Thermodynamics and optimal control of active systems

Mikkel Peter Andersen

Supervised by Karel Proesmans

Master's thesis

Niels Bohr Institute
University of Copenhagen
August 3, 2023

"Another reason you should turn
Your attention to the motes that drift and tumble in the light:
Such turmoil means that there are secret motions, out of sight,
That lie concealed in matter. For you'll see the motes careen
Off course, and then bounce back again, by means of blows unseen,
Drifting now in this direction, now that, on every side.
You may be sure this starts with atoms; they are what provide
The base of this unrest. For atoms are moving on their own,
Then small formations of them, nearest them in scale, are thrown
Into agitation by unseen atomic blows,
And these strike slightly larger clusters, and on and on it goes –
A movement that begins on the atomic level, by slight
Degrees ascends until it is perceptible to our sight,
So that we can behold the dust-motes dancing in the sun,
Although the blows that move them can't be seen by anyone."

Lucretius, "On the Nature of Things", First century BC

Abstract

Active systems convert energy from the environment into directed motion at the level of the individual constituents and are thus driven far from equilibrium. Examples include bacteria, molecular motors and flocks of birds. The potential for real-world applications from these systems is large and includes drug delivery [1] and nano/micro-robots with a multitude of uses [2, 3, 4], which in turn calls for the need to build a novel, predictive framework for the control of such systems. The control encompasses minimizing the dissipation of the system while driving it with a protocol from an initial to a final state. As these systems operate far from equilibrium, the framework of equilibrium thermodynamics is insufficient to build such an optimal control framework. Here we then utilize recent developments in stochastic thermodynamics and non-equilibrium response theory to build a framework that can predict the optimal protocols for minimizing the dissipated work of active systems which operate arbitrarily far from equilibrium. We start by presenting an in-depth analysis of the optimal control of a passive Brownian particle in a harmonic potential which provides the fundamentals for the rest of the thesis. Inspired by this approach, we then construct a framework which can predict the optimal protocol for the dissipated work for an active system. We find that the agreement between the exact, numerical solution and the approximated, analytical solution is remarkably good, even for small protocol durations. Specifically, we show that the dissipated work for an active Ornstein-Uhlenbeck particle (AOUP) follows the same dissipation as its passive counterpart, agreeing within the same regime of protocol durations, but with a different friction coefficient. This work extends and bridges the optimal control framework developed by Sivak and Crooks [5] to encompass active systems and the framework developed by Davis et al. [6] to also include the dissipated work, thus broadening the tools available to optimally control active and passive systems.

Acknowledgements

This thesis is the result of a year's work and would not have been possible without the expertise and guidance of my supervisor Karel Proesmans. It has been a real pleasure to work with you and take part in your fascinating research. A big thank you.

I would also like to give a heartfelt thank you to my family and friends who supported me during the process.

Last but not least, a thanks to Carsten Dominik for developing the CDLaTeX-package making it much, much faster and easier to write math and LaTeX in Emacs.

Contents

1 Introduction

Life inherently operates out of equilibrium. Whether it is a bacterium, a flock of birds, the firing of neurons in the brain or a molecular machine such as kinesin, what is central to all of them is that they take in energy from the environment and convert it to useful work while dissipating heat in the process. What happens when these biological systems stop this process is what we naturally refer to as death, or "the decay into thermodynamical equilibrium" as fittingly described by Schrödinger [7]. This tacitly implies that equilibrium thermodynamics is incapable of describing living systems, thus demanding a framework of non-equilibrium thermodynamics. Historically, the field of thermodynamics has been confined to describing macroscopic systems at equilibrium and the transitions between different states of equilibrium. However, during the last couple of decades the theoretical framework of stochastic thermodynamics has emerged and produced important and groundbreaking results for non-equilibrium systems, particularly small scale systems [8], [9]. These include for example colloidal particles, biological macromolecules and molecular motors [10], [5].

One step in trying to understand living systems is to elucidate the design principles governing them, specifically how these systems operate out of equilibrium most efficiently, for example by minimizing excessive dissipation of heat into the environment [5]. For instance the rotary molecular motor F1-ATPase has been shown to sometimes operate at 100% efficiency[11, 12]. These motors are often more efficient than human-made motors, prompting us to understand their design principles in order to engineer more energy-efficient devices.
Constructing a theoretical framework to understand and control such systems is the focus of this work, particularly how to optimally drive a process out of equilibrium with externally controllable protocols. We can imagine that an experimentalist has access to such protocols and can change them over time from an initial state to a final state and thereby drive the system between those two states. An example of a protocol could be the trap stiffness of an optical tweezer or the position of an optical bead trapped in the tweezer [13]. An *optimal protocol* is then the specific protocol varied over time that minimizes the dissipation (for example the work or heat) to the environment [13].

A specific class of non-equilibrium systems that has gained alot of attention recently is active matter [14, 15, 16, 17, 18, 19], which covers a broad range of active systems that operate far from thermodynamic equilibrium. What characterizes all of them is that the individual constituents of the system convert energy from the surroundings into useful work, propelling themselves forward or generating other types of directed motion. As a consequence these systems display emergent phenomena that are not found at equilibrium such as collective motion [20, 21], active turbulence [22] and motility-induced phase separation (MIPS) [23]. Examples span different length scales and include both natural and artificial systems such as social animal groups (flocking birds, human crowds) [24], swarming bacteria [25, 26], colloids activated by light [27] and cells moving during wound healing [28]. There is considerable interest in developing a framework for how to optimally control active matter, both for understanding the underlying principles and the physics of active systems, but also importantly for the potential real-world applications this may provide. These include drug delivery targeted at specific organs [1], nano/micro-robots with complex swarming behavior [2, 3, 4] as well as novel synthetic materials [29].

So far in the literature on active matter, two central questions have been: how far from equilibrium are active-matter systems and what can we do far from equilibrium? [30] The former entails finding quantita-

tive measures of irreversibility (as done in for example [17]), whereas the latter, which is also the focus of this thesis, is centered around applications and control of active matter.

In this work, we utilize stochastic thermodynamics and response theory to develop a theoretical framework which is capable of predicting the optimal protocols when driving both a passive and an active system. For the active system, we build upon the framework recently developed by Davis et. al [6]. We construct our framework in general terms and for testing our results on model systems, we specifically consider a passive Brownian particle in a harmonic potential as the model system for the passive case and an active Ornstein-Uhlenbeck particle (AOUP) [31, 32, 33, 34] as the model system for the active case.

We end this introduction with a few practical remarks:
The thesis is written for a target audience of fellow physics students who have a basic knowledge of thermodynamics, statistical mechanics, stochastic processes and probability theory. All code and calculations done with mathematical software are *not* included in the appendix, but can be found at GitHub at this link: MasterThesis

1.1 Outline

Aiming at presenting a broad understanding of the theoretical tools that underlie the subject, section 2 describes the basics of stochastic thermodynamics, the Langevin formalism, the fluctuation-dissipation relation (FDR) and linear response theory.
In section 3, we derive the optimal protocols for a passive system, further implementing this framework for a Brownian particle in a harmonic potential in one dimension in section 4. This system is then simulated with the optimal protocols in section 5 and compared with the analytical predictions.
We then move on to derive an optimal control framework for an active system in section 6. Following this, we then use the optimal control framework derived in the previous section on an active system, an active Ornstein-Uhlenbeck particle in one dimension, in section 7. Lastly, in section 8, we numerically simulate this particle and use the optimal protocols derived in the previous section to compare the numerically calculated work with the analytical prediction. We then discuss the results obtained in this thesis, the challenges and future perspectives in section 9, and then end the thesis with a conclusion in section 10.

2 Theory

We start by giving an overview of the theoretical framework used in this thesis.
First we present the basic concepts of stochastic thermodynamics and establish a formal introduction to the Langevin equation. From here we discuss the concepts of heat and work in stochastic thermodynamics leading to a description of linear response theory, from which we derive the fluctuation-dissipation theorem.

2.1 Basics of stochastic thermodynamics

During the last couple of decades, stochastic thermodynamics has emerged as a new theoretical framework describing the thermodynamics of small systems where, at this mesoscopic length scale, thermal fluctuations become relevant for the system. Usual concepts from equilibrium thermodynamics such as heat, work and entropy are defined at the individual level of the system, for example a particle immersed in a heat bath [10]. The particle will fluctuate around equilibrium values (for example the position) due to the thermal fluctuations arising from the bath resulting in random motion, i.e. stochastic behavior. Such a particle will trace out a specific trajectory due to this stochasticity and the heat, work and entropy are each assigned to this trajectory.

It is not necessary to assume that the individual particle is at equilibrium with the heat bath, which makes us able to describe non-equilibrium behavior. We do however need to assume that the heat bath is at equilibrium which is usually a fair assumption if the heat bath relaxes to equilibrium fast enough compared to the timescale of the dynamics of the particle, meaning that at this timescale the heat bath is approximately at equilibrium [10].

The particle can be driven out of equilibrium in different ways. One way is by external manipulations such as controlling it with an optical tweezer [10], which is what we will primarily consider in this thesis. Active systems, on the other hand, are furthermore driven out of equilibrium by their own self-propulsion which further increases the complexity of the system.
When we model a non-equilibrium system in stochastic thermodynamics, we primarily use two different frameworks depending on whether the state of the system is continuous (for example the position of the particle) or discrete (for example the configurational changes of a biological molecule, such as the reaction scheme of ATP hydrolysis) [10]. If the state is continuous one method is to use the so-called Langevin equation. A method for describing a discrete system is by using a master equation which gives the time evolution of the system in terms of the transition rates between discrete states. In this thesis, we will deal with systems described by continuous variables which makes the Langevin equation the natural choice.

2.2 The Langevin equation

The Langevin equation is a stochastic differential equation modelling the motion of a particle submerged in an aqueous solution. The particle will display so-called Brownian motion as it collides with the water molecules. To an observer it will look like the particle is jiggling around, which is something that can be seen by for example scattering pollen grains or other small particles on the surface of the solution, for example water or other solutions. The dynamics in 1 dimension can be described by:

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = -\zeta v + \eta(t) \tag{1}$$

where m is the mass of the particle, v its velocity and ζ is a constant and the last term, $\eta(t)$, is a random force which will be elaborated on below.

The particle experiences the forces on the right-hand side of the equation which are, respectively, a friction force proportional to the velocity v of the particle and a force $\eta(t)$ due to random kicks from the water molecules [35]. The Langevin equation can essentially be seen as the equivalent of Newton's 2nd law for this particular type of particle. If there is no net velocity, the particle gets kicked around uniformly which means it is fair to assume that the random force should be 0 on average, ie. we require that $\langle \eta(t) \rangle = 0$. Furthermore, as each collision is approximately instantaneous, the random force kicking the particle at time t and at another time t' are assumed to be uncorrelated, i.e. we require that $\langle \eta(t)\eta(t') \rangle = \Gamma\delta(t - t')$, where Γ is a constant which determines the strength of the random force [36]. This constant can be shown to be related to the friction, $\gamma = \frac{\zeta}{m}$, by solving the equation (1) and then calculating the mean square velocity of the particle and taking the long time limit $t \to \infty$, to obtain [36]

$$\lim_{t\to\infty} \left\langle v(t)^2 \right\rangle = \frac{\Gamma}{2m^2\gamma} \tag{2}$$

In this limit the system has reached equilibrium and the equipartition theorem then holds:

$$\lim_{t\to\infty} \frac{1}{2} m \left\langle v(t)^2 \right\rangle = \frac{1}{2} k_B T \tag{3}$$

Comparing (2) and (3) then leads to the relation

$$\frac{\Gamma}{2m\gamma} = k_B T \tag{4}$$

which relates the constant Γ that determines the fluctuations in the system with the constant γ which determines the friction (which always involves dissipation) [36]. This relation between fluctuations and dissipation is required for the system to reach thermal equilibrium at long times, as the dissipation tends to drive the system to a dead state while the fluctuations keep the system alive, thus balancing each other out and making it physically possible to reach an equilibrium distribution [35], [36]. This is a simple form of the fluctuation-dissipation theorem which we will encounter in more general terms later in the thesis.

The case that we will deal with most often involves the particle being subjected to an external force and if the particle is small enough (or if we consider long enough time scales), its motion will be overdamped as the viscous forces of the medium will dominate the inertial forces [37]. Modifying (1) accordingly by setting the left-hand side to zero and moving terms around, the overdamped Langevin equation will take the following form [10]:

$$v = \frac{\mathrm{d}x}{\mathrm{d}t} = \mu F(x,t) + \xi(t) = \mu\left(-\frac{\partial \epsilon(x,\lambda)}{\partial x} + f(x,t)\right) + \xi(t) \tag{5}$$

where the first term expanded from $F(x,t)$ is a conservative force (the derivative of a potential) and the second term a non-conservative force. Here the energy $\epsilon(x,\lambda)$ depends on the external protocol $\lambda(t)$, which is varied between two states λ_0 and λ_f. This is the quantity we will optimize later on. Furthermore, we define $\mu = 1/\zeta$ as the mobility and we rescale the original noise $\eta(t)$ as $\xi(t) = \frac{1}{\zeta}\eta(t)$. The correlations for the rescaled noise are now given as

$$\langle \xi(t)\xi(t')\rangle = \frac{\Gamma}{\zeta^2}\delta(t-t') = \frac{2k_BT}{\zeta}\delta(t-t') = 2\mu k_BT\delta(t-t') \tag{6}$$

Here we can also define $D = \mu k_B T$ which is known as the Einstein relation and relates the diffusion coefficient D with the mobility μ. Equation (5) describes the dynamics of a colloidal particle and can be regarded as the paradigmatic model for stochastic thermodynamics [9].

2.3 Heat and work in stochastic thermodynamics

Solutions to the Langevin equation in (1) generate stochastic trajectories, $\{\boldsymbol{x(t)}\}$, and in stochastic thermodynamics we define heat and work along these trajectories, which means that heat and work also will be stochastic variables [10]. Sekimoto [38] was one of the first to connect the Langevin equation with thermodynamics and to define heat and work along these trajectories, of which we will give a brief account in the following.

We start by considering the work along the trajectory which is made up of two contributions:

$$\mathrm{d}w = \frac{\partial\epsilon(x,\lambda)}{\partial\lambda}\mathrm{d}\lambda + f(x,t)\circ\mathrm{d}x \tag{7}$$

where $\circ$ denotes the Stratonovich product. The first term is the contribution to the incremental work due to a manipulation from the manipulation protocol λ and the second term is the contribution to the incremental work from an external non-conservative force (such as a driving force) [10]. Both represent forces that are external to the heat bath.

In order to formulate a 1st law of stochastic thermodynamics, we consider the incremental change in energy along a trajectory

$$\mathrm{d}\epsilon(x,\lambda) = \mathrm{d}w - \mathrm{d}q \tag{8}$$

By inserting (7) into (8) and using the definition of $\mathrm{d}\epsilon$ we have

$$\mathrm{d}\epsilon(x,\lambda) = \frac{\partial}{\partial x}\epsilon(x,\lambda)\circ\mathrm{d}x + \frac{\partial\epsilon(x,\lambda)}{\partial\lambda}\mathrm{d}\lambda \tag{9}$$

$$= \frac{\partial\epsilon(x,\lambda)}{\partial\lambda}\mathrm{d}\lambda + f(x,t)\circ\mathrm{d}x - \mathrm{d}q \tag{10}$$

$$\tag{11}$$

which tells us that the heat increment $\mathrm{d}q$ must then take the form

$$\mathrm{d}q = -\frac{\partial}{\partial x}\epsilon(x,\lambda)\circ\mathrm{d}x + f(x,t)\circ\mathrm{d}x = F(x,t)\circ\mathrm{d}x \tag{12}$$

These two expressions for the incremental heat and work (equations (12) and (7)) then constitute the 1st law of thermodynamics at trajectory-level. It is possible to derive expressions for the entropy production along a trajectory, however this is outside the scope of this work.

2.4 Linear Response Theory

A first step in describing systems out of equilibrium might be to try to "poke" a system that is initially in equilibrium and observe how it responds. In this way the system is gently pushed out of equilibrium, and we can then use suitable approximation techniques to understand the perturbed system, as long as the perturbations are small. This is the general idea in linear response theory. We perturb the system out of equilibrium and then do a series expansion in the perturbation variable and keep the linear terms.

If the system is perturbed by a weak, time-dependent field $f(t)$, we can describe it with a Hamiltonian given by

$$H(X) = H_0(X) - f(t)X, \tag{13}$$

where X is a given observable and $H_0(X)$ is the equilibrium Hamiltonian. The system is initially in thermal equilibrium and has the Hamiltonian H_0 and is then perturbed out of equilibrium by the field $f(t)$ at time t. At equilibrium the observable has the ensemble average $\langle X \rangle_0$. We are interested in describing the non-equilibrium ensemble average $\langle X \rangle$, which is caused by perturbing the system out of equilibrium.

In stochastic thermodynamics, the external field $f(t)$ is usually a so-called manipulation protocol $\lambda(t)$, which is an experimental protocol that an outside experimenter has access to. An example could be changing the stiffness of an optical tweezer, or moving the tweezer spatially.

Instead of the notation in (13), we use the following notation:

$$\epsilon_x(\lambda) = \epsilon_{x,0} - \sum_{\alpha=1}^{r} \lambda_\alpha X_{\alpha,x}, \tag{14}$$

where ϵ_x is the energy associated with the state x, the r different ways of manipulating the system is encoded in the summation and $X_{\alpha,x}$ is a given observable which is a function of the state x.
If we then expand an observable $\left\langle X_\beta(t) \right\rangle$ in powers of the protocol λ_α, we get:

$$\left\langle X_\beta(t) \right\rangle = \langle X_\alpha \rangle_0 + \int_{-\infty}^{\infty} \sum_{\alpha=1}^{r} \mathcal{K}_{\beta,\alpha}(t,t')\lambda_\alpha(t')\mathrm{d}t' + \mathcal{O}(\lambda_\alpha^2) \tag{15}$$

The first term $\langle X_\alpha \rangle_0$ is the zeroth order term which is independent of the protocol λ, while the second term is the first-order term which depends linearly on λ. Here $\mathcal{K}$ is the so-called *linear response function* which weighs the perturbation λ over the entire duration of the perturbation up until and including time t where we observe the system.

To gain another perspective on $\mathcal{K}$, consider for a moment the situation where we just have one protocol parameter and for simplicity take the time interval to be $-\infty$ to t so that (15) becomes

$$\langle X(t) \rangle = \langle X \rangle_0 + \int_{-\infty}^{t} \mathcal{K}(t,t')\lambda(t')\mathrm{d}t' + \mathcal{O}(\lambda^2) \tag{16}$$

Specifically, $\langle X(t)\rangle$ is a functional of $\lambda(t)$. Now we discretize time into N parts so that $t_N = Nh$ where h is the spacing between times. We then have a corresponding value of λ at each discretized timestep and write this as $\boldsymbol{\lambda} = \{\lambda(t)\} = \{\lambda(t_0), \lambda(t_1), \ldots, \lambda(t_N)\} = \{\lambda_0, \lambda_1, \ldots, \lambda_N\}$. Now using a Taylor expansion analogue for functionals [39, 40], we expand $\langle X(t)\rangle[\boldsymbol{\lambda}]$ around $\lambda = 0$ to obtain

$$\langle X(t)\rangle[\boldsymbol{\lambda}] = \langle X(t)\rangle[\{\lambda(t_0), \lambda(t_1), \ldots, \lambda(t_N)\}] \tag{17}$$

$$= \langle X(t)\rangle[\{\lambda(t=0)\}] + \sum_{i=-\infty}^{N} \left(\frac{\partial \langle X\rangle}{\partial \lambda_i}\right)_{\lambda_i=0} \lambda_i + \mathcal{O}(\lambda_i^2) \tag{18}$$

And taking the continuum limit $h \longrightarrow 0$, we obtain

$$\langle X(t)\rangle[\boldsymbol{\lambda}] = \langle X(t)\rangle_0 + \int_{-\infty}^{t} \left(\frac{\delta \langle X(t)\rangle}{\delta \lambda(t')}\right) \lambda(t')\mathrm{d}t' + \mathcal{O}(\lambda(t')^2) \tag{19}$$

where in this limit the partial derivatives turn into the functional derivative [41] and we can identify the linear response function from (16) as

$$\mathcal{K}(t, t') = \frac{\delta \langle X(t)\rangle}{\delta \lambda(t')} \tag{20}$$

This has the interpretation that the response function is the rate of change of $\langle X(t)\rangle$ due to the change in the external perturbation $\lambda(t')$ at time t' [39].

The linear response function $\mathcal{K}$ furthermore has some properties that are important to note:

- **Causality**: There can be no response from the system if the perturbation from the protocol has not been applied yet. We express this formally as $\mathcal{K}_{\beta,\alpha}(t, t') = 0$ for $t' > t$ [10]. This means we can adjust the integration limits in (15) to $-\infty$ to t, where t is the time we observe the system.
- **Time invariance**: The system must be invariant under time translations as only the time difference between applying the protocol and observing the system is relevant. We formalize this as

$$\mathcal{K}_{\beta,\alpha}(t, t') = \mathcal{K}_{\beta,\alpha}(t - t') \tag{21}$$

As we saw earlier in (4), there is a fundamental connection between fluctuations and dissipation in systems that are in equilibrium.
We will now derive a generalization of this concept in the framework of linear response theory. Specifically, we will show that the fluctuations in equilibrium (quantified by equilibrium correlation functions) are intimately connected to the non-equilibrium response of a system perturbed by a small external driving force in the form of a manipulation protocol [42].

2.4.1 Deriving the fluctuation-dissipation theorem

In the following we derive the fluctuation-dissipation theorem in the context of stochastic thermodynamics, following Peliti and Pigolotti [10].
We assume that the average $\langle X_\alpha \rangle_0 = 0$ for all α without loss of generality. It is possible to derive it with the averages, but we derive it in a more simple way here for the sake of brevity. In the following, $\langle \dots \rangle_0$ means an ensemble average with $\lambda = 0$ and $\langle \dots \rangle_{\{\lambda(t)\}}$ denotes a non-equilibrium ensemble average under the entire protocol history $\{\lambda(t)\}$.

We start by considering the manipulation protocol $\lambda(t) = \lambda_0 \theta(-t)$ where $\theta(t)$ is the Heaviside step function defined as

$$\theta(t) = \begin{cases} 1, & t > 0 \\ 0, & t \leq 0 \end{cases} \tag{22}$$

This type of protocol means that the manipulation acts as a constant λ_0 from $t = -\infty$ and is then turned off at $t = 0$. This implies that at $t = 0$ the probability distribution of the system is the Boltzmann distribution $p^{eq}(\lambda)$ for $\lambda = \lambda_0$ and with Hamiltonian $\epsilon_x(\lambda_0) = \epsilon_{x,0} - \sum_\alpha \lambda_{\alpha,0} X_{\alpha,x}$.

The Boltzmann distribution is given by

$$p_i^{eq} = \frac{\exp(-\epsilon_i / k_B T)}{Z} \tag{23}$$

where ϵ_i is the energy associated with microstate i and Z is the partition function given by

$$Z = \sum_i \exp(-\epsilon_i / k_B T) \tag{24}$$

The partition function is related to the free energy as

$$F = -k_B T \ln Z. \tag{25}$$

If we then write Z as

$$\exp(-F / k_B T) = Z, \tag{26}$$

we can write the Boltzmann distribution as

$$p_i^{eq} = \frac{\exp(-\epsilon_i / k_B T)}{\exp(-F / k_B T)} = \exp((-\epsilon_i + F)/k_B T) \tag{27}$$

Using this way of writing the distribution, we then have

$$p_x(t = 0; \{\lambda(t)\}) = p_x^{eq}(\lambda_0) = \exp\left(\frac{F(\lambda_0) - \epsilon_x(\lambda_0)}{k_B T}\right) = \exp\left(\frac{F(\lambda_0) - \epsilon_{x,0} + \sum_\alpha \lambda_{\alpha,0} X_{\alpha,x}}{k_B T}\right) \tag{28}$$

Here $p_x(t = 0; \{\lambda(t)\})$ means the distribution at $t = 0$ under the influence of the entire history of $\lambda(t)$ (which in this case simplifies to λ_0).

We now expand $F(\lambda_0)$ to first order in λ_0 as $F(\lambda_0) = F(0) + \sum_\alpha \lambda_\alpha \frac{\partial F(\lambda_0)}{\partial \lambda_\alpha}\Big|_{\lambda_0=0} + \mathcal{O}\left(\lambda_\alpha^2 \frac{\partial^2 F(\lambda_0)}{\partial \lambda_\alpha^2}\Big|_{\lambda_0=0}\right)$.
We then keep the linear terms, and at the same time note the following:

$$\frac{\partial F(\lambda_0)}{\partial \lambda_\alpha}\bigg|_{\lambda_0=0} = -k_B T \frac{\partial \ln Z(\lambda_0)}{\partial \lambda_\alpha}\bigg|_{\lambda_0=0} \tag{29}$$

$$= -\frac{k_B T}{Z}\frac{\mathrm{d}Z(\lambda_0)}{\mathrm{d}\lambda_\alpha}\bigg|_{\lambda_0=0} \tag{30}$$

$$= -\frac{k_B T}{Z}\frac{\mathrm{d}}{\mathrm{d}\lambda_\alpha}\sum_x \exp(-\epsilon_x(\lambda_0)/k_B T)\big|_{\lambda_0=0} \tag{31}$$

$$= -\frac{k_B T}{Z}\sum_x \frac{\mathrm{d}}{\mathrm{d}\lambda_\alpha}\exp\left((-\epsilon_{x,0} + \sum_\beta \lambda_{\beta,0} X_{\beta,x})/k_B T\right)\bigg|_{\lambda_0=0} \tag{32}$$

$$= -\frac{k_B T}{Z}\sum_x \exp(-\epsilon_x(\lambda_0)/k_B T)\left(\frac{\mathrm{d}}{\mathrm{d}\lambda_\alpha}\sum_\beta \lambda_\beta X_{\beta,x}/k_B T\right)\bigg|_{\lambda_0=0} \tag{33}$$

$$= -\frac{1}{Z}\sum_x \exp(-\epsilon_x(\lambda_0)/k_B T)\left(\sum_\beta \delta_{\beta,\alpha} X_{\beta,x}\right)\bigg|_{\lambda_0=0} \tag{34}$$

$$= -\frac{1}{Z}\sum_x \exp(-\epsilon_x(\lambda_0)/k_B T) X_{\alpha,x}\big|_{\lambda_0=0} \tag{35}$$

$$= -\langle X_\alpha \rangle_0 , \tag{36}$$

where we in (33) used the chain rule and $\delta_{\beta,\alpha}$ in (34) is the Kronecker delta. Now since we assumed this to be 0, this means we are just left with the $F(0)$ in the expansion.
We thus obtain

$$p_x(0;\{\lambda(t)\}) \approx \exp\left(\frac{F(0) - \epsilon_{x,0} + \sum_\alpha \lambda_{\alpha,0} X_{\alpha,x}}{k_B T}\right) \tag{37}$$

$$= \exp\left(\frac{F(0) - \epsilon_{x,0}}{k_B T}\right)\exp\left(\frac{\sum_\alpha \lambda_{\alpha,0} X_{\alpha,x}}{k_B T}\right) \tag{38}$$

$$\approx p_x^{eq}(t=0)\left(1 + \frac{1}{k_B T}\sum_\alpha \lambda_{\alpha,0} X_{\alpha,x}\right) \tag{39}$$

where we used the first order Taylor expansion $\exp(x) \approx 1 + x$ in (39).
We now find an expression for the probability distribution $p_x(t;\{\lambda(t)\})$ as we are interested in an expression for $\left\langle X_\beta(t)\right\rangle_{\{\lambda(t)\}}$.
Using what we just derived in (37), we get:

$$p_x(t) = \sum_{x'} G_{xx'}(t) p_{x'}(0;\{\lambda(t)\}) \approx \sum_{x'} G_{xx'}(t) p_{x'}^{eq}(0)\left(1 + \frac{1}{k_B T}\sum_\alpha \lambda_{\alpha,0} X_{\alpha,x'}\right) \tag{40}$$

Here $G_{xx'}(t)$ denotes the Green function $G_{xx'}(t,0) = p_{x;t|x';t=0}$, where we in general have $G_{xx'}(t,t') = p_{x;t|x';t'}$ [10].
We then have that

$$p_x(t) = \sum_{x'} p_{x;t|x';t'} p_{x'}(t') = \sum_{x'} G_{xx'}(t,t') p_{x'}(t') \tag{41}$$

If we have the initial condition $p_{x'}(t=0;\{\lambda(t)\})$ the solution of $p_x(t)$ is given by [10]:

$$p_x(t) = \sum_{x'} G_{xx'}(t) p_{x'}(t=0;\{\lambda(t)\}) \tag{42}$$

We furthermore have that

$$p_x^{eq}(0) = \sum_{x'} G_{xx'}(t) p_{x'}^{eq}(0) \tag{43}$$

since if the system starts in an equilibrium distribution and propagate forwards, it will still be in the equilibrium distribution at a later time.

Using this relation on the first term in (40), we get

$$p_x(t) \approx p_x^{eq}(0) + \frac{1}{k_B T} \sum_{x'} G_{xx'}(t) \sum_{\alpha} p_{x'}^{eq}(0) \lambda_{\alpha,0} X_{\alpha,x'} \tag{44}$$

We now turn to calculating (15). By definition the ensemble average is given by

$$\begin{aligned}
\left\langle X_\beta(t) \right\rangle_{\{\lambda(t)\}} &= \sum_x X_{\beta,x} p_x(t) && (45)\\
&\approx \sum_x X_{\beta,x} \left(p_x^{eq}(0) + \frac{1}{k_B T} \sum_{x'} G_{xx'}(t) \sum_{\alpha} p_{x'}^{eq}(0) \lambda_{\alpha,0} X_{\alpha,x'} \right) && (46)\\
&= \left\langle X_{\beta,x} \right\rangle_0 + \frac{1}{k_B T} \sum_x \sum_{x'} X_{\beta,x} G_{xx'}(t) \sum_{\alpha} p_{x'}^{eq}(0) \lambda_{\alpha,0} X_{\alpha,x'} && (47)\\
&= \frac{1}{k_B T} \sum_{\alpha} \lambda_{\alpha,0} \sum_x \sum_{x'} X_{\beta,x} G_{xx'}(t) X_{\alpha,x'} p_{x'}^{eq}(0) && (48)\\
&= \frac{1}{k_B T} \sum_{\alpha} \lambda_{\alpha,0} \sum_x \sum_{x'} X_{\beta,x} X_{\alpha,x'} p_{x;t|x';0} p_{x'}^{eq}(0) && (49)\\
&= \frac{1}{k_B T} \sum_{\alpha} \lambda_{\alpha,0} \sum_x \sum_{x'} X_{\beta,x} X_{\alpha,x'} p_{x;t,x';0} && (50)\\
&= \frac{1}{k_B T} \sum_{\alpha} \lambda_{\alpha,0} \left\langle X_\beta(t) X_\alpha(0) \right\rangle_0 && (51)
\end{aligned}$$

where we in (50) used the definition of conditional probability.
Comparing this expression with the original expression in (15), we can identify the following relation

$$\int_{-\infty}^{\infty} \sum_{\alpha} K_{\beta\alpha}(t-t') \lambda_\alpha(t') \mathrm{d}t' = \frac{1}{k_B T} \sum_{\alpha} \lambda_{\alpha,0} \left\langle X_{\beta,x}(t) X_{\alpha,x}(0) \right\rangle \tag{52}$$

Since we defined our protocol to be $\lambda(t) = \lambda_0\theta(-t)$, the LHS is 0 from $t = 0$ to $t = \infty$ and we get

$$\int_{-\infty}^{0} K_{\beta\alpha}(t-t')\mathrm{d}t' = \frac{1}{k_B T}\left\langle X_{\beta,x}(t)X_{\alpha,x}(0)\right\rangle \tag{53}$$

Using the substitution $t'' = t - t'$ we can rewrite the integral

$$\int_{-\infty}^{0} K_{\beta\alpha}(t-t') = \int_{\infty}^{t} K_{\beta\alpha}(t'')(-\mathrm{d}t'') = \int_{t}^{\infty} K_{\beta\alpha}(t'')\mathrm{d}t'' \tag{54}$$

Substituting this into (53) and differentiating both sides, we get

$$K_{\beta\alpha}(t) = -\frac{\theta(t)}{k_B T}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle X_\beta(t)X_\alpha(0)\right\rangle_0 = -\frac{\theta(t)}{k_B T}\frac{\mathrm{d}}{\mathrm{d}t}C_{\beta\alpha}(t) \tag{55}$$

where we introduce the time-dependent equilibrium correlation function, $C_{\beta\alpha}(t)$, and where the Heaviside function makes sure that $K_{\beta\alpha}(t)$ vanishes for $t < 0$.

This result is known as the *fluctuation-dissipation theorem* and relates spontaneous fluctuations in the equilibrium system at different times (as measured by the correlation function) with the response of a non-equilibrium system to an external manipulation (as measured by the response function) [10], [39] - in the linear regime (close to equilibrium). This important theorem gives us a way of predicting the non-equilibrium behavior of a system by probing the equilibrium fluctuations [43], which is an observable parameter throughout experimental setups.

We have presented the basic theoretical tools used in the thesis. In the next section, we will develop the framework for finding the optimal protocols for a system in the linear response regime. The Langevin equation will be the basis for the dynamics as we will start by considering the work produced along a trajectory of the system, and then use linear response theory to obtain an expression for the average work rate in terms of equilibrium correlation functions and the speeds of the protocols.

3 Optimal protocols

In this section, we derive the optimal protocol framework which serves as a basis for much of the work in the rest of the thesis. This entails considering the work along a trajectory of the system when the system is perturbed by an external protocol, and then using linear-response theory to obtain an expression for the average work rate.

When driving a system out of equilibrium, we are interested in finding the optimal ways of doing so in order to minimize excessive dissipation.
Driving a system out of equilibrium usually involves some kind of manipulation of the system. For the case of a Brownian particle in a harmonic potential, the manipulation could be shifting the harmonic potential across space or increasing the strength of the potential. There are many different ways of performing these manipulation protocols and the optimal protocols are defined as those that minimize some specific cost function. In our situation, we choose this cost function to be the average dissipated work [10], in line with previous work [5, 13, 44].

The setup is such that we consider a system that can be manipulated by a collection of potentially different protocols $\{\lambda_\alpha(t)\}$ with $\alpha \in \{1, \dots, r\}$. We start in equilibrium at time $t = t_0$ with protocol parameters $\lambda(t_0)$ and then drive the system with the protocols to a final state at $t = t_f$ with protocol parameters $\lambda(t_f)$. In equilibrium the macroscopic state of the system is specified by the Boltzmann distribution and is completely determined by the values of the protocol control parameters [5], for example $\lambda(t_0)$ or $\lambda(t_f)$. In conventional thermodynamics only quasistatic processes are fully describable [45], and when we consider realistic, finite-time processes and drive them between two states, some work will be dissipated to the environment as a consequence. This dissipated work is given by [10, 46, 47]

$$W_{diss} = W - \Delta F = W - (F(\lambda_f) - F(\lambda_0)) \geq 0 \tag{56}$$

Here the change in free energy is the change in energy between the initial and final states, given that the system has had time to reach equilibrium after the protocol has been switched off. Driving the system infinitely slowly (ie. quasistatically) the amount of work is then equal to the free energy difference which means that W_{diss} is then equal to 0 [10]. Out of equilibrium, on the other hand, the probability distribution is no longer the Boltzmann distribution as it depends on the full history of the protocol which we denote by $\{\lambda(t)\}$.

To derive an expression for the average dissipated work, we first consider the work increment (7). Instead of one manipulation protocol λ, we now have more than one, and we also assume that the work can only change due to the manipulation protocol:

$$\mathrm{d}w = \sum_\alpha \frac{\partial \epsilon(x, \lambda_\alpha)}{\partial \lambda_\alpha} \mathrm{d}\lambda_\alpha = \sum_\alpha \frac{\partial \epsilon(x, \lambda_\alpha)}{\partial \lambda_\alpha} \frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t} \mathrm{d}t. \tag{57}$$

Rearranging and taking the ensemble average, we have:

$$\frac{\mathrm{d}\,\langle w \rangle_{\{\lambda(t)\}}}{\mathrm{d}t} = \sum_\alpha \left\langle \frac{\partial \epsilon(x, \lambda_\alpha)}{\partial \lambda_\alpha} \right\rangle_{\{\lambda(t)\}} \dot{\lambda_\alpha} = \sum_\alpha \dot{\lambda_\alpha} \langle X_\alpha(\lambda(t)) \rangle_{\{\lambda(t)\}} \tag{58}$$

where $\langle \dots \rangle_{\{\lambda(t)\}}$ is the non-equilibrium ensemble average with the protocol history $\{\lambda(t)\}$, and we define the observable $X_\alpha \equiv \frac{\partial \epsilon(x, \lambda_\alpha)}{\partial \lambda_\alpha}$.

We now consider the average rate of dissipated work, which is the excess rate of work produced by the external manipulation compared to a reference system at equilibrium [10]:

$$\frac{\mathrm{d}\langle w^{diss}\rangle_{\{\lambda(t)\}}}{\mathrm{d}t} = \sum_{\alpha} \frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t} \langle \Delta X_\alpha(t)\rangle_{\{\lambda(t)\}} \tag{59}$$

where $\Delta X_\alpha(t) = X_\alpha(t) - \langle X_\alpha\rangle^{eq}_{\lambda(t)}$ is the deviation of $X_\alpha(t)$ from its equilibrium value under the manipulation $\lambda(t)$ and $\langle \dots \rangle^{eq}_{\lambda(t)}$ denotes an equilibrium ensemble average with the given value of λ.

We now use the linear response framework we derived earlier in section 2.4 to describe $\Delta X_\alpha(t)$.

We have

$$\langle \Delta X_\alpha(t)\rangle_{\{\lambda(t)\}} = \langle X_\alpha(t)\rangle_{\{\lambda(t)\}} - \langle X_\alpha\rangle^{eq}_{\lambda(t)} \tag{60}$$

$$= \sum_\beta \int_{-\infty}^{t} K_{\alpha\beta}(t-t')\lambda_\beta(t')\mathrm{d}t' - \sum_\beta \int_{-\infty}^{t} K_{\alpha\beta}(t-t')\lambda_\beta(t)\mathrm{d}t' \tag{61}$$

$$= \sum_\beta \int_{-\infty}^{t} K_{\alpha\beta}(t-t')(\lambda_\beta(t') - \lambda_\beta(t))\mathrm{d}t' \tag{62}$$

where

$$K_{\alpha\beta}(t-t') = -\frac{\theta(t-t')}{k_B T}\frac{\mathrm{d}}{\mathrm{d}t} C_{\alpha\beta}(t-t') \tag{63}$$

is the linear response function we derived in (55). We define the correlation function in this context as:

$$C_{\alpha\beta}(t) = \left\langle \Delta X_\alpha \Delta X_\beta \right\rangle^{eq}_{\lambda(t)} = \left\langle (X_\alpha(t) - \langle X_\alpha\rangle^{eq}_{\lambda(t)})(X_\beta(0) - \left\langle X_\beta \right\rangle^{eq}_{\lambda(t)}) \right\rangle^{eq}_{\lambda(t)} \tag{64}$$

where again $\langle \cdots \rangle^{eq}_{\lambda(t)}$ is the equilibrium average at constant control parameter and we also included the averages that we previously neglected in section 2.4.

Substituting eq. (62) into eq. (59), we get to linear order in the perturbation:

$$\frac{\mathrm{d}\langle w^{diss}\rangle_{\{\lambda(t)\}}}{\mathrm{d}t} = \sum_\alpha \frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t} \sum_\beta \int_{-\infty}^{t} K_{\alpha\beta}(t-t')(\lambda_\beta(t') - \lambda_\beta(t))\mathrm{d}t' \tag{65}$$

$$= \sum_\alpha \sum_\beta \frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t} \int_{-\infty}^{t} -\frac{\theta(t-t')}{k_B T}\frac{\mathrm{d}}{\mathrm{d}t} C_{\alpha\beta}(t-t')(\lambda_\beta(t') - \lambda_\beta(t))\mathrm{d}t' \tag{66}$$

$$= -\frac{1}{k_B T} \sum_\alpha \sum_\beta \frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t} \int_{-\infty}^{t} \frac{\mathrm{d}}{\mathrm{d}t} C_{\alpha\beta}(t-t')(\lambda_\beta(t') - \lambda_\beta(t))\mathrm{d}t' \tag{67}$$

Now using integration by parts we get:

$$\frac{\mathrm{d}\langle w^{diss}\rangle_{\{\lambda(t)\}}}{\mathrm{d}t} = -\frac{1}{k_B T}\sum_{\alpha}\sum_{\beta}\frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t}\Big(\Big[C_{\alpha\beta}(t-t')(\lambda_\beta(t')-\lambda_\beta(t))\Big]_{t'=-\infty}^{t'=t} \tag{68}$$

$$-\int_{-\infty}^{t} C_{\alpha\beta}(t-t')\frac{\mathrm{d}}{\mathrm{d}t'}(\lambda_\beta(t')-\lambda_\beta(t))\mathrm{d}t'\Big) \tag{69}$$

$$= \frac{1}{k_B T}\sum_{\alpha}\sum_{\beta}\frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t}\left(\int_{-\infty}^{t} C_{\alpha\beta}(t-t')\frac{\mathrm{d}\lambda_\beta(t')}{\mathrm{d}t'}\mathrm{d}t'\right) \tag{70}$$

where the boundary term at $t' = t$ vanishes trivially and the other term at $t' = -\infty$ vanishes since here there are no correlations between $X_\alpha(t)$ and $X_\beta(0)$ causing the correlation function $C_{\alpha\beta}$ to be zero. This we can fulfill if we for example assume that the system is initially in equilibrium at $t = -\infty$.

Changing variables to $t'' = t - t'$ such that $\mathrm{d}t' = -\mathrm{d}t''$ and evaluating the new boundaries at $t''(t)$ and then taking the limit $\lim_{t\to-\infty} t''(t)$ we get

$$\frac{\mathrm{d}\langle w^{diss}\rangle_{\{\lambda(t)\}}}{\mathrm{d}t} = \frac{1}{k_B T}\sum_{\alpha}\sum_{\beta}\frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t}\int_{0}^{\infty} C_{\alpha\beta}(t'')\frac{\mathrm{d}\lambda_\beta(t')}{\mathrm{d}t'}\bigg|_{t'=t-t''}\mathrm{d}t'' \tag{71}$$

Now Taylor expanding the control parameter velocities in (71) around t using the Taylor expansion in the form of $f(x-h) = f(x) - h\frac{\mathrm{d}f(x)}{\mathrm{d}x} + \frac{h^2}{2}\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} + \mathcal{O}(h^3)$ we get

$$\frac{\mathrm{d}\lambda_\beta(t')}{\mathrm{d}t'}\bigg|_{t'=t-t''} = \frac{\mathrm{d}\lambda_\beta(t')}{\mathrm{d}t'}\bigg|_{t'=t} - t''\frac{\mathrm{d}^2\lambda_\beta(t')}{\mathrm{d}t'^2}\bigg|_{t'=t} + \mathcal{O}\left(t''^2\frac{\mathrm{d}^3\lambda_\beta(t')}{\mathrm{d}t'^3}\bigg|_{t'=t}\right) \tag{72}$$

$$\tag{73}$$

Assuming that the control parameter velocities are slow and change on time scales slower than the correlation time of the system, we can keep the first term in the expansion and get:

$$\frac{\mathrm{d}\langle w^{diss}\rangle_{\{\lambda(t)\}}}{\mathrm{d}t} = \frac{1}{k_B T}\sum_{\alpha}\sum_{\beta}\frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t}\int_{0}^{\infty} C_{\alpha\beta}(t'')\frac{\mathrm{d}\lambda_\beta(t)}{\mathrm{d}t}\mathrm{d}t'' \tag{74}$$

$$= \sum_{\alpha}\sum_{\beta}\frac{\mathrm{d}\lambda_\alpha}{\mathrm{d}t}\zeta_{\alpha\beta}(\lambda(t))\frac{\mathrm{d}\lambda_\beta(t)}{\mathrm{d}t} \tag{75}$$

where we introduce the friction tensor [10], [5]

$$\zeta_{\alpha\beta}(\lambda(t)) = \frac{1}{k_B T}\int_{0}^{\infty} C_{\alpha\beta}(t'')\mathrm{d}t'' \tag{76}$$

which is a time-independent equilibrium matrix that is symmetric, positive semidefinite and varies smoothly except at macroscopic phase transitions [5].

To sum up, we have derived an optimal protocol framework by expressing the dissipation in terms of the friction tensor [5] and the experimentally controllable protocol velocities $\dot{\lambda}_\alpha(t)$, in the regime where the system obeys linear response theory. The optimal protocol can then be obtained by optimizing (75) by using the Euler-Lagrange equations. In the next section we will demonstrate how to use this framework on a simple model.

4 Passive Brownian particle in a harmonic potential

In this section we will consider in detail an example of using the optimal protocol framework, which we derived in the last section. The model system we are considering is the overdamped Langevin equation with a harmonic potential, i.e. a passive Brownian particle diffusing in a harmonic trap. This model can for example describe a small colloidal particle or piece of DNA being trapped in an optical tweezer where we have a controllable parameter κ which sets the strength of the optical trap and a parameter l which controls the position of the harmonic trap (and therefore on average, the particle).
The dynamics are given by:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\mu_P \kappa (x - l) + \xi(t), \tag{77}$$

where κ is the strength of the harmonic potential and l is the position. The noise is delta-correlated and 0 on average, ie. $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = 2\mu_P k_B T \delta(t - t')$.
The energy associated with the force from the harmonic potential is given by:

$$E(x, l, \kappa) = \frac{1}{2}\kappa (x - l)^2 \tag{78}$$

The relevant observables (as defined earlier in (58)) are in this case given by

$$X_l = \frac{\partial E(x, l, \kappa)}{\partial l} = -\kappa (x - l), \tag{79}$$

$$X_k = \frac{\partial E(x, l, \kappa)}{\partial k} = \frac{1}{2}(x - l)^2. \tag{80}$$

$$\tag{81}$$

We are interested in finding the optimal protocols for this system for which we need to calculate the friction coefficient from (76). As these depend on the correlation functions, we will start by obtaining expressions for them.

4.1 Calculation of the correlation functions

To calculate the correlations we start with (64). For this specific problem, this expression will be a $2x2$-matrix since we have 2 variables κ and l, so we will proceed by calculating this entry by entry. However before we do this, it is convenient to calculate some specific quantities beforehand. When doing the correlation matrix calculations, we will encounter terms like $\langle X_l \rangle^{eq} = -\kappa \langle (x - l) \rangle^{eq}$ and $\langle X_\kappa \rangle^{eq} = \frac{1}{2} \left\langle (x - l)^2 \right\rangle^{eq}$ and as we are considering a system in equilibrium the moments $\langle x \rangle^{eq}$, $\left\langle x^2 \right\rangle^{eq}$ etc. can be calculated as:

$$\langle x^n \rangle = \int_{-\infty}^{\infty} x^n p(x) dx \tag{82}$$

where

$$p(x) = \frac{1}{Z} \exp\left(-\frac{E(x, l, \kappa)}{k_B T} \right) = \frac{1}{Z} \exp\left(-\frac{1}{2} \frac{\kappa (x - l)^2}{k_B T} \right) \tag{83}$$

and Z is a normalization constant (the partition function). Imposing the normalization $\int_{-\infty}^{\infty} p(x)dx = 1$ we can find Z to be equal to $\sqrt{\frac{2\pi k_B T}{\kappa}}$. Writing this as $Z = \sqrt{2\pi}\sqrt{\frac{k_B T}{\kappa}} = \sqrt{2\pi}\sigma$, we can then identify the distribution in (83) as a normal distribution with mean l and standard deviation $\sigma = \sqrt{k_B T/\kappa}$:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{\sqrt{\kappa}(x-l)}{\sqrt{k_B T}}\right)^2\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{(x-l)}{\sigma}\right)^2\right) \tag{84}$$

One way to obtain the moments for the normal distribution is by using the moment-generating function for the normal distribution [10, 48]

$$M(t) = \langle \exp(tX) \rangle = \exp(lt) \exp\left(\frac{1}{2}(\sigma t)^2\right) \tag{85}$$

then the n'th moment is given by

$$M_n = \langle X^n \rangle = \left.\frac{\mathrm{d}^n M(t)}{\mathrm{d}t^n}\right|_{t=0} \tag{86}$$

We then have that the first 2 moments are given by:

$$\langle x \rangle = \left.\frac{\mathrm{d}M(t)}{\mathrm{d}t}\right|_{t=0} = \left.\frac{\mathrm{d}}{\mathrm{d}t} \exp\left(lt + \frac{1}{2}(\sigma t)^2\right)\right|_{t=0} \tag{87}$$

$$= \left.\exp\left(lt + \frac{1}{2}(\sigma t)^2\right)(l + \sigma^2 t)\right|_{t=0} \tag{88}$$

$$= l \tag{89}$$

and

$$\left\langle x^2 \right\rangle = \left.\frac{\mathrm{d}^2 M(t)}{\mathrm{d}t^2}\right|_{t=0} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\left(\exp\left(lt + \frac{1}{2}(\sigma t)^2\right)(l + \sigma^2 t)\right)\right|_{t=0} \tag{90}$$

$$= \left(M(t)\sigma^2 + (l + \sigma^2 t)M(t)(l + \sigma^2 t)\right)\big|_{t=0} \tag{91}$$

$$= \left.M(t)(\sigma^2 + (l + \sigma^2 t)^2)\right|_{t=0} \tag{92}$$

$$= (\sigma^2 + l^2) \tag{93}$$

$$= \left(\frac{k_B T}{\kappa} + l^2\right) \tag{94}$$

Moving on, the correlation matrix for this problem looks like this

$$C(l, \kappa, t) = \begin{pmatrix} C_{l,l}(t) & C_{l,\kappa}(t) \\ C_{\kappa,l}(t) & C_{\kappa,\kappa}(t) \end{pmatrix} \tag{95}$$

We start by calculating the first entry $C(t)_{l,l}$. We have by definition:

$$C(t)_{l,l} = \langle (X_l(t) - \langle X_l \rangle^{eq})(X_l(0) - \langle X_l \rangle^{eq}) \rangle^{eq} \tag{96}$$
$$= \langle X_l(t)(X_l(0) - X_l(t)\langle X_l \rangle^{eq} - X_l(0)\langle X_l \rangle^{eq} + (\langle X_l \rangle^{eq})^2 \rangle^{eq} \tag{97}$$
$$= \langle X_l(t)X_l(0) \rangle^{eq} - \langle X_l(t) \rangle^{eq} \langle X_l \rangle^{eq} \tag{98}$$
$$= \kappa^2 \langle (x(t) - l)(x(0) - l) \rangle^{eq} - \kappa^2 \langle (x(t) - l) \rangle^{eq} \langle (x - l) \rangle^{eq} \tag{99}$$
$$= \kappa^2 (\langle x(t)x(0) \rangle^{eq} - \langle x(t)l \rangle^{eq} - \langle x(0)l \rangle^{eq} + l^2) - \kappa^2 (\langle x(t) \rangle^{eq} \langle x \rangle^{eq} - \langle x(t)l \rangle^{eq} - \langle xl \rangle^{eq} + l^2) \tag{100}$$
$$= \kappa^2 \langle x(t)x(0) \rangle^{eq} - \kappa^2 \langle x(t) \rangle^{eq} \langle x \rangle^{eq} \tag{101}$$

where $\langle \dots \rangle^{eq}$ is an equilibrium ensemble average which means that we can use stationarity ($\langle x(t) \rangle = \langle x(t + \tau) \rangle$ where τ is an arbitrary time interval. This also implies that $\langle X_l(0) \rangle^{eq} = \langle X_l(t) \rangle^{eq} = \langle X_l \rangle^{eq}$, and that $\langle x \rangle^{eq} = l$ (which follows from (87)).

Differentiating this expression we get (the second term is not time-dependent and therefore disappears):

$$\frac{\mathrm{d}C(t)_{l,l}}{\mathrm{d}t} = \kappa^2 \frac{\mathrm{d}}{\mathrm{d}t} (\langle x(t)x(0) \rangle^{eq} \tag{102}$$

We now make use of Ito's formula [49] in the following form:

$$\mathrm{d}f = \left(\mu_t \frac{\partial f}{\partial x} + \frac{\sigma_t^2}{2} \frac{\partial^2 f}{\partial x^2} \right) \mathrm{d}t + \sigma_t \frac{\partial f}{\partial x} dW_t \tag{103}$$

where $f(X(t))$ is a function of the Ito process $X(t)$ satisfying the following SDE:

$$\mathrm{d}X(t) = \mu_t \mathrm{d}t + \sigma_t dW_t \tag{104}$$

where W_t is the Wiener process.

In our case, this Ito SDE corresponds to

$$\mathrm{d}x = \mu_t \mathrm{d}t + \sigma_t dW_t = -\mu_P \kappa (x - l)\mathrm{d}t + \sqrt{2\kappa\mu k_B T}\xi(t)\mathrm{d}t \tag{105}$$

making it possible to identify $\mu_t = -\mu_P \kappa (x - l)$ and $\sigma_t = \sqrt{2\kappa\mu k_B T}$.

To proceed further, we let $f = x(t)x(0)$ so that

$$\frac{\mathrm{d}f}{\mathrm{d}t} = -\kappa\mu(x(t) - l)\frac{\partial}{\partial x}(x(t)x(0)) + \kappa\mu k_B T \frac{\partial^2}{\partial x^2}(x(t)x(0) + \sqrt{2\kappa\mu k_B T}\frac{\partial}{\partial x}(x(t)x(0))\xi(t) \tag{106}$$
$$= -\kappa\mu(x(t)x(0) - lx(0)) + \sqrt{2\kappa\mu k_B T}x(0)\xi(t) \tag{107}$$

and $\xi(t) = \dfrac{\mathrm{d}W_t}{\mathrm{d}t}$ is the noise which is provided in the given form for notational convenience and is not mathematically rigorous.

Now taking the ensemble average, we then have

$$\frac{\mathrm{d}\langle f \rangle}{\mathrm{d}t} = -\kappa\mu(\langle x(t)x(0) \rangle - \langle lx(0) \rangle) = -\kappa\mu(\langle x(t)x(0) \rangle - \langle x(t) \rangle \langle x(0) \rangle) \tag{108}$$

where we used that $\langle x(0)\xi(t)\rangle^{eq} = \langle x(0)\rangle^{eq}\langle \xi(t)\rangle^{eq} = 0$ since a random force at time t must be uncorrelated with the position of the particle at $t = 0$ and the fact that $\langle \xi(t)\rangle = 0$.

Substituting this back into (102) we get:

$$\frac{\mathrm{d}C(t)_{l,l}}{\mathrm{d}t} = \kappa^2 \frac{\mathrm{d}}{\mathrm{d}t}\langle f\rangle^{eq} \tag{109}$$

$$= \kappa^2(-\kappa\mu(\langle x(t)x(0)\rangle^{eq} - \langle x(t)\rangle^{eq}\langle x(0)\rangle^{eq}) \tag{110}$$

$$= -\kappa\mu(\kappa^2\langle x(t)x(0)\rangle^{eq} - \kappa^2\langle x(t)\rangle^{eq}\langle x(0)\rangle^{eq}) \tag{111}$$

$$= -\kappa\mu C(t)_{l,l} \tag{112}$$

where the final line follows from (101).

This is a simple separable differential equation with an exponential solution given by

$$C(t)_{l,l} = C(0)_{l,l}\exp(-\mu\kappa|t|) \tag{113}$$

where we have taken t to be $|t|$ instead, as the correlation function needs to be symmetric.
The next challenge is to determine the constant $C(0)_{l,l}$. We get:

$$C(0)_{l,l} = \langle (X_l(0) - \langle X_l\rangle^{eq})(X_l(0) - \langle X_l\rangle^{eq})\rangle^{eq} \tag{114}$$

$$= \langle X_l(0)^2\rangle^{eq} + (\langle X_l\rangle^{eq})^2 - 2\langle X_l(0)\rangle^{eq}\langle X_l\rangle^{eq} \tag{115}$$

$$= \langle X_l(0)^2\rangle^{eq} - (\langle X_l\rangle^{eq})^2 \tag{116}$$

$$= \kappa^2\langle (x(0) - l)^2\rangle^{eq} - \kappa^2(\langle (x - l)\rangle^{eq})^2 \tag{117}$$

$$= \kappa^2(\langle x(0)^2\rangle^{eq} - (\langle x\rangle^{eq})^2) \tag{118}$$

$$= \kappa^2(\langle x(0)^2\rangle^{eq} - l^2) \tag{119}$$

$$= \kappa^2\left(l^2 + \frac{k_B T}{\kappa} - l^2\right) \tag{120}$$

$$= \kappa k_B T \tag{121}$$

where the second to last line follows from previous calculations of $\left\langle x^2\right\rangle$ in (90). All in all we have:

$$C(t)_{l,l} = \kappa k_B T\exp(-\mu\kappa|t|) \tag{122}$$

And now we calculate $C(t)_{\kappa,\kappa}$:

$$C(t)_{\kappa,\kappa} = \langle (X_\kappa(t) - \langle X_\kappa\rangle^{eq})(X_\kappa(0) - \langle X_\kappa\rangle^{eq})\rangle^{eq} \tag{123}$$

$$= \langle X_\kappa(t)(X_\kappa(0) - X_\kappa(t)\langle X_\kappa\rangle^{eq} - X_\kappa(0)\langle X_\kappa\rangle^{eq} + (\langle X_\kappa\rangle^{eq})^2\rangle^{eq} \tag{124}$$

$$= \langle X_\kappa(t)X_\kappa(0)\rangle^{eq} - \langle X_\kappa(t)\rangle^{eq}\langle X_\kappa\rangle^{eq} \tag{125}$$

$$= \left\langle \frac{1}{2}(x(t) - l)^2\frac{1}{2}(x(0) - l)^2\right\rangle^{eq} - \left\langle \frac{1}{2}(x(t) - l)^2\right\rangle^{eq}\left\langle \frac{1}{2}(x - l)^2\right\rangle^{eq} \tag{126}$$

$$= \frac{1}{4}\left\langle (x(t) - l)^2(x(0) - l)^2\right\rangle^{eq} - \frac{1}{4}\left\langle (x(t) - l)^2\right\rangle^{eq}\left\langle (x - l)^2\right\rangle^{eq} \tag{127}$$

$$= \frac{1}{4}\left\langle (x(t) - l)^2(x(0) - l)^2\right\rangle^{eq} - \frac{1}{4}\left(\left\langle (x(t) - l)^2\right\rangle^{eq}\right)^2 \tag{128}$$

Similarly to the case for $C(t)_{l,l}$ we can define $f = (x(t) - l)^2(x(0) - l)^2$ and using Ito's lemma, we get:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = -\kappa\mu(x(t) - l)\frac{\partial}{\partial x}\left[(x(t) - l)^2(x(0) - l)^2)\right] \tag{129}$$

$$+ \frac{\sigma_t^2}{2}\frac{\partial^2}{\partial x^2}\left[(x(t) - l)^2(x(0) - l)^2\right] + \sigma_t\frac{\partial}{\partial x}\left[(x(t) - l)^2(x(0) - l)^2\right]\xi(t) \tag{130}$$

$$= -\kappa\mu(x(t) - l)(x(0) - l)^2(2(x(t) - l)) + \frac{\sigma_t^2}{2}((x(0) - l)^2)2 \tag{131}$$

$$+ \sigma_t((x(0) - l)^2(2(x(t) - l))\xi(t) \tag{132}$$

$$= -2\kappa\mu(x(t) - l)^2(x(0) - l)^2) + \sigma_t^2(x(0) - l)^2 + 2\sigma_t((x(0) - l)^2(x(t) - l))\xi(t) \tag{133}$$

$$\tag{134}$$

We then have

$$\frac{\mathrm{d}\left\langle f\right\rangle}{\mathrm{d}t} = -2\kappa\mu\left\langle(x(t) - l)^2(x(0) - l)^2\right\rangle + \sigma_t^2\left\langle(x(0) - l)^2\right\rangle \tag{135}$$

We can expand the last term to get:

$$\left\langle(x(0) - l)^2\right\rangle = \left\langle x(0)^2 + l^2 - 2x(0)l\right\rangle \tag{136}$$

$$= \left\langle x(0)^2\right\rangle + l^2 - 2\left\langle x(0)\right\rangle l \tag{137}$$

$$= \left\langle x(0)^2\right\rangle - l^2 \tag{138}$$

$$= \left\langle x^2\right\rangle - \left\langle x\right\rangle^2 \tag{139}$$

where we used the fact that $\langle x(0)\rangle = \langle x(t)\rangle = \langle x\rangle = l$ in equilibrium.

Again using the moment calculations from (87) and (90), we then get:

$$\left\langle(x(0) - l)^2\right\rangle = \left\langle x^2\right\rangle - \left\langle x\right\rangle^2 = l^2 + \frac{k_BT}{\kappa} - l^2 = \frac{k_BT}{\kappa} \tag{140}$$

We have that $\sigma_t^2 = \left(\sqrt{2\mu k_BT}\right)^2 = 2\mu k_BT$, and thus we get

$$\frac{\mathrm{d}\left\langle f\right\rangle}{\mathrm{d}t} = -2\kappa\mu\left\langle(x(t) - l)^2(x(0) - l)^2\right\rangle + \sigma_t^2\left\langle(x(0) - l)^2\right\rangle \tag{141}$$

$$= -2\kappa\mu\left\langle(x(t) - l)^2(x(0) - l)^2\right\rangle + 2\mu k_BT\left(\frac{k_BT}{\kappa}\right) \tag{142}$$

And now using

$$\frac{\mathrm{d}C(t)_{\kappa,\kappa}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{4}\left\langle(x(t) - l)^2(x(0) - l)^2\right\rangle^{eq} - \frac{1}{4}\left(\left\langle(x(t) - l)^2\right\rangle^{eq}\right)^2\right) \tag{143}$$

$$= \frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle(x(t) - l)^2(x(0) - l)^2\right\rangle^{eq} \tag{144}$$

as the second term does not have any time dependence. Continuing on:

$$\frac{\mathrm{d}C(t)_{\kappa,\kappa}}{\mathrm{d}t} = \frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle (x(t)-l)^2(x(0)-l)^2\right\rangle^{eq} \tag{145}$$

$$= \frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}t}\langle f\rangle \tag{146}$$

$$= \frac{1}{4}\left(-2\kappa\mu\left\langle (x(t)-l)^2(x(0)-l)^2\right\rangle + 2\mu k_B T\left(\frac{k_B T}{\kappa}\right)\right) \tag{147}$$

$$= -2\kappa\mu\left(C(t)_{\kappa,\kappa} + \frac{1}{4}\left(\left\langle (x(t)-l)^2\right\rangle^{eq}\right)^2\right) + \frac{1}{4}2\mu k_B T\left(\frac{k_B T}{\kappa}\right) \tag{148}$$

$$= -2\kappa\mu C(t)_{\kappa,\kappa} - 2\kappa\mu\frac{1}{4}\left(\frac{k_B T}{\kappa}\right)^2 + \frac{1}{4}2\mu k_B T\left(\frac{k_B T}{\kappa}\right) \tag{149}$$

$$= -2\kappa\mu C(t)_{\kappa,\kappa} - \mu\frac{1}{2}\left(\frac{(k_B T)^2}{\kappa}\right) + \frac{1}{2}\mu\left(\frac{(k_B T)^2}{\kappa}\right) \tag{150}$$

$$= -2\kappa\mu C(t)_{\kappa,\kappa} \tag{151}$$

which can be solved similarly to the case for $C(t)_{l,l}$.
Now determining the constant $C(0)_{\kappa,\kappa}$, we get

$$C(0)_{\kappa,\kappa} = \langle (X_\kappa(0) - \langle X_\kappa\rangle^{eq})(X_\kappa(0) - \langle X_\kappa\rangle^{eq})\rangle^{eq} \tag{152}$$

$$= \langle (X_\kappa(0) - \langle X_\kappa\rangle^{eq})^2\rangle^{eq} \tag{153}$$

$$= \langle X_\kappa(0)^2\rangle^{eq} + (\langle X_\kappa\rangle^{eq})^2 - 2\langle X_l(0)\rangle^{eq}\langle X_\kappa\rangle^{eq} \tag{154}$$

$$= \langle X_\kappa(0)^2\rangle^{eq} - (\langle X_\kappa\rangle^{eq})^2 \tag{155}$$

$$= \frac{1}{4}\langle (x(0)-l)^4\rangle^{eq} - \left(\frac{1}{2}\langle (x(0)-l)^2\rangle^{eq}\right)^2 \tag{156}$$

We start by calculating the first term which can be expanded as:

$$\langle (x(0)-l)^4\rangle^{eq} = \langle x(0)^4 - 4x(0)^3 l + 6x(0)^2 l^2 - 4x(0)l^3 + l^4\rangle^{eq} \tag{157}$$

$$= \langle x(0)^4\rangle^{eq} - 4\left\langle x(0)^3\right\rangle^{eq} l + 6\left\langle x(0)^2\right\rangle^{eq} l^2 - 4\langle x(0)\rangle^{eq} l^3 + l^4 \tag{158}$$

Here we only need to calculate $\langle x(0)^4\rangle^{eq}$ and $\left\langle x(0)^3\right\rangle^{eq}$ as we already know the other moments. These can be found by the same approach introduced in (86), however are here just provided due to spatial concerns:

$$\langle x(0)^4\rangle^{eq} = \int_{-\infty}^{\infty} x^4 p(x)dx = l^4 + \frac{6k_B T l^2}{\kappa} + 3\left(\frac{k_B T}{\kappa}\right)^2 \tag{159}$$

and

$$\langle x(0)^3\rangle^{eq} = \int_{-\infty}^{\infty} x^3 p(x)dx = l^3 + \frac{3k_B T l}{\kappa} \tag{160}$$

We then get:

$$\langle (x(0)-l)^4 \rangle^{eq} = l^4 + \frac{6k_BTl^2}{\kappa} + 3\left(\frac{k_BT}{\kappa}\right)^2 - 4l\left(l^3 + \frac{3k_BTl}{\kappa}\right) + 6l^2\left(l^2 + \frac{k_BT}{\kappa}\right) - 4l^4 + l^4 \tag{161}$$

$$= \frac{6k_BTl^2}{\kappa} + 3\left(\frac{k_BT}{\kappa}\right)^2 - \frac{12k_BTl^2}{\kappa} + \frac{6k_BTl^2}{\kappa} \tag{162}$$

$$= 3\left(\frac{k_BT}{\kappa}\right)^2 \tag{163}$$

which means that equation (152) becomes:

$$C(0)_{\kappa,\kappa} = \frac{1}{4}\langle (x(0)-l)^4 \rangle^{eq} - \left(\frac{1}{2}\langle (x(0)-l)^2 \rangle^{eq}\right)^2 \tag{164}$$

$$= \frac{3}{4}\left(\frac{k_BT}{\kappa}\right)^2 - \frac{1}{4}\left(\frac{k_BT}{\kappa}\right)^2 \tag{165}$$

$$= \frac{1}{2}\left(\frac{k_BT}{\kappa}\right)^2 \tag{166}$$

thus arriving at

$$C(t)_{\kappa,\kappa} = \frac{1}{2}\left(\frac{k_BT}{\kappa}\right)^2 \exp(-2\kappa\mu|t|) \tag{167}$$

Now we have calculated the diagonal entries of the correlation matrix. We still need to calculate the off-diagonals:

$$C(t)_{l,\kappa} = \langle (X_l(t) - \langle X_l \rangle^{eq})(X_\kappa(0) - \langle X_\kappa \rangle^{eq}) \rangle^{eq} \tag{168}$$

$$= \langle X_l(t)X_\kappa(0) - X_l(t)\langle X_\kappa \rangle^{eq} - X_\kappa(0)\langle X_l \rangle^{eq} + \langle X_l \rangle^{eq} \langle X_\kappa \rangle^{eq} \rangle^{eq} \tag{169}$$

$$= \langle X_l(t)X_\kappa(0) \rangle^{eq} - \langle X_l(t) \rangle^{eq} \langle X_\kappa \rangle^{eq} - \langle X_\kappa(0) \rangle^{eq} \langle X_l \rangle^{eq} + \langle X_l \rangle^{eq} \langle X_\kappa \rangle^{eq} \tag{170}$$

$$= \langle X_l(t)X_\kappa(0) \rangle^{eq} - \langle X_l(t) \rangle^{eq} \langle X_\kappa \rangle^{eq} \tag{171}$$

$$= -\frac{1}{2}\kappa \left\langle (x(t)-l)(x(0)-l)^2 \right\rangle^{eq} + \frac{1}{2}\kappa \langle (x(t)-l) \rangle^{eq} \left\langle (x-l)^2 \right\rangle^{eq} \tag{172}$$

$$\tag{173}$$

We can get a differential equation for $C(t)_{l,\kappa}$ in the same way we did for the two diagonal terms by using Ito's lemma. But when we calculate the constant $C(t=0)_{l,\kappa}$ we see that it is equal to 0:

$$C(0)_{l,\kappa} = \langle (X_l(0) - \langle X_l \rangle^{eq})(X_\kappa(0) - \langle X_\kappa \rangle^{eq}) \rangle^{eq} \tag{174}$$
$$= \langle X_l(0)X_\kappa(0) - X_l(0)\langle X_\kappa \rangle^{eq} - X_\kappa(0)\langle X_l \rangle^{eq} + \langle X_l \rangle^{eq} \langle X_\kappa \rangle^{eq} \rangle^{eq} \tag{175}$$
$$= \langle X_l(0)X_\kappa(0) \rangle^{eq} - \langle X_l(0) \rangle^{eq} \langle X_\kappa \rangle^{eq} - \langle X_\kappa(0) \rangle^{eq} \langle X_l \rangle^{eq} + \langle X_l \rangle^{eq} \langle X_\kappa \rangle^{eq} \tag{176}$$
$$= \langle X_l(0)X_\kappa(0) \rangle^{eq} - \langle X_l(0) \rangle^{eq} \langle X_\kappa \rangle^{eq} \tag{177}$$
$$= -\frac{1}{2}\kappa \left\langle (x(0) - l)(x(0) - l)^2 \right\rangle^{eq} + \frac{1}{2}\kappa \langle (x(0) - l) \rangle^{eq} \left\langle (x - l)^2 \right\rangle^{eq} \tag{178}$$
$$= -\frac{1}{2}\kappa \left\langle (x(0) - l)^3 \right\rangle^{eq} + \frac{1}{2}\kappa \langle (x(0) - l) \rangle^{eq} \left\langle (x - l)^2 \right\rangle^{eq} \tag{179}$$
$$\tag{180}$$

The first term gives:

$$-\frac{1}{2}\kappa \left\langle (x(0)^3 - 3x(0)^2 l + 3x(0)l^2 - l^3) \right\rangle^{eq} = -\frac{1}{2}\kappa \left(\left\langle x(0)^3 \right\rangle^{eq} - 3 \left\langle x(0)^2 \right\rangle^{eq} l + 2l^3 \right) \tag{181}$$
$$= -\frac{1}{2}\kappa \left(l^3 + \frac{3k_B T l}{\kappa} - 3l \left(l^2 + \frac{k_B T}{\kappa} \right) + 2l^3 \right) \tag{182}$$
$$= -\frac{1}{2}\kappa \left(\frac{3k_B T l}{\kappa} - 3l \frac{k_B T}{\kappa} \right) \tag{183}$$
$$= 0 \tag{184}$$

And the second term also gives 0 since

$$\frac{1}{2}\kappa \langle (x(0) - l) \rangle^{eq} = \frac{1}{2}\kappa (\langle x(0) \rangle^{eq} - l) = 0 \tag{185}$$

This means that $C(0)_{l,\kappa} = 0$ which in turn implies that $C(t)_{l,\kappa} = 0$. Doing the same analysis on $C(t)_{\kappa,l}$ reveals the same pattern, just in reverse order, leading to $C(t)_{\kappa,l} = 0$ as well.

Summing up, the equilibrium correlation matrix is then given by:

$$C(l, \kappa, t) = \begin{pmatrix} \kappa k_B T \exp(-\mu\kappa|t|) & 0 \\ 0 & \frac{1}{2}\left(\frac{k_B T}{\kappa}\right)^2 \exp(-2\kappa\mu|t|) \end{pmatrix} \tag{186}$$

If we simulate a system given by the Langevin equation in (77) and let it reach equilibrium, we can calculate all the relevant quantities numerically (the entries of the correlation matrix) and thereby show that (186) holds. Plotting the numerically calculated quantities and the entries of (186) with $k_B T = 1$, we obtain the following relations:

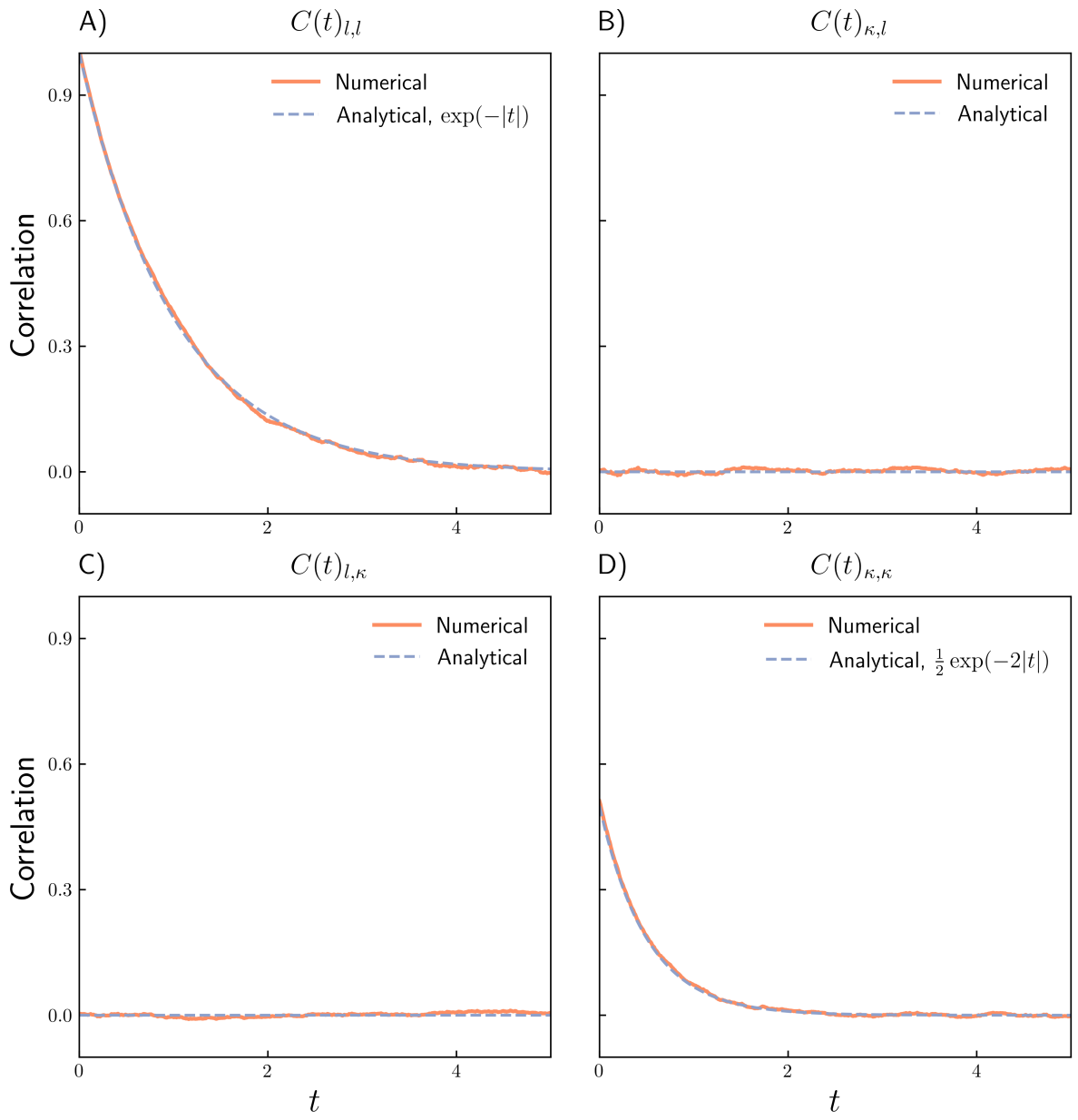

Figure 1: Comparison of the analytical expressions in the correlation matrix (186) vs. correlations obtained by numerical simulation. Simulation parameters are: $\mu = 1$, $k_B T = 1$, $\kappa = 1$, $l = 0$.

which verifies our analytical calculations of the correlation matrix.

4.2 Calculating the friction coefficient matrix

As we now have the correlation matrix, we can calculate the friction coefficient matrix introduced earlier in equation (76):

$$\zeta_{l,l}(t) = \frac{1}{k_B T}\int_0^\infty C_{l,l}(t)\mathrm{d}t = \frac{1}{k_B T}\int_0^\infty \kappa k_B T \exp(-\mu\kappa t)\mathrm{d}t \tag{187}$$

$$= \kappa\left[-\frac{\exp(-\mu\kappa t)}{\mu\kappa}\right]_{t=0}^{t=\infty} \tag{188}$$

$$= -\frac{1}{\mu}\left(0 - \exp(0)\right) \tag{189}$$

$$= \frac{1}{\mu} \tag{190}$$

and

$$\zeta_{\kappa,\kappa}(t) = \frac{1}{k_B T}\int_0^\infty C_{\kappa,\kappa}(t)\mathrm{d}t = \frac{1}{k_B T}\int_0^\infty \frac{1}{2}\left(\frac{k_B T}{\kappa}\right)^2 \exp(-2\kappa\mu t)\mathrm{d}t \tag{191}$$

$$= \frac{k_B T}{2\kappa^2}\left[-\frac{\exp(-2\mu\kappa t)}{2\mu\kappa}\right]_{t=0}^{t=\infty} \tag{192}$$

$$= -\frac{k_B T}{4\mu\kappa^3}\left(0 - \exp(0)\right) \tag{193}$$

$$= \frac{k_B T}{4\mu\kappa^3} \tag{194}$$

And the off-diagonals are equal to 0. The friction coefficient matrix is then given by:

$$\zeta(l,\kappa) = \begin{pmatrix} \frac{1}{\mu} & 0 \\ 0 & \frac{k_B T}{4\mu\kappa^3} \end{pmatrix} \tag{195}$$

The two results derived here (the correlation matrix and friction coefficient matrix) are consistent with the results in [10] and [5].

4.3 Finding the optimal protocols

We now turn to finding the optimal protocols for the passive Brownian particle in a harmonic trap, i.e. the control parameter protocols that minimize the dissipated work. We imagine that we start out in equilibrium with the control parameter given by $\lambda(t_0)$ and then change the control parameter until we reach a final state $\lambda(t_f)$ where the protocol must transition between these 2 states in a time interval $T = t_f - t_0$ [10], [5]. As this is an optimization problem, we solve the Euler-Lagrange equations [50] where the function to be minimized is the dissipated work rate, $f(\lambda(t), \dot{\lambda}) = \dot{\lambda}\zeta(\lambda(t))\dot{\lambda}$, where $\lambda(t)$ in the given system corresponds to $l(t)$ or $\kappa(t)$ but we keep it general for now.

We have

$$\frac{\partial}{\partial\lambda}f(t,\lambda,\dot{\lambda}) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial\dot{\lambda}}f(t,\lambda,\dot{\lambda}) = 0 \Rightarrow \tag{196}$$

$$\dot{\lambda}^2\zeta'(\lambda) - \frac{\mathrm{d}}{\mathrm{d}t}(2\dot{\lambda}\zeta(\lambda)) = 0 \tag{197}$$

$$\dot{\lambda}^2\zeta'(\lambda) - 2(\ddot{\lambda}\zeta(\lambda) + \dot{\lambda}\zeta'(\lambda)\dot{\lambda}) = 0 \tag{198}$$

$$-\dot{\lambda}^2\zeta'(\lambda) - 2\ddot{\lambda}\zeta(\lambda) = 0 \tag{199}$$

where we used the product rule and the chain rule in (198).
This is a second-order ODE which we solve with mathematical software. For the l-protocol we get:

$$-\frac{\mathrm{d}l(t)}{\mathrm{d}t}^2\frac{\mathrm{d}\zeta(l(t))}{\mathrm{d}l} - 2\frac{\mathrm{d}^2l(t)}{\mathrm{d}t^2}\zeta(l(t)) = 0 \tag{200}$$

We know the friction tensor for this problem from earlier. It is given by

$$\zeta(l,\kappa) = \begin{pmatrix} \frac{1}{\mu} & 0 \\ 0 & \frac{k_BT}{4\mu\kappa^3} \end{pmatrix} \tag{201}$$

Substituting $\zeta_{l,l} = \zeta(l(t)) = \frac{1}{\mu}$ into (200) we get

$$-\frac{2}{\mu}\frac{\mathrm{d}^2l(t)}{\mathrm{d}t^2} = 0 \tag{202}$$

which yields $l(t) = c_1t + c_2$ where c_1 and c_2 are integration constants to be determined by the boundary conditions. In the rest of this section, we take $t_0 = 0$ for simplicity. We then get that $l(t_0) = c_2$. For t_f we get $l(t_f) = c_1t_f + l(0)$ which gives $c_1 = (l(t_f) - l(0))/t_f$ so all in all we get:

$$l(t) = \frac{l_f - l_0}{t_f}t + l_0 \tag{203}$$

For the κ-protocol we get:

$$-\frac{\mathrm{d}\kappa(t)}{\mathrm{d}t}^2\frac{\mathrm{d}\zeta(\kappa(t))}{\mathrm{d}\kappa} - 2\frac{\mathrm{d}^2\kappa(t)}{\mathrm{d}t^2}\zeta(\kappa(t)) = 0 \Rightarrow \tag{204}$$

$$-\frac{\mathrm{d}\kappa(t)}{\mathrm{d}t}^2\frac{\mathrm{d}}{\mathrm{d}\kappa}\left(\frac{k_BT}{4\mu\kappa^3}\right) - 2\frac{\mathrm{d}^2\kappa(t)}{\mathrm{d}t^2}\left(\frac{k_BT}{4\mu\kappa^3}\right) = 0 \tag{205}$$

$$\frac{\mathrm{d}\kappa(t)}{\mathrm{d}t}^2\left(\frac{3k_BT}{4\mu\kappa^4}\right) - \frac{\mathrm{d}^2\kappa(t)}{\mathrm{d}t^2}\left(\frac{k_BT}{2\mu\kappa^3}\right) = 0 \tag{206}$$

$$\frac{3}{2}\frac{\mathrm{d}\kappa(t)}{\mathrm{d}t}^2 - \frac{\mathrm{d}^2\kappa(t)}{\mathrm{d}t^2}\kappa(t) = 0 \tag{207}$$

Solving this differential equation for $\kappa(t)$ we get

$$\kappa(t) = \frac{4}{(c_1t + c_2)^2} \tag{208}$$

where c_1 and c_2 are integration constants. Determining them we get:

$$\kappa(0) = \frac{4}{c_2^2} \Rightarrow c_2 = \pm\sqrt{\frac{4}{\kappa(0)}} = \pm\frac{2}{\sqrt{\kappa(0)}} \tag{209}$$

Determining c_1 we get:

$$\kappa(t_f) = \frac{4}{\left(c_1 t_f \pm \frac{2}{\sqrt{\kappa(0)}}\right)^2} \Rightarrow \tag{210}$$

$$\pm\sqrt{\frac{4}{\kappa(t_f)}} = c_1 t_f \pm \frac{2}{\sqrt{\kappa(0)}} \Rightarrow \tag{211}$$

$$\left(\sqrt{\frac{4}{\kappa(t_f)}} - \frac{2}{\sqrt{\kappa(0)}}\right)\frac{1}{t_f} = c_1 \quad \text{and} \quad \left(-\sqrt{\frac{4}{\kappa(t_f)}} + \frac{2}{\sqrt{\kappa(0)}}\right)\frac{1}{t_f} = c_1 \tag{212}$$

$$\tag{213}$$

so

$$c_1 = \pm\frac{2}{t_f}\left(\frac{1}{\sqrt{\kappa(t_f)}} - \frac{1}{\sqrt{\kappa(0)}}\right) \tag{214}$$

which means $\kappa(t)$ is given by

$$\kappa(t) = \frac{4}{\left(\pm\frac{2}{t_f}\left(\frac{1}{\sqrt{\kappa(t_f)}} - \frac{1}{\sqrt{\kappa(0)}}\right)t \pm \frac{2}{\sqrt{\kappa(0)}}\right)^2} \tag{215}$$

$$= \frac{1}{\left(\frac{1}{t_f}\left(\frac{1}{\sqrt{\kappa(t_f)}} - \frac{1}{\sqrt{\kappa(0)}}\right)t + \frac{1}{\sqrt{\kappa(0)}}\right)^2} \tag{216}$$

As we have the friction coefficient matrix and expressions for both the l and κ protocols, we can calculate the average dissipation rate for the optimal protocols. For l we get:

$$\frac{\mathrm{d}W^{diss}}{\mathrm{d}t} = \frac{\mathrm{d}l(t)}{\mathrm{d}t}^2 \zeta(l(t)) = \frac{c_1^2}{\mu} \tag{217}$$

where we used that $\dfrac{\mathrm{d}l(t)}{\mathrm{d}t} = c_1$. Now solving (217) we get:

$$W^{diss} = \int_{t=0}^{t=t_f} \frac{\mathrm{d}l(t)}{\mathrm{d}t}^2 \zeta(l(t))\mathrm{d}t = \int_{t=0}^{t=t_f} \left(\frac{l_f - l_0}{t_f}\right)^2 \frac{1}{\mu}\mathrm{d}t = \frac{1}{\mu}\left(\frac{l_f - l_0}{t_f}\right)^2 t_f = \frac{1}{\mu}\frac{(l_f - l_0)^2}{t_f} \tag{218}$$

For κ we get:

$$\frac{\mathrm{d}W^{diss}}{\mathrm{d}t} = \frac{\mathrm{d}\kappa(t)}{\mathrm{d}t}^2 \zeta(\kappa(t)) \tag{219}$$

$$= \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{4}{(c_1 t + c_2)^2}\right)\right)^2 \frac{k_B T}{4\mu\kappa(t)^3} \tag{220}$$

$$= \left(\frac{-8c_1}{(c_1 t + c_2)^3}\right)^2 \frac{k_B T}{4\mu\kappa(t)^3} \tag{221}$$

$$= \left(\frac{64c_1^2}{(c_1 t + c_2)^6}\right) \frac{k_B T}{4\mu} \frac{1}{\left(\frac{4}{(c_1 t + c_2)^2}\right)^3} \tag{222}$$

$$= \frac{k_B T}{4}\frac{c_1^2}{\mu} \tag{223}$$

Solving this differential equation we get:

$$W^{diss} = \frac{k_B T}{4}\frac{c_1^2}{\mu}\int_{t=0}^{t=t_f} dt = \frac{k_B T}{4}\frac{c_1^2}{\mu} t_f \tag{224}$$

$$= \frac{k_B T}{4\mu}\left(\pm\frac{2}{t_f}\left(\frac{1}{\sqrt{\kappa(t_f)}} - \frac{1}{\sqrt{\kappa(0)}}\right)\right)^2 t_f \tag{225}$$

$$= \frac{k_B T}{\mu}\left(\frac{1}{\sqrt{\kappa(t_f)}} - \frac{1}{\sqrt{\kappa(0)}}\right)^2 \frac{1}{t_f} \tag{226}$$

Finally equipped with both expressions for the dissipated work (for the l-protocol and the κ-protocol), we can test our analytical expressions against numerical simulations and verify our results, as will be the subject of the following chapter.

5 Numerical simulation of a passive Brownian particle in a harmonic potential

In this section, we numerically simulate a passive Brownian particle in 1D driven by the optimal protocols for the l- and κ-protocols, calculate the work for each process and compare with the analytically derived expressions from the previous section.

To simulate the system we use the Euler-Maruyama method to solve the Langevin equation numerically. We want to drive the system between the values λ_0 and λ_f which means that besides solving the SDE numerically, we also calculate the optimal protocols ((203) and (208)) at each time step, as they are an explicit part of the Langevin dynamics. We then also discretize the incremental heat, work and energy and calculate them numerically along each simulated trajectory of the Langevin equation using the previously introduced expressions (12), (7) and (8).
As the dissipated work is given by [10]

$$W^{diss} = W - (F_{\lambda_f} - F_{\lambda_0}) = W - \Delta F, \tag{227}$$

this means that when we simulate the system and calculate the work, we also need to subtract the free energy difference between the initial state of the system (where it is in equilibrium with protocol value λ_0) and the final state (where it is in equilibrium with protocol value λ_f). The free energy is given by

$$F = \langle E \rangle - T \langle S \rangle \tag{228}$$

where E is the internal energy, T the temperature and S the entropy. As $S = -k_B \sum p_i \ln(p_i)$, where p_i is the Boltzmann distribution, this is something we can easily calculate by using the Boltzmann distribution calculated in the 2 equilibrium states λ_0 and λ_f (see corresponding Maple documents at MasterThesisGitHub). We also simulate equilibrium reservoirs at the two values of λ in order to draw samples from these. These equilibrium distributions are also used as the initial conditions for the Langevin equation simulation.
We then do simulations for each of the 2 protocols, while keeping the other protocol constant throughout, specifically we take $l = 0$ while driving κ and take $\kappa = 1$ while driving l. We use the following parameters in the simulation:

Passive Brownian particle	λ_0	λ_f	dt	μ	N	$k_B T$
l-protocol	0	5	0.001	1	20000	1
κ-protocol	1	2	0.001	1	20000	1

Table 1: Parameter values for the simulation of a passive Brownian particle.

Below we plot the simulations for the two protocols using the values given in Table 1

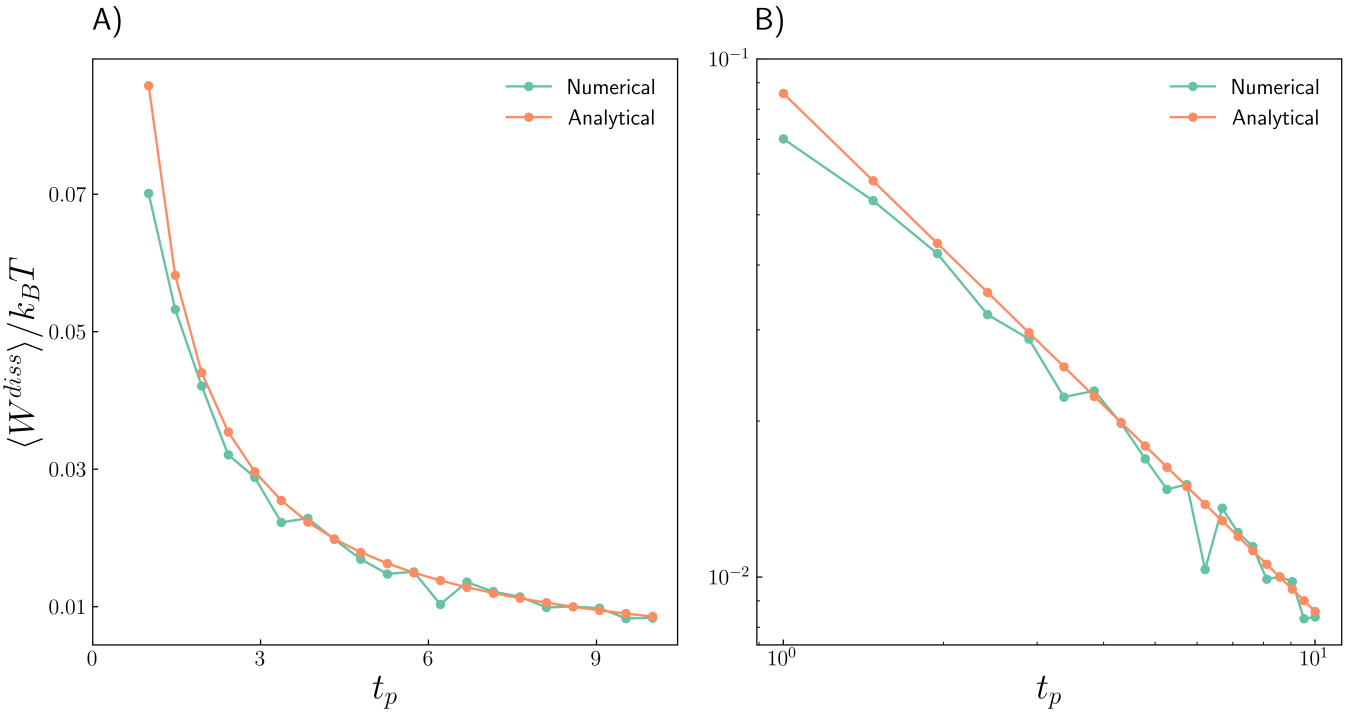

Figure 2: Passive Brownian particle in a one dimensional harmonic potential driven by the optimal κ-protocol that changes the strength of the potential. A) Comparison of the numerical and analytical solution of the average dissipated work B) Log-scale of A).

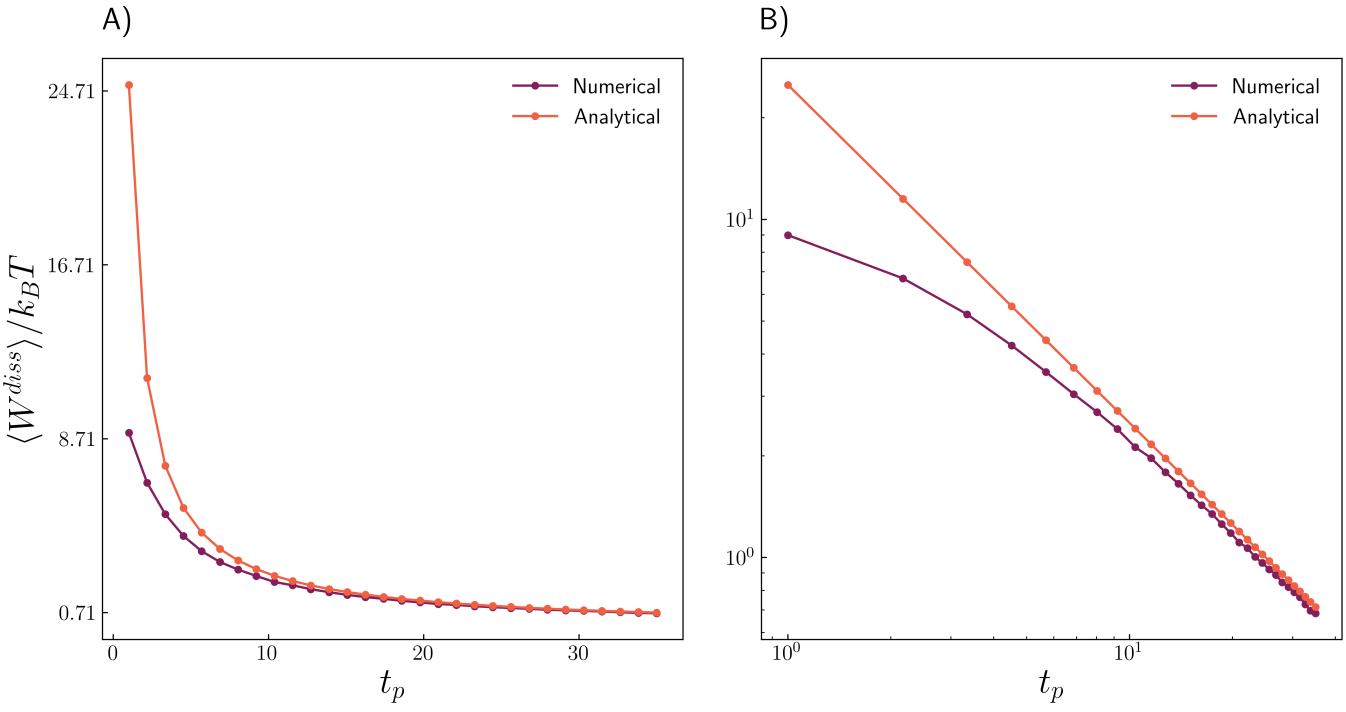

Figure 3: Passive Brownian particle in a one dimensional harmonic potential driven by the optimal l-protocol that shifts the potential in space. A) Comparison of the numerical and analytical solution of the average dissipated work B) Log-scale of A).

From the B) plots above, we can deduce that the the average dissipated work declines as $1/t_p$ since this would produce a line with slope equal to -1 on a log-log plot. Furthermore, we observe that the numerical (exact) solution agrees with the analytical (approximated) solution for slow protocols (i.e. long protocol durations). This means that driving the protocols infinitely slowly is the optimal protocol duration which dissipates the least amount of work. The exact and approximated solutions start to diverge for smaller protocol durations, since here the linear response framework (which we used to derive the optimal protocols) breaks down as it is only able to accommodate slow and weak driving. The results obtained above in Figures 2 and 3 are consistent with [5].

To obtain a comparison between an optimal protocol and a naive protocol, we furthermore simulate the system being driven by a naive protocol. A naive protocol is a sub-optimal protocol, for example for the κ-protocol, this would be a linear protocol, meaning that the protocol speed is constant. We implement such a protocol in the case where we change the strength of the harmonic potential (κ). We do this numerically by changing κ through discrete and evenly spaced values for example $\kappa_0 = 0.01$, $\kappa_1 = 0.02$, $\kappa_2 = 0.03$ and so forth. We use the values in Table 1 except this time we simulate $N = 10^5$ trajectories. We obtain the following plot

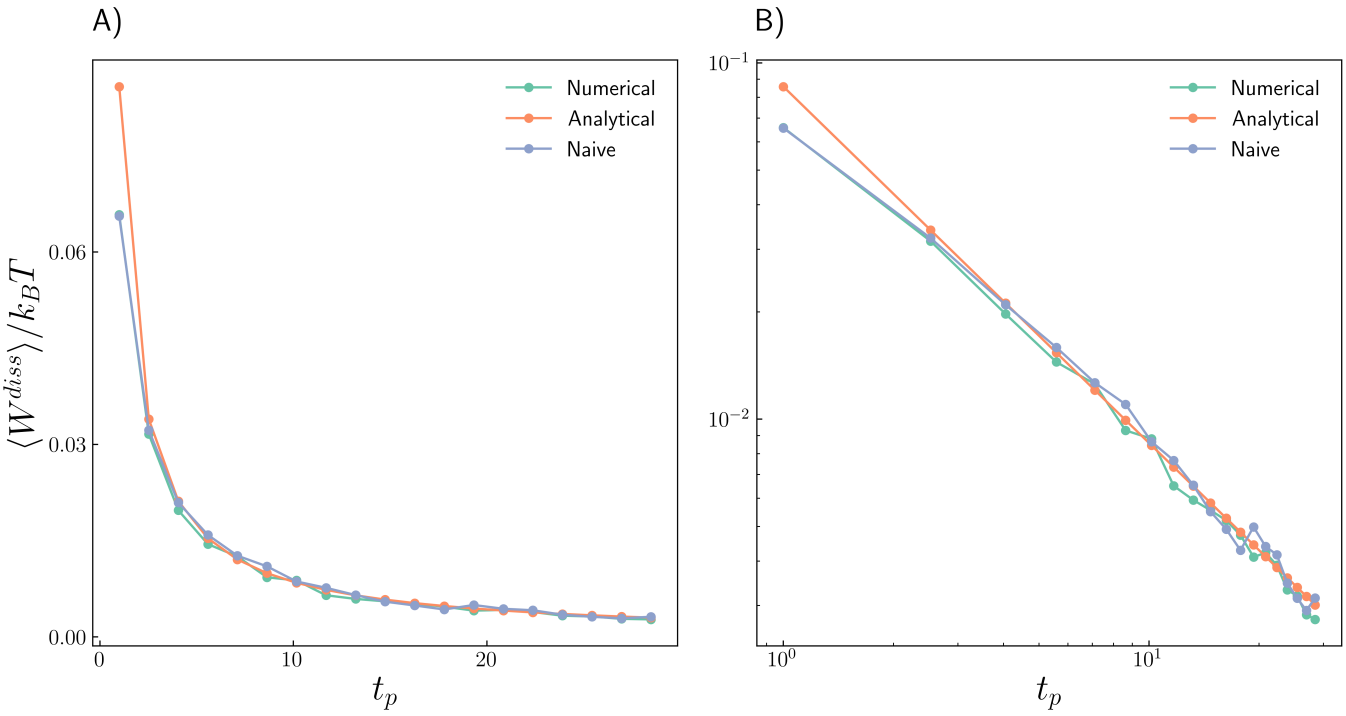

Figure 4: Passive Brownian particle in a one dimensional harmonic potential with κ being driven by a linear (naive) protocol and an optimal protocol. A) Comparison of the numerical and analytical solution of the average dissipated work for an optimal protocol (red and green) and a numerically calculated naive protocol (blue) B) Log-scale of A).

We would expect the system driven by the naive protocol to dissipate more work than the system driven by the optimal protocol. This is also what we observe in figure 4, however the difference between the dissipated work for the 2 systems is rather small. We reflect on why this might be later in the discussion.

Summing up, we have implemented and simulated numerically the optimal protocols for a passive Brownian particle and obtained agreement between the analytically derived and numerically simulated dissipated work, as well as considered the case of a naive protocol. Taking a step further in the direction of describ-

ing more complex systems, we derive an optimal protocol framework for active systems in the following section.

6 Optimal control for an active system

We now have a framework for handling passive systems close to equilibrium, and in this section we will derive an optimal protocol framework for an active system. We will consider active systems that are out of equilibrium partly by their own activity and partly by being driven by external protocols. We thus proceed in a way similar to the passive case.

Our system of interest is a modified version of the passive Brownian particle in the sense that we now have an active term in the equations of motion. This active term can be seen as an extra, active force that acts on the particle. It is stochastic in nature and amounts to a self-propulsion which means the particle propels itself forward, on average.

We consider an entire system N of these active particles and describe their motion with the following equation [6]

$$\frac{\mathrm{d}\boldsymbol{r_i}}{\mathrm{d}t} = -\mu\nabla_i\phi + \boldsymbol{v_i} + \sqrt{2D}\boldsymbol{\eta_i} \tag{229}$$

where $\boldsymbol{r}_i$ is the position of the i'th particle, $\phi(\alpha, \{\boldsymbol{r}\})$ is the potential energy which is a function of the protocol parameter α and all the particle positions $\{\boldsymbol{r}\}$. $\boldsymbol{\eta_i}$ is Gaussian vector noise with zero mean and unit variance. We also define the force $\boldsymbol{f}_i = -\nabla_i\phi + \boldsymbol{v_i}/\mu$.

Among the many models used to model active systems, 3 popular ones are obtained by different choices of the active force $\boldsymbol{v_i}$. Active Brownian particles (ABPs) and Run-and-Tumble particles (RTPs) have a constant magnitude of $\boldsymbol{v_i}$ while the orientation of the particle changes due to either rotational diffusion in ABPs or due to sudden changes in orientation that are Poisson-distributed in RTPs [15]. Here RTP is a popular choice for modelling bacteria dynamics as they often reorient themselves with their flagellum stochastically. The third option, which is our baseline for this work, is the active Ornstein-Uhlenbeck particle (AOUP) where $\boldsymbol{v_i}$ is a self-propulsion velocity modelled as a simple Ornstein-Uhlenbeck process:

$$\tau\frac{\mathrm{d}\boldsymbol{v_i}}{\mathrm{d}t} = -\boldsymbol{v_i} + \sqrt{2D_1}\boldsymbol{\zeta_i} \tag{230}$$

Here τ is the persistence time which is the characteristic time scale over which the particles retain their self-propulsion before changing direction [51] and the noise $\boldsymbol{\zeta_i}$ is Gaussian noise with zero mean and unit variance. The velocity correlations are given by $\langle v_{i,\alpha}(t)v_{i,\beta}(t')\rangle = \delta_{i,j}\delta_{\alpha,\beta}\frac{D_1}{\tau}\exp(-|t-t'|/\tau)$, where i, j indexes the particles and α, β the coordinates.

It is important to note the distinction between passive and active dynamics here: The fact that a passive Brownian particle is immersed in an aqueous solution (that also acts as a heat bath) gives rise to both the dissipation and the input of energy: the *dissipation* comes from the drag that the particle experiences due to the average force exerted by the water molecules, $-\frac{1}{\mu}\dot{\boldsymbol{r}}_\boldsymbol{i}$, while the energy input comes from the *fluctuations* due to the stochastic force, $\sqrt{2D}\boldsymbol{\eta_i}$, which is caused by collisions between the particle and the surrounding water molecules [15]. When the particle is in equilibrium with the heat bath, dissipation and injection of energy are related by the fluctuation-dissipation theorem through the Einstein relation $D = \mu k_B T$. However, for an active particle much of the injection of energy comes from the self-propulsion $\boldsymbol{v_i}$ and would thus

not satisfy the fluctuation-dissipation theorem. The assumption here is that $\boldsymbol{v_i}$ is *not* Gaussian white noise, otherwise the particle would indeed behave like a passive Brownian particle with a shift in temperature [15].

However keeping this in mind, much of the theory and methods from the passive case we considered in the previous section carry over to the analysis of the active system. To find the optimal protocols for the work for the active system, we thus follow similar steps as those we took in the passive case and start from the definition in (7):

$$\mathrm{d}w = \frac{\partial \phi(\alpha, \boldsymbol{r})}{\partial \alpha}\mathrm{d}\alpha = \frac{\partial \phi(\alpha, \boldsymbol{r})}{\partial \alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}t}\mathrm{d}t = \frac{\partial \phi(\alpha, \boldsymbol{r})}{\partial \alpha}\dot{\alpha}\mathrm{d}t \tag{231}$$

Integrating both sides and taking the ensemble average we get

$$\langle W \rangle = \int_0^{t_p} \left\langle \frac{\partial \phi(\alpha, \boldsymbol{r})}{\partial \alpha} \right\rangle \dot{\alpha}\mathrm{d}t \tag{232}$$

where t_p is the final time point of the protocol duration.
Similar to the expression we obtained in (62), where the expansion was linear, we now expand a given observable $X(t)$ to second order. We have [6]

$$\langle X(t) \rangle \approx \langle X \rangle_{ss} + \int_{-\infty}^{t} \Delta\alpha_{t,t'} R_1(X; t, t')\mathrm{d}t' \tag{233}$$

$$+ \int_{-\infty}^{t}\int_{-\infty}^{t'} \Delta\alpha_{t,t'}\Delta\alpha_{t,t''} R_2(X; t, t', t'')\mathrm{d}t'\mathrm{d}t'' \tag{234}$$

where $t'' \leq t' \leq t$, $\Delta\alpha_{t,t'} = \alpha(t) - \alpha(t')$ and R_1 and R_2 are the first-order and second-order response functions, respectively, defined from (19) as

$$R_1(X; t, t') = \left.\frac{\delta \langle X(t) \rangle}{\delta \Delta\alpha_{t,t'}}\right|_{\Delta\alpha \to 0}, \tag{235}$$

$$R_2(X; t, t', t'') = \left.\frac{\delta^2 \langle X(t) \rangle}{\delta \Delta\alpha_{t,t'}\delta \Delta\alpha_{t,t''}}\right|_{\Delta\alpha \to 0} \tag{236}$$

As we are now expanding around a non-equilibrium steady state, $\langle X \rangle_{ss}$ denotes a steady state ensemble average. By expanding the observable $\langle X(t) \rangle$ in this way, it is implicitly assumed that we drive the protocols slowly and weakly so that $\alpha(t)$ has at least continuous second-order derivatives.
Similar to the passive case where we expanded the protocol velocities $\dot{\lambda}(t)$, we now expand $\Delta\alpha_{t,t'}$ as

$$\Delta\alpha_{t,t'} = (t - t')\dot{\alpha}(t') + \frac{(t - t')^2}{2}\ddot{\alpha}(t') + \mathcal{O}((t - t')^3) \tag{237}$$

To get an expression for the average work (equation (232)), we then expand $\left\langle \frac{\partial \phi(\alpha, \boldsymbol{r})}{\partial \alpha} \right\rangle$ to first order in $\dot{\alpha}$ so the average work becomes of order $1/t_p$, as in the passive linear response case. We do this by using (233) and

(237) and we use the notation $\frac{\partial \phi(\alpha, \boldsymbol{r})}{\partial \alpha} = \partial_\alpha \phi$ for simplifcation.

We have:

$$\langle \partial_\alpha \phi \rangle \approx \langle \partial_\alpha \phi \rangle_{ss} + \int_{-\infty}^{t} \Delta\alpha_{t,t'} R_1(\partial_\alpha \phi; t, t') dt' \tag{238}$$

$$+ \int_{-\infty}^{t} \int_{-\infty}^{t'} \Delta\alpha_{t,t'} \Delta\alpha_{t,t''} R_2(\partial_\alpha \phi; t, t', t'') \mathrm{d}t'' \mathrm{d}t' \tag{239}$$

$$\approx \langle \partial_\alpha \phi \rangle_{ss} + \int_{-\infty}^{t} \left((t - t')\dot{\alpha}(t) + \frac{(t - t')^2}{2} \ddot{\alpha}(t) \right) R_1(\partial_\alpha \phi; t, t') dt' \tag{240}$$

$$+ \int_{-\infty}^{t} \int_{-\infty}^{t'} (t - t')(t - t'') \dot{\alpha}^2(t) R_2(\partial_\alpha \phi; t, t', t'') \mathrm{d}t'' \mathrm{d}t' \tag{241}$$

$$= \langle \partial_\alpha \phi \rangle_{ss} + \dot{\alpha}(t) f(\alpha) + \frac{1}{2} \ddot{\alpha}(t) g(\alpha) + \dot{\alpha}^2(t) h(\alpha) \tag{242}$$

where

$$f(\alpha) = \int_{-\infty}^{t} R_1(\partial_\alpha \phi; t, t')(t - t') \mathrm{d}t' \tag{243}$$

$$g(\alpha) = \int_{-\infty}^{t} R_1(\partial_\alpha \phi; t, t')(t - t')^2 \mathrm{d}t' \tag{244}$$

$$h(\alpha) = \int_{-\infty}^{t} \int_{-\infty}^{t'} R_2(\partial_\alpha \phi; t, t', t'')(t - t')(t - t'') \mathrm{d}t' \mathrm{d}t'' \tag{245}$$

This results in the following expression for the work

$$\langle W \rangle \approx \int_{0}^{t_p} \langle \partial_\alpha \phi \rangle_{ss} \dot{\alpha}(t) \mathrm{d}t + \int_{0}^{t_p} \left[\left(\dot{\alpha}(t) f(\alpha) + \frac{1}{2} \ddot{\alpha}(t) g(\alpha) + \dot{\alpha}^2(t) h(\alpha) \right) \dot{\alpha}(t) \right] \mathrm{d}t \tag{246}$$

$$= \int_{\alpha_0}^{\alpha_f} \langle \partial_\alpha \phi \rangle_{ss} \, \mathrm{d}\alpha + \int_{0}^{t_p} \left[\dot{\alpha}^2(t) f(\alpha) + \frac{1}{2} \dot{\alpha}(t) \ddot{\alpha}(t) g(\alpha) + \dot{\alpha}^3(t) h(\alpha) \right] \mathrm{d}t \tag{247}$$

where we rewrote the first term using the substitution

$$\int_{0}^{t_p} \langle \partial_\alpha \phi \rangle_{ss} \dot{\alpha}(t) \mathrm{d}t = \int_{\alpha(0)}^{\alpha(t_p)} \langle \partial_\alpha \phi \rangle_{ss} \, \mathrm{d}\alpha \tag{248}$$

We assume that the protocol is varying slowly, and keep terms up to the order of $\dot{\alpha}^2$ in the integral in (247). Considering the second term, we can write

$$\dot{\alpha}(t)\ddot{\alpha}(t) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\dot{\alpha}^2(t) \tag{249}$$

so we have for the second term

$$\frac{1}{2}\int_0^{t_p} \dot{\alpha}\ddot{\alpha}g(\alpha)dt = \frac{1}{4}\int_0^{t_p} \frac{\mathrm{d}}{\mathrm{d}t}\dot{\alpha}^2 g(\alpha)\mathrm{d}t \tag{250}$$

$$= \frac{1}{4}\left(\left[\dot{\alpha}^2 g(\alpha)\right]_0^{t_p} - \int_0^{t_p} \dot{\alpha}^2 \underbrace{\frac{\mathrm{d}}{\mathrm{d}t}g(\alpha)}_{=g'(\alpha)\dot{\alpha}}\mathrm{d}t\right) \tag{251}$$

Both terms in the above scale as $\dot{\alpha}^2$ and thus can be neglected, which means that we finally obtain

$$\langle W\rangle \approx \int_{\alpha_0}^{\alpha_f} \langle \partial_\alpha \phi\rangle_{ss}\,\mathrm{d}\alpha + \int_0^{t_p} \left[\dot{\alpha}^2(t) f(\alpha)\right]\mathrm{d}t \tag{252}$$

$$= B(\alpha_0, \alpha_{t_p}, \dot{\alpha}_0, \dot{\alpha}_{t_p}) + \int_0^{t_p} \mathcal{L}(\alpha, \dot{\alpha})\mathrm{d}t \tag{253}$$

We can now see that the first term is a boundary term (which we denote B) dependent on the initial and final state of the protocol (the free energy difference between the two states) while for the second term we can identify a Lagrangian $\mathcal{L}$ which explicitly depends on $\alpha(t)$ and $\dot{\alpha}(t)$, and which is the term we need to optimize.
Comparing with the passive case, we can identify $f(\alpha)$ as the friction coefficient (eq. (76)) and the boundary term as the "non-equilibrium free energy difference", which is not obvious how to interpret for a non-equilibrium ensemble.

Solving the Euler-Lagrange equations with the cost function $\mathcal{L}(\alpha, \dot{\alpha}) = \dot{\alpha}^2 f(\alpha)$ yields the differential equation

$$\dot{\alpha}^2 f'(\alpha) + 2\ddot{\alpha}f(\alpha) = 0 \tag{254}$$

exactly like the passive case.

We however solve this in a different way (following [6]) than in the passive case by using the fact that the Lagrangian is independent of time. We can then show that since the Lagrangian is independent of time, there exists a conserved quantity, which we call E, that in classical mechanics corresponds to the energy of the system [50]. This is a useful way of obtaining an expression for $\alpha(t)$. Note however that for our system it is not the energy as understood in the classical mechanics sense.

We are assuming that the Lagrangian $\mathcal{L}(\alpha, \dot{\alpha}) = \dot{\alpha}(t)^2 f(\alpha)$ is a solution to the Euler-Lagrange equations

$$\frac{\partial}{\partial \alpha}\mathcal{L}(\alpha,\dot{\alpha}) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{\alpha}}\mathcal{L}(\alpha,\dot{\alpha}) = 0 \tag{255}$$

which means that $\alpha(t)$ is a stationary point of the action

$$S[\alpha] = \int_0^{t_p} \mathcal{L}(\alpha,\dot{\alpha})\mathrm{d}t \tag{256}$$

The constant of motion E for this particular system is given by [50]

$$E = \frac{\partial \mathcal{L}}{\partial \dot{\alpha}}\frac{\mathrm{d}\alpha}{\mathrm{d}t} - \mathcal{L} \tag{257}$$

We want to show that this E is indeed constant along $\alpha(t)$. We have

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\alpha}}\frac{\mathrm{d}\alpha}{\mathrm{d}t}\right) - \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} \tag{258}$$

The second term is, using the chain rule, given by

$$\frac{\mathrm{d}\mathcal{L}(\alpha,\dot{\alpha})}{\mathrm{d}t} = \frac{\partial \mathcal{L}}{\partial \alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}t} + \frac{\partial \mathcal{L}}{\partial \dot{\alpha}}\frac{\mathrm{d}\dot{\alpha}}{\mathrm{d}t} + \frac{\partial \mathcal{L}}{\partial t} \tag{259}$$

$$\tag{260}$$

Substituting this into (258) we have

$$\begin{aligned}
\frac{\mathrm{d}E}{\mathrm{d}t} &= \frac{\partial \mathcal{L}}{\partial \dot{\alpha}}\ddot{\alpha} + \dot{\alpha}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\alpha}} - \left(\frac{\partial \mathcal{L}}{\partial \alpha}\dot{\alpha} + \frac{\partial \mathcal{L}}{\partial \dot{\alpha}}\ddot{\alpha} + \frac{\partial \mathcal{L}}{\partial t}\right) && (261)\\
&= \dot{\alpha}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\alpha}} - \frac{\partial \mathcal{L}}{\partial \alpha}\dot{\alpha} - \frac{\partial \mathcal{L}}{\partial t} && (262)\\
&= \dot{\alpha}\frac{\partial \mathcal{L}}{\partial \alpha} - \frac{\partial \mathcal{L}}{\partial \alpha}\dot{\alpha} - \frac{\partial \mathcal{L}}{\partial t} && (263)\\
&= -\frac{\partial \mathcal{L}}{\partial t} && (264)
\end{aligned}$$

where we used that $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\alpha}} = \frac{\partial \mathcal{L}}{\partial \alpha}$ from (255). As $\frac{\partial \mathcal{L}}{\partial t} = \frac{\partial}{\partial t}\dot{\alpha}^2 f(\alpha) = 0$, since the Lagrangian has no explicit time-dependence, this means that $\frac{\mathrm{d}E}{\mathrm{d}t} = 0$ implying that E is a constant.
The constant of motion E is from (257) given by

$$\begin{aligned}
E &= \frac{\partial \mathcal{L}}{\partial \dot{\alpha}}\dot{\alpha} - \mathcal{L} && (265)\\
&= \dot{\alpha}(2f(\alpha)\dot{\alpha}) - \dot{\alpha}^2 f(\alpha) && (266)\\
&= \dot{\alpha}^2 f(\alpha) && (267)
\end{aligned}$$

We can identify this as the kinetic energy of the system where $f(\alpha) = \frac{1}{2}m$.
We can then solve (265) for $\dot{\alpha}$ as

$$\dot{\alpha} = \pm\sqrt{\frac{E}{f(\alpha)}} \tag{268}$$

This is useful since $\dot{\alpha}$ now only depends on $f(\alpha)$, and we are free to choose E.
If we then separate the variables we get

$$\mathrm{d}\alpha = \pm\sqrt{\frac{E}{f(\alpha)}}\mathrm{d}t \tag{269}$$

giving us

$$\mathrm{d}t = \pm\sqrt{\frac{f(\alpha)}{E}}\mathrm{d}\alpha \tag{270}$$

Integrating both sides and assuming we start the protocol from $t = 0$ we get

$$t_p = \pm\int_{\alpha(0)}^{\alpha(t_p)} \sqrt{\frac{f(\alpha)}{E}}\mathrm{d}\alpha = \left|\int_{\alpha_0}^{\alpha_f} \sqrt{\frac{f(\alpha)}{E}}\mathrm{d}\alpha\right| \tag{271}$$

We now have expressions for calculating the optimal protocol and the optimal protocol durations with equations (268) and (271), respectively.

6.1 Calulating the response functions

Since $f(\alpha) = \int_{-\infty}^{t} R_1(\partial_\alpha\phi; t, t')(t-t')\mathrm{d}t'$ depends on the response function R_1 we need to obtain an expression for it in order to continue. As we discussed at the beginning of this chapter, an active system violates the classical fluctuation-dissipation theorem meaning that we cannot rely on linear response theory and appeal to the Boltzmann distribution as we did in the passive case. We thus take a different approach inspired by recent progress in non-equilibrium response theory [52, 53, 6].
By definition of the response function found earlier in (20) and (235) we have

$$R_1(X; t, t') = \left.\frac{\delta\langle X(t)\rangle}{\delta\Delta\alpha_{t,t'}}\right|_{\Delta\alpha\to 0} \tag{272}$$

for some observable $X(t)$.
Using a path-integral approach [9, 6], we write $X(t)$ as

$$\langle X(t)\rangle = N\int X(t)\exp\left(-\mathcal{S}\right)\mathcal{D}[\boldsymbol{\eta}, \boldsymbol{\upsilon}] \tag{273}$$

where $\mathcal{D}[\boldsymbol{\eta}, \boldsymbol{\upsilon}] = \prod_i P_\upsilon[\boldsymbol{\upsilon_i}]\mathcal{D}\boldsymbol{\eta_i}\mathcal{D}\boldsymbol{\upsilon_i}$, N is the normalization given by $N = \frac{1}{\int\exp(-\mathcal{S})\mathcal{D}[\boldsymbol{\eta},\boldsymbol{\upsilon}]}$ and $\mathcal{S} = \frac{1}{4D}\int_0^{t_p}(\dot{\boldsymbol{r}}_{\boldsymbol{i}} - \mu\boldsymbol{f_i})^2\mathrm{d}t$ is the associated action.
Substituting the path integral into (272), we get:

$$R_1 = \frac{\delta}{\delta \Delta\alpha_{t,t'}} \left(N \int X(t) \exp(-\mathcal{S}) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \right) \tag{274}$$

$$= \int X(t) \frac{\delta}{\delta \Delta\alpha} (N \exp(-\mathcal{S})) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \tag{275}$$

$$= \int X(t) \left(\frac{\delta N}{\delta \Delta\alpha} \exp(-\mathcal{S}) + N \frac{\delta \exp(-\mathcal{S})}{\delta \Delta\alpha} \right) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \tag{276}$$

Using the chain rule for functionals [54], the second term is equal to $N \frac{\delta \exp(-\mathcal{S})}{\delta \Delta\alpha} = N \exp(-\mathcal{S}) \left(\frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right)$, and the first term yields:

$$\frac{\delta N}{\delta \Delta\alpha} = \frac{\delta}{\delta \Delta\alpha} \left(\frac{1}{\int \exp(-\mathcal{S}) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}]} \right) \tag{277}$$

$$= -\frac{1}{(\int \exp(-\mathcal{S}) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}])^2} \frac{\delta}{\delta \Delta\alpha} \left(\int \exp(-\mathcal{S}) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \right) \tag{278}$$

$$= -N^2 \int \frac{\delta \exp(-\mathcal{S})}{\delta \Delta\alpha} \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \tag{279}$$

$$= -N^2 \int \exp(-\mathcal{S}) \left(-\frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \tag{280}$$

$$= N^2 \int \exp(-\mathcal{S}) \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \tag{281}$$

$$= N \left\langle \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right\rangle_{ss} \Bigg|_{\Delta\alpha \to 0} \tag{282}$$

Substituting these two terms into (276) we get:

$$R_1 = \int X(t) \left(\frac{\delta N}{\delta \Delta\alpha} \exp(-\mathcal{S}) + N \frac{\delta \exp(-\mathcal{S})}{\delta \Delta\alpha} \right) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \tag{283}$$

$$= \int X(t) \left(N \left\langle \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right\rangle_{ss} - N \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right) \exp(-\mathcal{S}) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \tag{284}$$

$$= \left\langle \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right\rangle_{ss} \int X(t) N \exp(-\mathcal{S}) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] - \int X(t) N \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \exp(-\mathcal{S}) \mathcal{D}[\boldsymbol{\eta}, \boldsymbol{v}] \tag{285}$$

$$= \left[\left\langle \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right\rangle_{ss} \langle X(t) \rangle - \left\langle X(t) \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right\rangle_{ss} \right]_{\Delta\alpha \to 0} \tag{286}$$

$$= -\left[\left\langle X(t) \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right\rangle_{ss} - \left\langle \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right\rangle_{ss} \langle X(t) \rangle \right]_{\Delta\alpha \to 0} \tag{287}$$

$$= -\left\langle \left\langle X(t) \frac{\delta \mathcal{S}}{\delta \Delta\alpha} \right\rangle \right\rangle \tag{288}$$

where we define $\langle \langle X(t) Y(t') \rangle \rangle = \langle X(t) Y(t') \rangle_{ss} - \langle X \rangle_{ss} \langle Y \rangle_{ss}$.

To get an expression for the functional derivative of $\mathcal{S}$ in (288), we consider $\mathcal{S}$ when we vary $\Delta\alpha$:

$$\mathcal{S}(\Delta\alpha + \delta\Delta\alpha) = \frac{1}{4D}\int_0^{t_p} (\dot{\boldsymbol{r}} - \mu \boldsymbol{f}[\Delta\alpha + \delta\Delta\alpha])^2 \, \mathrm{d}t' \tag{289}$$

and Taylor expanding $f[\Delta\alpha + \delta\Delta\alpha] = \boldsymbol{f}[\Delta\alpha] + \frac{\partial f}{\partial \Delta\alpha}\delta\Delta\alpha + \mathcal{O}(\delta\Delta\alpha^2)$ we get

$$\mathcal{S}(\Delta\alpha + \delta\Delta\alpha) = \frac{1}{4D}\int_0^{t_p} \left(\dot{\boldsymbol{r}} - \mu\left(\boldsymbol{f}[\Delta\alpha] + \frac{\partial \boldsymbol{f}}{\partial\Delta\alpha}\delta\Delta\alpha\right)\right)^2 \mathrm{d}t' \tag{290}$$

$$= \frac{1}{4D}\int_0^{t_p} \left(\dot{\boldsymbol{r}} - \mu\boldsymbol{f}[\Delta\alpha] + \mu\frac{\partial \boldsymbol{f}}{\partial\Delta\alpha}\delta\Delta\alpha\right)^2 \mathrm{d}t' \tag{291}$$

$$= \frac{1}{4D}\int_0^{t_p} \left((\dot{\boldsymbol{r}} - \mu\boldsymbol{f}[\Delta\alpha])^2 + \left(\mu\frac{\partial \boldsymbol{f}}{\partial\Delta\alpha}\delta\Delta\alpha\right)^2 + 2(\dot{\boldsymbol{r}} - \mu\boldsymbol{f}[\Delta\alpha]) \cdot \left(\mu\frac{\partial \boldsymbol{f}}{\partial\Delta\alpha}\delta\Delta\alpha\right)\right) \mathrm{d}t' \tag{292}$$

$$= \frac{1}{4D}\int_0^{t_p} (\dot{\boldsymbol{r}} - \mu\boldsymbol{f}[\Delta\alpha])^2 \mathrm{d}t' + \frac{\mu}{2D}\int_0^{t_p} \left((\dot{\boldsymbol{r}} - \mu\boldsymbol{f}[\Delta\alpha]) \cdot \frac{\partial \boldsymbol{f}}{\partial\Delta\alpha}\delta\Delta\alpha\right) \mathrm{d}t' + \frac{1}{4D}\int_0^{t_p} \left(\mu\frac{\partial \boldsymbol{f}}{\partial\Delta\alpha}\delta\Delta\alpha\right)^2 \mathrm{d}t' \tag{293}$$

$$= \mathcal{S}[\Delta\alpha] + \frac{\mu}{2D}\int_0^{t_p} \left((\dot{\boldsymbol{r}} - \mu\boldsymbol{f}[\Delta\alpha]) \cdot \frac{\partial \boldsymbol{f}}{\partial\Delta\alpha}\delta\Delta\alpha\right) \mathrm{d}t' + \mathcal{O}(\delta\Delta\alpha^2) \tag{294}$$

$$\tag{295}$$

Subtracting $\mathcal{S}[\Delta\alpha]$ and using the fact that $D = \mu k_B T$, we get

$$\mathcal{S}[\Delta\alpha + \delta\Delta\alpha] - \mathcal{S}[\Delta\alpha] = \frac{1}{2k_B T}\int_0^{t_p} \left((\dot{\boldsymbol{r}} - \mu\boldsymbol{f}[\Delta\alpha]) \cdot \frac{\partial \boldsymbol{f}}{\partial\Delta\alpha}\delta\Delta\alpha\right) \mathrm{d}t' + \mathcal{O}(\delta\Delta\alpha^2) \tag{296}$$

The functional derivative is defined through [41]

$$\delta J = \int_{x_1}^{x_2} \left(\frac{\delta J}{\delta y(x)}\right) \delta y(x) \mathrm{d}x \tag{297}$$

where δJ is the $\mathcal{O}(\epsilon)$-term of the variation $J[y + \epsilon\eta] - J[y]$, with ϵ a small constant and η an arbitrary function which vanishes at the boundaries. Correspondingly, for (296) we have

$$\delta\mathcal{S} = \frac{1}{2k_B T}\int_0^{t_p} \left((\dot{\boldsymbol{r}} - \mu\boldsymbol{f}[\Delta\alpha]) \cdot \frac{\partial \boldsymbol{f}}{\partial\Delta\alpha}\delta\Delta\alpha\right) \mathrm{d}t \tag{298}$$

making us able to identify the functional derivative of $\mathcal{S}$ as

$$\frac{\delta \mathcal{S}}{\delta \Delta\alpha(t')} = \frac{1}{2k_BT}\left(\frac{\partial \boldsymbol{f}}{\partial \Delta\alpha}\cdot(\dot{\boldsymbol{r}} - \mu \boldsymbol{f}[\Delta\alpha])\right)_{t'} \tag{299}$$

$$= \frac{1}{2k_BT}\left(\frac{\partial}{\partial\alpha}(-\nabla_i\phi + \frac{\boldsymbol{v}_i}{\mu})\cdot(\dot{\boldsymbol{r}} - \mu \boldsymbol{f}[\Delta\alpha])\right)_{t'} \tag{300}$$

$$= -\frac{1}{2k_BT}\left(\frac{\partial}{\partial\alpha}\nabla_i\phi\cdot(\dot{\boldsymbol{r}} - \mu \boldsymbol{f}[\Delta\alpha])\right)_{t'} \tag{301}$$

Substituting this into (288) and ignoring the higher order terms, we get for R_1:

$$R_1 = -\left\langle\left\langle X(t)\left(-\frac{1}{2k_BT}\left(\frac{\partial}{\partial\alpha}\nabla_i\phi\cdot(\dot{\boldsymbol{r}} - \mu \boldsymbol{f}[\alpha])\right)_{t'}\right)\right\rangle\right\rangle \tag{302}$$

$$= \frac{1}{2k_BT}\left(\left\langle\left\langle X(t)\left[\frac{\partial}{\partial\alpha}\nabla_i\phi\cdot\dot{\boldsymbol{r}}\right]_{t'}\right\rangle\right\rangle - \mu\left\langle\left\langle X(t)\left[\frac{\partial}{\partial\alpha}\nabla_i\phi\cdot\boldsymbol{f}\right]_{t'}\right\rangle\right\rangle\right) \tag{303}$$

$$= \frac{1}{2k_BT}\left(\frac{\mathrm{d}}{\mathrm{d}t'}\left\langle\left\langle X(t)\left[\frac{\partial}{\partial\alpha}\phi\right]_{t'}\right\rangle\right\rangle - \mu\left\langle\left\langle X(t)\left[\frac{\partial}{\partial\alpha}\nabla_i\phi\cdot\boldsymbol{f}\right]_{t'}\right\rangle\right\rangle\right) \tag{304}$$

where we used the chain rule [6]

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \sum_i \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t}\frac{\partial\phi}{\partial\boldsymbol{r}_i} + \frac{\partial\phi}{\partial\alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \boldsymbol{r}_i\cdot\nabla_i\phi \tag{305}$$

and where the last term containing $\dot{\alpha}$ is 0 as α is assumed to be constant here, since we are in the unperturbed, steady-state dynamics.
The result in (304) can be viewed as a generalization of the fluctuation-dissipation relation in (55) for an active system. The difference between the two is that the generalized FDT we have obtained here contains an extra term due to the force $\boldsymbol{f} = -\nabla\phi + \boldsymbol{v}/\mu$, which contains the active force $\boldsymbol{v}$ and thus corrects for the self-propulsion of the active particle. It can be shown that by taking $\boldsymbol{v} = \boldsymbol{0}$ that the standard fluctuation-dissipation theorem is then recovered [6].

We are now in a position to write the expression for the work more explicitly:

$$\langle W\rangle \approx B(\alpha_0, \alpha_{t_p}, \dot{\alpha}_0, \dot{\alpha}_{t_p}) + \int_0^{t_p}\left[\dot{\alpha}^2(t) f(\alpha)\right]\mathrm{d}t \tag{306}$$

$$= B(\alpha_0, \alpha_{t_p}, \dot{\alpha}_0, \dot{\alpha}_{t_p}) + \int_0^{t_p}\left[\dot{\alpha}^2(t)\int_{-\infty}^{t} R_1(\partial_\alpha\phi; t, t')(t-t')\mathrm{d}t'\right]\mathrm{d}t \tag{307}$$

$$= B(\alpha_0, \alpha_{t_p}, \dot{\alpha}_0, \dot{\alpha}_{t_p}) + \int_0^{t_p}\left[\dot{\alpha}^2(t)\int_{-\infty}^{t}\frac{1}{2k_BT}\left(\frac{\mathrm{d}}{\mathrm{d}t'}\left\langle\left\langle\left[\frac{\partial}{\partial\alpha}\phi\right]_t\left[\frac{\partial}{\partial\alpha}\phi\right]_{t'}\right\rangle\right\rangle\right.\right. \tag{308}$$

$$\left.\left. - \mu\left\langle\left\langle\left[\frac{\partial}{\partial\alpha}\phi\right]_t\left[\frac{\partial}{\partial\alpha}\nabla_i\phi\cdot\boldsymbol{f}\right]_{t'}\right\rangle\right\rangle\right)(t-t')\mathrm{d}t'\right]\mathrm{d}t$$

In this section, we have derived a general framework for the optimal control of active systems, using dynamical ensembles described by path integrals and methods from non-equilibrium response theory. This lead to a generalized fluctuation-dissipation relation in (302) and thus a different expression for the average work in (306) compared to the passive case. In the next section, we will apply this framework to an active particle.

7 Active Ornstein-Uhlenbeck particle in a harmonic potential

Having derived a framework for the optimal control of active systems, we are now in a position to apply this to the example of an active Ornstein-Uhlenbeck particle in a harmonic potential in one dimension. We calculate the optimal work done during the protocol $\alpha(t)$ step by step by first obtaining an expression for the active friction coefficient $f(\alpha)$, which involves calculating the response function R_1 and which in turn requires calculating the appropriate correlation functions.

The dynamics are given by

$$\dot{r} = \mu f + \sqrt{2D}\eta = -\mu\alpha r + v + \sqrt{2D}\eta \tag{309}$$

Figure 5 shows an illustration of the setup. The potential energy is given by $\phi(r) = 1/2\alpha r^2$ and is changed through α_0 to α_f which amounts to changing the strength of the potential.

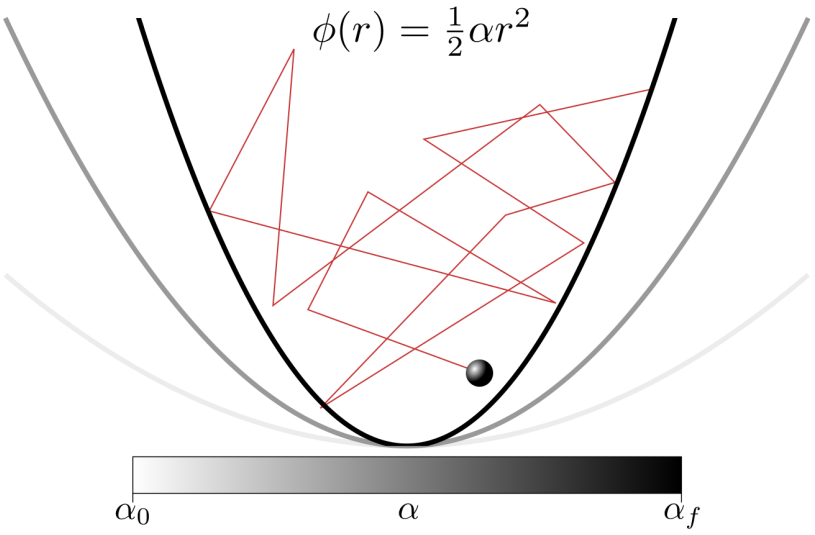

Figure 5: Illustration of the active dynamics: an active Ornstein-Uhlenbeck particle in a harmonic potential.

We have the following quantities that we need in the calculations:

$$f = -\alpha r + \frac{v}{\mu}, \tag{310}$$

$$\phi = \frac{1}{2}\alpha r^2, \tag{311}$$

$$\frac{\partial\phi}{\partial\alpha} = \frac{1}{2}r^2, \tag{312}$$

$$\nabla\frac{\partial\phi}{\partial\alpha} = \frac{\partial}{\partial r}\frac{\partial\phi}{\partial\alpha} = r \tag{313}$$

We start by calculating $f(\alpha) = \int_{-\infty}^{t} R_1(\partial_\alpha\phi; t, t')(t - t')\mathrm{d}t'$:

$$\int_{-\infty}^{t} R_1(\partial_\alpha\phi;t,t')(t-t')\mathrm{d}t' = \frac{1}{2k_BT}\int_{-\infty}^{t}\left(\frac{\mathrm{d}}{\mathrm{d}t'}\langle\!\langle\left[\partial_\alpha\phi\right]_t\left[\partial_\alpha\phi\right]_{t'}\rangle\!\rangle - \mu\langle\!\langle\left[\partial_\alpha\phi\right]_t\left[(\partial_r\partial_\alpha\phi)f\right]_{t'}\rangle\!\rangle\right)(t-t')\mathrm{d}t' \tag{314}$$

$$= \frac{1}{2k_BT}\int_{-\infty}^{t}\left(\frac{\mathrm{d}}{\mathrm{d}t'}\langle\!\langle\frac{1}{4}\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle - \mu\langle\!\langle\frac{1}{2}\left[r^2\right]_t\left[r\left(-\alpha r+\frac{v}{\mu}\right)\right]_{t'}\rangle\!\rangle\right)(t-t')\mathrm{d}t' \tag{315}$$

$$= \frac{1}{2k_BT}\int_{-\infty}^{t}\left(\frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}t'}\langle\!\langle\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle - \frac{1}{2}\mu\left(\langle\!\langle-\alpha\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle + \frac{1}{\mu}\langle\!\langle\left[r^2\right]_t\left[rv\right]_{t'}\right)\rangle\!\rangle\right)(t-t')\mathrm{d}t' \tag{316}$$

Applying integrating by parts to the first term, we get:

$$\frac{1}{2k_BT}\int_{-\infty}^{t}\frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}t'}\langle\!\langle\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle(t-t')\mathrm{d}t' = \frac{1}{8k_BT}\left(\left[\langle\!\langle\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle(t-t')\right]_{-\infty}^{t} - \int_{-\infty}^{t}\langle\!\langle\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle\frac{\mathrm{d}(t-t')}{\mathrm{d}t'}\mathrm{d}t'\right) \tag{317}$$

$$= \frac{1}{8k_BT}\int_{-\infty}^{t}\langle\!\langle\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle\mathrm{d}t' \tag{318}$$

$$\tag{319}$$

where the boundary terms vanish, the one at t trivially and the other one because by definition

$$\langle\!\langle\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle = \langle\left[r^2\right]_t\left[r^2\right]_{t'}\rangle_{ss} - \langle\left[r^2\right]_t\rangle_{ss}\langle\left[r^2\right]_{t'}\rangle_{ss} \tag{320}$$

and taking the limit $\lim_{t'\to-\infty}$, ie. for large separation between t and t', the correlations between $\left[r^2\right]_{t'}$ and $\left[r^2\right]_t$ are uncorrelated due to being in equilibrium. Thus we get

$$\lim_{t'\to-\infty}\langle\!\langle\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle = \langle\left[r^2\right]_t\rangle_{ss}\langle\left[r^2\right]_{t'}\rangle_{ss} - \langle\left[r^2\right]_t\rangle_{ss}\langle\left[r^2\right]_{t'}\rangle_{ss} = 0 \tag{321}$$

Using a change of variables, with $t'' = t - t'$ we can write:

$$\int_{-\infty}^{t}\langle\!\langle\left[r^2\right]_t\left[r^2\right]_{t'}\rangle\!\rangle\mathrm{d}t' = \int_{t''=t-(-\infty)}^{t''=t-t}\langle\!\langle\left[r^2\right]_t\left[r^2\right]_{t-t''}\rangle\!\rangle(-\mathrm{d}t'') \tag{322}$$

$$= -\int_{\infty}^{0}\langle\!\langle\left[r^2\right]_{t'+t''}\left[r^2\right]_{t'}\rangle\!\rangle\mathrm{d}t'' \tag{323}$$

$$= \int_{0}^{\infty}\langle\!\langle\left[r^2\right]_{t''}\left[r^2\right]_0\rangle\!\rangle\mathrm{d}t'' \tag{324}$$

where we used that $\frac{\mathrm{d}t''}{\mathrm{d}t'} = \frac{\mathrm{d}(t-t')}{\mathrm{d}t'} = -1$ such that $\mathrm{d}t'' = -\mathrm{d}t'$ and stationarity of the correlation function $\langle\langle \left[r^2\right]_{t'+t''} \left[r^2\right]_{t'} \rangle\rangle = \langle\langle \left[r^2\right]_{t''} \left[r^2\right]_0 \rangle\rangle$

So we get:

$$\int_{-\infty}^{t} R_1(\partial_\alpha \phi; t, t')(t-t')\mathrm{d}t' = \frac{1}{2k_B T} \int_{-\infty}^{t} \left(\frac{1}{4} \frac{\mathrm{d}}{\mathrm{d}t'} \langle\langle \left[r^2\right]_t \left[r^2\right]_{t'} \rangle\rangle - \frac{1}{2}\mu \left(\langle\langle -\alpha \left[r^2\right]_t \left[r^2\right]_{t'} \rangle\rangle + \frac{1}{\mu} \langle\langle \left[r^2\right]_t [rv]_{t'} \right) \rangle\rangle \right) (t-t')\mathrm{d}t' \tag{325}$$

$$= \frac{1}{8k_B T} \int_0^\infty \langle\langle \left[r^2\right]_{t''} \left[r^2\right]_0 \rangle\rangle \mathrm{d}t'' \tag{326}$$

$$- \frac{1}{4k_B T}\mu \int_{-\infty}^{t} \left(\langle\langle -\alpha \left[r^2\right]_t \left[r^2\right]_{t'} \rangle\rangle + \frac{1}{\mu} \langle\langle \left[r^2\right]_t [rv]_{t'} \right) \rangle\rangle (t-t')\mathrm{d}t' \tag{327}$$

$$= \frac{1}{8k_B T} \int_0^\infty \langle\langle \left[r^2\right]_{t''} \left[r^2\right]_0 \rangle\rangle \mathrm{d}t'' \tag{328}$$

$$- \frac{1}{4k_B T} \int_0^\infty \left(\langle\langle -\alpha\mu \left[r^2\right]_{t''} \left[r^2\right]_0 \rangle\rangle + \langle\langle \left[r^2\right]_{t''} [rv]_0 \rangle\rangle \right) t'' \mathrm{d}t'' \tag{329}$$

$$= \frac{1}{4k_B T} \int_0^\infty \left[\left(\frac{1}{2} + \alpha\mu t'' \right) \langle\langle \left[r^2\right]_{t''} \left[r^2\right]_0 \rangle\rangle - \langle\langle \left[r^2\right]_{t''} [rv]_0 \rangle\rangle t'' \right] \mathrm{d}t'' \tag{330}$$

In order to minimize the work, we need an expression for $f(\alpha)$ which involves calculating the correlation functions $\langle\langle \left[r^2\right]_{t''} \left[r^2\right]_0 \rangle\rangle$ and $\langle\langle \left[r^2\right]_{t''} [rv]_0 \rangle\rangle$.

Here we first need to solve the Langevin equation for the system given by (309). Various methods are viable given this is a linear stochastic differential equation. For example we have used variation of parameters, however not presented here due to space considerations.
The solution to (309) is given by [6]:

$$r(t) = \int_{-\infty}^{t} \exp(-\alpha\mu(t-t')) \left[v(t') + \sqrt{2D}\eta(t') \right] \mathrm{d}t' \tag{331}$$

We can now calculate the correlation functions:

$$\left\langle [r^2]_t [r^2]_s \right\rangle = \left\langle \int_{-\infty}^{t} \int_{-\infty}^{t} \exp(-\alpha\mu(t-t') - \alpha\mu(t-t'')) \left[\sqrt{2D}\eta(t') + v(t')\right] \left[\sqrt{2D}\eta(t'') + v(t'')\right] \mathrm{d}t'\mathrm{d}t'' \times \right. \tag{332}$$

$$\left. \int_{-\infty}^{s} \int_{-\infty}^{s} \exp(-\alpha\mu(s-t''') - \alpha\mu(s-t'''')) \left[\sqrt{2D}\eta(t''') + v(t''')\right] \left[\sqrt{2D}\eta(t'''') + v(t'''')\right] \mathrm{d}t'''\mathrm{d}t'''' \right\rangle \tag{333}$$

$$= \int_{-\infty}^{t} \mathrm{d}t_1 \int_{-\infty}^{t} \mathrm{d}t_2 \int_{-\infty}^{s} \mathrm{d}t_3 \int_{-\infty}^{s} \mathrm{d}t_4 \exp\left(\sum_{i=1}^{i=4} t_i - 2(t+s)\right) \times \tag{334}$$

$$\left\langle \left[\sqrt{2D}\eta_1 + v_1\right] \left[\sqrt{2D}\eta_2 + v_2\right] \left[\sqrt{2D}\eta_3 + v_3\right] \left[\sqrt{2D}\eta_4 + v_4\right] \right\rangle \tag{335}$$

where we simplified the notation using $v(t_i) = v_i$, $\eta(t_i) = \eta_i$ to replace the dummy variables $t', t'', ..$ [6]
Using the same method, we get similarly for the other correlation function:

$$\left\langle [rv]_t [r^2]_s \right\rangle = \int_{-\infty}^{t} \mathrm{d}t_1 \int_{-\infty}^{t} \mathrm{d}t_2 \int_{-\infty}^{s} \mathrm{d}t_3 \int_{-\infty}^{s} \exp\left(\sum_{i=1}^{i=3} t_i - (t+2s)\right) \times \tag{336}$$

$$\left\langle \left[\sqrt{2D}\eta_1 + v_1\right] v(t) \left[\sqrt{2D}\eta_2 + v_2\right] \left[\sqrt{2D}\eta_3 + v_3\right] \right\rangle \tag{337}$$

Calculating them involves the product

$$\left\langle \left[\sqrt{2D}\eta_1 + v_1\right] \left[\sqrt{2D}\eta_2 + v_2\right] \left[\sqrt{2D}\eta_3 + v_3\right] \left[\sqrt{2D}\eta_4 + v_4\right] \right\rangle = \tag{338}$$

$$\left\langle (\sqrt{2D}\eta_1\sqrt{2D}\eta_2 + \sqrt{2D}\eta_1 v_2 + \sqrt{2D}\eta_2 v_1 + v_1 v_2) \cdot (\sqrt{2D}\eta_3\sqrt{2D}\eta_4 + \sqrt{2D}\eta_3 v_4 + \sqrt{2D}\eta_4 v_3 + v_3 v_4) \right\rangle \tag{339}$$

$$\tag{340}$$

Here terms like $\langle \eta_1\eta_2\eta_3 v_4 \rangle = \langle \eta_1\eta_2 \rangle \langle \eta_3 v_4 \rangle = 0$ due to causality (i.e. $\langle \eta_3 v_4 \rangle = 0$).
In the following we let $k = \sqrt{2D}$, it can then be shown by straightforward algebra that the above product gives:

$$k^4 \langle \eta_1\eta_2\eta_3\eta_4 \rangle + k^2 \langle \eta_1\eta_2 v_3 v_4 \rangle + k^2 \langle \eta_1\eta_3 v_2 v_4 \rangle + k^2 \langle \eta_3\eta_4 v_2 v_3 \rangle \tag{341}$$

$$+k^2 \langle \eta_2\eta_3 v_1 v_4 \rangle + k^2 \langle \eta_2\eta_4 v_1 v_3 \rangle + k^2 \langle \eta_3\eta_4 v_1 v_2 \rangle + \langle v_1 v_2 v_3 v_4 \rangle \tag{342}$$

The last term we can rewrite using Isserlis' theorem [55]:

$$\langle v_1 v_2 v_3 v_4 \rangle = \langle v_1 v_2 \rangle \langle v_3 v_4 \rangle + \langle v_1 v_3 \rangle \langle v_2 v_4 \rangle + \langle v_1 v_4 \rangle \langle v_2 v_3 \rangle \tag{343}$$

which holds for Gaussian random variables with zero-mean.
Using this fact and that $\langle \eta_1\eta_2\eta_3\eta_4 \rangle = \langle \eta_1\eta_2 \rangle \langle \eta_3\eta_4 \rangle$ (because the noise is uncorrelated for times that are not overlapping).

We can then use Isserlis' theorem and the fact that $\langle \eta_1 \eta_2 \rangle = \delta(t_1 - t_2)$ and $\langle v(t)v(t') \rangle = \frac{D_1}{\tau} \exp(-|t-t'|/\tau)$ to calculate the correlation functions.
As an example and illustration of the method, the correlation function in (332) would take the form

$$\left\langle [r^2]_t [r^2]_s \right\rangle = \int_{-\infty}^{t} \mathrm{d}t_1 \int_{-\infty}^{t} \mathrm{d}t_2 \int_{-\infty}^{s} \mathrm{d}t_3 \int_{-\infty}^{s} \mathrm{d}t_4 \exp\left(t_1 + t_2 + t_3 + t_4 - 2(t+s)\right) \times \tag{344}$$

$$\left(k^4 \left(\langle \eta_1 \eta_2 \rangle \langle \eta_3 \eta_4 \rangle + \langle \eta_1 \eta_3 \rangle \langle \eta_2 \eta_4 \rangle + \langle \eta_1 \eta_4 \rangle \langle \eta_2 \eta_3 \rangle\right)\right. \tag{345}$$

$$+ k^2 \left(\langle \eta_1 \eta_2 \rangle \langle v_3 v_4 \rangle + \langle \eta_1 \eta_3 \rangle \langle v_2 v_4 \rangle + \langle \eta_1 \eta_4 \rangle \langle v_2 v_3 \rangle\right) \tag{346}$$

$$+ k^2 \left(\langle \eta_2 \eta_3 \rangle \langle v_1 v_4 \rangle + \langle \eta_2 \eta_4 \rangle \langle v_1 v_3 \rangle + \langle \eta_3 \eta_4 \rangle \langle v_1 v_2 \rangle\right) \tag{347}$$

$$\left. + \langle v_1 v_2 \rangle \langle v_3 v_4 \rangle + \langle v_1 v_3 \rangle \langle v_2 v_4 \rangle + \langle v_1 v_4 \rangle \langle v_2 v_3 \rangle\right) \tag{348}$$

$$= \int_{-\infty}^{t} \mathrm{d}t_1 \int_{-\infty}^{t} \mathrm{d}t_2 \int_{-\infty}^{s} \mathrm{d}t_3 \int_{-\infty}^{s} \mathrm{d}t_4 \exp\left(t_1 + t_2 + t_3 + t_4 - 2(t+s)\right) \times \tag{349}$$

$$\left(\sqrt{2D}^4 \left(\delta(t_1 - t_2)\delta(t_3 - t_4) + \delta(t_1 - t_3)\delta(t_2 - t_4) + \delta(t_1 - t_4)\delta(t_2 - t_3)\right)\right. \tag{350}$$

$$+ \sqrt{2D}^2 \left(\frac{D_1}{\tau} \exp(-|t_3 - t_4|/\tau)\delta(t_1 - t_2) + \frac{D_1}{\tau} \exp(-|t_2 - t_4|/\tau)\delta(t_1 - t_3) \right. \tag{351}$$

$$+ \frac{D_1}{\tau} \exp(-|t_2 - t_3|/\tau)\delta(t_1 - t_4) + \frac{D_1}{\tau} \exp(-|t_1 - t_4|/\tau)\delta(t_2 - t_3) \tag{352}$$

$$\left. + \frac{D_1}{\tau} \exp(-|t_1 - t_3|/\tau)\delta(t_2 - t_4) + \frac{D_1}{\tau} \exp(-|t_1 - t_2|/\tau)\delta(t_3 - t_4) \right) \tag{353}$$

$$+ \frac{D_1}{\tau}^2 \exp(-|t_1 - t_2|/\tau) \exp(-|t_3 - t_4|/\tau) \tag{354}$$

$$+ \frac{D_1}{\tau}^2 \exp(-|t_1 - t_3|/\tau) \exp(-|t_2 - t_4|/\tau) \tag{355}$$

$$\left. + \frac{D_1}{\tau}^2 \exp(-|t_1 - t_4|/\tau) \exp(-|t_2 - t_3|/\tau)\right) \tag{356}$$

and we would get similar expressions for $\left\langle [rv]_t [r^2]_s \right\rangle$ (see also [6]). All these have been calculated using mathematical software.

When calculating the correlation functions such as $\langle\langle \left[r^2\right]_t \left[r^2\right]_{t'} \rangle\rangle = \langle \left[r^2\right]_t \left[r^2\right]_{t'} \rangle_{ss} - \langle \left[r^2\right]_t \rangle_{ss} \langle \left[r^2\right]_{t'} \rangle_{ss}$ we also need the steady-state averages $\langle r^2 \rangle_{ss}$ and $\langle rv \rangle_{ss}$. Once again we can calculate these using Ito's lemma. Define $f = r^2$ to obtain

$$\frac{\mathrm{d}r^2}{\mathrm{d}t} = (-\mu\alpha r + v)\frac{\partial r^2}{\partial r} + \frac{\sqrt{2D}^2}{2}\frac{\partial^2 r^2}{\partial r^2} + \sqrt{2D}\frac{\partial r^2}{\partial r}\eta \tag{357}$$

$$= 2r(-\mu\alpha r + v) + 2D + 2\sqrt{2D}\eta \tag{358}$$

$$= 2(-\mu\alpha r^2 + rv + D + \sqrt{2D}\eta) \tag{359}$$

now dividing by 2 and taking the ensemble average we get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle r^2\right\rangle = -\mu\alpha\left\langle r^2\right\rangle + \langle rv\rangle + D \tag{360}$$

Similarly for $f = rv$ we obtain by using the multidimensional Ito formula [56]

$$\frac{\mathrm{d}rv}{\mathrm{d}t} = (-\mu\alpha r + v)\frac{\partial rv}{\partial r} - \frac{v}{\tau}\frac{\partial rv}{\partial v} + \frac{\sqrt{2D}^2}{2}\frac{\partial^2 rv}{\partial r^2} + \frac{\sqrt{2D_1}^2}{2}\frac{\partial^2 rv}{\partial v^2} + \sqrt{2D}\frac{\partial rv}{\partial r}\eta + \sqrt{2D}\frac{\partial rv}{\partial v}\zeta \tag{361}$$

$$= -\mu\alpha rv + v^2 - \frac{rv}{\tau} + \sqrt{2D}v\eta + \sqrt{2D}r\zeta \tag{362}$$

And when taking the ensemble average we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle rv\rangle = -\mu\alpha\langle rv\rangle + \frac{D_1}{\tau} - \frac{\langle rv\rangle}{\tau} = -\langle rv\rangle\left(\mu\alpha + \frac{1}{\tau}\right) + \frac{D_1}{\tau} \tag{363}$$

We then have in steady-state that (363) and (360) each are equal to 0 which yields

$$\langle rv\rangle_{ss} = \frac{D_1}{\tau(\mu\alpha + \frac{1}{\tau})} = \frac{D_1}{(\tau\mu\alpha + 1)} \tag{364}$$

and

$$\left\langle r^2\right\rangle_{ss} = \frac{\langle rv\rangle}{\mu\alpha} + \frac{D}{\mu\alpha} = \frac{D_1}{(\tau\mu\alpha + 1)\mu\alpha} + \frac{D}{\mu\alpha} = \frac{1}{\mu\alpha}\left(\frac{D_1}{(\tau\mu\alpha + 1)} + D\right) \tag{365}$$

To check that the analytically derived correlation functions are correct, we calculate them numerically and compare with the analytical results:

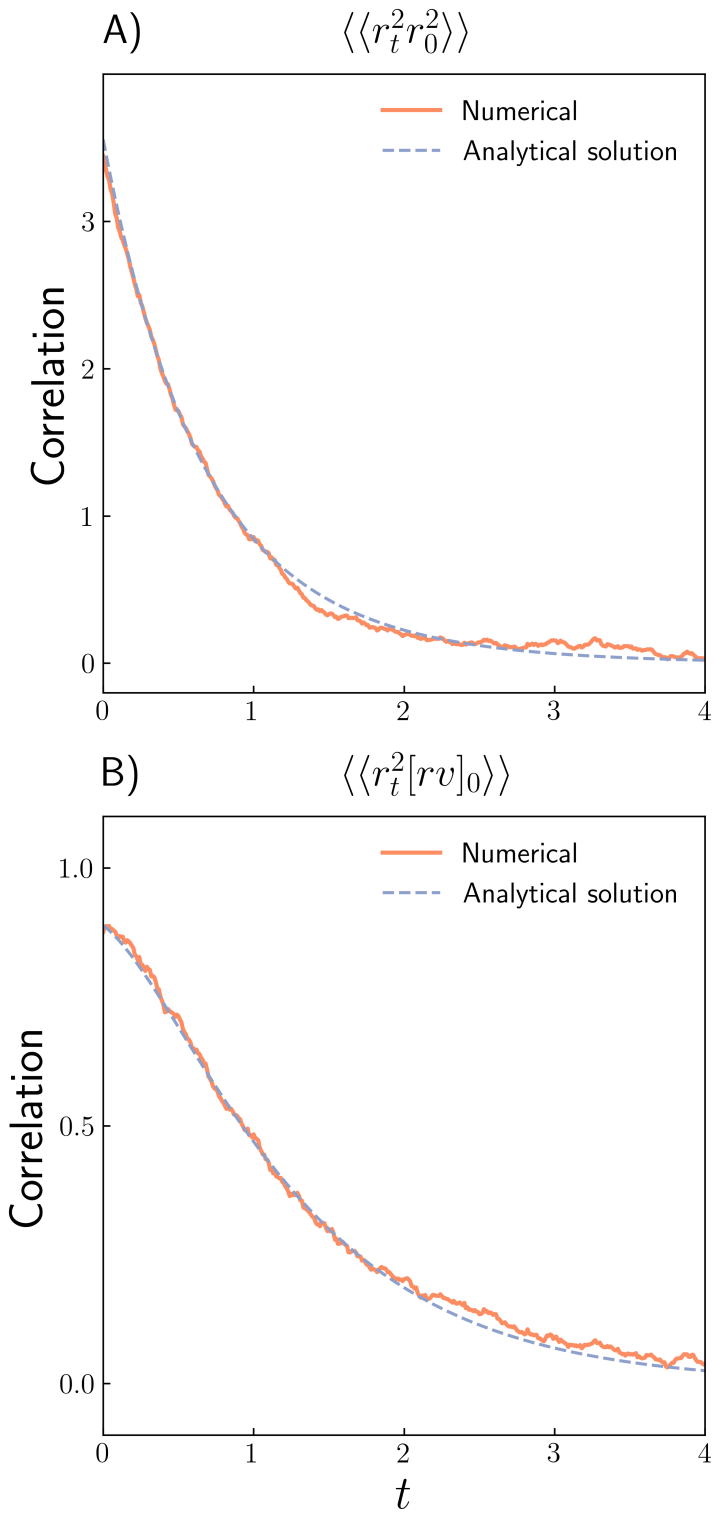

Figure 6: The 2 correlation functions used in the friction coefficient $f(\alpha)$ compared with the solution obtained from numerical simulations, averaged over $N = 10000$ trajectories. A) Correlation function $\langle\langle [r^2]_t [r^2]_0 \rangle\rangle$ vs. time. B) Correlation function $\langle\langle [r^2]_t [rv]_0 \rangle\rangle$ vs. time.

By visual inspection, we find that the calculated analytical correlations are largely in agreement with the numerical simulations, considering the number of averaged trajectories is $n = 10000$.
Continuing on, the expression we end up with for $f(\alpha)$ is:

$$f(\alpha) = \int_{-\infty}^{t} R_1(\partial_\alpha \phi; t, t')(t - t')\mathrm{d}t' \tag{366}$$

$$= \frac{1}{4k_B T} \int_0^{\infty} \left[\left(\frac{1}{2} + \alpha\mu t'' \right) \langle\langle \left[r^2\right]_{t''} \left[r^2\right]_0 \rangle\rangle - \langle\langle \left[r^2\right]_{t''} \left[rv\right]_0 \rangle\rangle t'' \right] \mathrm{d}t'' \tag{367}$$

$$= \frac{D_1(\alpha^3\tau^3\mu^3 D + 3\tau^2(D + \frac{4D_1}{3})\mu^2\alpha^2 + 3\mu\tau(D + D_1)\alpha + D + D_1)}{4\alpha^3\mu^3 k_B T(\alpha\mu\tau + 1)^3} \tag{368}$$

Considering the 2 limits $\tau \to \infty$ and $\tau \to 0$, we can recover the passive dynamics and use this to verify if the friction coefficient reduces to what we would expect in the passive case. Both limits reproduce a passive Brownian particle but with different strengths for the noise. Taking the limit $\tau \to 0$, we get for the noise

$$\lim_{\tau\to 0} \langle v_{i,\alpha}(t) v_{j,\beta}(t') \rangle = \lim_{\tau\to 0} \delta_{i,j}\delta_{\alpha,\beta} \frac{D_1}{\tau} \exp(-|t - t'|/\tau) = \delta_{i,j}\delta_{\alpha,\beta} 2D_1 \delta(t - t') \tag{369}$$

where we used the Dirac delta function representation $\delta(t) = \lim_{\tau\to 0} \frac{1}{2\tau} \exp(-|t|/\tau)$, which means that $\boldsymbol{v_i}$ becomes a zero-mean Gaussian white noise [57] and the dynamics of the original active Ornstein-Uhlenbeck particle in (229) becomes a passive Brownian particle with noise strength $D + D_1$.
Taking the other limit where $\tau \to \infty$, we obtain the passive Brownian particle dynamics with noise strength D and which also reproduces the friction coefficient for the passive case:

$$\lim_{\tau\to\infty} f(\alpha) = \lim_{\tau\to\infty} \frac{D_1(\alpha^3\tau^3\mu^3 D + 3\tau^2(D + \frac{4D_1}{3})\mu^2\alpha^2 + 3\mu\tau(D + D_1)\alpha + D + D_1)}{4\alpha^3\mu^3 k_B T(\alpha\mu\tau + 1)^3} \tag{370}$$

$$= \frac{D^2}{4\alpha^3\mu^3 k_B T} \tag{371}$$

$$= \frac{(k_B T\mu)^2}{4\alpha^3\mu^3 k_B T} \tag{372}$$

$$= \frac{k_B T}{4\mu\alpha^3} \tag{373}$$

which is the same expression for the friction coefficient we obtained in equation (191).

We then use (368) to calculate the work where $f(\alpha)$ is used in the Lagrangian term. The work is then given by

$$\langle W \rangle \approx \int_{\alpha_0}^{\alpha_f} \langle \partial_\alpha \phi \rangle_{ss} \, \mathrm{d}\alpha + \int_0^{t_p} \left[\dot{\alpha}^2(t) f(\alpha) \right] \mathrm{d}t \tag{374}$$

$$= \frac{1}{2} \int_{\alpha_0}^{\alpha_f} \left\langle r^2 \right\rangle_{ss} \mathrm{d}\alpha + \int_0^{t_p} \left[\dot{\alpha}^2(t) \frac{D_1(\alpha^3\tau^3\mu^3 D + 3\tau^2(D + \frac{4D_1}{3})\mu^2\alpha^2 + 3\mu\tau(D + D_1)\alpha + D + D_1)}{4\alpha^3\mu^3 k_B T(\alpha\mu\tau + 1)^3} \right] \mathrm{d}t \tag{375}$$

$$= \frac{1}{2\mu} \int_{\alpha_0}^{\alpha_f} \frac{1}{\alpha} \left(\frac{D_1}{(\tau\mu\alpha + 1)} + D \right) \mathrm{d}\alpha + \int_0^{t_p} \left[\dot{\alpha}^2(t) \frac{D_1(\alpha^3\tau^3\mu^3 D + 3\tau^2(D + \frac{4D_1}{3})\mu^2\alpha^2 + 3\mu\tau(D + D_1)\alpha + D + D_1)}{4\alpha^3\mu^3 k_B T(\alpha\mu\tau + 1)^3} \right] \mathrm{d}t \tag{376}$$

To sum up, in this section we have illustrated how to use the optimal control framework for an active system (which we derived in section 6) on a model system consisting of an active Ornstein-Uhlenbeck particle in a harmonic potential in 1D. We calculated the necessary quantities in the friction coefficient $f(\alpha)$, importantly we showed how to obtain expressions for the correlation functions that enter into $f(\alpha)$. We also calculated the correlation functions numerically in order to check that the analytically derived correlation functions were indeed correct. Furthermore, we also checked the limiting case of $\tau \to \infty$ to verify that the friction coefficient in this case reduced to the friction coefficient corresponding to the passive case, as we would expect. Lastly, having obtained $f(\alpha)$ we then obtained an analytical expression for the work. In the next section we simulate the system and verify whether our analytical expression for the work agrees with the numerical solution.

8 Numerical simulation of an active Ornstein-Uhlenbeck particle in a harmonic potential

We now turn to simulating the active Ornstein-Uhlenbeck particle in a harmonic potential in 1D. As we now have an analytical expression for the work from (376), we can compare this with the numerically calculated work. We use the same Euler-Maruyama scheme as in the passive case and discretize the dynamics accordingly. When simulating the system we then use (269) and (271) to numerically calculate $\alpha(t)$ and obtain the optimal protocol duration t_p by choosing an appropriate value of E.
Using the following parameters for the simulations

active Ornstein-Uhlenbeck particle	α_0	α_f	dt	μ	μ_1	D	D_1	τ	N
α-protocol	1	2	0.001	1	1	1	1	2	500,000

Table 2: Parameter values for the simulation of an active Ornstein-Uhlenbeck particle

we start by calculating $f(\alpha)$ which gives

$$f(\alpha) = \frac{4\alpha^3 + 14\alpha^2 + 6\alpha + 1}{2\alpha^3(2\alpha + 1)} \tag{377}$$

The work in (376) then takes the form

$$\langle W \rangle \approx \frac{1}{2} \int_{\alpha_0}^{\alpha_f} \frac{1}{\alpha} \left(\frac{1}{(2\alpha + 1)} + 1 \right) \mathrm{d}\alpha + \int_0^{t_p} \left[\dot{\alpha}^2(t) \frac{4\alpha^3 + 14\alpha^2 + 6\alpha + 1}{2\alpha^3(2\alpha + 1)} \right] \mathrm{d}t \tag{378}$$

Using that we already know the solution for $\alpha(t)$ is given by $\alpha(t) = \frac{64c_1^2}{(c_1 t + c_2)^6}$ by virtue of the fact that the Euler-Lagrange equations take the same form as in the passive case, we get for the work

$$\langle W \rangle = 0.4377343691 + \frac{0.1494582537}{t_p} \tag{379}$$

As the protocol durations are determined through the constant E in (271), we partition E into 3 appropriate intervals to get a sufficiently long protocol duration when simulating the system. We then plot the solution in (379) alongside the numerical solution simulated with the parameters given in Table 2. The results are presented in the following plot

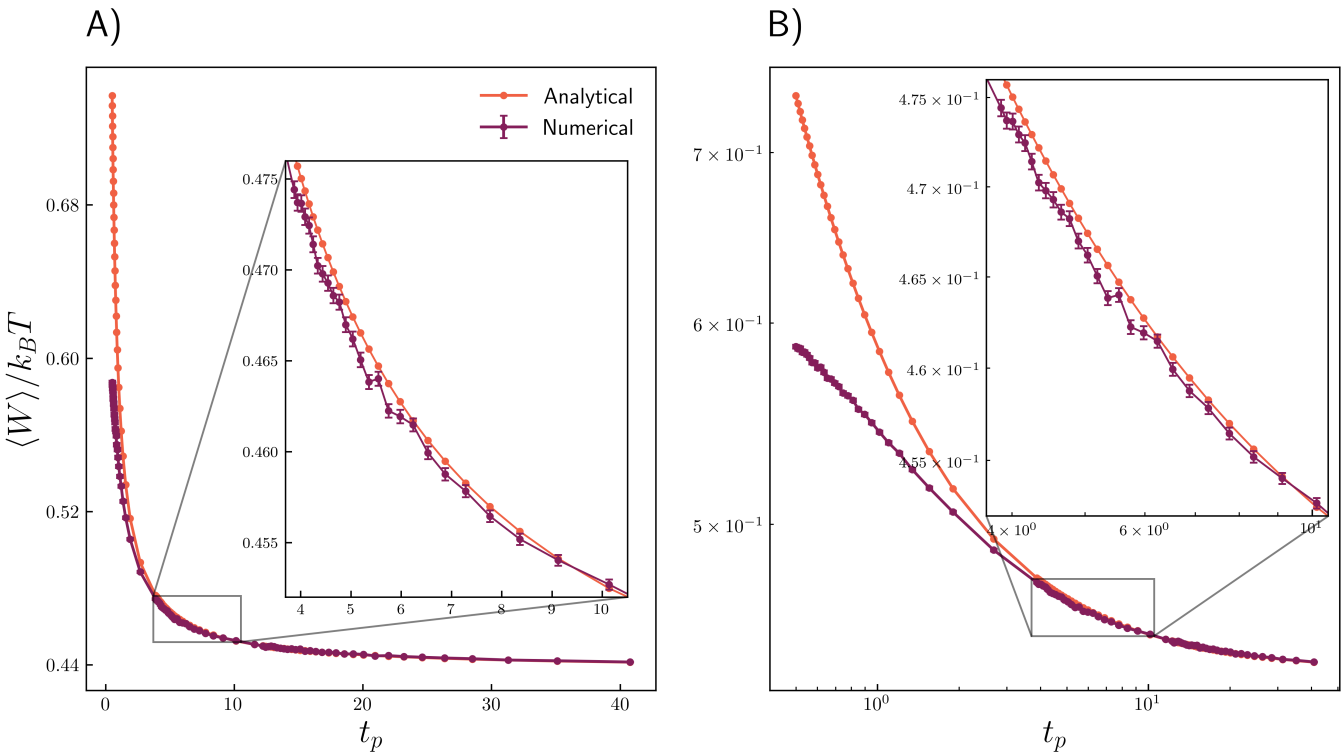

Figure 7: Average work for an active particle in a one dimensional harmonic potential. A) Average work vs. protocol duration for an active Ornstein-Uhlenbeck particle in 1D in a harmonic potential driven by the α-protocol that changes the strength of the potential. The inset is zoomed in on the part where the numerical and analytical solutions start to diverge from each other. Error bars are 1 standard deviation scaled by $1/\sqrt{N}$ where N is the number of simulations. B) Log-log plot of A).

As we can see from the plot, for long protocol durations the numerical and approximated (analytical) solutions converge, as in the passive case. For shorter protocol times with $t_p \sim 8$, the two solutions start to separate. This is as expected since one of our assumptions when deriving the analytical expressions was slow driving of the protocols. From this we can also deduce that the dissipated work (the second term in (374)) declines as $1/t_p$ like the passive case, which is more interesting and leads to some considerations. In an active system one would expect that the self-propulsion causes the system to dissipate heat to the environment but not work, as the self-propulsion is a force that arises from the particle. This implies that when driving the system with the protocol $\alpha(t)$, the dissipated work would come primarily from the external protocol and would thus lead to the same behavior as in the passive case. Considering these physical characteristics of the system, this result is as expected. In [6] they consider the heat instead of the work and show that the average heat grows when $t_p \to \infty$, contrary to the results for the average work we have obtained here. This is because the system dissipates heat all the time due to the activity of the particle.

In the below plot, we also show the difference between the analytical and numerical solutions, again confirming that it is around $t_p \sim (5, 10)$ that the approximation holds.

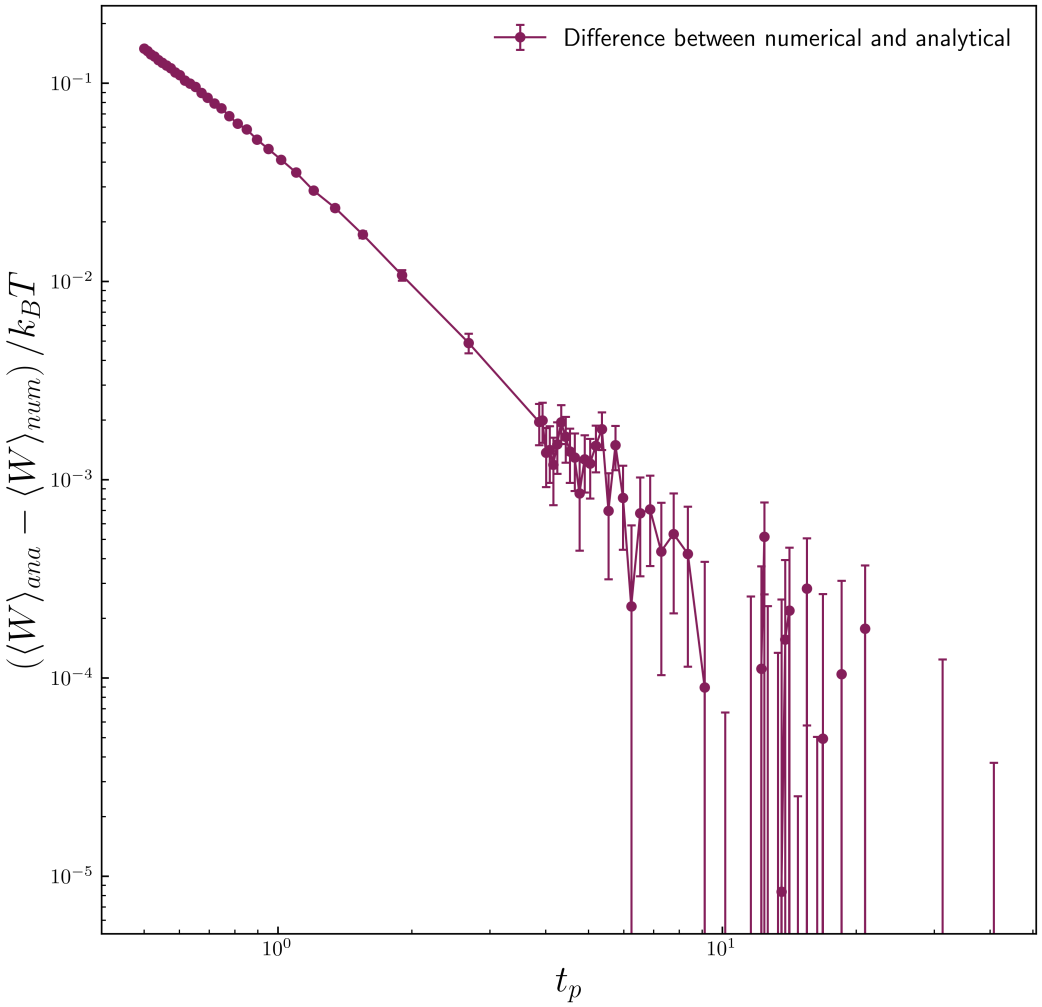

Figure 8: The difference between the numerical and analytical solutions of the average work vs. protocol duration. Table 2 specifies the parameters used.

We have in this section simulated an active system in the form of an active Ornstein-Uhlenbeck particle in a harmonic potential in 1D and calculated the work associated with driving the protocol $\alpha(t)$ from $\alpha_0 = 1$ to $\alpha_f = 2$ and then plotted the numerically calculated work vs. the analytical expression for the work. We found good agreement between the two solutions in the slow-driving regime. In the next section, we discuss the results we have obtained in this thesis and sketch some perspectives for future work.

9 Discussion

In this work we have presented frameworks for working with optimal protocols both for passive and active systems. We have simulated these and compared the theoretical findings with the numerical investigations. This has led to results on the optimal control of these systems, giving rise to questions and discussions. In the following we will address the most prominent of these questions, as well as the challenges and directions for future research.

The results we have obtained in this work for the passive and active systems is a systematic construction of a theoretical framework which extends and builds upon the framework developed in [5] and [6] to obtain the optimal protocols for minimizing the work (dissipated work in the passive case) during the driving of the system between two protocol values λ_0 and λ_f. In both the active and the passive case, the optimal protocol is the longest one, ie. $t_p \to \infty$, also called the quasi-static limit [9, 6]. This is not immediately obvious for the active case, but for the passive case, this should be expected. The reason why is that, as we know from equilibrium thermodynamics, driving the system quasi-statically results in an energy change given by the change in free energy between the initial and final states of the process, whereas a finite-time process will always incur some dissipation [45, 47].

What is surprising is how robust both results seem to be against varying the protocol duration. We would expect a large deviation for smaller protocol times as this amounts to driving the protocol fast, but as illustrated in Figures 2 and 7, even for protocol durations $t_p \sim 8 - 10$, these are still reasonable approximations. The assumptions we made when deriving the theoretical framework were slow and smooth driving of the protocols (i.e. that we drive the protocols slowly compared to the characteristic time scale with which the system relaxes naturally) and, additionally in the passive case, the near-equilibrium requirement of linear response. It is therefore to be expected that the results diverge for smaller protocol durations (faster driving), but it is remarkable that the difference between the exact and approximated solutions is not larger than observed. Indeed, it has been shown recently that optimal protocols obtained by linear response theory perform well outside their expected range of validity [44, 46, 58]. Interestingly, and reported even more recently, the passive framework has been applied to the exact same active system considered in this work, showing that protocols obtained from the passive framework surpass the performance of naive protocols [57]. This again confirms the robustness of the passive framework.

As we have built the active framework with the same tools as the passive framework [5] using response theory, this robustness carries over to the active framework which displays surprisingly good agreement between the numerical solution and analytical expression for the work, also for shorter protocol durations (seen for both the passive and active cases for the same protocol (the strength of the harmonic trap) in Figures 2 and 7).
Theoretically, it should be possible to include higher-order response functions to obtain agreement with even faster driving, though it would quickly become cumbersome to calculate the response functions analytically. This is one of the downsides of our approach, being constrained to (linear) response theory and thus sacrificing generality for convenience. In relation to this, the example of a Brownian particle in a harmonic potential which we used to showcase the optimal protocol framework in the passive case is the only known analytical solution for this type of optimal control [44]. This and the fact that the analytical expressions for the higher-order response functions become hard to evaluate raises the question of whether future

work for active systems should focus on numerical solutions, and not analytical solutions.
On the topic of analytical solutions, in hindsight the calculations we did in section 4 could have been simplified and done in smarter ways by for example using the substitution $y = x - l$ to obtain the new stochastic variable y which would have simplified calculations significantly. Furthermore, the off-diagonal terms in the correlation matrix (186) could have been realized to be equal to zero by inspection of equation (168) which contains an odd product of Gaussian random variables and are thus, by Isserlis' theorem [55], equal to zero.

The noticeable difference in how much the dissipated work fluctuates in Figures 2 and 3 when driving the two different protocols l and κ, we attribute to the fact that the scale of the dissipated work is different between the two protocols, partly due to the fact that the protocol duration is different between the two cases. The l-protocol does, however, also dissipate a factor of about $10 - 15$ times more work than the κ-protocol, for equal protocol durations of length unity (compare 2 with 9 in Appendix A). The difference in fluctuations is then obscured by this scale and there is thus no difference in the stochastic behavior between the two systems.

There is a difference in how well the optimal protocols for κ and l perform, most notably seen by comparing Figures 2 and 4. The numerical dissipated work for the κ-protocol agrees well with the analytical solution already around $t_p \sim 5$, whereas the same is true for the l-protocol beyond $t_p \sim 10$. This indicates that the robustness depends on the type of protocol and that the slow driving assumption may be relaxed for some protocols.

The naive protocol we considered in the passive case in 4 showed that the work dissipated using the optimal protocol is smaller than what it would be for a corresponding system with a protocol driven naively. The reason that the difference in the dissipated work between the two systems driven by the optimal and naive protocols is rather small, we attribute to the fact that the functional form of the optimal protocol might be close to the naive (linear) protocol. For example, the reason we did not showcase the naive protocol for l (changing the position of the harmonic potential), is because the naive protocol in this case is also the optimal protocol. It would be interesting for future work to implement a naive protocol for the active system in order to be able to obtain a benchmark for how well the optimal protocol performs in this case.

The fact that the dissipated work in the active system decays as $1/t_p$ similarly to the passive system may at first glance seem counter-intuitive, but the actual active force from 229 never enters into the work expression in 231. The work dissipated by the system only comes from the external work performed by the protocol, and the difference between the two results (active vs. passive) in figure 2 and 7 (in which the dissipated work is approximately $2 - 3$ times larger in the active case), we attribute to the fact that an active system is "harder" to drive, as it also contains internal dissipation, i.e. the particle constantly changing direction due to the activity. The parameters of the two simulations are the same. The difference between the passive and the active case is essentially that the friction coefficients (equations (76) and (243)) are different between the two systems (as well as the boundary conditions), each yielding different correlation functions because of different dynamics.

Simulating the systems (both passive and active), is a computationally heavy task and the corresponding code takes a considerable time to run if run sequentially and for large ensembles. If given more time, we would like to have run the simulations for even larger ensembles in order to obtain larger statistics and thus smoother curves for the results. We appreciate that further work would benefit from optimizing the computations by implementing to a low-level language or to run the code in parallel, making it feasible to investigate larger systems. In this thesis, we chose to calculate the exact expressions for the work using tra-

jectories of the Langevin equation and then calculating the work using the definitions in equations (8) - (12). We could have obtained the exact expression using equation (246) instead, as done by [6] (for the heat instead of the work). This might be more computionally efficient. Furthermore, obtaining the optimal protocol duration using equation (271) when simulating the active system entails solving the integral numerically. We used the function `quad` from the `scipy.integrate`-library, but other more efficient numerical integration techniques might be a better choice.

For the passive case, methods exist to obtain fast and strong driving (see [44] for an overview of the regimes of thermodynamic control), but it is not immediately clear how these would work for active systems. Since active systems generally exhibit richer behavior than their passive counterparts, such as emergent, collective effects and non-equilibrium phase transitions (for example motility-induced phase separation [23]), the assumption of weak driving in the active framework may be a hindrance to describe dynamics near phase-transitions as it is possible that the protocol is no longer smooth when crossing phase boundaries [6]. Furthermore, phase transitions like motility-induced phase separation happen for large collections of particles and many active matter systems are modelled hydrodynamically, begging the question of whether an optimal control framework could be extended and generalized to a coarse-grained field theory which would be an interesting line for future work. Indeed, it is an open research question how to connect thermodynamics and active field theory, as there is no obvious way to define how coarse-grained active systems exchange energy with their environment [59]. One approach is presented in [59] by embedding active field theory within linear irreversible thermodynamics as developed by Onsager and Prigogine since here there is a natural definition of entropy production and dissipation at the hydrodynamic level [59]. For future work, this may be one starting point for the development of an optimal control framework for active systems described by active field theory. Additionally, as the active Ornstein-Uhlenbeck particle model is a highly simplified model of active matter, packing all the complexities of self-propulsion into a single term v, it would be interesting for future work to use our framework for other active systems displaying more complex behavior, such as active Brownian particles or run-and-tumble particles.

For the active framework, in the spirit of previous works [5, 45, 13, 46, 57], we have chosen to optimize the work. We could have chosen to optimize for example the heat, as done by [6]. Here the functional form of the dissipation is different due to extra terms coming from the heat which stems from the activity of the particle.
Furthermore, we have considered overdamped dynamics in this work, but it should be possible to generalize the framework to include inertial effects as the dynamics in 229 would be changed according to (1). Additionally, we used one protocol $\alpha(t)$ when deriving the active framework, but it should also be possible to generalize it to multiple different protocols as we did in the passive case.

The passive framework has so far been used to optimize the dissipation in two model systems, theoretically in the molecular motor F1-ATPase [60] and experimentally in DNA-hairpin pulling experiments [61]. In both cases, it has been found that the designed protocols obtained from the optimal protocol framework were much more efficient than naive (undesigned) ones. To our knowledge, no experimental designs have been conducted yet of the optimal protocols for an active system. Though as suggested by [57], it should be possible by modifying a DNA-hairpin pulling experiment by letting the hairpin beads experience noise from electrodes coupled to a resistor and an amplifier.

10 Conclusion

To conclude, we have developed a robust, predictive framework for obtaining the optimal protocols for the work for active systems. We have applied this to a model system, an active Ornstein-Uhlenbeck particle in one dimension, by first calculating the friction coefficient, the response function and the associated correlation functions and then comparing the work following a protocol obtained from the analytical calculations with the work obtained from numerical simulations, following the same protocol. The two solutions display a surprising level of agreement over a large interval of protocol durations despite the assumptions of slow and weak driving of the protocols. Furthermore, we have also built an optimal control framework in the case of a passive Brownian particle, which was originally presented in Sivak and Crooks [5]. This then served as the base and inspiration for creating an optimal control framework for active systems. In the passive case, we calculated the relevant correlation functions, associated friction coefficients and from these the optimal protocols. We then compared the analytically obtained expressions for the dissipated work with numerical simulations, obtaining results in agreement with the literature [5, 46, 10]. We have thus bridged the frameworks presented in [5] and [6] and extended the framework presented in [6] to be able to describe the work alongside the heat, suggesting the possibility for a unifying theoretical framework for the optimal control of passive and active systems. Lastly, we have also sketched perspectives and directions for future research, suggesting further investigations into the possiblities of developing an optimal control framework for active matter systems described by active field theory.

A Additional plots

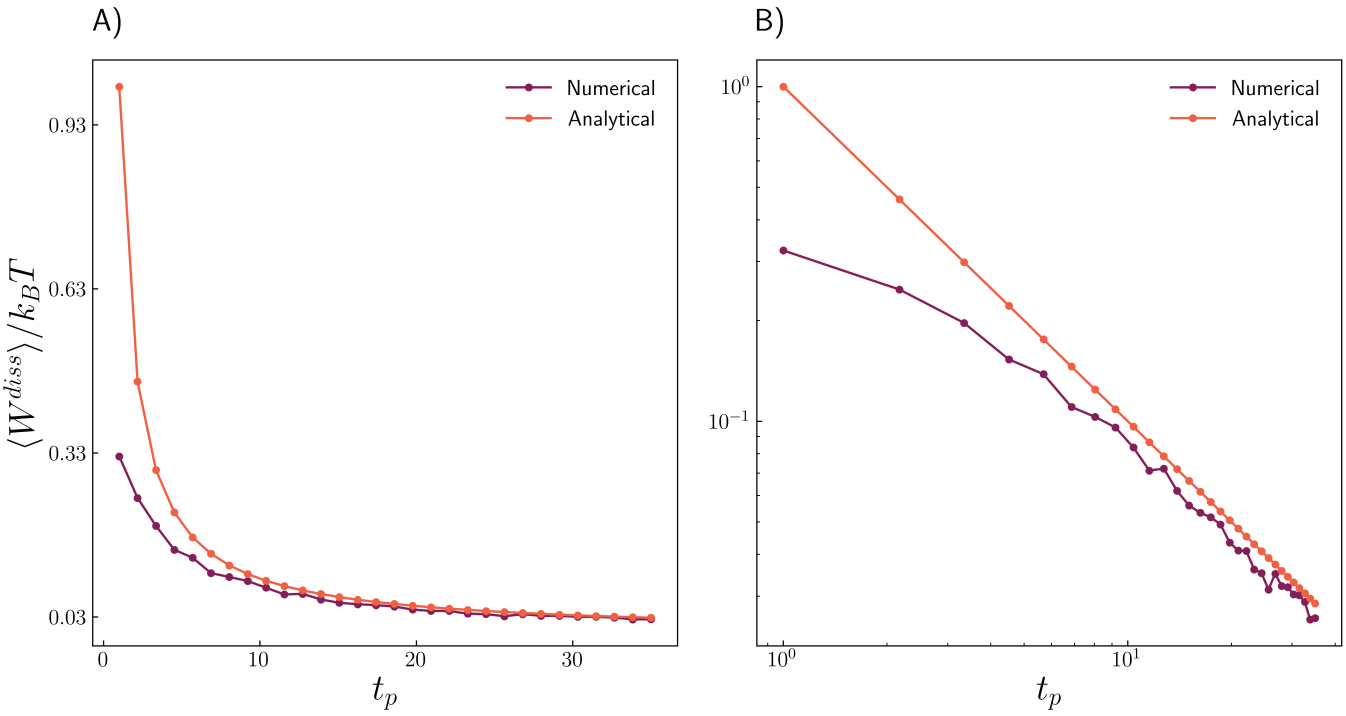

Figure 9: Passive Brownian particle in a one dimensional harmonic potential driven by the optimal l-protocol from $l_0 = 0$ to $l_f = 1$. A) Comparison of the numerical and analytical solution of the average dissipated work B) Log-scale of A).

References

[1] Arijit Ghosh, Weinan Xu, Neha Gupta, and David H Gracias. Active matter therapeutics. *Nano Today*, 31:100836, 2020.

[2] Dongdong Jin and Li Zhang. Collective behaviors of magnetic active matter: Recent progress toward reconfigurable, adaptive, and multifunctional swarming micro/nanorobots. *Accounts of Chemical Research*, 55(1):98–109, 2021.

[3] Yulei Fu, Hengao Yu, Xinli Zhang, Paolo Malgaretti, Vimal Kishore, and Wendong Wang. Microscopic swarms: From active matter physics to biomedical and environmental applications. *Micromachines*, 13(2):295, 2022.

[4] Berk Yigit, Yunus Alapan, and Metin Sitti. Programmable collective behavior in dynamically self-assembled mobile microrobotic swarms. *Advanced Science*, 6(6):1801837, 2019.

[5] David A Sivak and Gavin E Crooks. Thermodynamic metrics and optimal paths. *Physical review letters*, 108(19):190602, 2012.

[6] Luke K Davis, Karel Proesmans, and Étienne Fodor. Active matter under control: Insights from response theory. *arXiv preprint arXiv:2305.11078*, 2023.

[7] Erwin Schrodinger. *What is life?: With mind and matter and autobiographical sketches*. Cambridge University Press, 2012.

[8] Christian Van den Broeck and Massimiliano Esposito. Ensemble and trajectory thermodynamics: A brierange of validf introduction. *Physica A: Statistical Mechanics and its Applications*, 418:6–16, 2015.

[9] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on progress in physics*, 75(12):126001, 2012.

[10] Luca Peliti and Simone Pigolotti. *Stochastic Thermodynamics: An Introduction*. Princeton University Press, 2021.

[11] Ryohei Yasuda, Hiroyuki Noji, Kazuhiko Kinosita, and Masasuke Yoshida. F1-atpase is a highly efficient molecular motor that rotates with discrete 120 steps. *Cell*, 93(7):1117–1124, 1998.

[12] Shoichi Toyabe, Takahiro Watanabe-Nakayama, Tetsuaki Okamoto, Seishi Kudo, and Eiro Muneyuki. Thermodynamic efficiency and mechanochemical coupling of f1-atpase. *Proceedings of the National Academy of Sciences*, 108(44):17951–17956, 2011.

[13] Grant M Rotskoff and Gavin E Crooks. Optimal control in nonequilibrium systems: Dynamic riemannian geometry of the ising model. *Physical Review E*, 92(6):060102, 2015.

[14] Étienne Fodor, Robert L Jack, and Michael E Cates. Irreversibility and biased ensembles in active matter: Insights from stochastic thermodynamics. *Annual Review of Condensed Matter Physics*, 13:215–238, 2022.

[15] Jérémy O'Byrne, Yariv Kafri, Julien Tailleur, and Frédéric van Wijland. Time irreversibility in active matter, from micro to macro. *Nature Reviews Physics*, 4(3):167–183, 2022.

[16] Gerhard Gompper, Roland G Winkler, Thomas Speck, Alexandre Solon, Cesare Nardini, Fernando Peruani, Hartmut Löwen, Ramin Golestanian, U Benjamin Kaupp, Luis Alvarez, et al. The 2020 motile active matter roadmap. *Journal of Physics: Condensed Matter*, 32(19):193001, 2020.

[17] Étienne Fodor, Cesare Nardini, Michael E Cates, Julien Tailleur, Paolo Visco, and Frédéric Van Wijland. How far from equilibrium is active matter? *Physical review letters*, 117(3):038103, 2016.

[18] Étienne Fodor and M Cristina Marchetti. The statistical physics of active matter: From self-catalytic colloids to living cells. *Physica A: Statistical Mechanics and its Applications*, 504:106–120, 2018.

[19] M Cristina Marchetti, Jean-François Joanny, Sriram Ramaswamy, Tanniemola B Liverpool, Jacques Prost, Madan Rao, and R Aditi Simha. Hydrodynamics of soft active matter. *Reviews of modern physics*, 85(3):1143, 2013.

[20] Andrea Cavagna, Lorenzo Del Castello, Irene Giardina, Tomas Grigera, Asja Jelic, Stefania Melillo, Thierry Mora, Leonardo Parisi, Edmondo Silvestri, Massimiliano Viale, et al. Flocking and turning: a new model for self-organized collective motion. *Journal of Statistical Physics*, 158:601–627, 2015.

[21] John Toner, Yuhai Tu, and Sriram Ramaswamy. Hydrodynamics and phases of flocks. *Annals of Physics*, 318(1):170–244, 2005.

[22] Ricard Alert, Jaume Casademunt, and Jean-François Joanny. Active turbulence. *Annual Review of Condensed Matter Physics*, 13:143–170, 2022.

[23] Michael E Cates and Julien Tailleur. Motility-induced phase separation. *Annu. Rev. Condens. Matter Phys.*, 6(1):219–244, 2015.

[24] Andrea Cavagna and Irene Giardina. Bird flocks as condensed matter. *Annu. Rev. Condens. Matter Phys.*, 5(1):183–207, 2014.

[25] Michael E Cates. Diffusive transport without detailed balance in motile bacteria: does microbiology need statistical physics? *Reports on Progress in Physics*, 75(4):042601, 2012.

[26] Jens Elgeti, Roland G Winkler, and Gerhard Gompper. Physics of microswimmers—single particle motion and collective behavior: a review. *Reports on progress in physics*, 78(5):056601, 2015.

[27] Jeremie Palacci, Stefano Sacanna, Asher Preska Steinberg, David J Pine, and Paul M Chaikin. Living crystals of light-activated colloidal surfers. *Science*, 339(6122):936–940, 2013.

[28] Mathieu Poujade, Erwan Grasland-Mongrain, A Hertzog, J Jouanneau, Philippe Chavrier, Benoît Ladoux, Axel Buguin, and Pascal Silberzan. Collective migration of an epithelial monolayer in response to a model wound. *Proceedings of the National Academy of Sciences*, 104(41):15988–15993, 2007.

[29] Daniel Needleman and Zvonimir Dogic. Active matter at the interface between materials science and cell biology. *Nature reviews materials*, 2(9):1–14, 2017.

[30] Mark J Bowick, Nikta Fakhri, M Cristina Marchetti, and Sriram Ramaswamy. Symmetry, thermodynamics, and topology in active matter. *Physical Review X*, 12(1):010501, 2022.

[31] Luis L Bonilla. Active ornstein-uhlenbeck particles. *Physical Review E*, 100(2):022601, 2019.

[32] GH Philipp Nguyen, René Wittmann, and Hartmut Löwen. Active ornstein–uhlenbeck model for self-propelled particles with inertia. *Journal of Physics: Condensed Matter*, 34(3):035101, 2021.

[33] David Martin, Jérémy O'Byrne, Michael E Cates, Étienne Fodor, Cesare Nardini, Julien Tailleur, and Frédéric Van Wijland. Statistical mechanics of active ornstein-uhlenbeck particles. *Physical Review E*, 103(3):032607, 2021.

[34] Lennart Dabelow and Ralf Eichhorn. Irreversibility in active matter: General framework for active ornstein-uhlenbeck particles. *Frontiers in Physics*, page 516, 2021.

[35] Robert Zwanzig. *Nonequilibrium statistical mechanics*. Oxford university press, 2001.

[36] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

[37] Edward M Purcell. Life at low reynolds number. *American journal of physics*, 45(1):3–11, 1977.

[38] Ken Sekimoto. Langevin equation and thermodynamics. *Progress of Theoretical Physics Supplement*, 130:17–27, 1998.

[39] David Chandler. *Introduction to modern statistical mechanics*, volume 5. 1987.

[40] Eberhard Engel. *Density functional theory*. Springer, 2011.

[41] Michael Stone and Paul Goldbart. *Mathematics for physics: a guided tour for graduate students*. Cambridge University Press, 2009.

[42] Raj Kumar Pathria. *Statistical mechanics*. Elsevier, 2016.

[43] Linda E Reichl. *A modern course in statistical physics*. American Association of Physics Teachers, 1999.

[44] Steven Blaber and David A Sivak. Optimal control in stochastic thermodynamics. *Journal of Physics Communications*, 7(3):033001, 2023.

[45] Sebastian Deffner and Marcus VS Bonança. Thermodynamic control—an old paradigm with new applications. *Europhysics Letters*, 131(2):20001, 2020.

[46] Lucas P Kamizaki, Marcus VS Bonança, and Sérgio R Muniz. Performance of optimal linear-response processes in driven brownian motion far from equilibrium. *Physical Review E*, 106(6):064123, 2022.

[47] Marcus VS Bonança and Sebastian Deffner. Optimal driving of isothermal processes close to equilibrium. *The Journal of chemical physics*, 140(24), 2014.

[48] Kenneth Franklin Riley and Michael Paul Hobson. *Essential mathematical methods for the physical sciences*. Cambridge University Press, 2011.

[49] Crispin W Gardiner et al. *Handbook of stochastic methods*, volume 3. springer Berlin, 1985.

[50] H. Goldstein, J.L. Safko, and C.P. Poole. *Classical Mechanics: Pearson New International Edition*. Pearson Education, 2014.

[51] Francisco J Sevilla, Rosalío F Rodríguez, and Juan Ruben Gomez-Solano. Generalized ornstein-uhlenbeck model for active motion. *Physical Review E*, 100(3):032123, 2019.

[52] Lorenzo Caprini, Andrea Puglisi, and Alessandro Sarracino. Fluctuation–dissipation relations in active matter systems. *Symmetry*, 13(1):81, 2021.

[53] Christian Maes. Response theory: a trajectory-based approach. *Frontiers in Physics*, 8:229, 2020.

[54] Walter Greiner and Joachim Reinhardt. *Field quantization*. Springer Science & Business Media, 1996.

[55] Leon Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1/2):134–139, 1918.

[56] Édgar Roldán, Izaak Neri, Raphael Chetrite, Shamik Gupta, Simone Pigolotti, Frank Jülicher, and Ken Sekimoto. Martingales for physicists. *arXiv preprint arXiv:2210.09983*, 2022.

[57] Deepak Gupta, Sabine HL Klapp, and David A Sivak. Efficient control protocols for an active ornstein-uhlenbeck particle. *arXiv preprint arXiv:2304.12926*, 2023.

[58] Marcus VS Bonança and Sebastian Deffner. Minimal dissipation in processes far from equilibrium. *Physical Review E*, 98(4):042103, 2018.

[59] Tomer Markovich, Étienne Fodor, Elsen Tjhung, and Michael E Cates. Thermodynamics of active field theories: Energetic cost of coupling to reservoirs. *Physical Review X*, 11(2):021057, 2021.

[60] Deepak Gupta, Steven J Large, Shoichi Toyabe, and David A Sivak. Optimal control of the f1-atpase molecular motor. *The Journal of Physical Chemistry Letters*, 13:11844–11849, 2022.

[61] Sara Tafoya, Steven J Large, Shixin Liu, Carlos Bustamante, and David A Sivak. Using a system's equilibrium behavior to reduce its energy dissipation in nonequilibrium processes. *Proceedings of the National Academy of Sciences*, 116(13):5920–5924, 2019.