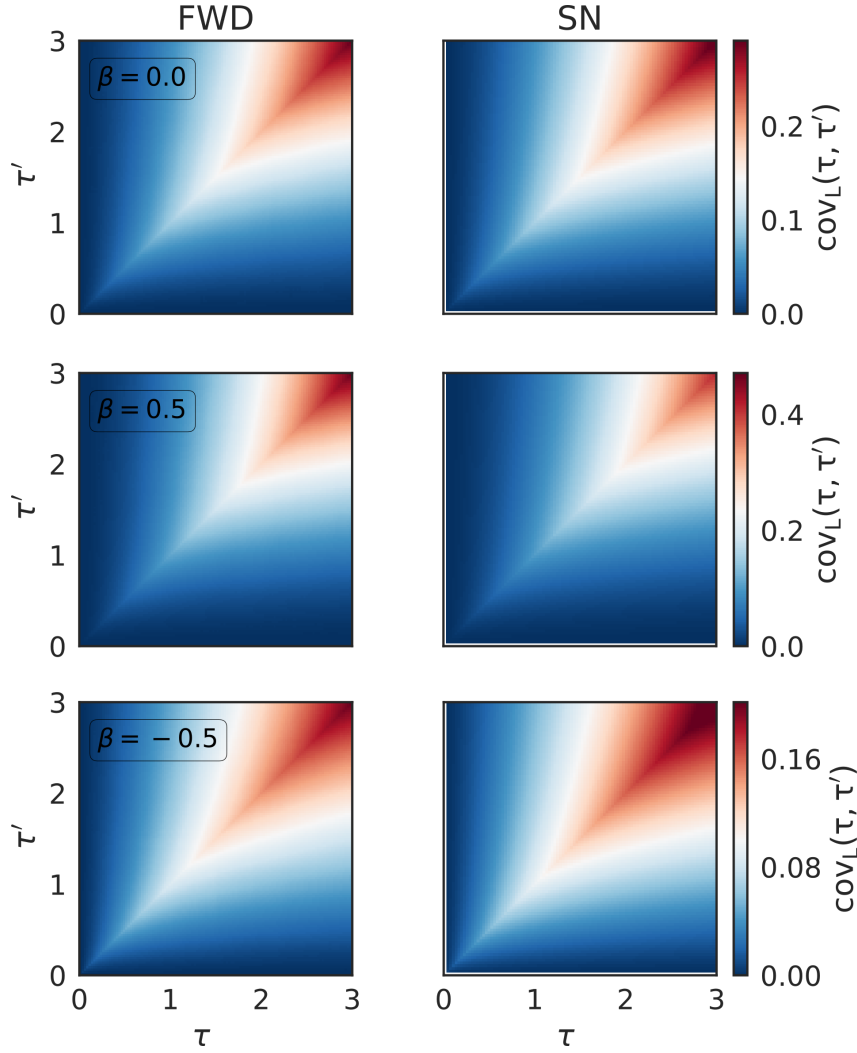


Figure 7.6: The two-time covariance for force driven dynamics with multiplicative noise can be predicted from a small noise expression. Comparison of the forward (FWD) and reverse time TSA small noise (SN) covariance for $f(L) = -\frac{\gamma}{L}$ and multiplicative noise $D(L) = DL^\beta$, with $\beta = 0, 0.5, -0.5$. The statistics of the forward dynamics are based on 20000 trajectories that start at $\widehat{L}_0 = 20$. The small noise covariance is stated in Eq. (7.49). The used parameters are $\gamma = 1$ and $D = 0.2$.

We next evaluate the first order equation Eq. (7.53), which, compared to Eq. (4.177), is only changed in the noise. It describes an Ornstein-Uhlenbeck process

$$dL_1 = L_1 k(L_0(\tau)) d\tau + c(L_0(\tau)) dW_\tau \quad (7.55)$$

with time dependent noise

$$c(L_0(\tau)) = L_0^{\frac{\beta}{2}} = ((1 - \alpha)\gamma\tau)^{\frac{\beta}{2-2\alpha}}, \quad (7.56)$$

where we used the explicit solution for L_0 stated in Eq. (4.180). The formal solution to Eq. (7.55) then reads

$$L_1(\tau) = \int_0^\tau c(L_0(\tau')) e^{\int_{\tau'}^\tau k(L_0(s)) ds} dW_{\tau'}, \quad (7.57)$$

or explicitly written

$$= \int_0^\tau ((1-\alpha)\gamma\tau')^{\frac{\beta}{2-2\alpha}} \left(\frac{\tau}{\tau'}\right)^{-\frac{\alpha}{\alpha-1}} dW_{\tau'}, \quad (7.58)$$

with $k(L_0(s))$ taken from Eq. (4.183) and $c(L_0(\tau'))$ from Eq. (7.56). As above and for simplicity, we assume the target state at $L(0) = 0$ and thus $L_1(0) = 0$.

We next generalize L_1 to include the noise dependent $c(L_0(\tau))$ term. Using the definition in Eq. (4.186), the variance evaluates to

$$\sigma_L^2(\tau) = D\langle L_1^2(\tau) \rangle = D \int_0^\tau ((1-\alpha)\gamma\tau')^{\frac{\beta}{1-\alpha}} \left(\frac{\tau}{\tau'}\right)^{-\frac{2\alpha}{\alpha-1}} d\tau' = \frac{D}{\gamma} \frac{((1-\alpha)\gamma\tau)^{\frac{1-\alpha+\beta}{1-\alpha}}}{1-3\alpha+\beta} \quad \text{for } \alpha < 0. \quad (7.59)$$

Similar to the variance, the two time covariance is built from the product of two $L_1(\tau)$ terms. Evaluating these terms we find for the covariance

$$\begin{aligned} C(\tau, \tau') &= D\langle L_1(\tau)L_1(\tau') \rangle = D \int_0^{\min(\tau, \tau')} ((1-\alpha)\gamma s)^{\frac{\beta}{1-\alpha}} \left(\frac{\tau}{s}\right)^{-\frac{\alpha}{\alpha-1}} \left(\frac{\tau'}{s}\right)^{-\frac{\alpha}{\alpha-1}} ds \\ &= \frac{D}{\gamma} \begin{cases} \frac{((1-\alpha)\gamma\tau')^{-\frac{\alpha}{\alpha-1}} ((1-\alpha)\gamma\tau)^{\frac{1-2\alpha+\beta}{1-\alpha}}}{1-3\alpha+\beta} & \text{for } \tau < \tau' \\ \frac{((1-\alpha)\gamma\tau)^{-\frac{\alpha}{\alpha-1}} ((1-\alpha)\gamma\tau')^{\frac{1-2\alpha+\beta}{1-\alpha}}}{1-3\alpha+\beta} & \text{for } \tau > \tau' \\ \frac{((1-\alpha)\gamma\tau)^{\frac{1-\alpha+\beta}{1-\alpha}}}{1-3\alpha+\beta} & \text{for } \tau = \tau'. \end{cases} \quad \text{and } \alpha < 0 \quad (7.60) \end{aligned}$$

In the last step we solve the averaged SDE of the second order contribution to $L(\tau)$ Eq. (7.54) to obtain the order D contribution to the mean. As in section 4.6, it is a linear ODE

$$\frac{d\langle L_2 \rangle}{d\tau} = \frac{da(L_0)}{dL_0} \langle L_2 \rangle + \frac{1}{2} \frac{d^2 a(L_0)}{dL_0^2} \langle L_1^2 \rangle - (\alpha - \beta) L_0^{\beta-1}, \quad (7.61)$$

with the last term slightly modified, compared to Eq. (4.192), to match the extension to multiplicative noise. Solving this ODE we find

$$\langle L_2(\tau) \rangle = \frac{(7\alpha^2 - \alpha(8\beta + 3) + 2\beta(\beta + 1)) ((1-\alpha)\gamma\tau)^{\frac{\alpha-\beta}{\alpha-1}}}{2\gamma(2\alpha - \beta)(3\alpha - \beta - 1)} \quad \text{for } \alpha < 0. \quad (7.62)$$

The sought final result stated in Eq. (7.47) is obtained by adding the found D -order result to the zero-th order contribution stated in Eq. (4.180).

Combining heterogeneity and multiplicative noise for force driven dynamics

The small noise moments of force driven TSA dynamics with multiplicative noise generalize to heterogeneous dynamics (i.e. with variations between trajectories in the γ -parameter) in exactly the same way as those with constant D . For the mean

$$\bar{L}(\tau) = \langle \gamma^{\frac{1}{1-\alpha}} \rangle ((1-\alpha)\tau)^{\frac{1}{1-\alpha}} + \langle \gamma^{\frac{1-\beta}{\alpha-1}} \rangle D \frac{(7\alpha^2 - \alpha(8\beta + 3) + 2\beta(\beta + 1)) ((1-\alpha)\tau)^{\frac{\alpha-\beta}{\alpha-1}}}{2(2\alpha - \beta)(3\alpha - \beta - 1)} \quad (7.63)$$

we again exploit Eq. (7.3), which tells us to average over the homogeneous mean Eq. (4.165) with respect to the distribution $P(\gamma)$.

For the variance we find three terms up to order D . The first is the variance of the deterministic contribution to the dynamics due to varying realizations of γ already stated in Eq. (7.8).

The second is obtained by averaging over the homogeneous variance component in the multiplicative case Eq. (7.48) with respect to γ . The third couples the zeroth order to the order D term and is different from zero only in the heterogeneous case. Together we find

$$\begin{aligned} \sigma_L^2(\tau) = & \left(\langle \gamma^{\frac{2}{1-\alpha}} \rangle - \langle \gamma^{\frac{1}{1-\alpha}} \rangle^2 \right) ((1-\alpha)\tau)^{\frac{2}{1-\alpha}} + \langle \gamma^{\frac{\beta}{1-\alpha}} \rangle D \frac{(1-\alpha)((1-\alpha)\tau)^{\frac{\beta}{1-\alpha}}}{1-3\alpha+\beta} \tau \\ & + \left(\langle \gamma^{\frac{\beta}{1-\alpha}} \rangle - \langle \gamma^{\frac{1}{1-\alpha}} \rangle \langle \gamma^{\frac{1-\beta}{\alpha-1}} \rangle \right) D \frac{(7\alpha^2 - \alpha(8\beta+3) + 2\beta(\beta+1)) ((1-\alpha)\tau)^{1-\frac{\beta}{\alpha-1}}}{(2\alpha-\beta)(3\alpha-\beta-1)}. \end{aligned} \quad (7.64)$$

The small noise covariance

$$\begin{aligned} C_L^{\text{het}}(\tau, \tau') = & \left(\langle \gamma^{\frac{2}{1-\alpha}} \rangle - \langle \gamma^{\frac{1}{1-\alpha}} \rangle^2 \right) (1-\alpha)^{\frac{2}{1-\alpha}} (\tau\tau')^{\frac{1}{1-\alpha}} \\ & + \langle \gamma^{\frac{\beta}{1-\alpha}} \rangle D \begin{cases} \frac{((1-\alpha)\tau')^{-\frac{\alpha}{\alpha-1}} ((1-\alpha)\tau)^{\frac{1-2\alpha+\beta}{1-\alpha}}}{1-3\alpha+\beta} & \text{for } \tau < \tau' \\ \frac{((1-\alpha)\tau)^{-\frac{\alpha}{\alpha-1}} ((1-\alpha)\tau')^{\frac{1-2\alpha+\beta}{1-\alpha}}}{1-3\alpha+\beta} & \text{for } \tau > \tau' \\ \frac{((1-\alpha)\tau)^{\frac{1-\alpha+\beta}{1-\alpha}}}{1-3\alpha+\beta} & \text{for } \tau = \tau' \end{cases} \\ & + \left(\langle \gamma^{\frac{\beta}{1-\alpha}} \rangle - \langle \gamma^{\frac{1}{1-\alpha}} \rangle \langle \gamma^{\frac{1-\beta}{\alpha-1}} \rangle \right) D \frac{(7\alpha^2 - \alpha(8\beta+3) + 2\beta(\beta+1)) (1-\alpha)^{1-\frac{\beta}{\alpha-1}}}{2(2\alpha-\beta)(3\alpha-\beta-1)} \times \\ & \times \left(\tau^{\frac{1}{1-\alpha}} \tau'^{\frac{\alpha-\beta}{\alpha-1}} + \tau'^{\frac{1}{1-\alpha}} \tau^{\frac{\alpha-\beta}{\alpha-1}} \right) \end{aligned} \quad (7.65)$$

similarly combines the extension to heterogeneity stated in Eq. (7.9) with the extension to multiplicative noise Eq. (7.49).

7.3 Ensemble equivalence of forward and reverse time ensembles

We here discuss the forward in time partners to the dynamics so far exclusively studied in reverse time. We then show that the generated ensembles are indistinguishable and that the dynamics derive from the same starting point.

Analyzing an ensemble of sample paths in forward or reverse time clearly does not change the ensemble itself. What changes are the dynamical equations that are capable to reproduce the data. This observation is best illustrated for a bridge process which does not start or end at a boundary. In forward time a bridge process is largely unconstrained close to its delta-initial conditions and evolves mostly according to the dynamical law of the unconstrained process. It is however strongly confined close to its target state where the entropy of the system must be reduced to reach a single state. For the reverse-time bridge it is exactly opposite. It first evolves mostly freely according to the sign inverted force law and is then strongly constrained to reach the forward initial conditions. In mathematical terms the forward bridge is defined as¹⁰⁴

$$d\widehat{L}(t) = \left(f(\widehat{L}) + D(\widehat{L}) \frac{\partial}{\partial \widehat{L}} \log \left(P(\widehat{L}_f, t_f | \widehat{L}, t) \right) \right) dt + \sqrt{D(\widehat{L})} dW_t, \quad (7.66)$$

or written as Fokker-Planck equation

$$\begin{aligned} \frac{\partial P(\widehat{L}, t | \widehat{L}_f, t_f, \widehat{L}_0, t_0)}{\partial t} = & - \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) + D(\widehat{L}) \frac{\partial}{\partial \widehat{L}} \log P(\widehat{L}_f, t_f | \widehat{L}, t) \right) P(\widehat{L}, t | \widehat{L}_f, t_f, \widehat{L}_0, t_0) \right] \\ & + \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) P(\widehat{L}, t | \widehat{L}_f, t_f, \widehat{L}_0, t_0). \end{aligned} \quad (7.67)$$

Comparing these expressions to the reverse time bridge with multiplicative noise stated in Eq. (7.15), we find that both display the same diffusion term and have to be interpreted in Ito sense. They differ however in their drift term. While $f(L)$ is expectedly only sign inverted, we find a different expression for the guiding force. For the forward dynamics the guiding force depends on the solution of the backward Fokker-Planck with its conditioning on its final condition. For the reverse time case we use the solution of the forward Fokker-Planck with its dependence on the forward initial condition. Furthermore, the reverse dynamics features a $D(L)$ both inside and outside the logarithm, while in the forward bridge the $D(L)$ only occurs outside of the logarithm. While conceptually obvious, to the best of our knowledge, this link between forward and reverse time Langevin bridges has not formally been established. In this section we show that forward and reverse time dynamics are different but describe the same ensemble.

The idea of ensemble equivalence also extends to dynamics which end at a target state which is an absorbing boundary. Majumdar and Orland¹⁰⁴ suggested to derive such a dynamics by taking the limit of Eq. (7.67) with respect to $\widehat{L}_f \rightarrow \widehat{L}_{ts}$, to obtain

$$\begin{aligned} \frac{\partial P(\widehat{L}, t | \widehat{L}_{ts}, t_f, \widehat{L}_0, t_0)}{\partial t} &= - \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) + D(\widehat{L}) \lim_{\widehat{L}_f \rightarrow \widehat{L}_{ts}} \frac{\partial}{\partial \widehat{L}} \log P(\widehat{L}_f, t_f | \widehat{L}, t) \right) P(\widehat{L}, t | \widehat{L}_{ts}, t_f, \widehat{L}_0, t_0) \right] \\ &+ \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) P(\widehat{L}, t | \widehat{L}_{ts}, t_f, \widehat{L}_0, t_0) . \end{aligned} \quad (7.68)$$

This approach however suffers under the limitation that the guiding force must first be evaluated with respect to an arbitrary $\widehat{L}_f \neq \widehat{L}_{ts}$, with the limit taken subsequently. Note, that taking the limit inside the logarithm is wrong and would lead to $\lim_{L_f \rightarrow 0} P(\widehat{L}_f, t_f | \widehat{L}, t) = 0$, using the definition of absorbing boundary conditions.

Along the lines of our previous derivation however a different approach is possible where $P(\widehat{L}_f, t_f | \widehat{L}, t)$ is replaced by the hitting time distribution $\rho_{\widehat{L}_{ts}}(t_f | L, t)$ during the construction of the respective Fokker-Planck bridge as done in section 4.2. The Fokker-Planck bridge with target state at an absorbing boundary reads

$$\begin{aligned} \frac{\partial R_{\widehat{L}_{ts}}(\widehat{L}, t | t_f, \widehat{L}_0, t_0)}{\partial t} &= - \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) + D(\widehat{L}) \frac{\partial}{\partial \widehat{L}} \log \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) \right) R_{\widehat{L}_{ts}}(\widehat{L}, t | t_f, \widehat{L}_0, t_0) \right] \\ &+ \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) R_{\widehat{L}_{ts}}(\widehat{L}, t | t_f, \widehat{L}_0, t_0) . \end{aligned} \quad (7.69)$$

Below in this section we show that the dynamics of Eq. (7.68) and Eq. (7.69) are indeed equivalent using a l'Hospital like argument. This equivalence also holds by the same argument when constructing explicit solutions for forward bridges. The limit approach however breaks down when constructing the non-normalized bridges used in the construction of TSA ensembles in section 4.2. Here, the product transition probabilities with absorbing boundary conditions $P(\widehat{L}_f, t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0)$, on which the derivation is based, would evaluate to zero for all times.

The generalization of ensemble equivalence of forward and reverse time ensembles to TSA ensembles therefore follows the construction approach laid out in section 4.2. Inspecting the resulting equation

$$\begin{aligned} \frac{\partial}{\partial t} R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f; \widehat{L}_0) &= P^{\text{in}}(\widehat{L}) \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) \\ &- \frac{\partial}{\partial \widehat{L}} \left(\left[f(\widehat{L}) + D(\widehat{L}) \frac{\partial}{\partial \widehat{L}} \log \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) \right] R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f; \widehat{L}_0) \right) + \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) R_{\widehat{L}_{ts}}(\widehat{L}, t; t_f; \widehat{L}_0) , \end{aligned} \quad (7.70)$$

which we will derive below in this section, we find that it is formally equivalent to a forward bridge with source term equivalent in its strength to the sink of the reverse time TSA dynamics. While mathematically consistent, this form of the TSA dynamics is less powerful than its reverse time version when used for data inference or synthesization close to the target state. First of all and most obvious when synthesizing data, the process starts infinitely far away from the target state with probability mass zero. To reach the target state in finite time, already here an approximation is necessary. We suggest to choose the minimal simulation time $t_f - t_0$ such that, given the amount of planned sample paths n_{ens} , on average no sample path yet exists at t_0 . This value can be determined from the over space and time integrated source, i.e. the integrated sample path “production” rate. We start simulations where on average no sample path has yet been produced. Second, the guiding term is – unlike for the reverse time case – explicitly time dependent. With only very few dynamics known where $\rho_{\widehat{L}_{\text{ts}}}(t_f|\widehat{L}, t)$ can be stated explicitly, the power of Eq. (7.70) probably lies more in its conceptual appeal than as a starting point for forward inference schemes.

Derivation of ensemble equivalence of forward and reverse time bridges

The derivation of both the forward and reverse time bridge process starts from the same joint conditional probability $P(\widehat{L}, t; \widehat{L}_f, t_f | \widehat{L}_0, t_0)$ already used in section 4.2 and section 7.2 for multiplicative noise. For readability we again ignore the dependency on the initial conditions. The two derivations then start to deviate in the choice of how to apply the chain rule for the second time on the D -dependent terms Eq. (4.15). We here start our derivation from a restatement of the decisive term

$$\frac{1}{2} \frac{\partial}{\partial \widehat{L}} \left(D(\widehat{L})P(\widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) - P(\widehat{L}_f, t_f | \widehat{L}, t) \frac{\partial}{\partial \widehat{L}} D(\widehat{L})P(\widehat{L}, t) \right) = \quad (7.71)$$

generalized to include multiplicative noise (i.e. with $D(\widehat{L})$ attached to $P(\widehat{L}, t)$ according to section 7.2). Unlike in the derivation for the reverse time bridge, we now apply the product rule to the second term and find

$$\begin{aligned} &= \frac{1}{2} \frac{\partial}{\partial \widehat{L}} \left(D(\widehat{L})P(\widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) - \frac{\partial}{\partial \widehat{L}} \left(D(\widehat{L})P(\widehat{L}, t) P(\widehat{L}_f, t_f | \widehat{L}, t) \right) \right. \\ &\quad \left. + D(\widehat{L})P(\widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) \right) . \end{aligned} \quad (7.72)$$

Summing identical terms and combining $P(\widehat{L}_f, t_f | \widehat{L}, t)$ and $P(\widehat{L}, t)$ then yields

$$= \frac{\partial}{\partial \widehat{L}} \left(D(\widehat{L})P(\widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) \right) - \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} \left(D(\widehat{L})P(\widehat{L}, t) P(\widehat{L}_f, t_f | \widehat{L}, t) \right) . \quad (7.73)$$

Keeping the collection of all $f(L)$ dependent terms unchanged compared to the derivation of the reverse time bridge, we end up with the evolution equation

$$\begin{aligned} - \frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} &= \frac{\partial}{\partial \widehat{L}} \left[f(\widehat{L})P(\widehat{L}, t, \widehat{L}_f, t_f) + D(\widehat{L})P(\widehat{L}, t) \frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t) \right] \\ &\quad - \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L})P(\widehat{L}, t, \widehat{L}_f, t_f) . \end{aligned} \quad (7.74)$$

Compared to Eq. (7.13), which states the corresponding version of the reverse time bridge, the terms in the guiding term (including its sign) have changed position with respect to the

derivative. The diffusion term comes with a minus sign in front of it. Extending the guiding term both in numerator and denominator with $P(\widehat{L}_f, t_f | \widehat{L}, t)$ and combining $P(\widehat{L}_f, t_f | \widehat{L}, t)$ with $P(\widehat{L}, t)$, we end up with

$$-\frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} = \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) + \frac{D(\widehat{L})}{P(\widehat{L}_f, t_f | \widehat{L}, t)} \frac{\partial P(\widehat{L}_f, t_f | \widehat{L}, t)}{\partial \widehat{L}} \right) P(\widehat{L}, t, \widehat{L}_f, t_f) \right] - \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) P(\widehat{L}, t, \widehat{L}_f, t_f). \quad (7.75)$$

Rewriting the guiding force in its well known logarithmic form we end up with the final bridge dynamic

$$-\frac{\partial P(\widehat{L}, t, \widehat{L}_f, t_f)}{\partial t} = \frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) + D(\widehat{L}) \frac{\partial}{\partial \widehat{L}} \log P(\widehat{L}_f, t_f | \widehat{L}, t) \right) P(\widehat{L}, t, \widehat{L}_f, t_f) \right] - \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) P(\widehat{L}, t, \widehat{L}_f, t_f), \quad (7.76)$$

which is stated in Eq. (7.67) in its more familiar form multiplied by minus one. There the dependence on (L_0, t_0) is explicitly stated and the whole term is normalized by $P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)$.

Taking the target state limit for the final state of bridge processes

In a recent paper, Majumdar and Orland implicitly assume that bridges processes with target states can be obtained from the limit $\widehat{L}_f \rightarrow \widehat{L}_{ts}$ applied to the guiding force of bridge processes Eq. (7.68), which do not end at the target state¹⁰⁴. Interestingly, this seemingly contradicts our result from section 4.2, where we have shown that the guiding force must be evaluated with respect to the hitting time if the process ends at a target state. We here show that both formulations

$$\frac{\partial}{\partial \widehat{L}} \log \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) = \lim_{\widehat{L}_f \rightarrow \widehat{L}_{ts}} \frac{\partial}{\partial \widehat{L}} \log P(\widehat{L}_f, t_f | \widehat{L}, t) \quad (7.77)$$

are equivalent using the rule of l'Hospital. For cases where both $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t)$ and $P(\widehat{L}_f, t_f | \widehat{L}, t)$ are known this identity can easily be checked. For example, for the random walk case with $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t)$ defined in Eq. (4.90), the solution of the left hand side can directly be read off, and we find the same result as stated in Majumdar and Orland¹⁰⁴. Furthermore, with the guiding forces equivalent in the limit $\widehat{L}_f \rightarrow \widehat{L}_{ts}$, also the respective bridges as defined in section 4.2 are equivalent in this limit and

$$\frac{P(\widehat{L}, t | \widehat{L}_0, t_0) \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t)}{\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}_0, t_0)} = \lim_{\widehat{L}_f \rightarrow \widehat{L}_{ts}} \frac{P(\widehat{L}, t | \widehat{L}_0, t_0) P(\widehat{L}_f, t_f | \widehat{L}, t)}{P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)} \quad (7.78)$$

must hold.

The derivation of both statements starts with the observation, that the hitting time, using its definition Eq. (4.26), can be rewritten as

$$\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) = - \int_{\widehat{L}_{ts}}^{\infty} d\widehat{L}_f \frac{\partial}{\partial t_f} P(\widehat{L}_f, t_f | \widehat{L}, t) = \frac{1}{2} \lim_{\widehat{L}_f \rightarrow \widehat{L}_{ts}} \frac{\partial}{\partial \widehat{L}_f} D(\widehat{L}_f) P(\widehat{L}_f, t_f | \widehat{L}, t). \quad (7.79)$$

The second identity relates the hitting time distribution to the flux through the absorbing boundary assuming a natural boundary at ∞ . It is obtained by substituting the forward Fokker-Planck equation into the first identity, and carrying out the integration. Performing the limit $\widehat{L}_f \rightarrow \widehat{L}_{ts}$ is equivalent to evaluating the resulting flux at the absorbing boundary \widehat{L}_{ts} .

Now, we start the derivation from the left hand side of Eq. (7.77), evaluate the logarithm

$$\frac{\partial}{\partial \widehat{L}} \log \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t) = \frac{\frac{\partial}{\partial \widehat{L}} \rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t)}{\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t)} \quad (7.80)$$

and substitute $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t)$ by the flux through the boundary Eq. (7.79)

$$= \frac{\lim_{\widehat{L}_f \rightarrow \widehat{L}_{ts}} \frac{\partial}{\partial \widehat{L}} \frac{\partial}{\partial \widehat{L}_f} D(\widehat{L}_f) P(\widehat{L}_f, t_f | \widehat{L}, t)}{\lim_{\widehat{L}_f \rightarrow \widehat{L}_{ts}} \frac{\partial}{\partial \widehat{L}_f} D(\widehat{L}_f) P(\widehat{L}_f, t_f | \widehat{L}, t)}. \quad (7.81)$$

Taking the limit outside of the fraction and applying the rule of l'Hospital with respect to the \widehat{L}_f derivate, yields

$$= \lim_{\widehat{L}_f \rightarrow \widehat{L}_{ts}} \frac{\frac{\partial}{\partial \widehat{L}} D(\widehat{L}_f) P(\widehat{L}_f, t_f | \widehat{L}, t)}{D(\widehat{L}_f) P(\widehat{L}_f, t_f | \widehat{L}, t)}. \quad (7.82)$$

We are now allowed to cancel the diffusion terms $D(\widehat{L}_f)$ and obtain

$$= \lim_{\widehat{L}_f \rightarrow \widehat{L}_{ts}} \frac{\frac{\partial}{\partial \widehat{L}} P(\widehat{L}_f, t_f | \widehat{L}, t)}{P(\widehat{L}_f, t_f | \widehat{L}, t)}, \quad (7.83)$$

which simplifies to the seeked expression for the guiding force

$$= \lim_{\widehat{L}_f \rightarrow \widehat{L}_{ts}} \frac{\partial}{\partial \widehat{L}} \log P(\widehat{L}_f, t_f | \widehat{L}, t). \quad (7.84)$$

Using exactly the same argument, but leaving out the derivative with respect to \widehat{L} and substituting $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}, t)$ by $\rho_{\widehat{L}_{ts}}(t_f | \widehat{L}_0, t_0)$ in the denominator of Eq. (7.80), the ensemble equivalence Eq. (7.78) follows directly.

This correspondence Eq. (7.78) however also shows that the TSA ensemble, which is decaying with reverse time τ , can only be correctly described with our approach, as it is based on the non-normalized product $P(\widehat{L}, t | \widehat{L}_0, t_0) P(\widehat{L}_f, t_f | \widehat{L}, t)$, which is zero when absorbing boundary conditions are assumed. The case of Majumdar and Orland is therefore only a special case and not suitable to e.g. directly calculate TSA ensembles from its definition Eq. (4.43).

7.4 Insights on meander processes

A stochastic process which up to a time t_f does not reach its starting point is called a meander process. A simple way to construct such a process is to simulate many random walks and take only those which have not returned to their starting point up to time t_f .

A different way to construct such a process was introduced by Majumdar and Orland¹⁰⁴. The authors claim the dynamics of such a process can be constructed from the forward bridge equation Eq. (7.66). The transition probability $P(\widehat{L}_f, t_f | \widehat{L}, t)$ inside the guiding term is the solution of the backward Fokker-Planck equation Eq. (4.6) with absorbing boundary conditions Eq. (4.25). According to Majumdar and Orland¹⁰⁴ the meander process

$$d\widehat{L}(t) = \left(f(\widehat{L}(t)) + D \frac{\partial}{\partial \widehat{L}} \log \left(\int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}, t) \right) \right) dt + \sqrt{D} dW_t \quad (7.85)$$

is obtained by integrating over $P(\widehat{L}_f, t_f | \widehat{L}, t)$ inside the logarithm and with respect to the endpoint \widehat{L}_f of the dynamics of a single bridge. In this section, we show that Majumdar's construction of meander processes¹⁰⁴ can thoroughly be derived from our novel understanding of targeted bridge processes.

Different to Majumdar's construction, and learning from our perspective on constructing ensembles from bridge processes, the meander process is given as a weighted ensemble of sub-processes

$$P(\widehat{L}, t | \widehat{L}_0, t_0, t_f) = \int_0^\infty d\widehat{L}_f W(\widehat{L}_f) P(\widehat{L}, t | \widehat{L}_0, t_0, \widehat{L}_f, t_f) \quad (7.86)$$

i.e. the average over bridge processes with respect to

$$W(\widehat{L}_f) = \frac{P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)}{\int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)}, \quad (7.87)$$

which is the probability mass preserved distribution of final positions that can be reached with a meander (and which can be calculated from the forward Fokker-Planck using absorbing boundary conditions and subsequent normalization). The derivation provided below follows analog to the derivation of the TSA ensemble in section 4.2. In a final step, where both approaches are again equivalent, the limit that takes the initial condition to the absorbing boundary $\widehat{L}_0 \rightarrow \widehat{L}_{\text{abs}}$ is taken. Due to the normalization in the construction of the bridge process with respect to $P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)$ it is ensured that the limit exists (see section 7.3 for a similar argument.) We show the derivation below in this section. For full generality we directly state the derivation with multiplicative noise. For Eq. (7.85) this simply implies to replace D with $D(\widehat{L})$.

Derivation of the meander process from sub-ensembles

We start our derivation of the meander process by taking the time-derivative on both sides of Eq. (7.86). We obtain

$$\frac{\partial}{\partial t} P(\widehat{L}, t | \widehat{L}_0, t_0, t_f) = \int_0^\infty d\widehat{L}_f W(\widehat{L}_f) \frac{\partial}{\partial t} P(\widehat{L}, t | \widehat{L}_0, t_0, \widehat{L}_f, t_f) \quad (7.88)$$

Replacing the forward bridge $P(\widehat{L}, t | \widehat{L}_0, t_0, \widehat{L}_f, t_f)$ by its definition Eq. (4.8) we arrive at

$$= \int_0^\infty d\widehat{L}_f W(\widehat{L}_f) \frac{\partial}{\partial t} \left(\frac{P(\widehat{L}_f, t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0)}{P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)} \right). \quad (7.89)$$

We use the definition of the weight $W(\widehat{L}_f)$ Eq. (7.87) and cancel terms to find

$$= \frac{1}{\int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)} \int_0^\infty d\widehat{L}_f \frac{\partial}{\partial t} \left(P(\widehat{L}_f, t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0) \right). \quad (7.90)$$

We continue our derivation only with the integral expression

$$\int_0^\infty d\widehat{L}_f \frac{\partial}{\partial t} \left(P(\widehat{L}_f, t_f | \widehat{L}, t) P(\widehat{L}, t | \widehat{L}_0, t_0) \right) \quad (7.91)$$

and multiply with its normalization $\frac{1}{\int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)}$ at the very end of the derivation. The solution to the integrand in terms of drift and diffusion terms is known from recapitulating the

derivation of the forward bridge process in section 7.3. We use the result stated in Eq. (7.74) and integrate over \widehat{L}_f to arrive at

$$\begin{aligned} & \int_0^\infty d\widehat{L}_f \frac{\partial}{\partial t} \left(P(\widehat{L}_f, t_f | \widehat{L}, t) P(\widehat{L}_t, t | \widehat{L}_0, t_0) \right) = \\ & = \frac{\partial}{\partial \widehat{L}} \left[-f(\widehat{L}) \int_0^\infty d\widehat{L}_f P(\widehat{L}, t, \widehat{L}_f, t_f | \widehat{L}_0, t_0) - D(\widehat{L}) P(\widehat{L}, t | \widehat{L}_0, t_0) \frac{\partial}{\partial \widehat{L}} \int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}, t) \right] \\ & + \frac{D}{2} \frac{\partial^2}{\partial \widehat{L}^2} \int_0^\infty d\widehat{L}_f P(\widehat{L}, t, \widehat{L}_f, t_f | \widehat{L}_0, t_0) . \end{aligned} \quad (7.92)$$

In the next step we transform the second term of the drift into a guiding force. To do this we multiply $P(\widehat{L}, t | \widehat{L}_0, t_0)$ by $\int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}, t)$ both in numerator and denominator. Rewritten as a logarithmic term we obtain

$$\begin{aligned} & = \frac{\partial}{\partial \widehat{L}} \left[-f(\widehat{L}) \int_0^\infty d\widehat{L}_f P(\widehat{L}, t, \widehat{L}_f, t_f | \widehat{L}_0, t_0) \right. \\ & \quad \left. - D(\widehat{L}) \int_0^\infty d\widehat{L}_f P(\widehat{L}, t, \widehat{L}_f, t_f | \widehat{L}_0, t_0) \frac{\partial}{\partial \widehat{L}} \log \int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}, t) \right] \\ & + \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) \int_0^\infty d\widehat{L}_f P(\widehat{L}, t, \widehat{L}_f, t_f | \widehat{L}_0, t_0) . \end{aligned} \quad (7.93)$$

To conclude the derivation of the equivalence of averaged bridges and a construction of meander processes based on the integration of the guiding probability distribution, we divide by the normalization $\int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}_0, t_0)$ and obtain the sought Fokker-Planck equation

$$\begin{aligned} \frac{\partial}{\partial t} P(\widehat{L}, t | \widehat{L}_0, t_0, t_f) & = -\frac{\partial}{\partial \widehat{L}} \left[\left(f(\widehat{L}) + D(\widehat{L}) \frac{\partial}{\partial \widehat{L}} \log \int_0^\infty d\widehat{L}_f P(\widehat{L}_f, t_f | \widehat{L}, t) \right) P(\widehat{L}, t | \widehat{L}_0, t_0, t_f) \right] \\ & + \frac{1}{2} \frac{\partial^2}{\partial \widehat{L}^2} D(\widehat{L}) P(\widehat{L}, t | \widehat{L}_0, t_0, t_f) , \end{aligned} \quad (7.94)$$

which is equivalent to the Langevin expression (7.85) known from Majumdar and Orland¹⁰⁴. The existence of the limit which brings the initial condition to the boundary $\widehat{L}_0 \rightarrow \widehat{L}_{\text{abs}}$ is ensured by the normalization of the bridge.

Conclusion

In this work I derive a consistent mathematical theory that allows to infer biological dynamics in reverse time starting from a target state. I apply this data driven approach to *Drosophila* embryogenesis to infer an effective model of cell contact constriction during germband extension mediated by cell intercalation. Target state alignment creates spurious forces. The TSA ensemble theory allows to separate these forces from the true underlying biological dynamics.

Different from previous work, the here derived theory of target state aligned ensembles allows to predict summary statistics of non-linear, non-equilibrium dynamics without reference to probably unknown initial conditions. Applied to *Drosophila* germband extension this statistical approach establishes a visco-elastic model of junction shrinkage.

The TSA framework is an exact mathematical description of sample path alignment to a target state. Using sample paths of finite length and working in reverse time the initial conditions eventually matter. I show how the influence of forward initial conditions on TSA ensemble statistics can be systematically assessed.

In a controlled approximation I derive the reverse time SDE valid for well separated initial and final conditions. The obtained result allows to formulate a reverse time path ensemble inference scheme and to recover forward dynamics close to a target state from reverse time trajectories. For forward power laws I derive the exact reverse time forces to be used in the inference. I use this result to derive the universal behavior of force and noise dominated dynamics close to the target state. Expressions for mean, variance and two-time covariance are stated. I demonstrate the excellent quality of this approach for a toy model of cytokinetic ring constriction.

With the theory at hand a large class of biological processes can now be analyzed. The TSA approach allows to infer effective dynamics from experimentally obtained target state aligned sample paths. The derived analytical moments provide a heuristic approach to judge whether or not tested models are capable of generating the experimentally observed statistics. For example, in the here studied case of germband extension, we find that the experimentally observed two-time covariance of cell contact length, can under no circumstances be generated by a homogeneous, time independent model.

The approach to rule out whole model classes based on their ensemble statistics can be extended to models with multiplicative noise. I show that heterogeneity in the realizations of individual sample paths leaves a detectable finger print in the covariance. The last source of insecurity is the measurement noise. I show that TSA ensemble statistics can be used to statistically separate it from the data.

Combining theoretical reasoning based on ensemble statistics and model inference I characterize a visco-elastic model as the most simple effective dynamic to reproduce all accessible statistics of junction shrinkage in the experiment. In this model elastic fluctuations around an equilibrium cell contact length on short timescales are combined with a time dependent change

of the equilibrium length on long timescales. Using a path ensemble inference technique, I can simultaneously predict the initial equilibrium cell contact length, and the global parameters such as the timescale of elastic contact fluctuations and the noise strength. I can further access the quality of the model for each individual sample path and characterize where the model is over simplistic.

Data driven model inference of the effective dynamics of junction shrinkage is only one step towards a full understanding of this complex process. Future work will combine the ensemble approach with recordings of the junction length at different z-values, i.e. measurements at different depths below the tissue surface. Deriving different effective models for different z-levels should eventually provide us with a model of the three dimensional dynamics of cell intercalation. It will be of further interest to simultaneously record the junction length with the dynamics of for the constriction relevant molecules such as myosin, of regulatory molecules such as Rho and Rok, of molecules involved in junction remodeling such as E-cadherin and of molecules assumed to act as a ratchet such as Rab 35. The TSA ensemble inference approach provides the framework to step by step include more details in the model and monitor their impact on the predicted ensemble statistics.

The core unit of all derivations in this text are bridge processes, i.e. a stochastic dynamic with point initial and final conditions. As a by-product of the TSA derivation I show that although bridge processes describe the same ensemble in forward and reverse time, they differ in their generating dynamics. I further show that meander processes, i.e. stochastic processes which at least up to a time t do not return to their initial conditions, can be accumulated from an ensemble average over bridge processes.

In summary, with additional insights in several directions, the TSA ensemble approach provides a versatile tool for the analysis of directional dynamics. It provides the means to study biological dynamics as ensembles but avoids perturbing the system under study through the introduction of artificial initial conditions. It provides the natural framework to study dynamics close to target states.

Bibliography

- [1] M Abramowitz and IA Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Vol. 55. Courier Corporation, 1964.
- [2] Y Ait-Sahalia. “Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach”. In: *Econometrica* **70** (2002), pp. 223–262.
- [3] Y Ait-Sahalia et al. “Closed-form likelihood expansions for multivariate diffusions”. In: *The Annals of Statistics* **36** (2008), pp. 906–937.
- [4] F Amat et al. “Efficient processing and analysis of large-scale light-sheet microscopy data”. In: *Nature protocols* **10** (2015), p. 1679.
- [5] BD Anderson. “Reverse-time diffusion equation models”. In: *Stochastic Processes and their Applications* **12** (1982), pp. 313–326.
- [6] D Arcizet et al. “Temporal analysis of active and passive transport in living cells”. In: *Physical review letters* **101** (2008), p. 248103.
- [7] F Balzarotti et al. “Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes”. In: *Science* **355** (2017), pp. 606–612.
- [8] K Bambardekar et al. “Direct laser manipulation reveals the mechanics of cell contacts in vivo”. In: *Proceedings of the National Academy of Sciences* **112** (2015), pp. 1416–1421.
- [9] A Berezhkovskii and A Szabo. “Time scale separation leads to position-dependent diffusion along a slow coordinate”. In: *The Journal of chemical physics* **135** (2011), p. 074108.
- [10] OG Berg, RB Winter, and PH Von Hippel. “Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory”. In: *Biochemistry* **20** (1981), pp. 6929–6948.
- [11] AJ Berglund. “Statistics of camera-based single-particle tracking”. In: *Physical Review E* **82** (2010), p. 011917.
- [12] C Bertet, L Sulak, and T Lecuit. “Myosin-dependent junction remodelling controls planar cell intercalation and axis elongation”. In: *Nature* **429** (2004), p. 667.
- [13] A Beskos et al. “Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion)”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **68** (2006), pp. 333–382.
- [14] PJ Bickel, BJ Kleijn, et al. “The semiparametric Bernstein–von Mises theorem”. In: *The Annals of Statistics* **40** (2012), pp. 206–237.

- [15] GB Blanchard, J Étienne, and N Gorfinkiel. “From pulsatile apicomedial contractility to effective epithelial mechanics”. In: *Current opinion in genetics & development* **51** (2018), pp. 78–87.
- [16] JT Blankenship et al. “Multicellular rosette formation links planar cell polarity to tissue morphogenesis”. In: *Developmental cell* **11** (2006), pp. 459–470.
- [17] CP Brangwynne et al. “Intracellular transport by active diffusion”. In: *Trends in cell biology* **19** (2009), pp. 423–427.
- [18] A Bray. “Random walks in logarithmic and power-law potentials, nonuniversal persistence, and vortex dynamics in the two-dimensional XY model”. In: *Physical Review E* **62** (2000), p. 103.
- [19] BW Brunton, MM Botvinick, and CD Brody. “Rats and humans can optimally accumulate evidence for decision-making”. In: *Science* **340** (2013), pp. 95–98.
- [20] J Buchner et al. “X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue”. In: *Astronomy & Astrophysics* **564** (2014), A125.
- [21] S Budnar and AS Yap. “A mechanobiological perspective on cadherins and the actin-myosin cytoskeleton”. In: *F1000prime reports* **5** (2013).
- [22] S Burov et al. “Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking”. In: *Physical Chemistry Chemical Physics* **13** (2011), pp. 1800–1812.
- [23] JA Campos-Ortega and V Hartenstein. “Stages of *Drosophila* embryogenesis”. In: *The embryonic Development of Drosophila melanogaster*. Springer, 1997, pp. 9–102.
- [24] A Caspi, R Granek, and M Elbaum. “Diffusion and directed motion in cellular transport”. In: *Physical Review E* **66** (2002), p. 011916.
- [25] I Castillo, R Nickl, et al. “On the Bernstein–von Mises phenomenon for nonparametric Bayes procedures”. In: *The Annals of Statistics* **42** (2014), pp. 1941–1969.
- [26] M Chalfie et al. “Green fluorescent protein as a marker for gene expression”. In: *Science* **263** (1994), pp. 802–805.
- [27] AD Chalmers and J Slack. “The *Xenopus* tadpole gut: fate maps and morphogenetic movements”. In: *Development* **127** (2000), pp. 381–392.
- [28] G Charras and AS Yap. “Tensile forces and mechanotransduction at cell–cell junctions”. In: *Current Biology* **28** (2018), R445–R457.
- [29] B-C Chen et al. “Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution”. In: *Science* **346** (2014), p. 1257998.
- [30] VY Chernyak, M Chertkov, and C Jarzynski. “Path-integral analysis of fluctuation theorems for general Langevin processes”. In: *Journal of Statistical Mechanics: Theory and Experiment* **2006** (2006), P08001.
- [31] AK Churchland et al. “Variance as a signature of neural computations during decision making”. In: *Neuron* **69** (2011), pp. 818–831.
- [32] R Clément et al. “Viscoelastic dissipation stabilizes cell shape changes during tissue morphogenesis”. In: *Current Biology* **27** (2017), pp. 3132–3142.
- [33] AD Co et al. “Stochastic timing in gene expression for simple regulatory strategies”. In: *Nucleic acids research* **45** (2016), pp. 1069–1078.

-
- [34] D Collin et al. “Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies”. In: *Nature* **437** (2005), p. 231.
- [35] C Collinet et al. “Local and tissue-scale forces drive oriented junction growth during tissue extension”. In: *Nature cell biology* **17** (2015), p. 1247.
- [36] C Dargatz. “Bayesian inference for diffusion processes with applications in life sciences”. PhD thesis. lmu, 2010.
- [37] AG De Las Bayonas et al. “Distinct RhoGEFs activate apical and junctional actomyosin contractility under control of G proteins during epithelial morphogenesis”. In: *bioRxiv* (2019), p. 566919.
- [38] N De Souza. *Pulling on single molecules*. 2012.
- [39] D Devenport. “Tissue morphodynamics: translating planar polarity cues into polarized cell behaviors”. In: *Seminars in cell & developmental biology*. Vol. 55. Elsevier. 2016, pp. 99–110.
- [40] K Dierkes et al. “Spontaneous oscillations of elastic contractile materials with turnover”. In: *Physical review letters* **113** (2014), p. 148102.
- [41] GB Durham and AR Gallant. “Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes”. In: *Journal of Business & Economic Statistics* **20** (2002), pp. 297–338.
- [42] A Einstein. “Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen”. In: *Annalen der physik* **322** (1905), pp. 549–560.
- [43] M El Beheiry, M Dahan, and J-B Masson. “InferenceMAP: mapping of single-molecule dynamics with Bayesian inference”. In: *Nature methods* **12** (2015), p. 594.
- [44] O Elerian et al. “A note on the existence of a closed form conditional transition density for the Milstein scheme”. In: *Economics discussion paper* (1998), W18.
- [45] R Erban and SJ Chapman. “Reactive boundary conditions for stochastic simulations of reaction–diffusion processes”. In: *Physical Biology* **4** (2007), p. 16.
- [46] R Farhadifar et al. “The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing”. In: *Current Biology* **17** (2007), pp. 2095–2104.
- [47] R Fernandez-Gonzalez et al. “Myosin II dynamics are regulated by tension in intercalating cells”. In: *Developmental cell* **17** (2009), pp. 736–743.
- [48] F Feroz, M Hobson, and M Bridges. “MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics”. In: *Monthly Notices of the Royal Astronomical Society* **398** (2009), pp. 1601–1614.
- [49] F Feroz et al. “Importance nested sampling and the MultiNest algorithm”. In: *arXiv preprint arXiv:1306.2144* (2013).
- [50] F Feroz and MP Hobson. “Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses”. In: *Monthly Notices of the Royal Astronomical Society* **384** (2008), pp. 449–463.
- [51] K Fogelmark et al. “Fitting a function to time-dependent ensemble averaged data”. In: *Scientific reports* **8** (2018), p. 6984.
- [52] H Föllmer. “Time reversal on Wiener space”. In: *Stochastic Processes in Mathematics and Physics*. Springer, 1986, pp. 119–129.

- [53] R Friedrich and J Peinke. “Description of a turbulent cascade by a Fokker-Planck equation”. In: *Physical Review Letters* **78** (1997), p. 863.
- [54] R Gao et al. “Cortical column and whole-brain imaging with molecular contrast and nanoscale resolution”. In: *Science* **363** (2019), eaau8302.
- [55] CW Gardiner. “Handbook of stochastic methods for physics, chemistry and the natural sciences, vol. 13 of”. In: *Springer series in synergetics* (1985).
- [56] MA Gelbart et al. “Volume conservation principle involved in cell lengthening and nucleus movement during tissue morphogenesis”. In: *Proceedings of the National Academy of Sciences* **109** (2012), pp. 19298–19303.
- [57] KR Ghusinga, JJ Dennehy, and A Singh. “First-passage time approach to controlling noise in the timing of intracellular events”. In: *Proceedings of the National Academy of Sciences* **114** (2017), pp. 693–698.
- [58] DT Gillespie. “The chemical Langevin equation”. In: *The Journal of Chemical Physics* **113** (2000), pp. 297–306.
- [59] D Gilmour, M Rembold, and M Leptin. “From morphogen to morphogenesis and back”. In: *Nature* **541** (2017), p. 311.
- [60] J Gorman et al. “Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair”. In: *Proceedings of the National Academy of Sciences* **109** (2012), E3074–E3083.
- [61] R Gräf, J Rietdorf, and T Zimmermann. “Live cell spinning disk microscopy”. In: *Microscopy techniques*. Springer, 2005, pp. 57–75.
- [62] C Guillot and T Lecuit. “Mechanics of epithelial tissue homeostasis and morphogenesis”. In: *Science* **340** (2013), pp. 1185–1189.
- [63] S Gupta et al. “Temporal precision of regulated gene expression”. In: *PLoS computational biology* **14** (2018), e1006201.
- [64] M Haering et al. “Automated segmentation of epithelial tissue using cycle-consistent generative adversarial networks”. In: *bioRxiv* (2018), p. 311373.
- [65] DP Hanes and JD Schall. “Neural control of voluntary movement initiation”. In: *Science* **274** (1996), pp. 427–430.
- [66] TD Hanks et al. “Distinct relationships of parietal and prefrontal cortices to evidence accumulation”. In: *Nature* **520** (2015), pp. 220–223.
- [67] J Hettich and JCM Gebhardt. “Transcription factor target site search and gene regulation in a background of unspecific binding sites”. In: *Journal of theoretical biology* **454** (2018), pp. 91–101.
- [68] F Höfling and T Franosch. “Anomalous transport in the crowded world of biological cells”. In: *Reports on Progress in Physics* **76** (2013), p. 046602.
- [69] D Holcman, N Hoze, and Z Schuss. “Analysis and interpretation of superresolution single-particle trajectories”. In: *Biophysical journal* **109** (2015), pp. 1761–1771.
- [70] D Holcman, A Marchewka, and Z Schuss. “Survival probability of diffusion with trapping in cellular neurobiology”. In: *Physical Review E* **72** (2005), p. 031910.
- [71] V Holý and P Tomanová. “Estimation of Ornstein-Uhlenbeck Process Using Ultra-High-Frequency Data with Application to Intraday Pairs Trading Strategy”. In: *arXiv preprint arXiv:1811.09312* (2018).

-
- [72] N Hozé and D Holcman. “Statistical methods for large ensembles of super-resolution stochastic single particle trajectories in cell biology”. In: *Annual Review of Statistics and Its Application* **4** (2017), pp. 189–223.
- [73] RJ Huebner and JB Wallingford. “Coming to consensus: a unifying model emerges for convergent extension”. In: *Developmental cell* **46** (2018), pp. 389–396.
- [74] L Hufnagel et al. “On the mechanism of wing size determination in fly development”. In: *Proceedings of the National Academy of Sciences* **104** (2007), pp. 3835–3840.
- [75] J Huisken and DY Stainier. “Selective plane illumination microscopy techniques in developmental biology”. In: *Development* **136** (2009), pp. 1963–1975.
- [76] J Huisken et al. “Optical sectioning deep inside live embryos by selective plane illumination microscopy”. In: *Science* **305** (2004), pp. 1007–1009.
- [77] G Hummer and A Szabo. “Free energy reconstruction from nonequilibrium single-molecule pulling experiments”. In: *Proceedings of the National Academy of Sciences* **98** (2001), pp. 3658–3661.
- [78] AS Hurn, J Jeisman, and KA Lindsay. “Seeing the wood for the trees: A critical evaluation of methods to estimate the parameters of stochastic differential equations”. In: *Journal of Financial Econometrics* **5** (2007), pp. 390–455.
- [79] KD Irvine and E Wieschaus. “Cell intercalation during *Drosophila* germband extension and its regulation by pair-rule segmentation genes”. In: *Development* **120** (1994), pp. 827–841.
- [80] J Jaeger, J Reinitz, et al. “*Drosophila* blastoderm patterning”. In: *Current opinion in genetics & development* **22** (2012), pp. 533–541.
- [81] CE Jewett et al. “Planar polarized Rab35 functions as an oscillatory ratchet during cell intercalation in the *Drosophila* epithelium”. In: *Nature communications* **8** (2017), p. 476.
- [82] GR Kale et al. “Distinct contributions of tensile and shear stress on E-cadherin levels during morphogenesis”. In: *Nature communications* **9** (2018), p. 5021.
- [83] PJ Keller et al. “Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy”. In: *science* **322** (2008), pp. 1065–1069.
- [84] R Keller et al. “The function and mechanism of convergent extension during gastrulation of *Xenopus laevis*”. In: *Development* **89** (1985), pp. 185–209.
- [85] R Keller. “Shaping the vertebrate body plan by polarized embryonic cell movements”. In: *Science* **298** (2002), pp. 1950–1954.
- [86] R Keller et al. “Mechanisms of convergence and extension by cell intercalation”. In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **355** (2000), pp. 897–922.
- [87] S Kerridge et al. “Modular activation of Rho1 by GPCR signalling imparts polarized myosin II activation during morphogenesis”. In: *Nature cell biology* **18** (2016), p. 261.
- [88] M Kessler. “Estimation of an ergodic diffusion from discrete observations”. In: *Scandinavian Journal of Statistics* **24** (1997), pp. 211–229.
- [89] BJK Kleijn, AW Van der Vaart, et al. “The Bernstein-von-Mises theorem under misspecification”. In: *Electronic Journal of Statistics* **6** (2012), pp. 354–381.
- [90] D Kong. “Dynamics of Cell Contacts During Cell Intercalation in Epithelial Tissue Elongation of (*Drosophila*) Embryos”. PhD thesis. Georg-August-Universität Göttingen, 2017.

- [91] D Kong, F Wolf, and J Grobthans. “Forces directing germ-band extension in *Drosophila* embryos”. In: *Mechanisms of development* **144** (2017), pp. 11–22.
- [92] U Krzic et al. “Multiview light-sheet microscope for rapid in toto imaging”. In: *Nature methods* **9** (2012), p. 730.
- [93] T Lecuit, P-F Lenne, and E Munro. “Force generation, transmission, and integration during cell and tissue morphogenesis”. In: *Annual review of cell and developmental biology* **27** (2011), pp. 157–184.
- [94] T Lecuit and AS Yap. “E-cadherin junctions as active mechanical integrators in tissue dynamics”. In: *Nature cell biology* **17** (2015), p. 533.
- [95] A Lee et al. “Unraveling the thousand word picture: an introduction to super-resolution data analysis”. In: *Chemical reviews* **117** (2017), pp. 7276–7330.
- [96] M Leptin and B Grunewald. “Cell shape changes during gastrulation in *Drosophila*”. In: *Development* **110** (1990), pp. 73–84.
- [97] M Leptin. “*Drosophila* gastrulation: from pattern formation to morphogenesis”. In: *Annual review of cell and developmental biology* **11** (1995), pp. 189–212.
- [98] R Levayer, A Pelissier-Monier, and T Lecuit. “Spatial regulation of Dia and Myosin-II by RhoGEF2 controls initiation of E-cadherin endocytosis during epithelial morphogenesis”. In: *Nature cell biology* **13** (2011), p. 529.
- [99] SS Lienkamp et al. “Vertebrate kidney tubules elongate using a planar cell polarity-dependent, rosette-based mechanism of convergent extension”. In: *Nature genetics* **44** (2012), p. 1382.
- [100] J Liphardt et al. “Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski’s equality”. In: *Science* **296** (2002), pp. 1832–1835.
- [101] J Lippincott-Schwartz and GH Patterson. “Development and use of fluorescent protein markers in living cells”. In: *Science* **300** (2003), pp. 87–91.
- [102] T-l Liu et al. “Observing the cell in its native state: Imaging subcellular dynamics in multicellular organisms”. In: *Science* **360** (2018), eaaq1392.
- [103] C Maes and K Netočný. “Time-reversal and entropy”. In: *Journal of statistical physics* **110** (2003), pp. 269–310.
- [104] SN Majumdar and H Orland. “Effective Langevin equations for constrained stochastic processes”. In: *Journal of Statistical Mechanics: Theory and Experiment* **2015** (2015), P06039.
- [105] C Manzo and MF Garcia-Parajo. “A review of progress in single particle tracking: from methods to biophysical insights”. In: *Reports on progress in physics* **78** (2015), p. 124601.
- [106] AC Martin, M Kaschube, and EF Wieschaus. “Pulsed contractions of an actin–myosin network drive apical constriction”. In: *Nature* **457** (2009), p. 495.
- [107] E Martin, U Behn, and G Germano. “First-passage and first-exit times of a Bessel-like stochastic process”. In: *Physical Review E* **83** (2011), p. 051115.
- [108] J-B Masson, M Dahan, and A Triller. “Pitfalls in the quantitative analysis of random walks and the mapping of single-molecule dynamics at the cellular scale”. In: *arXiv preprint arXiv:1502.07285* (2015).
- [109] J-B Masson et al. “Mapping the energy and diffusion landscapes of membrane proteins at the cell surface using high-density single-molecule imaging and Bayesian inference: application to the multiscale dynamics of glycine receptors in the neuronal membrane”. In: *Biophysical journal* **106** (2014), pp. 74–83.

-
- [110] S de Matos Simões, A Mainieri, and JA Zallen. “Rho GTPase and Shroom direct planar polarized actomyosin contractility during convergent extension”. In: *J Cell Biol* **204** (2014), pp. 575–589.
- [111] S de Matos Simões et al. “Rho-kinase directs Bazooka/Par-3 planar polarity during Drosophila axis elongation”. In: *Developmental cell* **19** (2010), pp. 377–388.
- [112] M Merkel et al. “Triangles bridge the scales: quantifying cellular contributions to tissue deformation”. In: *Physical Review E* **95** (2017), p. 032401.
- [113] Y Meroz and IM Sokolov. “A toolbox for determining subdiffusive mechanisms”. In: *Physics Reports* **573** (2015), pp. 1–29.
- [114] R Metzler et al. “Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking”. In: *Physical Chemistry Chemical Physics* **16** (2014), pp. 24128–24164.
- [115] H Meyer, T Voigtmann, and T Schilling. “On the non-stationary generalized Langevin equation”. In: *The Journal of chemical physics* **147** (2017), p. 214110.
- [116] X Michalet and AJ Berglund. “Optimal diffusion coefficient estimation in single-particle tracking”. In: *Physical Review E* **85** (2012), p. 061916.
- [117] H Mori. “Transport, collective motion, and Brownian motion”. In: *Progress of theoretical physics* **33** (1965), pp. 423–455.
- [118] A Munjal et al. “A self-organized biomechanical network drives shape changes during tissue morphogenesis”. In: *Nature* **524** (2015), p. 351.
- [119] I Nachman, A Regev, and S Ramanathan. “Dissecting timing variability in yeast meiosis”. In: *Cell* **131** (2007), pp. 544–556.
- [120] N Noll et al. “Active tension network model suggests an exotic mechanical state realized in epithelial tissues”. In: *Nature physics* **13** (2017), p. 1221.
- [121] I Nordlund. “New determination of the Avogadro constant from the Brownian motion of small spheres of mercury suspended in water”. In: *Z. Phys. Chem* **87** (1914), pp. 40–62.
- [122] K Norregaard et al. “Manipulation and motion of organelles and single molecules in living cells”. In: *Chemical reviews* **117** (2017), pp. 4342–4375.
- [123] H Oda and S Tsukita. “Real-time imaging of cell-cell adherens junctions reveals that Drosophila mesoderm invagination begins with two phases of apical constriction of cells”. In: *Journal of cell science* **114** (2001), pp. 493–501.
- [124] T Ozaki. “Statistical identification of storage models with application to stochastic hydrology 1”. In: *JAWRA Journal of the American Water Resources Association* **21** (1985), pp. 663–675.
- [125] N Özkaya and M Nordin. “Mechanical properties of biological tissues”. In: *Fundamentals of Biomechanics*. Springer, 1999, pp. 195–218.
- [126] AC Paré et al. “A positional Toll receptor code directs convergent extension in Drosophila”. In: *Nature* **515** (2014), p. 523.
- [127] S Pauls, B Geldmacher-Voss, and JA Campos-Ortega. “A zebrafish histone variant H2A. F/Z and a transgenic H2A. F/Z: GFP fusion protein for in vivo studies of embryonic development”. In: *Development genes and evolution* **211** (2001), pp. 603–610.
- [128] J Perrin. “L agitation moleculaire et le mouvement brownien”. In: *Comptes rendus hebdomadaires des seances de l academie des sciences* **146** (1908), pp. 967–970.

- [129] TD Pollard. “Mechanics of cytokinesis in eukaryotes”. In: *Current opinion in cell biology* **22** (2010), pp. 50–56.
- [130] TD Pollard. “Nine unanswered questions about cytokinesis”. In: *J Cell Biol* **216** (2017), pp. 3007–3016.
- [131] M Popović et al. “Active dynamics of tissue shear flow”. In: *New Journal of Physics* **19** (2017), p. 033006.
- [132] RM Power and J Huisken. “A guide to light-sheet fluorescence microscopy for multiscale imaging”. In: *Nature methods* **14** (2017), p. 360.
- [133] H Qian, MP Sheetz, and EL Elson. “Single particle tracking. Analysis of diffusion and flow in two-dimensional systems”. In: *Biophysical journal* **60** (1991), pp. 910–921.
- [134] R Ratcliff and G McKoon. “The diffusion decision model: theory and data for two-choice decision tasks”. In: *Neural computation* **20** (2008), pp. 873–922.
- [135] R Ratcliff et al. “Diffusion decision model: current issues and history”. In: *Trends in cognitive sciences* **20** (2016), pp. 260–281.
- [136] M Rauzi, P-F Lenne, and T Lecuit. “Planar polarized actomyosin contractile flows control epithelial junction remodelling”. In: *Nature* **468** (2010), p. 1110.
- [137] M Rauzi et al. “Nature and anisotropy of cortical forces orienting *Drosophila* tissue morphogenesis”. In: *Nature cell biology* **10** (2008), p. 1401.
- [138] S Redner. *A guide to first-passage processes*. Cambridge University Press, 2001.
- [139] LM Ricciardi and S Sato. “First-passage-time density and moments of the Ornstein-Uhlenbeck process”. In: *Journal of Applied Probability* **25** (1988), pp. 43–57.
- [140] A Robson, K Burrage, and MC Leake. “Inferring diffusion in single live cells at the single-molecule level”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* **368** (2013), p. 20120029.
- [141] JD Roitman and MN Shadlen. “Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task”. In: *Journal of neuroscience* **22** (2002), pp. 9475–9489.
- [142] G dos Santos et al. “FlyBase: introduction of the *Drosophila melanogaster* Release 6 reference genome assembly and large-scale migration of genome annotations”. In: *Nucleic acids research* **43** (2014), pp. D690–D697.
- [143] MJ Saxton and K Jacobson. “Single-particle tracking: applications to membrane dynamics”. In: *Annual review of biophysics and biomolecular structure* **26** (1997), pp. 373–399.
- [144] J Schindelin et al. “Fiji: an open-source platform for biological-image analysis”. In: *Nature methods* **9** (2012), p. 676.
- [145] CA Schneider, WS Rasband, and KW Eliceiri. “NIH Image to ImageJ: 25 years of image analysis”. In: *Nature methods* **9** (2012), p. 671.
- [146] T Schötz, RA Neher, and U Gerland. “Target search on a dynamic DNA molecule”. In: *Physical Review E* **84** (2011), p. 051911.
- [147] Z Schuss. *Brownian dynamics at boundaries and interfaces*. Springer, 2015.
- [148] U Seifert. “Entropy production along a stochastic trajectory and an integral fluctuation theorem”. In: *Physical review letters* **95** (2005), p. 040602.
- [149] U Seifert. “Stochastic thermodynamics, fluctuation theorems and molecular machines”. In: *Reports on progress in physics* **75** (2012), p. 126001.

-
- [150] K Sekimoto. “Langevin equation and thermodynamics”. In: *Progress of Theoretical Physics Supplement* **130** (1998), pp. 17–27.
- [151] H Shen et al. “Single particle tracking: from theory to biophysical applications”. In: *Chemical Reviews* **117** (2017), pp. 7331–7376.
- [152] T Shimozawa et al. “Improving spinning disk confocal microscopy by preventing pinhole cross-talk for intravital imaging”. In: *Proceedings of the National Academy of Sciences* **110** (2013), pp. 3399–3404.
- [153] A Shindo. “Models of convergent extension during morphogenesis”. In: *Wiley Interdisciplinary Reviews: Developmental Biology* **7** (2018), e293.
- [154] LC Siang, R Fernandez-Gonzalez, and JJ Feng. “Modeling cell intercalation during *Drosophila* germband extension”. In: *Physical biology* **15** (2018), p. 066008.
- [155] H Siedentopf and R Zsigmondy. “Über sichtbarmachung und größenbestimmung ultramikroskopischer teilchen, mit besonderer anwendung auf goldrubingläser”. In: *Annalen der Physik* **315** (1902), pp. 1–39.
- [156] S Siegert, R Friedrich, and J Peinke. “Analysis of data sets of stochastic systems”. In: *arXiv preprint cond-mat/9803250* (1998).
- [157] J Skilling et al. “Nested sampling for general Bayesian computation”. In: *Bayesian analysis* **1** (2006), pp. 833–859.
- [158] I Sokolov. “Itô, Stratonovich, Hänggi and all the rest: The thermodynamics of interpretation”. In: *Chemical Physics* **375** (2010), pp. 359–363.
- [159] IM Sokolov. “Statistical mechanics of entropic forces: disassembling a toy”. In: *European Journal of Physics* **31** (2010), p. 1353.
- [160] C Sommer et al. “Ilastik: Interactive learning and segmentation toolkit”. In: *2011 IEEE international symposium on biomedical imaging: From nano to macro*. IEEE. 2011, pp. 230–233.
- [161] AK Spencer, BA Siddiqui, and JH Thomas. “Cell shape change and invagination of the cephalic furrow involves reorganization of F-actin”. In: *Developmental biology* **402** (2015), pp. 192–207.
- [162] DJ Stephens and VJ Allan. “Light microscopy techniques for live cell imaging”. In: *science* **300** (2003), pp. 82–86.
- [163] D Sweeton et al. “Gastrulation in *Drosophila*: the formation of the ventral furrow and posterior midgut invaginations”. In: *Development* **112** (1991), pp. 775–789.
- [164] K Thorn. “A quick guide to light microscopy in cell biology”. In: *Molecular biology of the cell* **27** (2016), pp. 219–222.
- [165] R Tomer et al. “Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy”. In: *Nature methods* **9** (2012), p. 755.
- [166] C Tudela et al. “TGF-beta3 is required for the adhesion and intercalation of medial edge epithelial cells during palate fusion.” In: *International Journal of Developmental Biology* **46** (2002), pp. 333–336.
- [167] S Tuerkcan, A Alexandrou, and J-B Masson. “A Bayesian inference scheme to extract diffusivity and potential fields from confined single-molecule trajectories”. In: *Biophysical journal* **102** (2012), pp. 2288–2298.
- [168] S Turner, M Cabodi, and H Craighead. “Confinement-induced entropic recoil of single DNA molecules in a nanofluidic structure”. In: *Physical review letters* **88** (2002), p. 128103.

- [169] N Van Kampen. “Itô versus stratonovich”. In: *Journal of Statistical Physics* **24** (1981), pp. 175–187.
- [170] T Vanderleest. “Quantitative Aspects of Interface Remodeling during Germband Extension”. In: (2019).
- [171] CL Vestergaard, PC Blainey, and H Flyvbjerg. “Optimal estimation of diffusion coefficients from single-particle trajectories”. In: *Physical Review E* **89** (2014), p. 022726.
- [172] W Vogt. “Die Einrollung und Streelung der Urmundlippen bei Triton nach Versuchen mit einer neuen Methode embryonaler transplplantation”. In: *Verh Zool Ges* **27** (1922), pp. 49–51.
- [173] AH Voie, D Burns, and F Spelman. “Orthogonal-plane fluorescence optical sectioning: three-dimensional imaging of macroscopic biological specimens”. In: *Journal of microscopy* **170** (1993), pp. 229–236.
- [174] M Von Smoluchowski. “Zur kinetischen theorie der brownischen molekularbewegung und der suspensionen”. In: *Annalen der physik* **326** (1906), pp. 756–780.
- [175] Y Wang and DM Blei. “Frequentist consistency of variational Bayes”. In: *Journal of the American Statistical Association* (2018), pp. 1–15.
- [176] M Weber, M Mickoleit, and J Huisken. “Multilayer mounting for long-term light sheet microscopy of zebrafish”. In: *JoVE (Journal of Visualized Experiments)* (2014), e51119.
- [177] G Wilemski and M Fixman. “General theory of diffusion-controlled reactions”. In: *The Journal of chemical physics* **58** (1973), pp. 4009–4019.
- [178] M Williams et al. “Distinct apical and basolateral mechanisms drive planar cell polarity-dependent convergent extension of the mouse neural plate”. In: *Developmental cell* **29** (2014), pp. 34–46.
- [179] V Wollrab et al. “Still and rotating myosin clusters determine cytokinetic ring constriction”. In: *Nature communications* **7** (2016).
- [180] JA Zallen and JT Blankenship. “Multicellular dynamics during epithelial elongation”. In: *Seminars in cell & developmental biology*. Vol. 19. Elsevier. 2008, pp. 263–270.
- [181] JA Zallen and E Wieschaus. “Patterned gene expression directs bipolar planar polarity in *Drosophila*”. In: *Developmental cell* **6** (2004), pp. 343–355.
- [182] Q Zhang, J Brujić, and E Vanden-Eijnden. “Reconstructing free energy profiles from nonequilibrium relaxation trajectories”. In: *Journal of Statistical Physics* **144** (2011), pp. 344–366.
- [183] Y Zhang et al. “The glucosyltransferase Xiantuan of the endoplasmic reticulum specifically affects E-Cadherin expression and is required for gastrulation movements in *Drosophila*”. In: *Developmental biology* **390** (2014), pp. 208–220.
- [184] A Zumdieck et al. “Stress generation and filament turnover during actin ring constriction”. In: *PloS one* **2** (2007), e696.
- [185] R Zwanzig. “Memory effects in irreversible thermodynamics”. In: *Physical Review* **124** (1961), p. 983.

Acknowledgements and CV

Acknowledgements

I am grateful for the support of many people. First of all Fred Wolf and Stephan Eule without whom this project would not have been possible. I am truly in debt to their scientific guidance, suggestions and comments. I am especially grateful for Stephan Eule's continuous scientific support and his unofficial role as a supervisor he took on. I also want to thank Jörg Großhans for his scientific guidance in the fly project, his explanations and his judgments of a physicist's ideas on biological problems. I am further thankful for all the experimental data provided by Deqing Kong and his willingness to perform new experiments when new questions arose. I also want to thank Lars Reichel who already was on the move when I started the project but whose data extraction pipeline lay the foundations to my work. I am particularly grateful to Matthias Häring who took over Lars' position and always took his time to judge and comment on new ideas I had. For proofreading and valuable suggestions for improvement, I am deeply indebted to Matthias Häring and Erik Schultheis. Finally, I want to thank all current and former members of the Wolf group, who gave critical input on my work.