Internship Report

Efficiency Fluctuations of the Brownian Gyrorat

Lennart Dabelow
4 April – 11 November 2016

NORDITA (Nordic Institute for Theoretical Physics)
Roslagstullsbacken 23 · SE-106 91 Stockholm, Sweden

Advisor: Ralf Eichhorn

We investigate the thermodynamics and efficiency fluctuations of a nano-scale heat engine. Our system consists of a two-dimensional Brownian particle trapped by a parabolic potential and interacting with two heatbaths at different temperatures. We show that various anisotropies between these constituents cause the particle to exert a torque on the environment. This intrinsic torque can be used to perform work against an external force and thus to build a microscopic heat engine. Inspired by recent works on the universality of the efficiency distribution of such small-scale machines in the long-time limit, we derive the efficiency large deviation function for the system under study. This is achieved by a contraction of the scaled joint cumulant generating function of work and heat, which in turn is expressed as a functional determinant using path integral techniques. We present two different methods to evaluate this functional determinant, both of which require some sort of approximation, but are shown to be consistent. We then extract the corresponding efficiency large deviation function and also contrast it with independent computer simulations. Finally, we compare our result to the aforementioned universal theory proposed earlier, identify similarities and explain apparent deviations.
## CONTENTS

1 Introduction 3

2 Nanoscale Heat Engine 6
   2.1 Time Evolution and Steady State ................................. 7
   2.2 Average Work and Heat Transfer .................................... 8
   2.3 Nanoscale Heat Engine and Pump .................................... 11
   2.4 Entropy Production .................................................. 14

3 Efficiency Fluctuations 15
   3.1 Path Integral Formalism .............................................. 16
   3.2 Detailed Fluctuation Theorem ....................................... 17
   3.3 Large Deviation Principle ........................................... 18
   3.4 Cumulant Generating Functions of Stochastic Functionals: Closed Trajectory Approximation .............................. 21
   3.5 Large Deviation Function for Efficiency Fluctuations .............. 30
   3.6 Cumulant Generating Functions of Stochastic Functionals: Green’s Matrix Method ............................................ 34
   3.7 Comparison and Discussion ........................................... 39

4 Conclusion 46

Acknowledgements 49

A Solution of the General Ornstein-Uhlenbeck SDE 50

B Derivation of the Path Weight 52

C Construction of Green’s matrices for systems of second order ODEs 54

D Simulation Parameters 60

References 61
1 INTRODUCTION

Heat engines are an indispensable element of our everyday lives. The vast majority of energy we consume is produced, and has been so since the Industrial Revolution, by such machines. We find them in our power plants, our cars, our lawnmowers, and—in a reversed version—in our fridge at home as well. These devices are usually described in terms of classical thermodynamics [1], and their basic principle is the conversion of heat to work (or vice versa for refrigerators). Despite the tremendous size and complexity of these machines, the theory of classical thermodynamics requires only a small number of variables such as temperature, entropy, energy, pressure, ... to make surprisingly accurate predictions about their behavior. More precisely, all this is because of their tremendous size and complexity and owing to the law of large numbers. All fluctuations of the microscopic constituents effectively cancel each other, and the system is completely determined by average values.

In this report, we will study the properties of a particular heat engine known as the Brownian gyrator [2]. Contrary to the above mentioned everyday examples, the typical dimensions of this system are on the order of tens to hundreds of nanometers. At these scales, fluctuations are crucial and characteristic, and the behavior of the system may deviate significantly from its average during the course of observation. The simplest, yet standard example for the type of systems we will be concerned with is a pollen grain in water. In fact, this was precisely the object under study for Scottish botanist Robert Brown, who observed in 1827 that his pollens traveled around in the liquid in a seemingly random fashion [3]. It took almost 80 years until Albert Einstein was able to give a satisfying theoretical explanation for this so-called Brownian motion [4]. Nowadays we understand that the water molecules are constantly jiggling around their equilibrium positions due to their thermal energy. This way, they consistently bump into the pollen and thus push it around.

Although nanometer scales are small compared to our everyday experience, they are still far away from the regime where the microscopic laws of quantum physics are analytically and numerically tractable. A cube of length 10 nm still contains about $10^5$ atoms, and describing all their dynamics and interactions individually is obviously hopeless. Rather, a formalism is needed that is favorable to this mesoscopic scale between the quantum and classical worlds. At this stage, stochastic thermodynamics [5, 6] comes into play. Speaking in terms of the above example, the key idea is to restrict the dynamical description to the pollen grain, and to model the interactions with the water molecules by fluctuating forces in terms of random processes. This way, the corresponding equations of motion become stochastic as well, and the analysis requires a stochastic generalization of calculus [7–9]. Fluctuating quantities are then described mathematically by random variables and their probability distributions.

The aim of this work is to study the long-time efficiency fluctuations of the Brownian gyrator, which we are going to introduce in detail below. The efficiency of a heat engine operating between two thermostats is the ratio of its power output and heat input from the hot reservoir, thus quantifying how much of the energy extracted from a heat source is converted into mechanical work. The second law of thermodynamics provides an upper bound for the achievable efficiency of classical engines, which is called the Carnot efficiency and is always less than 1. However, single realizations of mesoscopic engines may reach arbitrary efficiencies with finite probability. Our interest in these
efficiency fluctuations mostly stems from recent discoveries of universality properties satisfied by their large-time distribution [10, 11]. Strictly speaking, these results do not apply to the system under study, but our results bear some resemblance nonetheless, and we would like investigate similarities and differences.

The report is organized as follows. In the remainder of this section, we will introduce our model and point out some possible experimental realizations. Section 2 will then be concerned with its solution and the macroscopic thermodynamic properties. In particular, we will classify the different operational regimes where the system works as a heat engine, heat pump, or dud engine, respectively. We will also analyze the macroscopic efficiency and entropy production. The latter quantity is known to satisfy certain fluctuation theorems that connect entropy and time-irreversibility, as we will explore in the beginning of Section 3. Thereafter, we will evaluate the scaled cumulant generating function of work and heat along trajectories, borrowing tools from statistical field theory. Using the theory of large deviations [12, 13], this will allow us to investigate fluctuations of the efficiency in the large-time limit and compare with recently found universality properties [14–16]. Finally, we will summarize our results in Section 4.

Model. The system of interest consists of a two-dimensional Brownian particle with position $x = (x_1, x_2)$ and velocity $v = (v_1, v_2)$. A sketch of the model is shown in Fig. 1. The particle is confined by a parabolic potential $U(x)$ with principal axes $u_1$ and $u_2$, which is generally tilted by an angle $\alpha$ with respect to the coordinate axes:

$$U(x) = \frac{1}{2} x^\top R_\alpha^\top u R_\alpha x \quad \text{with} \quad R_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{and} \quad u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \quad (1.1)$$

Such a potential may be created by an optical trap [17], for example, but could also represent an arbitrary potential near a local minimum. It exerts a force $F^U = -\nabla U$ on the particle.
Furthermore, the particle is immersed in a fluid environment and interacts with the molecules in the medium. This interaction is twofold. On the one hand, there is a general trend to dissipate kinetic energy via friction, described by a frictional force \( F_{\text{diss}}^i = -\gamma_i \dot{x}_i \) with friction coefficients \( \gamma_1 \) and \( \gamma_2 \) along the two coordinate axes. This dissipated heat is damped into two reservoirs at temperatures \( T_1 \) and \( T_2 \), respectively. The concept of an oriented temperature may appear odd at first sight, but we will provide physical examples for the system under study below.

On the other hand, the thermally excited molecules in the medium jiggle around their equilibrium positions and thereby bump into the particle, leading to pushes in seemingly random directions. This effect can be modeled by a force \( F_{\text{fluc}}^i = \sqrt{2k_B T_i} \gamma_i \xi_i \), where \( \xi_i(t) \) denotes a standard Gaussian white noise with vanishing mean and correlation \( \langle \xi_i(t)\xi_j(t') \rangle = \delta_{ij} \delta(t-t') \). The amplitude of this fluctuating force depends on the temperatures because the higher the thermal energy, the stronger we expect an average push to be. Furthermore, the amplitude also depends on the friction coefficients \( \gamma_i \), emphasizing the intimate relation between fluctuation and dissipation. Its precise form is in fact prescribed by the fluctuation-dissipation theorem [4, 18–20], which rigorously establishes the connection between the two phenomena.

The so far suggested model was first studied by Filliger and Reimann [2], who found that due to the asymmetry of the thermal and restoring forces, the particle rotates around the origin on average, thereby exterts a torque on the environment and thus can work as a microscopic heat engine. In this report, we are going to study the properties of this nanoscale device. To quantify the work done, we therefore introduce an additional external force

\[
F_{\text{ext}}^i(x) = -f_{\text{ext}}^i \varepsilon x, \quad \text{with } \varepsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},
\]

being the two-dimensional antisymmetric tensor. The parameter \( f_{\text{ext}}^i \) is an arbitrary constant determining the strength of the force. This force is obviously not conservative and can thus not simply be absorbed in the potential \( U \). Moreover, it leads to an external torque \( M_{\text{ext}} = |x \times F_{\text{ext}}^i(x)| = f_{\text{ext}}^i(x_1^2 + x_2^2) \).

Collecting all forces acting, we arrive at the equations of motion

\[
\begin{align*}
\dot{x}_i(t) &= v_i(t), \\
m \dot{v}_i(t) &= -\gamma_i v_i(t) - \frac{\partial U(x(t))}{\partial x_i} + F_{\text{ext}}^i(x(t)) + \sqrt{2k_B T_i} \gamma_i \xi_i(t).
\end{align*}
\]

(1.3a)

(1.3b)

We remark that the Gaussian white noise \( \xi(t) \) should be understood more thoroughly as the derivative of a standard Wiener process \( \omega(t) \), so that formally \( d\omega(t) = \xi(t) dt \). Note that as opposed to \( \xi(t) \), the differential \( d\omega(t) \) has a rigorous mathematical interpretation, but the former notation is more common in physics. We will switch at will between these two notations for the white noise process.

As mentioned above, our model is based on [2]. In particular, this reference suggests some possible experimental realization such as a particle in a usual thermal bath with an additionally heat source irradiating only along one of the coordinate axes, or an electrically charged particle between two plate capacitors coupled to resistors at different temperatures. However, the equations of motion (1.3) govern a much larger
class of physical systems [21–23]. For example, one can also think of two 1D particles coupled by a spring, which live in different reservoirs, or a purely electrical circuit, where the currents through two resistors become the dynamical variables.

**Overdamped approximation.** We will restrict our discussion in the following to the so-called overdamped approximation [24–26], which basically assumes that frictional forces are much stronger than inertial effects, so that any gain in kinetic energy is immediately dissipated. In this limit, the velocities are taken to be equilibrated with the thermal baths at all times, and the dynamical variables are reduced to the position vector. Equations (1.3) thus become

\[
\gamma_i \dot{x}_i(t) = -\frac{\partial U(x)}{\partial x_i} + F_i^{\text{ext}}(x) + \sqrt{2k_B T_i \gamma_i} \xi_i(t). \tag{1.4}
\]

Despite its simplicity, this assumption is generally valid for a broad class of systems and yields remarkably precise results. As an illustration of the involved scales, consider a pollen grain (radius $2 \mu$m, density $1 \text{ g/cm}^3$) in water (viscosity $8.9 \times 10^{-4} \text{ Pa s}$). Assuming Stokes friction, we obtain $\gamma = 33.6 \text{ mg/ms}$. The particle’s velocity will thus reduce by $1/e$ within $\frac{2\pi}{\gamma} = 0.5 \mu$s. Starting with an initial velocity of $1 \text{ mm/s}$, the particle would travel a distance of just $2 \text{ nm}$ if there were no fluctuating forces, which is a $1000$th of its radius. Thus, the assumption of dominating friction is more than justified.

For notational convenience, we are going to rewrite (1.4) slightly. Note first that both the potential and the external force are linear in the position, so that we can write the total force as $F(x) = -Kx$ with $K = R^T uR + f^{\text{ext}} \epsilon$ being a $(2 \times 2)$-matrix. Furthermore, we divide the equations by the friction coefficients. This way we obtain

\[
\dot{x}(t) = -Ax(t) + B\xi(t), \tag{1.5}
\]

where the constant matrices $A$ and $B$ are given by

\[
A = \begin{pmatrix} K_{11}/\gamma_1 & K_{12}/\gamma_1 \\ K_{21}/\gamma_2 & K_{22}/\gamma_2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \sqrt{2k_B T_1/\gamma_1} & 0 \\ 0 & \sqrt{2k_B T_2/\gamma_2} \end{pmatrix}, \tag{1.6}
\]

respectively. Equation (1.5) is the stochastic differential equation (SDE) of a general Ornstein-Uhlenbeck process.

### 2 NANOSCALE HEAT ENGINE

Having defined the system of interest, we aim at investigating its thermodynamic properties. Before studying the exchange of work and heat between the particle and the reservoirs, we will first characterize the system’s dynamics and establish its nonequilibrium steady-state (NESS) distribution. Thereafter, we will survey its macroscopic thermodynamics. The study of fluctuations will be postponed to Section 3.
2.1 Time Evolution and Steady State

**Propagator.** The SDE (1.5) is equivalent to a Fokker-Planck equation [9] for the transition probabilities or propagator \( p(t, x | t_0, x_0) \),

\[
\partial_t p(t, x | t_0, x_0) = \sum_{i,j} \left\{ A_{ij} \partial_{x_i} [x_j p(t, x | t_0, x_0)] + D_{ij} \partial_{x_i} \partial_{x_j} p(t, x | t_0, x_0) \right\},
\]

with the diffusion matrix \( D = \frac{1}{2} BB^\top \). The propagator gives the probability to find the particle in an infinitesimal volume element \( d^2x \) around \( x \) at time \( t \) given that it was at \( x_0 \) at an earlier time \( t_0 \). Since the system is Markovian, its statistics are fully determined by \( p \) and some initial distribution \( p_0(x_0) \).

The Fokker-Planck equation (2.1) can be solved exactly [7, 9]. We provide a detailed derivation in Appendix A. The propagator \( p(t, x | t_0, x_0) \) is given by

\[
p(t, x | 0, x_0) = \exp \left\{ -\frac{1}{2} [x - e^{-tA} x_0]^\top \Sigma^{-1}(t) [x - e^{-tA} x_0] \right\} \sum (2\pi)^2 \det \Sigma(t),
\]

where the covariance matrix is

\[
\Sigma(t) = \Sigma(\infty) - e^{-tA} \Sigma(\infty) e^{-tA^\top}
\]

and \( \Sigma(\infty) \) is obtained as the solution of

\[
A \Sigma(\infty) + \Sigma(\infty) A^\top = 2D.
\]

We immediately see that \( x(t) \) is a time-homogeneous Gaussian process. For our system with parameters (1.6), we find

\[
\Sigma(\infty) = \frac{1}{\text{tr} A \det A} \begin{pmatrix}
D_2 A_{12}^2 + D_1 (A_{22}^2 + \det A) & -D_1 A_{21} A_{22} - D_2 A_{11} A_{12} \\
-D_1 A_{21} A_{22} - D_2 A_{11} A_{12} & D_1 A_{21}^2 + D_2 (A_{11}^2 + \det A)
\end{pmatrix},
\]

where \( D_i = k_B T_i / \gamma_i \) are the entries of the diagonal matrix \( D \).

**Steady state.** From the solution (2.2), we understand that the system reaches a steady state in the limit \( t \to \infty \) if the matrix \( A \) is positive definite. In this case, the limit distribution is

\[
p_\infty(x) = \frac{e^{-\frac{1}{2} x^\top \Sigma^{-1}(\infty) x}}{(2\pi)^2 \det \Sigma(\infty)}.
\]

As can be readily inferred from inspecting its eigenvalues, the matrix \( A \) is indeed positive definite here as long as \( u_1, u_2 > 0 \). The characteristic relaxation time to the steady state is \( \tau_\infty = (\det A)^{-1/2} = \sqrt{7_1 \gamma_2 / [u_1 u_2 + (f_{\text{ext}})^2]} \), which is typically on the order of 0.1 ms...10 ms. In the rest of this report, we will assume that the system is in the steady state unless otherwise stated explicitly.
**Figure 2:** (a) Examples of the angular component of several trajectories; (b) mean change of the angle with time for 1000 realizations. Parameters are chosen according to set B (cf. Appendix D).

**Intrinsic torque and rotational motion.** As we mentioned in the introduction, it has been observed before in [2] that this system exhibits an intrinsic torque even in the case that $F_{\text{ext}} = 0$. By averaging with the steady-state distribution (2.6) we can reproduce the result stated there straightforwardly to find

$$
\langle M^{\text{int}} \rangle = \langle |x(t) \times F_U(t)| \rangle = -\frac{k_B(T_1 - T_2)(u_1 - u_2)\sin(2\alpha)}{u_1 + u_2 - \frac{\gamma_1 - \gamma_2}{\gamma_1 + \gamma_2}(u_1 - u_2)\cos(2\alpha)}.
$$

(2.7)

This result nicely illustrates the interplay of the various symmetry breaking mechanisms: In order for an intrinsic torque to exist, we need two reservoirs at different temperatures ($T_1 \neq T_2$) and a potential with asymmetric principal axes ($u_1 \neq u_2$) that is tilted with respect to the radiating directions of the heatbaths ($0 < \alpha < \pi/2$). More generally, this torque is a manifestation of the nonequilibrium character of the steady state. Despite the stationarity of the distribution $p_{\infty}$, there is a non-vanishing probability current [27]

$$
j_{\infty}(x) = -(Ax + D\nabla)p_{\infty}(x).
$$

(2.8)

The rotational motion can also be assessed by introducing polar coordinates $(r, \phi)$ such that $x_1 = r \cos \phi$ and $x_2 = r \sin \phi$. Considering the angular component, a trend towards a certain direction of motion is already visible at the level of individual trajectories, as shown in Fig. 2a. It is clearly recognizable when considering an average over several trajectories, depicted in Fig. 2b.

**2.2 Average Work and Heat Transfer**

The fact that the system is out of equilibrium and the torque resulting from this can be used to extract energy from the system by letting it work against the external torque exerted by the driving force $F_{\text{ext}}$. In a first step, we will analyze the average power supplied to or extracted from the system by the various constituents. This will allow us to characterize the operational regimes of the system as a heat engine, heat pump, or dud engine.
Stochastic energetics and the first law. The equations of motion (1.4) specify several forces acting on the particle, each of which can be associated with a physical component of the system. In particular, we have the mechanical forces \( F^U = -\nabla U \) and \( F^{\text{ext}} \), as well as forces linked to the interaction with the medium and reservoirs, namely the frictional force \( F^{\text{diss}} = -\gamma \dot{x} \) and the thermal collisions \( F^{\text{fluc}} = \sqrt{2k_B T \gamma} \xi \).

The equations of motion merely state the balance of these forces. Quantifying the work performed by the mechanical forces is straightforward. Given a trajectory \( x = [x(t)]_{t=0}^\tau \), we simply integrate the corresponding force \( F(x(t)) \) along the path followed by the particle:

\[
W = \int_0^\tau F(x(t)) \circ d\mathbf{x}(t) \quad (2.9)
\]

Here ‘\( \circ \)’ denotes the Stratonovich product \([7, 26]\). With the above definition of the forces, \( W \) becomes the work performed on the particle, meaning that it is positive when energy flows into the system. Due to the conservative nature of the potential force, we find \( W^U = U(x(0)) - U(x(\tau)) = -\Delta U \) for the energy gained from a possible potential difference between beginning and end points.

It is presumably less obvious how the heat flow from the reservoirs to the system can be assessed. Following Sekimoto’s stochastic energetics approach \([26, 28]\), we identify this heat with the work performed by the dissipating and thermally fluctuating forces:

\[
Q_i = \int_0^\tau -\gamma_i \dot{x}_i(t) + \sqrt{2k_B T_i \gamma_i} \xi_i(t) \circ d\mathbf{x}_i(t) \quad (2.10)
\]

The basic idea is that the energy the particle loses in each coordinate direction due to friction is dissipated as heat into the corresponding reservoir, thus this contribution is counted with a negative sign. On the other hand, energy is exchanged during the random collisions of the particle with the molecules in the medium. The particle’s energy increases if the push occurs along the direction of motion. We imagine that this energy is taken from the heat baths, thus “cooling down” the molecules in the medium, though this does not affect the reservoirs’ temperatures due to their assumed infinity. Similarly, if the push occurs against the particle’s direction of motion, energy is damped into the reservoirs. Fundamentally, then, both dissipation and thermal fluctuations are linked to interactions between the particle and the molecules in the medium, a relation that is considered by the fluctuation-dissipation theorem \([19, 20]\).

The interpretation of (2.10) as the exchanged heat can be further justified if we use the equations of motion (1.4) and rewrite (2.10) in differential form as

\[
dQ_i = \sum_j K_{ij} x_j \circ dx_i = \left[ \frac{\partial U}{\partial x_i} - F_i^{\text{ext}}(x) \right] \circ dx_i \quad (2.11)
\]

This relation is in fact just a mesoscopic version of the first law of thermodynamics,

\[
dU = dQ_1 + dQ_2 + dW \quad (2.12)
\]

The internal energy here is the potential energy of the particle, remembering that any change in kinetic energy is immediately dissipated in the overdamped limit. The above definition of heat thus guarantees conservation of energy at the level of fluctuations.
Note that we measure exchanged work and heat flowing \textit{towards} the system as illustrated in Fig. 3a. This is to say, work and heat are positive if they increase the (instantaneous) kinetic energy of the particle.

\textbf{Work along trajectories.} For a generic force $\mathbf{F}(\mathbf{x})$ acting on our particle, the work performed along an infinitesimal line element $d\mathbf{x}_i$ in the $i$-direction is given by

$$dW_i = F_i(\mathbf{x}) \circ d\mathbf{x}_i = F_i(\mathbf{x}) d\mathbf{x}_i + \sum_j \frac{\partial F_i}{\partial x_j} D_{ij} dt.$$ \hfill (2.13)

In transforming the Stratonovich product, we made use of the Itô rules [7, 29] and the equations of motion (1.4). To obtain the total work supplied to the particle, we have to integrate along the stochastic trajectory $\mathbf{x}(t)$ to obtain

$$W_i(\tau) = \int_0^\tau F_i(\mathbf{x}(t)) d\mathbf{x}_i(t)$$

$$= \int_0^\tau \sum_j \left[ -F_i(\mathbf{x}(t)) A_{ij} x_j(t) + \frac{\partial F_i}{\partial x_j} D_{ij} \right] dt + \int_0^\tau \sum_j F_i(\mathbf{x}(t)) B_{ij} d\omega_j(t).$$ \hfill (2.14)

Taking the ensemble average leads to

$$\langle W_i(\tau) \rangle = \int_0^\tau \sum_j \left\langle -F_i(\mathbf{x}(t)) A_{ij} x_j(t) + \frac{\partial F_i}{\partial x_j} D_{ij} \right\rangle dt.$$ \hfill (2.15)

Seeing as work and heat are time-extensive quantities, we will also consider the time-averaged work (power) $w_i(\tau) = W_i(\tau)/\tau$.

\textbf{External work.} Let us now turn to our particular model and consider the contributions from the external force $\mathbf{F}^\text{ext}(\mathbf{x}) = -f^\text{ext} \varepsilon \mathbf{x}$ (cf. Equation 1.2). First note that the Stratonovich drift $D_{ij} \frac{\partial F_i^\text{ext}}{\partial x_j}$ vanishes because $D_{ij}$ is symmetric whereas $\varepsilon_{ij}$ is antisymmetric under the exchange of indices. Moreover, remember that we are working in the steady state, so that the integrand in (2.15) is independent of time. Therefore, the time-averaged externally supplied work in the $i$-th direction is obtained as

$$\langle w_i^\text{ext} \rangle = f^\text{ext} \sum_{j,k} A_{ij} \varepsilon_{ik} \langle x_j x_k \rangle.$$ \hfill (2.16)

Plugging in the covariances obtained from (2.5), we find

$$\langle w_{1,2}^\text{ext} \rangle = \frac{2k_B(T_1 + T_2)}{(u_1 + u_2)(\gamma_1 + \gamma_2)} \left[ f^\text{ext} \mp \frac{(u_1 - u_2)}{2} \sin(2\alpha) \right] \left[ f^\text{ext} + \frac{T_1 - T_2}{u_1 - u_2} \frac{u_1 - u_2}{u_1 + u_2} \sin(2\alpha) \right]$$

$$- C_S \begin{pmatrix} \frac{u_1 - u_2}{u_1 + u_2} \gamma_1 - \gamma_2 \cos(2\alpha) \end{pmatrix}$$ \hfill (2.17)

with the upper sign corresponding to the first and the lower sign corresponding to the second index. The prefactor we marked with $C_S$ is merely a positive scaling factor of dimension $(\text{length})^2(\text{mass})^{-1}(\text{time})$ that does not affect the thermodynamic properties of the system at all. We see that the rates depend quadratically on the strength $f^\text{ext}$.
of the external force. The only difference between the two directions \( i = 1 \) and \( i = 2 \) is the sign of the first root (‘−’ corresponding to \( i = 1 \), ‘+’ to \( i = 2 \)). Splitting up the work into contributions along coordinate axes is rather artificial in the case of the external force. The total externally supplied work is therefore of much greater interest:

\[
\langle w^{\text{ext}} \rangle = \langle w_1^{\text{ext}} \rangle + \langle w_2^{\text{ext}} \rangle = C_S \frac{f^{\text{ext}}}{} \left[ f^{\text{ext}} \frac{T_1 - T_2}{T_1 + T_2} \frac{u_1 - u_2}{2} \sin(2\alpha) \right] .
\]

(2.18)

**Potential energy.** The average work supplied by the potential can be computed in a similar way. Starting from (2.15) with \( \mathbf{F} U(x) = -\nabla U(x) = -\mathbf{R} \mathbf{u} \mathbf{R} x \) as defined in (1.1), we find

\[
\langle w_i^{\text{pot}} \rangle = \sum_{j,k} A_{ij} (\mathbf{R} \mathbf{u} \mathbf{R})_{ik} \langle x_j x_k \rangle - D_i (\mathbf{R} \mathbf{u} \mathbf{R})_{ii} = 0 .
\]

(2.19)

This is not so surprising after all, given that the potential leads to a conservative force and thus the work exerted on the particle depends only on the initial and final position of the particle. In the steady state, however, the particle’s expected positions in the beginning and end of the trajectory are the same, so that the average work from conservative forces vanishes.

Regarding the long-time behavior of the system and its asymptotic thermodynamics, the fluctuations of the potential energy thus play only a minor role. We will therefore suppress the superscript ‘ext’ when speaking of the work supplied by the external force, \( i.e. \) we identify \( w \equiv w^{\text{ext}} \), \( W \equiv W^{\text{ext}} \), etc.

**Exchanged heat.** Similar to the mechanical work, we can evaluate the heat flowing out of the reservoirs as defined by (2.10), which is equivalent to (2.11). Remembering that there was no average potential energy change in either direction, we immediately find that \( \langle q_i \rangle = -\langle w_i \rangle \) with \( q_i = Q_i / \tau \) denoting the time-intensive heat exchange rate. We repeat the result for completeness:

\[
\langle q_{1,2} \rangle = -C_S \frac{f^{\text{ext}} \frac{u_1 - u_2}{2} \sin(2\alpha)}{1 - \frac{u_1 - u_2}{u_1 + u_2} \frac{\eta_1 - \eta_2}{\eta_1 + \eta_2} \cos(2\alpha)} .
\]

(2.20)

Also note the obvious balance relation

\[
\langle w \rangle + \langle q_1 \rangle + \langle q_2 \rangle = 0
\]

(2.21)

satisfied by the average work/heat exchange rates.

### 2.3 Nanoscale Heat Engine and Pump

Having computed the average contributions from the various channels of energy exchange between particle and environment shown in Fig. 3a, we would like to study the different modes of operation of this heat engine model.
Operational modes. As a function of the amplitude $f_{\text{ext}}$ of the external force, the system exhibits four different regimes, separated by the zeros of heat flow to the cold reservoir and the externally supplied power. A typical phase diagram is depicted in Fig. 3b. W.l.o.g. we may assume in the following that $T_1$ denotes the temperature of the hot reservoir and $T_2$ that of the cold one, i.e. $T_1 > T_2$. Moreover, we will assume $u_1 > u_2$. The general situation can still be recovered by variation of the angle $\alpha$. Just note that the picture given by Fig. 3b will appear reflected with respect to the y-axis when $\sin(2\alpha)$ turns negative.

All three flows, $\langle q_1 \rangle$, $\langle q_2 \rangle$, and $\langle w \rangle$, share a common root at

$$f_{\text{ext}}^0 = -\frac{T_1 - T_2}{T_1 + T_2} \frac{u_1 - u_2}{2} \sin(2\alpha), \quad (2.22)$$

which corresponds precisely to the point where the external torque compensates the intrinsic torque of the system, i.e. $\langle |\mathbf{x} \times \mathbf{F}| \rangle = \langle |\mathbf{x} \times \mathbf{F}_{\text{ext}}| \rangle$. At this point, the system does, in fact, nothing.

To the left of it, we find a regime where heat is extracted from the cold reservoir ($\langle q_2 \rangle > 0$) and pumped to the hot reservoir, consuming external power. The system thus operates as a heat pump or refrigerator. This regime ends at the second root of $\langle q_2 \rangle$,

$$f_{\text{ext}}^R = -\frac{(u_1 - u_2)}{2} \sin(2\alpha). \quad (2.23)$$

Beyond this point, the external work trivially heats both the hot and cold reservoirs, operating as a “dud engine”. It is interesting to note that the critical force $f_{\text{ext}}^R$ does not depend on the temperatures of the reservoirs, but only on the (mechanical) properties of the potential.

The other interesting region lies to the right of the zero flow point $f_{\text{ext}}^0$. Here, the
system operates as a heat engine and works against the external force (\(\langle w \rangle < 0\)), thus producing a net output power \(\langle -w \rangle\). This regime ends at \(f_{ext}^R = 0\) where the external force vanishes. Beyond this point, we observe a trivial heat transfer from the hot to the cold reservoir until the second root of \(\langle q_1 \rangle\) is reached, from where on both reservoirs are heated again.

**Efficiciencies.** The standard measure of efficiency for a heat pump is its coefficient of performance \(\bar{\beta}\) [1], comparing the heat extracted from the cold reservoir (2.20) with the consumed external power (2.18). In the region \(f_{ext} = f_{ext}^R \ldots f_{ext}^0\), we find

\[
\bar{\beta} = \frac{\langle q_2 \rangle}{\langle w \rangle} = -\frac{1}{2} \left( 1 + \frac{(u_1 - u_2) \sin(2\alpha)}{2 f_{ext}} \right) \geq 0. \tag{2.24}
\]

It approaches the thermodynamically maximum possible value \(\bar{\beta}_{\text{max}} = T_2/(T_1 - T_2)\) [1] near the zero flow point \(f_{ext}^0\). However, the cooling rate tends to zero here, too, so it may not be beneficial to run the pump too close to this point.

For the heat engine efficiency \(\bar{\eta}\), which is ratio of available power (2.18) and extracted heat from the hot reservoir (2.20), we obtain

\[
\bar{\eta} = \frac{\langle -w \rangle}{\langle q_1 \rangle} = \frac{2 f_{ext}}{K_{12}} = \frac{2 f_{ext}}{f_{ext} - \frac{(u_1 - u_2)}{2} \sin(2\alpha) \geq 0}. \tag{2.25}
\]

This notion of efficiency is a macroscopic one, applicable to classical heat engines. The second law of thermodynamics implies that it is bounded by the Carnot efficiency \(\eta_C = 1 - T_2/T_1\) [1, 6, 30], i.e. \(\bar{\eta} \leq \eta_C\). Remarkably, the explicit formula for \(\bar{\eta}\) shows no dependence on the temperatures \(T_1\) and \(T_2\). Note, however, that (2.25) is of course only meaningful in the regime where the system acts as a heat engine, which naturally depends on the temperatures as can be seen from (2.18). Carnot efficiency is reached (macroscopically)—similarly to the maximum COP of the heat pump—as the system approaches the operational turning point \(f_{ext}^0\), where, of course, the output power tends to zero, too.

The performance of small-scale systems such as the Brownian gyrator will generally deviate significantly from these macroscopic predictions, and a major aim of this report is precisely to investigate these fluctuations. To this end, we will define an efficiency on the level of individual trajectories as \(\eta = -w/q_1\) and study its stochastic properties. We will attack this task in Section 3, but already present an example of the distribution of \(\eta\) obtained from a simulation of the system in Fig. 4.

As can be seen, a large spectrum of efficiencies is realized by single trajectories. The distribution is peaked at the macroscopic efficiency \(\bar{\eta}\). However, there is a significant probability to observe negative values (meaning that energy is supplied to the system from the external working force) and, more interestingly, also to find efficiencies higher than the Carnot efficiency \(\eta_C\). In general, the distribution will become narrower as the observation time is increased, and accumulate its mass around the macroscopic efficiency \(\bar{\eta}\).

Let us finally remark that whenever \(T_1 \neq T_2\), \(u_1 \neq u_2\), and \(\sin(2\alpha) \neq 0\), we will find these three different modes of operation. This can be seen from the fact that in
Figure 4: Example of the efficiency distribution after time $\tau = 5 \approx 14.4 \tau_\infty$ obtained from a simulation with parameters $A$ (cf. Appendix D).

this case, $0 = f_{E}^{\text{ext}} < |f_{0}^{\text{ext}}| < |f_{R}^{\text{ext}}|$. It also seems noteworthy that neither efficiency measure depends on the friction coefficients $\gamma_i$.

2.4 Entropy Production

Entropy production plays a crucial role in the study of stochastic thermodynamics. As a matter of fact, understanding entropy relates to (broken) time reversal symmetry in our everyday experience and the question of its microscopic origin. Numerous results linking entropy production to time reversibility have been derived within recent years [27, 31–34]. In particular, within the framework of stochastic processes governed by Langevin equations, entropy production can be related to ratios of path probabilities of individual trajectories, leading to fluctuation theorems. For the system under study, we will explore these in more detail in Section 3.2. Here, we will follow a more phenomenological approach first, associating entropy production in the environment with the heat flow towards the reservoirs. However, we will show explicitly later that the two notions coincide in our case.

Reservoirs. Generalizing the second law of macroscopic thermodynamics to mesoscopic scales, we may connect entropy changes in the reservoirs to the heat flowing towards them [32]:

$$
\text{d}S_R = -\frac{\text{d}Q_1}{T_1} - \frac{\text{d}Q_2}{T_2}.
$$

(2.26)

This definition promotes entropy from a (macroscopic) state function to a fluctuating quantity similar to work, heat, etc. It should not be surprising that the second law, stating the increase of entropy in any irreversible process, may be violated at the level of individual trajectories. We will come back to this in Section 3.2. However, the second law still holds in its original form on the ensemble level: $\langle \Delta S_R \rangle \geq 0$.

The differential heat flowing into the reservoirs was found in (2.11). Integrating along the trajectory $\mathbf{x}$ leads to an entropy production of

$$
\Delta S_R = -\int_0^T \text{d}t \sum_{i,j} T_i^{-1} K_{ij} x_j(t) \circ \text{d}x_i(t) = -\int_0^T \mathbf{x}^T(t) \mathbf{K}^T T^{-1} \circ \text{d}\mathbf{x}(t)
$$

(2.27)
in the environment. Plugging in the equations of motion (1.4), we obtain
\[ \Delta S_R = \int_0^\tau \left[ x^\top(t) K^\top T^{-1} A x(t) - k_B \text{tr} A \right] \, dt - \int_0^\tau x^\top(t) K^\top T^{-1} B d\omega(t), \tag{2.28} \]
where we used \( B B^\top = 2D = 2k_B T \gamma^{-1}, \ A^\top = K^\top T^{-1}, \) and the Stratonovich integral was converted to an equivalent Itô integral. In the steady state, the average entropy production thus becomes
\[ \langle \Delta S_R \rangle = \tau \left[ \langle x^\top K^\top T^{-1} A x \rangle - k_B \text{tr} A \right] = \tau \text{tr} \left[ \Sigma(\infty) K^\top T^{-1} A - k_B A \right], \tag{2.29} \]
where \( \Sigma(\infty) \) is the steady-state covariance matrix as before.

**System.** A distinguishing feature of the stochastic formalism is that we can assign a system entropy to each individual state the particle may visit along its path. Following Seifert [31, 32], this system entropy is given by \( S_{\text{sys}}(t) = -k_B \ln p(t, x), \) where \( p(t, x) = p(t, x \mid t_0, x_0)p_0(x_0) \) is the probability to find the particle in an infinitesimal neighborhood of \( x \) at time \( t. \) This quantity generally depends on the initial distribution \( p_0. \) That definition of \( \Delta S_{\text{sys}} \) is closely related to the statistical interpretation of entropy as a measure of the number of accessible microstates for a given configuration. Averaging over the ensemble, one recovers the macroscopic entropy
\[ \langle S_{\text{sys}}(t) \rangle = -k_B \int d^2x \ p(t, x) \ln p(t, x). \]

The change in system entropy along the trajectory \( x \) is thus
\[ \Delta S_{\text{sys}} = k_B \ln \frac{p(0, x(0))}{p(\tau, x(\tau))}. \tag{2.30} \]
Starting from the steady state, we have \( p(t, x) = p_\infty(x) \) independent of \( t. \) Therefore, there is no entropy production in the system on average, i.e. \( \langle \Delta S_{\text{sys}} \rangle = 0. \) In any case, we note that the entropy production in the system is asymptotically time-intensive such that \( \Delta S_{\text{sys}} \sim \text{const}, \) whereas the entropy production in the medium grows as \( \Delta S_R \sim \tau \) according to (2.29) in the large-time limit. For the total entropy production,
\[ \Delta S_{\text{tot}} = \Delta S_R + \Delta S_{\text{sys}}, \tag{2.31} \]
contributions from the environment thus dominate as \( \tau \to \infty. \)

This completes our survey of the ensemble thermodynamics of the Brownian gyrator. The results at this level are very similar to a macroscopic, classical heat engine. In the next section, we will lower the scale of observation dramatically and investigate the behavior of this nanoscale heat engine at a level where fluctuations become crucially important.

**3 EFFICIENCY FLUCTUATIONS**

The main goal of this section is to study the efficiency distribution of the Brownian gyrator as a heat engine in the long-time limit. While the efficiency is a fixed characteristic property for macroscopic machines, it is subject to vigorous fluctuations on the mesoscopic scale, similar to heat and work as its defining constituents.
We are mainly interested in the *large deviation function* (LDF), denoted $J(\eta)$, which describes the exponential decay of the probability density $p_\tau(\eta)$ as $\tau \to \infty$, such that $p_\tau(\eta) \sim e^{-\tau J(\eta)}$. This interest largely stems from recent observations by Verley, Willaert, *et al.*, who discovered some universal properties of $J$ for a large class of mesoscopic engines [10, 11]. However, their derivation is limited to engines with a finite state space, and although examples exist [35], it remains unclear how the theory generalizes to continuous systems and whether a similar universal shape of the LDF exists. A model similar to ours is studied in [36], and even though the approach there is quite different, the results are similar. We will compare and discuss our findings with both these results and the predictions of the “universal theory” in Section 3.7 after deriving the LDF in Section 3.5.

The method we will use is based on the path integral formalism for diffusion processes. We will therefore introduce the main concepts of this framework in Section 3.1. These tools also allow to draw the aforementioned connection between entropy production and irreversibility of trajectories, a relation we will briefly address in Section 3.2. The large deviation function $J(\eta)$ is related to the *cumulant generating function* (CGF) of work and heat by a Legendre transformation, a fact that we will explore in Section 3.3. Thereafter, we will show in Section 3.4 how the path integral allows the computation of CGFs for observables that are functionals of trajectories such as work (2.14) and heat (2.10). With the work/heat joint CGF at hand, we will be all set to compute the desired LDF in Section 3.5. Thereafter, we will step back and compute the CGF of work and heat again by a different, but related method in Section 3.6 to refine the approximation found in Section 3.4. Finally, we will compare these different results with numerical simulations as well as the universal theory [10, 11] and other recent findings [36] in Section 3.7.

### 3.1 Path Integral Formalism

The foundations of the path integral for diffusive stochastic processes were developed by Onsager and Machlup back in 1953 [37, 38]. The basic idea is to assign a probability density $p[\mathbf{x}]$ on the space of trajectories, *i.e.* a functional assigning a weight to each individual path $\mathbf{x} = [\mathbf{x}(t)]_{t=0}^\tau$. This *path weight* $p$ is formally defined as the continuum limit of the probability density for a discretized trajectory.

**General framework.** In the case of a Markovian system, the (conditional) path probability density can be written as a chain of propagators,

$$
p[\mathbf{x} | \mathbf{x}(0)] = \lim_{N \to \infty} \prod_{k=0}^{N-1} \mathbb{P}(\mathbf{t}_{k+1}, \mathbf{x}_{k+1} | \mathbf{t}_k, \mathbf{x}_k),
$$

where $t_k = k\tau/N$ and $\mathbf{x}_k = \mathbf{x}(t_k)$. This allows us to compute expected values of (functions of) arbitrary stochastic functionals $Y_\tau[\mathbf{x}] := \int_0^\tau y(t, \mathbf{x}(t)) \, dt$:

$$
\langle f(Y_\tau) \rangle = \int D\mathbf{x} \, f(Y_\tau[\mathbf{x}]) \, p[\mathbf{x} | \mathbf{x}(0)] \, p_0(\mathbf{x}(0)).
$$
Here $p_0$ denotes some initial distribution as before, and the integration measure

$$D\mathbf{x} = \lim_{N \to \infty} \prod_{k=0}^{N} d^d x_k,$$

(3.3)
is again the continuum limit of discretized path integrations. Technically, the limit $N \to \infty$ is usually enforced after performing the integration in cases where this is analytically possible. The connection to the single-point probability density, for instance, is given by

$$p(t, \mathbf{x}') = \langle \delta(\mathbf{x}(t) - \mathbf{x}') \rangle.$$  

**Path weight for the Brownian gyrator.** Using the definition (3.1) of the path weight, its precise form for our model system can be derived from the Langevin equation (1.4). We present the details in Appendix B and merely state the result here:

$$p[\mathbf{x} | \mathbf{x}(0)] \propto \exp \left\{ -\frac{1}{2} \int_0^\tau dt \ [\dot{x}(t) + A x(t)]^T (2D)^{-1} [\dot{x}(t) + A x(t)] \right\}. \quad (3.4)$$

The matrices $A$ and $D$ are the same as in (1.6) and (2.1). A (formally divergent) constant we omitted here assures normalization.

### 3.2 Detailed Fluctuation Theorem

Fluctuation theorems in the context of stochastic thermodynamics establish a link between the probability to observe a certain trajectory and extensive quantities produced along them, such as entropy, work, or heat. Many such relations have been found over the past couple of years [32, 33]. We will only discuss the so-called *detailed fluctuation theorem* (DFT) for the total entropy production here. It provides a rather fundamental interpretation of entropy as a measure of irreversibility and sheds some light on the fact that entropy always increases in macroscopic systems. In its integrated form, it yields a generalization of the second law of thermodynamics [32].

**Entropy production and fluctuation theorems.** The DFT we will show for the system under study relates the probabilities to observe a trajectory $\mathbf{x}$ and its time-reversed counterpart $\mathbf{x}$ with $\mathbf{x}(t) = \mathbf{x}(\tau - t)$ to the total entropy production along $\mathbf{x}$ as given by (2.31). It was first observed by Gallavotti and Cohen [39] and simplified and extended by Kurchan, Lebowitz and Spohn [40, 41]. In its most general form, formulated by Seifert [31], it reads

$$\frac{p[\mathbf{x}]}{\tilde{p}[\mathbf{x}]} = e^{\Delta S_{\text{tot}}[\mathbf{x}] / k_B}, \quad (3.5)$$

where $p$ is the path weight for the forward process and $\tilde{p}$ describes the path probability in a corresponding time-reversed protocol. For our system with no time-dependent external control parameters, we have $p = \tilde{p}$. The relation (3.5) may be decomposed into contribution from system and environment, respectively [32]. This leads to

$$\frac{p[\mathbf{x} | \mathbf{x}(0)]}{\tilde{p}[\mathbf{x} | \mathbf{x}(0)]} = e^{\Delta S_R / k_B} \quad \text{and} \quad \frac{p(0, \mathbf{x}(0))}{\tilde{p}(\tau, \mathbf{x}(\tau))} = e^{\Delta S_{\text{sys}} / k_B}, \quad (3.6)$$
the latter merely restating the definition (2.30) of the system entropy production. In
the steady state, it further simplifies because \( p(0, x) = p(\tau, x) = p_\infty(x) \).

**Integrated fluctuation theorem and the second law.** The DFT (3.5) implies
a so-called integrated fluctuation theorem (IFT) refining the second law of thermody-
namics. Using formal path integral manipulations, its derivation is straightforward.
Consider

\[
\langle e^{-\Delta S_{\text{tot}}/k_B} \rangle = \int D\hat{x} p[\hat{x}] e^{-\Delta S_{\text{tot}}[\hat{x}]/k_B} = \int D\hat{x} \tilde{p}[\hat{x}],
\]

where the last equality follows from the DFT (3.5). The remaining quantity is just an
integral over all realizations of the probability density \( \tilde{p} \), so that we are left with

\[
\langle e^{-\Delta S_{\text{tot}}/k_B} \rangle = 1
\]

by normalization. This is the IFT. Jensen’s inequality [42, 43] states that for a convex
function \( f \) and a random variable \( X \), \( \langle f(X) \rangle \geq f(\langle X \rangle) \). Therefore, (3.7) implies the
second law of thermodynamics, \( \langle \Delta S_{\text{tot}} \rangle \geq 0 \).

**Verification of the DFT.** For our model system, the path probabilities are given
by (3.4). For the reversed trajectory, we have \( \hat{x}(t) = x(\tau - t) \) and \( \hat{x}(t) = -\hat{x}(\tau - t) \).
Therefore, we obtain

\[
p[\hat{x} | \hat{x}(0)] \propto \exp \left\{ -\frac{1}{2} \int_0^\tau [\hat{x}(\tau - t) + Ax(\tau - t)]^T (2D)^{-1} [-\hat{x}(\tau - t) + Ax(\tau - t)] \, dt \right\}.
\]

Substituting \( \hat{\tau} = \tau - t \) leads to

\[
p[\hat{x} | x(\tau)] \propto \exp \left\{ -\frac{1}{2} \int_0^\tau \left[ \hat{x}(\hat{\tau}) + A\hat{x}(\hat{\tau}) \right]^T (2D)^{-1} \left[ -\hat{x}(\hat{\tau}) + A\hat{x}(\hat{\tau}) \right] \, d\hat{\tau} \right\},
\]

where we also substituted the final point of the forward trajectory for the initial point
of the backward trajectory. Relating the path weights for forward and backward
trajectories, we thus find

\[
\frac{p[x | x(0)]}{p[\hat{x} | x(\tau)]} = \exp \left\{ -2 \int_0^\tau x^T(t)A^T(2D)^{-1}\hat{x}(t) \, dt \right\}. \tag{3.8}
\]

Using \( 2A^T(2D)^{-1} = K^T \gamma^{-1} D^{-1} = K^T T^{-1}/k_B \), this becomes

\[
\frac{p[x | x(0)]}{p[\hat{x} | x(\tau)]} = \exp \left\{ -\frac{1}{k_B} \int_0^\tau x^T(t)K^T T^{-1} \circ dx(t) \right\} = e^{\Delta S_R/k_B},
\]

where we plugged in (2.27) to obtain the last equality. Obviously, then, the Brownian
gyrator satisfies the DFT (3.5) as expected.

**3.3 Large Deviation Principle**

Probability distributions of physical systems often depend on some scale parameter
that controls the behavior of extensive thermodynamic quantities such as energy, en-

tropy, heat, etc. Typical scale parameters are, for instance, the number of particles or the observation time. When the scale is increased, extensive quantities usually grow likewise on average. To study fluctuations, it is therefore useful to consider densities of these observables, i.e. we divide them by the scale such that they become intensive. The scale then roughly determines the “macroscopicity” of the system in the sense that the effect of fluctuations diminishes for these intensive quantities as the scale increases. For example, deriving classical thermodynamics in statistical mechanics usually involves the thermodynamic limit \( N, V \to \infty \), when fluctuating quantities like entropy or energy become state variables.

Large deviation theory describes the exponential decay of probability measures that depend on such a scale parameter \( \tau \), and an intensive quantity \( y_{\tau} = Y_{\tau}/\tau \) whose probability density follows \( p_{\tau}(y) \sim e^{-\tau I(y)} \) is said to satisfy a large deviation principle [12]. The function \( I \) is called the large deviation function (LDF) of \( y_{\tau} \). It satisfies \( I(y) \geq 0 \) and \( I(\langle y_{\tau} \rangle) = 0 \), which is a variant of the ergodic theorem stating that \( y_{\tau} \) approaches the ensemble average \( \langle y_{\tau} \rangle \) as \( \tau \to \infty \). In our case, the scale parameter \( \tau \) will be the duration and \( y_{\tau} \) some time-averaged functional of the particle’s trajectory.

**LDFs from CGFs.** An important result known as the Gärtner-Ellis theorem [12, 44, 45] relates the LDF of a time-averaged observable \( y_{\tau} = Y_{\tau}/\tau \) to the cumulant generating function (CGF) of its time-extensive counterpart. For a stochastic functional

\[
Y_{\tau}[x] = \int_0^\tau y(x(t)) \, dx(t),
\]

the CGF is defined as

\[
g(\tau, \lambda) := \ln \langle e^{\lambda Y_{\tau}} \rangle,
\]

where \( \langle \cdots \rangle \) denotes the ensemble average over all trajectories. The CGF encodes the full distribution of \( Y_{\tau} \) in the form of its moments or, more precisely, cumulants [7]. These can be extracted by differentiation:

\[
\langle (Y_{\tau})^n \rangle_c = \left. \frac{\partial^n g(\tau, \lambda)}{\partial \lambda^n} \right|_{\lambda=0},
\]

where \( \langle (\cdots)^n \rangle_c \) denotes precisely this \( n \)-th cumulant (connected correlator). From the CGF \( g \), we can derive the scaled cumulant generating function

\[
\hat{g}(\lambda) := \lim_{\tau \to \infty} \frac{1}{\tau} g(\tau, \lambda)
\]

as its infinite time average. The Gärtner-Ellis theorem [12] then states that if \( \hat{g} \) exists and is differentiable for all \( \lambda \in \mathbb{R} \), then the time average \( y_{\tau} = Y_{\tau}/\tau \) satisfies a large deviation principle \( p_{\tau}(y) \sim e^{-\tau I(y)} \). Moreover, its LDF is given by the Legendre transform of the scaled CGF:

\[
I(y) = \sup_{\lambda} [\lambda y - \hat{g}(\lambda)].
\]

We will not give a detailed proof of this relation, but the idea is to apply the Laplace (or saddle-point) approximation to the definition of the scaled CGF. Observe that for
Figure 5: Universal efficiency LDF discovered in [10, 11] for machines with finite state space satisfying a detailed fluctuation theorem. Characteristic features: exactly one minimum at $\eta$, exactly one maximum at $\eta^* = \eta_C$ (steady state), same limiting value as $\eta \to \pm \infty$. Figure taken from [11].

For large values of $\tau$, 
\[ e^{\tau \hat{g}(\lambda)} \sim \int dy \, e^{-\tau I(y) + \lambda \tau y} = \int dy \, e^{-\tau [I(y) - \lambda y]} . \]

As $\tau$ is increased, the integral on the right-hand side is dominated by the minimum of $I(y) - \lambda y$, or equivalently the maximum of $\lambda y - I(y)$, meaning that $\hat{g}$ is the Legendre transform of $I$. But at least for convex functions, the Legendre transform is its own inverse, so that (3.12) is suggested.

**Efficiency LDF.** In the next section, we will derive the scaled joint CGF of work $W$ and heat $Q_1$ for our Brownian gyrator heat engine. Using the Gärtner-Ellis relation (3.12), we can then compute the corresponding LDF $I(q_1, w)$, which in turn gives us access to the LDF $J(\eta)$ of their ratio $\eta = -w/q_1$ [11, 12]. To see this, we reproduce the derivation presented in [11]. The probability distributions are related via

\[ p_r(\eta) = \int dq_1 \, dw \, p_r(q_1, w) \delta \left( \eta - \frac{-w}{q_1} \right) = \int dq_1 \, |q_1| \, p_r(q_1, -\eta q_1) . \]  

(3.13)

In the limit $\tau \to \infty$, we can write $p_r(q_1, -\eta q_1) \sim e^{-\tau I(q_1, -\eta q_1)}$, and the integral is dominated by the value of $q_1$ that minimizes $I$. In other words, the saddle-point or Laplace approximation becomes exact, and we can substitute the minimizing value $q_1^{\text{min}}$:

\[ p_r(\eta) \sim |q_1^{\text{min}}| \, e^{-\tau I(q_1^{\text{min}}, -\eta q_1^{\text{min}})} . \]

Therefore, the efficiency LDF is given by

\[ J(\eta) = \min_{q_1} I(q_1, -\eta q_1) . \]  

(3.14)

Since $I$ vanishes at the average work and heat values, we immediately see $J(\bar{\eta}) = 0$, meaning that the macroscopic efficiency $\bar{\eta} = \langle w \rangle / \langle q_1 \rangle$ is the most likely one in the long-time limit.
Verley et al. derived an intriguing variety of further properties [10, 11] of the LDF. Assuming the detailed fluctuation theorem (3.5) and a finite state space, they then argue that \( J(\eta) \) follows a universal shape, sketched in Fig. 5. The assumption of finite state space, which is clearly violated by the Brownian gyrator system, is needed in order that fluctuations of the system entropy \( \Delta S_{\text{sys}} \) and energy \( \Delta U \) are negligible compared to entropy fluctuations in the reservoirs as \( \tau \to \infty \). In other words, universality should still be observed if energy and entropy fluctuations of the system are bounded.

The universal shape exhibits exactly one minimum at \( \bar{\eta} \), exactly one maximum at some other value \( \eta^* \), and approaches the same value \( J_\infty \) as \( \eta \to \pm \infty \). At the steady state, they further show that \( \eta^* = \eta_C \), so that the Carnot efficiency becomes the least likely value for long times. Other recent findings [36] claim that these results do not hold for a system similar to the one studied here. In particular, they do not find a unique maximum at Carnot efficiency. We will come back to these observations in Section 3.7 after deriving an expression for the efficiency LDF of the Brownian gyrator.

### 3.4 Cumulant Generating Functions of Stochastic Functionals: Closed Trajectory Approximation

In the previous section, we observed that the Gärtner-Ellis theorem offers a convenient way to compute LDFs from scaled CGFs. Since work and heat (as well as the efficiency as their ratio) are defined as functionals of trajectories, we will need a way to compute the cumulant generating function of such integrated observables. In particular, we would like to evaluate

\[
g(\tau, \lambda_Q, \lambda_W) := \ln \langle e^{\lambda_Q Q_{1\tau} + \lambda_W W_{\tau}} \rangle, \tag{3.15}\]

where

\[
W_{\tau}[x] = -\sum_{i,j} \int_0^\tau f^{\text{ext}}_{ij} x_j(t) \circ dx_i(t) \quad \text{and} \quad Q_{1\tau}[x] = \sum_i \int_0^\tau K_{1i} x_i \circ dx_1(t) \tag{3.16}
\]

are the work performed by the external force and the heat extracted from the hot reservoir, as defined in (1.2), (2.14), (2.10) and (2.11) above.

**CGF as a functional determinant.** The method we are going to use to calculate (3.15) is based on the path integral formalism. We will restrict for now to a single functional \( Y_\tau \) and compute \( g(\tau, \lambda) = \ln \langle e^{\lambda Y_\tau} \rangle \). The method can be generalized to several functionals straightforwardly. Furthermore, we observe that both \( W_\tau \) and \( Q_{1\tau} \) can be written in the form

\[
Y_\tau[x] = \sum_{i,j} \int_0^\tau Y_{ij} x_j(t) \circ dx_i(t), \tag{3.17}
\]

where \( Y = (Y_{ij}) \) is some (2 \( \times \) 2)-matrix characterizing the functional. In other words, we will only consider linear functionals.
Using (3.2) and the path weight (3.4), we can express $g$ as

$$e^{g(\tau,\lambda)} \propto \int \mathcal{D}x \exp \left\{ \int_0^\tau dt \left[ -\frac{1}{2} (\dot{x}_t + Ax_t)^T (2D)^{-1} (\dot{x}_t + Ax_t) + \lambda \dot{x}_t^T Y x_t \right] \right\},$$

(3.18)

where we introduced the short-hand notation $x_t = x(t)$ and prepared the system in the steady state by choosing (2.6) as our initial distribution. We also omitted normalization factors because they can always be restored later from the condition $g(\tau,0) = 0$. We now focus on the exponent and separate off boundary terms. To do so, note that for a symmetric matrix $M$, the integrand $x_t M \dot{x}_t dt = d \left( \frac{1}{2} x_t M x_t \right)$ is a total differential, so that the integral depends only on the beginning and end points of the trajectory. Therefore, we decompose all matrices coupling $x_t$ and $\dot{x}_t$ into symmetric and antisymmetric components and evaluate the integrals for the symmetric parts. Moreover, we partially integrate the term quadratic in $\dot{x}_t$, which transfers one of the derivatives and also gives an additional boundary term. All in all, this leads to

$$e^{g(\tau,\lambda)} \propto \int \mathcal{D}x \exp \left\{ \int_0^\tau dt \left[ \frac{1}{2} x_t (2D)^{-1} \dot{x}_t - \frac{1}{2} x_t^T A^T (2D)^{-1} A x_t - \frac{1}{2} x_t \left( A^T (2D)^{-1} - (2D)^{-1} A + \lambda Y - \lambda Y^T \right) x_t \right] \right\}$$

(3.19)

$$- \frac{1}{2} x_t^T (2D)^{-1} \dot{x}_t + \frac{1}{2} x_t^0 (2D)^{-1} \dot{x}_0$$

$$- \frac{1}{2} x_t^T \left( \frac{A^T (2D)^{-1} + (2D)^{-1} A}{2} - \lambda \frac{Y + Y^T}{2} \right) x_t$$

$$- \frac{1}{2} x_t^0 \left( \Sigma(\infty)^{-1} - \frac{A^T (2D)^{-1} + (2D)^{-1} A}{2} + \lambda \frac{Y + Y^T}{2} \right) x_t. \right\}$$

This expression can be written rather compactly as

$$e^{g(\tau,\lambda)} \propto \int \mathcal{D}x \exp \left\{ -\frac{1}{2} \int_0^\tau dt \int_0^\tau dt' x(t)^T V_\lambda(t,t') x(t') \right\}$$

(3.20)

by introducing an operator

$$V_\lambda(t,t') = \delta(t - t') \dot{V}_\lambda(t') + \delta(t - \tau) \delta(t' - \tau) \dot{V}_\lambda^T(t') + \delta(t) \delta(t') \dot{V}_\lambda^0(t')$$

(3.21a)

with

$$\dot{V}_\lambda(t) = -(2D)^{-1} \partial_t^2 + \left[ A^T (2D)^{-1} - (2D)^{-1} A + \lambda \left( Y - Y^T \right) \right] \partial_t + A^T (2D)^{-1} A$$

(3.21b)

and

$$\dot{V}_\lambda^T(t) = (2D)^{-1} \partial_t + \left( A^T (2D)^{-1} + (2D)^{-1} A \right) - \lambda \left( Y + Y^T \right),$$

(3.21c)

$$\dot{V}_\lambda^0(t) = -(2D)^{-1} \partial_t - \left( A^T (2D)^{-1} + (2D)^{-1} A \right) + \lambda \left( Y + Y^T \right) + \Sigma^{-1}(\infty).$$

(3.21d)

The right-hand side in (3.20) is just a functional variant of a Gaussian integral, i.e.
a Gaussian integral with an infinite dimensional matrix kernel $V_\lambda$ that has both an ordinary matrix part and a functional ($t$-space) part. Formally, it reduces to a determinant

$$e^{g(\tau, \lambda)} = (\text{Det} \det V_\lambda)^{-1/2},$$

(3.22)

where ‘det’ denotes the usual matrix determinant and ‘Det’ stands for the functional determinant of the differential operator. Such expressions frequently occur in field theoretic formulations of quantum or statistical physics [46–48], and these communities have developed various ways to deal with them in different situations. Here, we will evaluate it in two different ways, both of which unfortunately require some sort of approximation. Our first method uses a Fourier expansion and resummation of the operator trace. It has the advantage that the result is rather explicit, but neglects boundary contribution which are subleading in the limit of interest when $\tau \to \infty$. In contrast, our second method uses the inverse operator or Green’s matrix of $V_\lambda$ and is in principle exact, but requires the solution of high-dimensional systems of linear equations, which we can only carry out numerically, preventing us from taking the limit $\tau \to \infty$ directly.

**Kernel diagonalization.** The determinant of a matrix or more generally an operator is the product of its eigenvalues. One way to evaluate (3.22) further is thus to diagonalize the operator $V_\lambda$. To this end, we expand the trajectories $x(t)$ in a Fourier series:

$$x(t) = \sum_{n \in \mathbb{Z}} \hat{x}_n e^{-i\omega_n t}, \quad \omega_n = \frac{2\pi n}{\tau}. \quad (3.23)$$

A few remarks regarding the rigor of this transformation are in order. By Fourier-expanding the trajectories, we silently assume them to be periodic, meaning that we restrict the functional integral (3.20) to closed trajectories. For this reason, the $\lambda$-dependences of the boundary terms, encoded in $V_0^\lambda$ and $V_\tau^\lambda$, cancel in this approximation, as can be seen directly in (3.19) for $x_0 = x_\tau$.

In general, this is a serious limitation and the resulting CGF may deviate dramatically from the true solution. However, remember that we are ultimately only interested in the limit $\tau \to \infty$ and the scaled CGF $\hat{g}$. In this limit, the Fourier expansion becomes a continuous Fourier transformation, and effects from the boundary diminish.

Furthermore, we can employ the ergodic theorem and argue that infinite time averages converge to ensemble averages. This means that as $\tau \to \infty$, all trajectories represent the ensemble equally well, so that a CGF obtained from a reduced set of trajectories should still converge to the full CGF unless the associated cut in the space of trajectories is pathological. We will further support this claim with our second method for evaluating the functional determinant as well as numerical simulations below, showing that the closed trajectory CGF generates cumulants that are fairly close to the true values for quite small, finite durations $\tau$.

Plugging the Fourier expansion (3.23) into (3.21) and absorbing the boundary terms
into normalization (since they do not depend on $\lambda$ anymore), the exponent becomes
\[
-\frac{1}{2} \sum_{n,n'} \int_0^\tau dt \hat{x}_{n'}^T \left[ \omega_n^2 (2D)^{-1} + A^T (2D)^{-1} A \right. \\
\left. - i \omega_n \left( A^T (2D)^{-1} - (2D)^{-1} A + \lambda (Y - Y^T) \right) \right] \hat{x}_n e^{-i(\omega_n + \omega_{n'})t}.
\]

Note that only the antisymmetric component of the matrix $Y$ survives here and couples to the variable $\lambda$ since the symmetric component is precisely the boundary term that cancels for closed trajectories.

The integral can be carried out using the Fourier identity
\[
\int_0^\tau e^{-i(\omega_n + \omega_{n'})t} = \tau \delta_{n,-n'}.
\]
Moreover, since the trajectories $x(t)$ are real, we have that $\bar{x}_{-n} = \bar{x}_n^*$, the * denoting the complex conjugate. With these relations, the above integral becomes
\[
-\frac{1}{2} \sum_n \hat{x}_n^T \tau \left\{ \omega_n^2 (2D)^{-1} + i \omega_n \left[ (2D)^{-1} A - A^T (2D)^{-1} - \lambda (Y - Y^T) \right] + A^T (2D)^{-1} A \right\} \hat{x}_n,
\]
where $^\dagger$ stands for the Hermitian conjugate. From this, we can read off the Fourier-space representation of the operator $\hat{V}_\lambda$:
\[
\hat{V}_{\lambda,n} = \tau \left\{ \omega_n^2 (2D)^{-1} + i \omega_n \left[ (2D)^{-1} A - A^T (2D)^{-1} - \lambda (Y - Y^T) \right] + A^T (2D)^{-1} A \right\}.
\]

Consider now again (3.22). To arrive at $g(\tau, \lambda)$, we take the logarithm of this expression, so that
\[
g(\tau, \lambda) = \ln \left[ (\text{Det det } V_\lambda)^{-1/2} \right] = -\frac{1}{2} \text{Tr ln det } V_\lambda.
\]

Here ‘Tr’ is a functional trace in trajectory space, whereas ‘det’ still denotes the ordinary matrix determinant. Since determinant and trace are independent of the chosen basis, we can substitute (3.24) into this and obtain
\[
g(\tau, \lambda) \simeq -\frac{1}{2} \sum_n \ln \text{det} \left( \tau \left\{ \omega_n^2 (2D)^{-1} + A^T (2D)^{-1} A \right. \\
\left. - i \omega_n \left[ (2D)^{-1} A + A^T (2D)^{-1} - \lambda (Y - Y^T) \right] \right\} \right),
\]
expressing the trace as a sum over all eigenvalues. Note that we do not write an equality sign owing to the closed trajectory approximation.

**Resummation of the operator trace.** It remains to evaluate the functional trace in (3.26). The infinite sum over the frequencies $\omega_n$ is formally divergent, but we will be able to resum it by means of contour integral in the complex $\omega$-plane. Infinite sums of the above type, which run over a set of discrete Fourier frequencies, are known as *Matsubara series*, and the corresponding $\omega_n$ are called *Matsubara frequencies* after Takeo Matsubara [48]. They frequently occur in thermal quantum field theory at finite temperature, where bosonic and fermionic fields are periodic or antiperiodic, respectively, in the inverse temperature so that correlation functions can be expressed in terms of Fourier-space traces.
We will first simplify (3.26) slightly by computing the matrix determinant. For each \( \omega_n \), the matrix in parentheses is of order \( \omega_n^2 \) and has dimensions \( 2 \times 2 \). The determinant is thus a fourth-order polynomial in \( \omega_n \). We can factorize this polynomial and write (3.26) as

\[
 g(\tau, \lambda) \simeq -\frac{1}{2} \sum_n \ln \prod_{k=1}^{4} \left[ \sqrt{\tau} \left( \det 2D \right)^{-1/4} \left( i\omega_n - \hat{\omega}_k \right) \right],
\]

where \( -i\hat{\omega}_k \) are the four roots of

\[
 \det \left\{ \omega^2(2D)^{-1} + i\omega \left( (2D)^{-1} A - A^T(2D)^{-1} - \lambda(Y - Y^T) \right) + A^T(2D)^{-1} A \right\} = 0.
\]

(3.28)

Up to a constant that we will be able to restore from the normalization condition \( g(\tau, 0) = 0 \), the approximated CGF thus reads

\[
 g(\tau, \lambda) \simeq -\frac{1}{2} \sum_n \ln \left( i\omega_n - \hat{\omega}_k \right).
\]

(3.29)

In this equation we clearly see that a naive summation over \( n \) will diverge. Nevertheless, we can express it as a contour integral and find the value of the asymptotic series by analytic continuation. Since the procedure is the same for all four roots, we drop the superscript \( k \) and subscript \( \lambda \) for now and consider

\[
 s(\hat{\omega}) = \sum_n \ln(i\omega_n - \hat{\omega}).
\]

Differentiating with respect to \( \hat{\omega} \) (and shamelessly exchanging the order of summation and differentiation), we obtain

\[
 s'(\hat{\omega}) = -\sum_n \frac{1}{i\omega_n - \hat{\omega}}.
\]

(3.30)

We define a complex-valued function

\[
 h(z) := \frac{e^{\eta z}}{z - \hat{\omega}}
\]

with \( \eta > 0 \) being a convergence generating factor that we will eventually let to zero. With this definition, \( s'(\hat{\omega}) = -\lim_{\eta \to 0} \sum_n h(i\omega_n) \). We would like to express this sum as a sum of residues of some other function with poles at \( z = i\omega_n = 2\pi n/\tau \). We can artificially add such simple poles by multiplying \( h \) with \( 1/(1 - e^{\tau z}) \), since \( 1 - e^{\tau z} = 0 \) iff \( z = i\omega_n \). The product then has an additional simple pole from \( h \) at \( \hat{\omega} \), which we assume for now to have nonvanishing real part. The resulting overall pole structure is depicted by the black points in Fig. 6.

If we integrate this function along the contour \( C_I \) sketched in Fig. 6a, we find

\[
 \oint_{C_I} \frac{h(z)}{1 - e^{\tau z}} \, dz = 2\pi i \sum_n \text{Res}_{z=i\omega_n} \frac{h(z)}{1 - e^{\tau z}} = -\frac{2\pi i}{\tau} \sum_n h(i\omega_n).
\]

Now observe that

\[
 \frac{h(z)}{1 - e^{\tau z}} \sim \begin{cases}
 -e^{(\eta - \tau)z} & \text{as } Re \, z \to +\infty \\
 -e^{\eta z} & \text{as } Re \, z \to -\infty
\end{cases}
\]

25
Figure 6: (a) Contour of the integral to collect residues at the imaginary Matsubara frequencies $i\omega_n$; (b) the same contour now closed the opposite way with the integrand vanishing as $\text{Re } z \to \pm \infty$.

Therefore, we can close the contour in the opposite way at infinity as shown in Fig. 6b without picking up extra contributions because the integrand falls off faster than $1/z$ on the additional arcs. This yields

$$\sum_n h(i\omega_n) = -\frac{\tau}{2\pi i} \oint_{C_I} \frac{h(z)}{1-e^{\tau z}} \, dz = -\frac{\tau}{2\pi i} \left[ \oint_{C_O^{(+)}} + \oint_{C_O^{(-)}} \right] \frac{h(z)}{1-e^{\tau z}} \, dz.$$ 

With this choice of the contour, we only pick up the residue at the remaining pole at $\hat{\omega}$ coming from $h$, so that

$$\sum_n h(i\omega_n) = +\tau \text{ Res}_{z \to \hat{\omega}} h(z) = \frac{\tau e^{\eta\hat{\omega}}}{1-e^{\tau\hat{\omega}}}.$$ 

Remembering (3.30) and (3.31), we find $s'(\hat{\omega}) = \tau/(1-e^{\tau\hat{\omega}})$. After integration, we are left with the desired result:

$$\sum_n \ln(i\omega_n - \hat{\omega}) = \ln \left( e^{-\tau\hat{\omega}} - 1 \right).$$

(3.32)

Substituting into (3.29), we finally obtain

$$g(\tau, \lambda) \simeq -\frac{1}{2} \sum_{k=1}^4 \ln \left( e^{-\tau\hat{\omega}_\lambda^{(k)}} - 1 \right) + \text{normalization}$$

(3.33)

with the normalization constant determined by $g(\tau, 0) = 0$.

Given this result, we should keep in mind that we assumed the roots $\hat{\omega}_\lambda^{(k)}$ to have $\text{Re } \hat{\omega}_\lambda^{(k)} \neq 0$ when evaluating the functional trace. Only then can the Matsubara series be rewritten as a contour integral because this assumption separates the poles at the imaginary Matsubara frequencies $i\omega_n$ from the determinant pole at $\hat{\omega}$. If this is not the case, the resummation breaks down, a fact that also manifests itself in a discontinuity of the logarithm in (3.33). Analyzing the sum (3.29) from the beginning, we must take its value to be infinity in this case, i.e. $g(\tau, \lambda) = \infty$ for all $\lambda$ that render any of the real parts of $\hat{\omega}_\lambda^{(k)}$ vanishing. This effectively provides a cutoff to the CGF (3.33). We will see below that an additional cutoff is obtained from the symmetric contributions.
of the functional $\mathbf{Y}_\tau$, i.e. from the symmetric part of the matrix $\mathbf{Y}$, which we have neglected by Fourier-transforming and thus restricting to closed trajectories.

The values of the cumulants of $\mathbf{Y}_\tau$, however, only depend on $g(\tau, \lambda)$ in the vicinity of $\lambda = 0$, as can be seen from (3.10). We will see below that the domain of convergence includes the origin. Therefore, we can assess at this point already how well the cumulants are reproduced by our approximated CGF (3.33).

**Comparison to simulation.** In order to estimate the quality of the approximated CGF (3.33), we compare the cumulants generated by it to numerical simulations. For our purposes, we are mostly interested in the CGFs of work and heat. The $\mathbf{Y}$ matrices take the forms

$$
\mathbf{Y}_W = -f_{\text{ext}} \mathbf{\mathcal{E}} = \begin{pmatrix} 0 & -f_{\text{ext}} \\ f_{\text{ext}} & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{Y}_{Q_1} = \begin{pmatrix} K_{11} & K_{12} \\ 0 & 0 \end{pmatrix},
$$

(3.34)

respectively, for $W_\tau$ and $Q_{1\tau}$ given in (3.16). An example of the first four moments of these two functionals (or rather their time averages) obtained from a numerical simulation is shown in Fig. 7. The figure also depicts the cumulants computed from the corresponding approximated CGF (3.33).

We observe that the matrix $\mathbf{Y}_W$ is totally antisymmetric such that the Fourier-transformed operator $\hat{\mathbf{V}}_\lambda$ (3.24) contains all information about $W$. As one would expect, its cumulants are indeed well reproduced by the approximated CGF. Only for very small observation times $\tau$ does it seem to behave weirdly. This is, however, an artifact of
the closed trajectory approximation since paths with equal starting and end points become increasingly unlikely on short times and thus do not represent the ensemble appropriately.

The approximation appears to be a lot worse for the heat flow $Q_{1,\tau}$, which is of course caused by the fact that the matrix $Y_{Q_1}$ is represented only rather crudely by its antisymmetric component. The plots thus show considerable deviations between numerical results and analytic approximations, in particular for the even moments. Nevertheless, the approximation tends to improve as the observation time is increased. This is in accordance with our previous argument of decreasing importance of initial and final conditions in the long-time limit. We will show below that all moments of symmetric observables approach a finite constant as $\tau \to \infty$, so that the moments of time-averaged symmetric functionals indeed vanish.

**Symmetric contributions and domain of convergence.** Analyzing the resummation procedure to evaluate the functional trace above, we observed that it breaks down if $\text{Re} \hat{\omega}_\lambda^{(k)} = 0$ for any of the four roots $\hat{\omega}_\lambda^{(k)}$. In this case, the CGF (3.33) diverges, leading to a $\lambda$-dependent cutoff for $g(\tau, \lambda)$. An additional cutoff might arise from the boundary terms $V_0^\lambda$ and $V_\tau^\lambda$, and in particular from the symmetric contributions of the matrix $Y$. We argued before that our approximation using closed trajectories, which effectively cancels the symmetric part of $\lambda Y$, becomes exact in the limit $\tau \to \infty$ due to the ergodic theorem. We will now show that if we consider an observable $Y_\tau$ like (3.17) with a purely symmetric matrix $Y_{\text{sym}}$, the corresponding generating function $g_{\text{sym}}(\tau, \lambda)$ is indeed bounded in a neighborhood of $\lambda = 0$. This means that all moments approach constant values as $\tau \to \infty$, and consequently do not contribute to the time average in this limit. However, we will also see that $g_{\text{sym}}$ diverges for $\lambda$ outside this neighborhood. When this is the case, the assumption of negligible $Y_{\text{sym}}$ is not justified any longer. Rather, we obtain $g_{\text{sym}}(\tau, \lambda) = \infty$ then, too.

To underpin this reasoning with some mathematical rigor, let us consider the functional (3.17) with symmetric $Y \to Y_{\text{sym}}$. As mentioned above, the integral is then over a total differential and we obtain

$$Y_{\tau,\text{sym}}[x] = \int_0^\tau d\tau(t) \circ Y_{\text{sym}} x(t) = \frac{1}{2} x^T(t) Y_{\text{sym}} x(t) |_{t=0}^\tau. \quad (3.35)$$

The path integral for the CGF (3.18) thus reduces to an average with the two-point probability distribution function $p(\tau, x_\tau; 0, x_0) = p(\tau, x_\tau | 0, x_0) p(0, x_0) = e^{\frac{1}{2}(x_\tau^T Y_{\text{sym}} x_\tau - x_0^T Y_{\text{sym}} x_0)}$. (3.36)

These are two Gaussian integrals, and although their evaluation is quite cumbersome, they can be carried out straightforwardly. One finds

$$g_{\text{sym}}(\tau, \lambda) = -\frac{1}{2} \ln \det \left[ \left(1 - \lambda \Sigma(\tau) Y_{\text{sym}} \right) \left(1 + \lambda \Sigma(\infty) Y_{\text{sym}} \right) + H_\tau^{-1} \left(1 - (1 - \lambda H_\tau \Sigma(\infty) e^{-\tau A^T Y_{\text{sym}} e^{-\tau A}})^{-1}\right) \right] \quad (3.37)$$
Figure 8: Time-averaged cumulant generating functions of the symmetrized heat flow from reservoir 1 for different observation times \( \tau \). System parameters follow set A (cf. Appendix D).

with

\[
H_\tau = e^{\tau A} \Sigma(\infty) e^{\tau A^T} \Sigma^{-1}(\infty) - 1. 
\]  

(3.38)

To obtain the corresponding scaled CGF \( \hat{g}_{\text{sym}} \), we first observe that \( H_\tau \sim e^{\tau A} e^{\tau A^T} \) as \( \tau \to \infty \). Therefore, the term in square brackets in (3.37) is of order 1 in this limit, and multiplied by \( H_\tau^{-1} \) the product vanishes. This implies

\[
g_{\text{sym}}(\infty, \lambda) = -\frac{1}{2} \ln \det \left\{ 1 - \lambda^2 [\Sigma(\infty) Y_{\text{sym}}^2] \right\}. 
\]  

(3.39)

Both the finite-time CGF (3.37) and the infinite-time limit (3.39) are well-defined as long as the determinant is positive, which is guaranteed for small values of \( \lambda \). However, there will be values of \( \lambda \) such that the determinant becomes zero or takes on negative values. In this case, the logarithm is no longer well-defined, and we cannot just neglect the symmetric contributions naively.

The case of \( Y = Y_{Q_1} \) (cf. Equation 3.34) is of particular interest because its symmetric contributions will lead to a divergent joint CGF of work and heat if the corresponding \( \lambda \)-parameter becomes too large. The time-averaged CGF for the symmetrized heat functional \( Q_{\text{sym}}^{\tau} \) is shown in Fig. 8a for three different choices of \( \tau \). The graphs suggest that \( g_{\text{sym}}(\tau, \lambda)/\tau \) indeed decays to zero in a certain neighborhood of the origin, but diverges as a critical value (\( \lambda \approx 0.5 \) in the units chosen) is reached.

All in all, we thus have two bounds for the parameter \( \lambda \) beyond which \( g(\tau, \lambda) \) diverges. One of them arises from the resummation procedure, the other one from the symmetric contributions of the functional \( Y_{\tau} \) that enter the boundary components of the operator \( V_\lambda \). Therefore, we introduce a domain of convergence \( \Omega_\tau \) that depends on both the determinant roots \( \hat{\omega}_\lambda^{(k)} \) and the CGF \( g_{\text{sym}} \) of the symmetrized version of the functional \( Y_{\tau} \):

\[
\Omega_\tau := \{ \lambda : \text{Re} \hat{\omega}_\lambda^{(k)} \neq 0 \ (k = 1, \ldots, 4) \text{ and } g_{\text{sym}}(\tau, \lambda) < \infty \}. 
\]  

(3.40)

Using a characteristic function

\[
\chi_{\Omega_\tau}(\lambda) = \begin{cases} 
1 & \text{if } \lambda \in \Omega_\tau, \\
\infty & \text{otherwise}, 
\end{cases}
\]  

(3.41)
we can then refine our approximation of the CGF $g$ by writing

$$g(\tau, \lambda) \simeq -\frac{1}{2} \chi(\lambda) \sum_{k=1}^{4} \ln \left( e^{-\tau \tilde{\omega}_{\lambda}^{(k)}} - 1 \right) + \text{normalization}. \quad (3.42)$$

**Scaled CGF.** In order to apply the Gärtner-Ellis theorem and compute large deviation functions, we actually need the scaled CGF $\tilde{g}(\lambda)$ obtained from $g(\tau, \lambda)/\tau$ in the limit $\tau \to \infty$. We therefore consider (3.42) once again. There are two cases to be distinguished to assess the behavior of the logarithm in the infinite $\tau$ limit:

- if $\text{Re} \tilde{\omega}_{\lambda}^{(k)} > 0$: $\ln \left( e^{-\tau \tilde{\omega}_{\lambda}^{(k)}} - 1 \right) \sim i\pi - e^{-\tau \tilde{\omega}_{\lambda}^{(k)}}$ (principal branch);
- if $\text{Re} \tilde{\omega}_{\lambda}^{(k)} < 0$: $\ln \left( e^{-\tau \tilde{\omega}_{\lambda}^{(k)}} - 1 \right) \sim -\tau \tilde{\omega}_{\lambda}^{(k)}$.

In the first case, however, the corresponding term vanishes in (3.42) if divided by $\tau$ and sending $\tau \to \infty$. Hence, the scaled CGF $\tilde{g}(\lambda) = \lim_{\tau \to \infty} g(\tau, \lambda)/\tau$ is given by

$$\tilde{g}(\lambda) = \frac{1}{2} \chi(\lambda) \sum_{k: \text{Re} \tilde{\omega}_{\lambda}^{(k)} < 0} \tilde{\omega}_{\lambda}^{(k)} + \text{normalization} \quad (3.43)$$

in the closed-trajectory approximation. Before we present the second method to evaluate the functional determinant (3.22), which in contrast to the just derived result accounts for boundary contributions, but can only be computed numerically, we will use (3.43) to determine the efficiency LDF of the Brownian gyrator from it. As we will see below, the result in fact appears to be sound and in accordance with the numerical method as well as simulations.

### 3.5 Large Deviation Function for Efficiency Fluctuations

After this lengthy and tedious derivation of an approximate cumulant generating function for stochastic functionals, we finally have all tools required to compute the efficiency large deviation function. Remember that the efficiency LDF can be obtained by contracting the joint LDF of work and heat fluctuations according to (3.14). This object in turn can be computed as the Legendre transform of the their scaled joint CGF. Combining these two relations, we can compute the efficiency LDF directly from the scaled joint CGF of work and heat [11], leading to

$$J(\eta) = -\inf_{\lambda_W} \tilde{g}(\eta \lambda_W, \lambda_W). \quad (3.44)$$

The foundation of this derivation was the Gärtner-Ellis theorem presented in Section 3.3, which establishes the relation of $I(q, w)$ and $\tilde{g}(\lambda_Q, \lambda_W)$ by a Legendre transformation. We remind ourselves that one of the prerequisites of this theorem was that $\tilde{g}$ be differentiable for all $\lambda_Q$ and $\lambda_W$. This condition is clearly violated if the domain of convergence of $\tilde{g}$ is finite, which is indeed the case here as we will see below. Nevertheless, it will pay off to carry on and still use the Gärtner-Ellis theorem and its consequences, since the results obtained appear to be sound in the end.
Scaled joint CGF of work and heat. The method to compute scaled CGFs described in the previous section can be generalized immediately to several observables. All we have to do is add the $Y$-matrices corresponding to the different functionals to the operator $V$ in (3.21) or directly to the determinant (3.28) that defines the roots $\hat{\omega}_{\lambda Q,\lambda W}^{(k)}$. Therefore, we replace $\lambda Y$ by $\lambda Q Y_{Q_1} + \lambda W Y_W$ in the expressions of the previous section. To be explicit, we define $-i\hat{\omega}_{\lambda Q,\lambda W}^{(k)}$ to be the four roots of

$$\det \left\{ \omega^2(2D)^{-1} + i\omega \left[ (2D)^{-1} A - A^T (2D)^{-1} \right] - \lambda Q (Y_{Q_1} - Y_{Q_1}^T) - \lambda W Y_W \right\} + A^T (2D)^{-1} A = 0,$$

where $Y_{Q_1}$ and $Y_W$ were defined in (3.34). Note that we need not antisymmetrize $Y_W$ since it already is antisymmetric. Similarly, we obtain the domain of convergence $\Omega$ by considering the location of the roots $\hat{\omega}_{\lambda Q,\lambda W}^{(k)}$ in the complex plane and contributions from the symmetric part of $Y_{Q_1}$:

$$\Omega = \{ (\lambda Q, \lambda W) : \Re \hat{\omega}_{\lambda Q,\lambda W}^{(k)} \neq 0 \ (k = 1, \ldots, 4) \ \text{and} \ \hat{g}_{Q_1}^{\text{sym}}(\lambda Q) < \infty \}.$$ (3.46)

The scaled joint CGF of work and heat can thus be written as

$$\hat{g}(\lambda Q, \lambda W) = \frac{1}{2} \chi_{\Omega}(\lambda Q, \lambda W) \sum_{k : \Re \hat{\omega}_{\lambda Q,\lambda W}^{(k)} < 0} \hat{\omega}_{\lambda Q,\lambda W}^{(k)} + \text{normalization}$$ (3.47)

with normalization following from $\hat{g}(0,0) = 0$. The scaled joint CGF obtained this way for parameters $\Lambda$ (cf. Appendix D) is plotted in Fig. 9. The plot actually shows the contours of the unrestricted resummation result similar to (3.33) and outlines the domain of convergence $\Omega$ beyond which the function diverges in green. The color code chosen is such that “warmer” colors represent greater values of $\hat{g}$, meaning that the function values increase from the minimum $g_{\text{min}} < 0$ along the dark blue line towards the red region outside of $\Omega$.

We first of all notice that $\hat{g}$ is indeed convex as expected. It is, however, degenerate in the sense that the contour lines are all parallel, i.e. the function is constant along lines $\lambda Q - \overline{\eta} \lambda W = 0$, where $\overline{\eta}$ is the macroscopic efficiency. This peculiarity can be traced back to the closed trajectory approximation, which implied that only the antisymmetric components of the $Y$-matrices survived in (3.46). But an antisymmetric matrix in two dimensions has only one degree of freedom, which is $Y_{Q_1}^{\text{asym}} = -Y_{W_1}^{\text{asym}}$. The parameters $\lambda Q$ and $\lambda W$ thus enter (3.46) only in the combination

$$(Y_{Q_1}^{\text{asym}})_{12} \lambda Q + (Y_W)^{\text{asym}}_{12} \lambda W = \frac{1}{2} K_{12} \lambda Q - f_{\text{ext}} \lambda W \left[ f_{\text{ext}} + u_2 - u_1 \right. \left. \sin(2\alpha) \right] (\lambda Q - \overline{\eta} \lambda W),$$

as follows from (3.34) and (2.25).

Contraction to efficiency LDF. The procedure to obtain the efficiency LDF from the scaled joint CGF is encoded in (3.44). It states that for each efficiency $\eta$, the corresponding value $J(\eta)$ is obtained by minimizing $\hat{g}$ along the line $\lambda Q = \eta \lambda W$ and taking the negative of this minimum. Thus for each $\eta$, we can draw a line of slope...
In either case, for \( \eta = 0 \) corresponding to a vertical line, we inevitably pick up the global minimum \( g_{\text{min}} \) of \( \hat{g} \). Increasing the efficiency means rotating our minimization line with slope \( 1/\eta \) in a clockwise direction. Looking at the figure, we see that we keep obtaining the negative of the global minimum \( g_{\text{min}} \) until the line reaches the intersection of the global minimum contour with the right \( \lambda_Q \)-cutoff. The corresponding efficiency is denoted \( \eta_R \). From thereon, \( J(\eta) \) is determined by the value of \( -\hat{g} \) at the right cutoff, until we reach \( \bar{\eta} \). Here, the minimization line is parallel to the contour lines of \( \hat{g} \). Because the line goes through the origin and \( \hat{g}(0,0) = 0 \), we get \( J(\bar{\eta}) = 0 \). For \( \eta > \bar{\eta} \), the minimum is then attained at the left cutoff, implying that \( J(\eta) \) grows again. For the case shown in Fig. 10a, this continues and \( J(\eta) \) approaches the value from minimizing along the horizontal line \( \lambda_W = 0 \), corresponding to efficiency \( \eta = \pm \infty \). In the other case, Fig. 10b, the global minimum is picked up again for values \( \eta \geq \eta_{R2} \), which corresponds to the point where the \( g_{\text{min}} \) contour line intersects the left cutoff.

For negative efficiencies we can proceed in a similar fashion. Consider first the situation as in Fig. 10a. Starting again from \( \eta = 0 \), but rotating counterclockwise, we are left with \( J(\eta) = -g_{\text{min}} \) until the intersection point of the \( g_{\text{min}} \) contour with the left \( \lambda_Q \)-cutoff is crossed at efficiency \( \eta_L \). Thereafter, \( J(\eta) \) picks up the value at the intersection with the left cutoff and decreases monotonically towards the same limiting value \( J(\infty) \). In the case of Fig. 10b, on the contrary, there is no efficiency line
Figure 10: Computation of the efficiency LDF from the scaled joint CGF of work and heat. The top panels show a sketch of the construction of $J(\eta)$ from $\hat{g}(\lambda_Q, \lambda_W)$ according to (3.44), and the bottom panels depict the resulting efficiency LDF. In (a) parameters were chosen according to set A (cf. Appendix D), whereas in (b) set B was used. These two choices represent the two scenarios possible. Detailed explanation in the text.
that can pass through the intersection of the $g_{\text{min}}$ contour and the left $\lambda Q$ cutoff because this intersection lies outside the domain of convergence. Therefore, we obtain $J(\eta) = -g_{\text{min}}$ for all $\eta < 0$ here.

The resulting shape in case (a) (bottom panel of Fig. 10a) is to some extent similar to the universal shape described in Section 3.3 (Fig. 5), but there are also two striking differences. The first one is that it appears to be reflected because the maximum now turns up to the left of the macroscopic efficiency $\bar{\eta}$ rather than at Carnot efficiency $\eta_{C}$, which is always greater than $\bar{\eta}$. Secondly, there is in fact a whole region of maximum values given by the plateau of $J(\eta)$ around $\eta = 0$ as opposed to a single maximum predicted by the universal theory. In the second case (b), there seems to be less resemblance at first glance, but the LDF obtained from our approximation (bottom panel of Fig. 10b) is technically also just a truncated version of the reflected universal shape, the difference being that the cut seems to be taken below $J(\infty)$. We will elaborate in more detail on these distinguishing features in Section 3.7. Before doing so, however, we will work out our second method to evaluate the functional determinant (3.22) and thereby refine our result.

### 3.6 Cumulant Generating Functions of Stochastic Functionals: Green’s Matrix Method

This section is devoted to an alternative evaluation of the functional determinant (3.22) of the differential operator $V_\lambda$ defined in (3.21). The previous method only considered the ordinary component (3.21b) and was unable to capture the boundary contributions (3.21c) and (3.21d) because the “diagonalization” by means of Fourier expansion only works in the subspace of closed trajectories. These boundary terms had to be incorporated in a hand-waving way a posteriori by an independent analysis leading to a limited domain of convergence of the bulk result. This deficiency will be overcome by the method suggested in this section. The price we will have to pay, though, is increased complexity that will eventually force us to resort to numerical calculations.

**Operator inverse and Green’s matrix.** Instead of working with the operator $V_\lambda$ directly, we will now investigate its inverse $\Gamma_\lambda := V_\lambda^{-1}$, which is defined in a functional sense again via the relation

$$
\int_0^\tau dt' \, V_\lambda(t, t') \Gamma_\lambda(t', t'') = \delta(t - t'')
$$

(3.48)

with $1$ denoting the 2-dimensional unit matrix. As before, this is formally just an extension of the ordinary matrix inverse to an infinite number of dimensions. In the theory of differential equations, $\Gamma_\lambda$ is called a Green’s matrix for the differential operator $V_\lambda$. It allows one to compute particular solutions to the ordinary differential equation $\bar{V}_\lambda(t) y(t) = f(t)$ for arbitrary inhomogeneities $f$.

Going back to (3.22), it is clear that we can equally express the functional determinant there in terms of the inverse operator as

$$
e^{g(\tau, \lambda)} = (\text{Det} \, V_\lambda)^{-1/2} = (\text{Det} \, \Gamma_\lambda)^{1/2}.
$$

(3.49)
The CGF $g(\tau, \lambda)$ then becomes

$$
g(\tau, \lambda) = \frac{1}{2} \text{Tr} \ln \det \Gamma_{\lambda} \simeq \frac{\pi}{2} \int_0^\tau dt \ln \det \Gamma_{\lambda}(t, t),
$$

where the functional trace is just the sum of the “diagonal” entries of $\ln \det \Gamma_{\lambda}$.

**Constructing Green’s matrix.** The operator $V_{\lambda}$ in (3.21) consists of an ordinary component $\bar{V}_{\lambda}$ and two boundary components $V_{0,\lambda}$ and $V_{\tau,\lambda}$. To construct its inverse, we proceed in two steps: First, we compute Green’s matrix $\bar{\Gamma}_{\lambda}$ for the ordinary component, satisfying homogeneous boundary conditions, i.e.,

$$
\bar{V}_{\lambda}(t) \bar{\Gamma}_{\lambda}(t, t') = \delta(t - t') \quad \text{with} \quad \Gamma(0, t') = \Gamma(\tau, t') = 0.
$$

Second, we add to this a solution $\Gamma_{\lambda}^H$ of the associated homogeneous equation that fixes the boundary terms, i.e.,

$$
\bar{V}_{\lambda}(t) \Gamma_{\lambda}^H(t, t') = 0
$$

such that

$$
V_{0,\lambda}(t) \left[ \bar{\Gamma}_{\lambda}(t, t') + \Gamma_{\lambda}^H(t, t') \right] = 0 \quad \text{and} \quad V_{\tau,\lambda}(t) \left[ \bar{\Gamma}_{\lambda}(t, t') + \Gamma_{\lambda}^H(t, t') \right] = 0. \quad (3.53)
$$

For the ordinary component $\bar{\Gamma}_{\lambda}$, comprehensive instructions for its construction are provided in Appendix C. To connect to the notation used there, note that the operator $\bar{V}_{\lambda}$ may be written in the form

$$
\bar{V}_{\lambda}(t) = R \partial_t^2 + S \partial_t + T
$$

with

$$
R = -(2D)^{-1},
$$
$$
S = A^T(2D)^{-1} - (2D)^{-1} A + \lambda \left( Y - Y^T \right),
$$
$$
T = A^T(2D)^{-1} A \quad (3.55c)
$$

Following the construction procedure, we then find that $\bar{\Gamma}_{\lambda}$ can be written in the form (C.17) as

$$
\bar{\Gamma}_{\lambda}(t, t') = U_{\lambda}(t) \left[ \Theta(t' - t) C_{\lambda}^<(t', \tau) + \Theta(t - t') C_{\lambda}^>(t', \tau) \right],
$$

where $\Theta$ denotes the Heaviside step function, and the 16 components of the $(4 \times 2)$-matrices $C^<$ and $C^>$ are obtained as the solutions of the 16 linear equations (C.14)–(C.16), namely

$$
U_{\lambda}(0) C_{\lambda}^<(t', \tau) = 0, \quad (3.57a)
$$
$$
U_{\lambda}(\tau) C_{\lambda}^>(t', \tau) = 0, \quad (3.57b)
$$
$$
U_{\lambda}(t') \left[ C_{\lambda}^>(t', \tau) - C_{\lambda}^<(t', \tau) \right] = 0, \quad (3.57c)
$$
$$
\dot{U}_{\lambda}(t') \left[ C_{\lambda}^>(t', \tau) - C_{\lambda}^<(t', \tau) \right] = R^{-1}. \quad (3.57d)
Here, the \((2 \times 4)\)-matrix-valued function \(U_\lambda\) is a fundamental matrix of the operator \(\bar{V}_\lambda\), meaning that its columns form a complete set of fundamental solutions of the homogeneous problem \(\bar{V}_\lambda(t) u(t) = 0\). Its construction for the present case of a second-order operator with constant coefficients is also given in Appendix C, the solution taking the form (C.27).

The linear system of 16 equations (3.57) in fact decomposes into two systems of 8 equations each. Writing them out explicitly, we find

\[
\begin{bmatrix}
U_{11}^{(0)} & U_{12}^{(0)} & U_{13}^{(0)} & U_{14}^{(0)} & 0 & 0 & 0 & 0 \\
U_{21}^{(0)} & U_{22}^{(0)} & U_{23}^{(0)} & U_{24}^{(0)} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & U_{11}^{(\tau)} & U_{12}^{(\tau)} & U_{13}^{(\tau)} & U_{14}^{(\tau)} \\
0 & 0 & 0 & 0 & U_{21}^{(\tau)} & U_{22}^{(\tau)} & U_{23}^{(\tau)} & U_{24}^{(\tau)} \\
-U_{11}^{(\tau')} & -U_{12}^{(\tau')} & -U_{13}^{(\tau')} & -U_{14}^{(\tau')} & U_{11}^{(\tau)} & U_{12}^{(\tau)} & U_{13}^{(\tau)} & U_{14}^{(\tau)} \\
-U_{21}^{(\tau')} & -U_{22}^{(\tau')} & -U_{23}^{(\tau')} & -U_{24}^{(\tau')} & U_{21}^{(\tau)} & U_{22}^{(\tau)} & U_{23}^{(\tau)} & U_{24}^{(\tau)} \\
-\dot{U}_{11}^{(\tau')} & -\dot{U}_{12}^{(\tau')} & -\dot{U}_{13}^{(\tau')} & -\dot{U}_{14}^{(\tau')} & \dot{U}_{11}^{(\tau)} & \dot{U}_{12}^{(\tau)} & \dot{U}_{13}^{(\tau)} & \dot{U}_{14}^{(\tau)} \\
-\dot{U}_{21}^{(\tau')} & -\dot{U}_{22}^{(\tau')} & -\dot{U}_{23}^{(\tau')} & -\dot{U}_{24}^{(\tau')} & \dot{U}_{21}^{(\tau)} & \dot{U}_{22}^{(\tau)} & \dot{U}_{23}^{(\tau)} & \dot{U}_{24}^{(\tau)} \\
\end{bmatrix}
\begin{bmatrix}
C_{11}^< \\
C_{12}^< \\
C_{13}^< \\
C_{14}^< \\
C_{21}^< \\
C_{22}^< \\
C_{23}^< \\
C_{24}^< \\
\end{bmatrix} =
\begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
\end{bmatrix},
\tag{3.58}
\end{aligned}
\]

where \(* = 1, 2\) for the first and second system, respectively. For compactness of notation, we also moved the function argument of the fundamental matrix to the superscript and suppressed the \(\lambda\) dependence, \(i.e.\ U_{ij}^{(0)} = (U_\lambda)_{ij}(t)\) in the above coefficient matrix. Due to the complexity of the fundamental matrix and the high dimensionality of the system, it is not feasible to write down a general solution of these equations, and we were not able to solve the system symbolically in a reasonable amount of time. This is thus the point where we have to fall back to numerical methods despite the fact that the transformation itself is in principle exact.

Once the ordinary inverse \(\Gamma_\lambda\) has been obtained as (3.56) with \(C^<\) and \(C^>\) following from (3.58), we can proceed to the second step of our construction. The additional contribution \(\bar{\Gamma}_H\) to the total inverse \(\Gamma_\lambda\) is supposed to vanish when applying \(\bar{V}_\lambda\) to it according to (3.52). Therefore, we make an ansatz of the form

\[
\bar{\Gamma}_H(t,t') = U_\lambda(t) C_H(t',\tau)
\tag{3.59}
\]

with \(C_H(t',\tau)\) being a \((4 \times 2)\)-matrix to be determined by the boundary conditions (3.53). Since \(U_\lambda\) is a fundamental matrix, we have indeed \(\bar{V}_\lambda(t) \bar{\Gamma}_H(t,t') = 0\) as required. Imposing the boundary conditions (3.53) then leads to 8 equations for the 8 unknown coefficients of \(C_H\), namely

\[
\begin{bmatrix}
-R\dot{U}_\lambda(\tau) + \frac{S + S^T}{2} U_\lambda(\tau) \\
R\dot{U}_\lambda(0) + \left( \Sigma(\infty)^{-1} - \frac{S + S^T}{2} \right) U_\lambda(0)
\end{bmatrix}
\begin{bmatrix}
C_H^< \\
C_H^>
\end{bmatrix} =
\begin{bmatrix}
R\dot{U}_\lambda(\tau) C^> \\
-R\dot{U}_\lambda(0) C^<
\end{bmatrix},
\tag{3.60a}
\]

These equations can also be decomposed into two systems of 4 equations each. Writ-
ing
\[ M^{(0)} := R\tilde{U}_\lambda(0) + \left(\Sigma(\infty)^{-1} - \frac{S + S^T}{2}\right)U_\lambda(0), \]  
\[ M^{(\tau)} := -R\tilde{U}_\lambda(\tau) + \frac{S + S^T}{2}U_\lambda(\tau), \]
as well as
\[ E^{(0)} := -R\tilde{U}_\lambda(0)C^< \quad \text{and} \quad E^{(\tau)} := R\tilde{U}_\lambda(\tau)C^>, \]
we obtain
\[ \begin{bmatrix} M_{11}^{(\tau)} & M_{12}^{(\tau)} & M_{13}^{(\tau)} & M_{14}^{(\tau)} \\ M_{21}^{(\tau)} & M_{22}^{(\tau)} & M_{23}^{(\tau)} & M_{24}^{(\tau)} \\ M_{11}^{(0)} & M_{12}^{(0)} & M_{13}^{(0)} & M_{14}^{(0)} \\ M_{21}^{(0)} & M_{22}^{(0)} & M_{23}^{(0)} & M_{24}^{(0)} \end{bmatrix} \begin{bmatrix} C_{1s}^H \\ C_{2s}^H \\ C_{3s}^H \\ C_{4s}^H \end{bmatrix} = \begin{bmatrix} E_{1s}^{(\tau)} \\ E_{2s}^{(\tau)} \\ E_{1s}^{(0)} \\ E_{2s}^{(0)} \end{bmatrix} \]  
(3.64)
with \( * = 1,2 \) denoting the two systems again. It is equally impractical to express the solutions here explicitly, thus we dealt with this system numerically, too. All in all, we then find that Green’s matrix, the inverse of \( V_\lambda \), is given by
\[ \Gamma_\lambda(t,t') = U_\lambda(t) \left[ \Theta(t' - t) C_<^< (t', \tau) + \Theta(t - t') C_<^> (t', \tau) + C_>^H (t', \tau) \right] \]  
(3.65)
with \( C^< \) and \( C^> \) obtained from (3.58) and \( C^H \) from (3.64).

**Numerical evaluation of the CGF.** For the Brownian gyrator under study, we numerically computed Green’s matrix as described in the previous paragraph and then evaluated the functional trace according to (3.50). Remember that the operator \( V_{\lambda Q,\lambda W} \) of interest is obtained from (3.21) by replacing \( \lambda Y \rightarrow \lambda_Q Y_Q + \lambda_W Y_W \), where \( Y_Q \) and \( Y_W \) were given in (3.34). Consequently,
\[ g(\tau, \lambda_Q, \lambda_W) = \frac{\pi}{2} \int_0^\tau dt \ln \det \Gamma_{\lambda Q,\lambda W}(t, t) + \text{normalization}. \]  
(3.66)
Examples for various integration times \( \tau \) are shown in Fig. 11. Comparing to the result for the scaled CGF found using the closed trajectory approximation and shown in Fig. 9, we first notice that the above observed degeneracy of parallel contour lines seems to be confirmed. As \( \tau \) is increased, \( g(\tau, \lambda_Q, \lambda_W)/\tau \) is stretched along the axis of slope \( \lambda_W/\lambda_Q = 1/\bar{\eta} \) determined by the macroscopic efficiency \( \bar{\eta} \).

The white areas in Fig. 11 correspond to regions where the CGF becomes imaginary or the systems of linear equations (3.58) and (3.64) for the coefficient matrices become degenerate and unsolvable. This confirms the second structural element of the scaled CGF found in the closed trajectory approximation, namely the limited domain of convergence. However, the cutoffs appear slightly different now. There are still two cutoffs parallel to the slope of macroscopic efficiency, but there position is closer to the origin than in the previous solution. The other two cutoffs are now tilted with respect to the \( \lambda_W \)-axis whereas they appeared parallel to it before. Remember that in the closed trajectory approximation, these were obtained from the symmetric parts of the \( Y \)-matrices, i.e. the boundary terms alone. Since \( Y_W \) does not have a symmetric
Figure 11: Contour plot of the (scaled) CGF \( g(\tau, \lambda_Q, \lambda_W)/\tau \) for times (a) \( \tau = 10 \), (b) \( \tau = 30 \), (c) \( \tau = 100 \). As \( \tau \) is increased, the contour lines become more and more parallel, seemingly approaching the CGF of Fig. 9 found by the closed trajectory approximation. The cutoffs, however, are slightly deformed. Parameters are chosen according to set A (cf. Appendix D).

part, this cutoff could not depend on \( \lambda_W \) in this case. However, in the closed trajectory approximation we had to consider symmetric and antisymmetric contributions independently. Therefore, the tilted cutoffs experienced by the CGF as obtained from Green’s matrix can presumably be attributed to a mixing of symmetric and antisymmetric contributions. After all, they are manifestly not stochastically independent, such that modifications at this level were expected.

The situation is very similar for the second type of configurations (represented by parameters B, Appendix D). For this reason, we do not explicitly display the CGF resulting from the Green’s matrix method here, but merely note that there is a similar tendency for contour lines to become parallel as \( \tau \to \infty \) as well as a slightly sheared domain of convergence.

Efficiency LDF. With the joint CGF for work and heat at hand, we would like to proceed and compute the resulting efficiency LDF in order to compare it to the one obtained from the closed trajectory approximation. Unfortunately, the fact that we had to evaluate Green’s matrix numerically means that the limit \( \tau \to \infty \) is not directly accessible anymore. Rather, we have to compute finite-time estimates and extrapolate to infinity in the end. Therefore, we define

\[
J(\tau, \eta) := -\inf_{\lambda_W} \left[ \frac{1}{\tau} g(\tau, \eta \lambda_W, \lambda_W) \right]
\]

as a finite-time equivalent of (3.44). As \( \tau \to \infty \), we obtain \( J(\tau, \eta) \to J(\eta) \), which is the desired LDF. The dashed curves in Fig. 12 show examples for finite \( \tau \) for the two model configurations.

In order to perform the extrapolation to infinite \( \tau \), let us analyze the asymptotic behavior of \( J(\tau, \eta) \). Looking at the definition (3.67), we see that it is determined by the asymptotics of \( g(\tau, \lambda_Q, \lambda_W) \), which in turn is controlled by the behavior of the functions \( Q_{1\tau} \) and \( W_{\tau} \). For a generic functional of the form (3.17), we expect a linear growth with the integration time to leading order, which may be complemented by a subleading logarithmically growing term and constant contributions [35]. Therefore,
we make an ansatz of the form

\[ J(\tau, \eta) \simeq -\hat{a}(\eta) \frac{1}{\tau} + \hat{b}(\eta) \frac{\ln \tau}{\tau} + \hat{J}(\eta), \tag{3.68} \]

where \( \hat{a} \) models the constant contribution, \( \hat{b} \) the logarithmically growing part, and \( \hat{J} \) the linearly growing contribution, which is the only one remaining in the limit \( \tau \to \infty \) and will be our estimate for \( J(\eta) \).

Using three finite-time LDFs, we can then fit the model (3.68) to the numerically obtained functions and find the parameters \( \hat{a}(\eta), \hat{b}(\eta), \) and \( \hat{J}(\eta) \). The result is shown by the solid curves in Fig. 12. For comparison, we show two fits, the dark red one obtained from relatively small integration times and the purple one from larger times. The fits agree well in the region of “usual” efficiencies (\( \eta \approx 0...1 \)), but deviations are visible for the tails. For a trustworthy estimate we should, of course, choose the underlying \( \tau \) values as large as possible, meaning that we expect the second (purple) fits to be closer to the exact result.

We also remark that the obtained results appear to be in good qualitative agreement with the shapes obtained from the closed trajectory approximation (cf. Fig. 10). We will compare the two result in more detail in the next section, where we also supplement them with simulation data.

### 3.7 Comparison and Discussion

Both our methods to compute the efficiency LDF of the Brownian gyrator involved some sort of approximation. For the first one, presented in Sections 3.4 and 3.5, we restricted to closed trajectories and thereby effectively neglected boundary contributions. The second one, explained in Section 3.6, was in principle exact, but due to the complexity of the involved transformation, we were forced to evaluate the CGF of work and heat numerically and thus had to extrapolate to infinite integration times. Before
we compare the findings, we would like to introduce yet another estimate obtained from computer simulations.

**Simulation I.** To obtain an independent numerical estimate of the efficiency LDF, we simulated the system (1.4) starting from the steady state (2.6) and recorded samples of trajectories of various lengths $\tau$ along with the work $W$ and heat flow $Q_1$. From these, we computed an empirical distribution $\hat{p}_\tau(\eta)$ as a histogram of the associated trajectory efficiencies $\eta = -W/Q_1$. We already showed an example of such a distribution in Fig. 4.

To extract an estimate $\hat{J}(\eta)$ of the efficiency LDF, we followed a procedure similar to the one described in [35] and in analogy with the above extrapolation procedure used with the Green’s matrix method. We assume that $p_\tau(\eta)$ can be approximated as $p_\tau(\eta) = A(\eta)\tau^{-\hat{a}(\eta)}e^{-\tau\hat{J}(\eta)}$ for sufficiently large $\tau$. For each $\eta$, we then fit the empirical distribution $\hat{p}_\tau$ as a function of $\tau$ to the following model, similar to (3.68):

$$-rac{\ln \hat{p}_\tau(\eta)}{\tau} = -\hat{a}(\eta)\frac{1}{\tau} + \hat{b}(\eta)\ln \frac{\tau}{\tau} + \hat{J}(\eta),$$  

(3.69)

where $\hat{a}(\eta)$, $\hat{b}(\eta)$, and $\hat{J}(\eta)$ are the parameters to be adjusted such that the least-square residuals are minimized. This is a generalization of the method suggested in [35], where they use only three time points and adjust $\hat{a}(\eta)$, $\hat{b}(\eta)$ and $\hat{J}(\eta)$ such that the ansatz interpolates between them. At the end of the day, we are only interested in the value of $\hat{J}(\eta)$, which provides an extrapolation from the finite-$\tau$ data to $\tau \to \infty$.

In Fig. 13, we show the obtained fit along with the analytical approximations from the previous sections for parameter set A from Appendix D, corresponding to case (a) in the previous section and Figs. 10 and 12. Empirical distributions were obtained for 14 time points from $\tau = 5$ to $\tau = 77$ using between 5 and 32 million trajectories and time steps from $\Delta t = 3.5 \times 10^{-4}$ to $\Delta t = 5.5 \times 10^{-4}$ for each. The number of trajectories was gradually increased with $\tau$ in order to obtain reasonable coverage of the increasing rare event region for larger observation times. The time step was also carefully increased to reduce the total simulation time.

We notice that the simulation confirms the general shape of the LDF. The largest deviations occur in the tails and around the maximum at and around $\eta \simeq 0$, where the error bars are naturally largest, too, because efficiencies in this region are least likely to be observed. Remarkably, the tails found in the simulation are best matched by the closed trajectory approximation (orange curve), whereas the minimum likelihood region is better described by the Green’s matrix extrapolation (purple curve). We have to keep in mind, however, that the simulation data itself were obtained from an extrapolation to $\tau = \infty$, so that we assume them to suffer from similar finite-time effects as the Green’s matrix extrapolation. In fact, these effects are expected to be larger for the simulations because we extrapolate from noisy data and smaller times.

**Refined approximation.** Seeing as the CGF computed by the Green’s matrix method (cf. Fig. 11) seemed to converge to the result found by the closed trajectory approximation (cf. Fig. 9), we also tried to combine the “strengths” of both approximations. To this end, remember that the closed trajectory approximation took full account of the ordinary component of the Gaussian kernel operator $V_\lambda$ of the CGF and allowed an entirely analytic treatment. We were therefore able to compute the limit
Figure 13: Large deviation function of efficiency fluctuations for parameters A (cf. Appendix D). The solid curves show the analytically obtained approximations: Closed trajectory approximation (orange), Green’s matrix extrapolation from $\tau = 90, 95, 100$ (purple), and closed trajectory approximation with cutoffs taken from the Green’s matrix method (green). The data points represent the parameter estimate $J(\eta)$ from fitting (3.69) to empirical distributions for $p_\tau(\eta)$ obtained from sample trajectories with $\tau = 5, 10, \ldots, 70, 72, 77$. The macroscopic and Carnot efficiencies are also marked with dashed lines.
\( \tau \to \infty \) in an exact way, leading to the degenerate shape of the CGF with parallel contour lines. By the Green’s matrix method, on the other hand, we could include the boundary components of \( V_\lambda \), but lost the ability to take the limit \( \tau \to \infty \) analytically. All the above observations support the claim that the “bulk part” of the CGF has the degenerate form found from the closed trajectory approximation, whereas the boundary contributions control the domain of convergence and thus the region where the minimization (3.44) to extract the efficiency LDF should be carried out. Therefore, we suggest that this efficiency LDF can in fact be computed by minimizing the CGF obtained from the closed trajectory approximation in the domain found from the Green’s matrix method. The result is plotted in green along with the other approximations in Fig. 13.

We observe that the result obtained from this combination of closed trajectory and Green’s matrix methods in fact appears to agree less well with the simulation data than either approximation does individually. However, we would like to point out again that the simulation data are expected to show finite-time effects. Comparing the purple and red curves in Fig. 12a, we see that extrapolations from smaller times seem to underestimate the tails of the efficiency LDF. At the same time, they overshoot the plateau of minimum likelihood slightly at the left and right bounds. Correcting for these artifacts in the simulation result, one might indeed end up with the green curve in Fig. 13. Nonetheless, it should be mentioned that the two approximations of the CGF may possibly be incompatible in the sense that the interplay of bulk shape and domain of convergence is more subtle, which would invalidate the suggested combination, although this does not seem to be the case.

Simulation II. Let us come back to the simulation results and present our representative case for the second type of parameters leading to an efficiency LDF as in Figs. 10b and 12b, where we chose parameters \( B \) from Appendix D. Fig. 14 depicts a plot similar to the previous one for these parameters. The simulation extrapolation was performed by fitting the model (3.69) to 12 different time points between \( \tau = 21 \) and \( \tau = 76 \). For each value of \( \tau \), the empirical distribution \( \hat{p}_\tau \) was estimated from 50 million trajectories with a time step \( \Delta t = 2 \times 10^{-3} \).

As for the other parameter set, the general shapes of the simulation and our two analytic approximations agree reasonably well. In fact, there is almost no difference visible between the closed trajectory approximation and the Green’s matrix extrapolation. Interestingly, however, the combined approximation with the closed-trajectory CGF restricted to the Green’s matrix domain of convergence settles the plateau a bit lower than the two individual approximations. Considering Fig. 12b, we understand that the tails appear to be overestimated in this case when extrapolating from smaller times, so that the real plateau presumably lies below the one suggested by the simulation data.

Comparison to the universal theory and other works. Our interest in the efficiency LDF of the Brownian gyrator was largely caused by the discovery of a universal theory of efficiency fluctuations for finite-state machines by Verley and co-workers [10, 11]. We briefly summarized their findings already at the end of Section 3.3. In particular, they obtained a universal shape for the efficiency LDF, depicted in Fig. 5.

The Brownian gyrator’s efficiency LDF resembles this universal shape to some extent. First of all, we also find a unique minimum with \( J(\eta) = 0 \) at the macroscopic efficiency
Figure 14: Large deviation function of efficiency fluctuations for parameters B (cf. Appendix D). The solid curves show the analytically obtained approximations: Closed trajectory approximation (orange), Green’s matrix extrapolation from $\tau = 90, 95, 100$ (purple), and closed trajectory approximation with cutoffs taken from the Green’s matrix method (green). The data points represent the parameter estimate $J(\eta)$ from fitting (3.69) to empirical distributions for $p_\tau(\eta)$ obtained from sample trajectories with $\tau = 21, 26, \ldots, 76$. The dashed lines mark the macroscopic and Carnot efficiencies.
\( \eta = \bar{\eta} \), which becomes the only efficiency with finite probability in the limit \( \tau \to \infty \). Furthermore, the limits \( \eta \to \pm \infty \) are equal, which is a direct consequence of the construction procedure (Fig. 10).

For our first type of configurations (parameters A, Appendix D), we also find that \( J(\eta) \) becomes maximal at intermediate efficiencies. These regions of maximal \( J \) correspond to efficiencies that are least likely to be observed as \( \tau \to \infty \) in a large deviation sense. There exists, however, a crucial difference between our findings and the universal theory: The least likely efficiency of the Brownian gyrator is unequivocally not the Carnot efficiency. Instead, we observe a whole range of least likely efficiencies. Moreover, this range always includes the value \( \eta = 0 \) due to the degenerate shape of the joint CGF of work and heat.

This holds true for the second type of configurations (parameters B, Appendix D), too, but the deviations from the universal theory are even more striking here. The reason for this is that the minimum contour of the work/heat CGF intersects the line \( \lambda_W = 0 \) within the domain of convergence as opposed to outside of it in the case of parameters A (cf. Fig. 10). This implies that \( J(\eta) \) takes the maximum value at \( \eta = \pm \infty \), too, such that the plateau of least likely efficiencies extends across the whole line of \( \eta \)-values except for an interval around the macroscopic efficiency, as explained in Section 3.5. As before, the Carnot efficiency does not play any special role for the efficiency LDF.

Before we try to explain the observed deviations, we would like to summarize again how they come about mathematically. There are basically two reasons or peculiarities of the joint CGF of work and heat in the present case that lead to the curious differences. The reasoning behind the universal shape in Fig. 5 assumes the work/heat joint CGF to be smooth and convex in an infinite domain [11]. The second property of convexity is also obeyed by the Brownian gyrator CGF if one sets \( g(\tau, \lambda_Q, \lambda_W) = \infty \) outside of the domain of convergence. However, the cutoffs obviously render the function discontinuous. Moreover, the derivation of the universal theory in [11] does not consider the case of a degenerate CGF with parallel contour lines. Such a function is still convex, but minimizing along lines of constant efficiency leads to a degenerate LDF, too. It is the joint effect of these two discrepancies between our CGF and the assumptions made by Verley and co-workers that lead to a different picture for the efficiency LDF.

Deviations of this kind were also found in another recent work by Park et al. [36] who investigated the efficiency fluctuations of a similar model to ours. Instead of computing functional determinants, they used an eigenvalue expansion of the Fokker-Planck operator to extract the scaled CGF in the limit \( \tau \to \infty \). The structure they found is basically equal to our result. The scaled CGF is degenerate with parallel contour lines, but converges in a finite region of the \( \lambda_Q-\lambda_W \)-plane only. The shape reported for the resulting efficiency LDF is similar to our configuration B. However, Park et al. state that their analytic result describes simulations starting from fixed initial position, although they do not explicitly derive this proposition. Judging from the observations we made in this report, this is most likely a manifestation of the chosen cutoff, since boundary contributions and thus initial conditions merely affected the domain of convergence in all our findings.
Potential explanations for the deviations. In the just mentioned work [36], the authors also claim that the differences between their result and the universal theory are explained by the unboundedness of system energy and entropy. Indeed, for the derivation of their universal theory, Verley et al. explicitly assume that the system has a finite state space [11], such that the energy change and entropy production in the system are negligible compared to the total performed work and heat flows to the reservoirs as $\tau \to \infty$. Using the first law (2.12), the total entropy production (cf. Section 2.4) may be expressed as

$$\Delta S_{\text{tot}} = -\frac{Q_1}{T_1} - \frac{Q_2}{T_2} + \Delta S_{\text{sys}} = Q_1 \left( \frac{1}{T_2} - \frac{1}{T_1} \right) + \frac{W}{T_2} - \frac{\Delta U}{T_2} + \Delta S_{\text{sys}}, \quad (3.70)$$

where $\Delta U = U(x(\tau)) - U(x(0))$ denotes the change in system energy along the trajectory $x$. Heat and work are extensive quantities and grow linearly with $\tau$ (to leading order). But if both $\Delta U$ and $\Delta S_{\text{sys}}$ are bounded, as it is the case for systems with a finite state space, then these contributions become indeed negligible as $\tau \to \infty$ and the total entropy production is asymptotically given by the contributions from $Q_1$ and $W$. This is the assumption made in [10, 11], and from this they use the fluctuation theorem for the total entropy production to relate the maximum of the efficiency LDF to the Carnot efficiency for machines operating at a steady state.

Since the system energy and entropy are obviously unbounded for the Brownian gyrorator as well as the model studied in [36], this may possibly explain the observed discrepancies of least likely efficiencies that are not Carnot. This reasoning is further supported by other recent findings [16] where the authors study the influence of a third, non-negligible channel of entropy production, which could in our case model the last two terms in (3.70) that are disregarded in the universal theory. It was found there that general shape of the efficiency LDF remains similar to Fig. 5, but the maximum may now occur at an arbitrary efficiency $\eta_{\text{max}}$ that is unrelated to the Carnot efficiency. This result is obtained by introducing an additional “efficiency” that in our model relates the third entropy channel to the heat flow $Q_1$, such that the efficiency LDF becomes a function of two arguments. The quantity of interest, namely the LDF for the “actual” efficiency, is then a cut of this two-dimensional function obtained for typical configuration of the third channel [16].

This may indeed explain why the Carnot efficiency is not singled out as the least likely one in the efficiency LDF of the Brownian gyrorator. It would be interesting to check this assumption by means of an enhanced model where the energy and entropy fluctuations in the system become bounded, for example by introducing a wall that definitely confines the particle to a finite region of space. This model would presumably not be analytically tractable any longer, but can still be investigated by simulations.

What the extended theory of a hidden entropy channel does not account for, though, is the appearance of a whole plateau of maximum efficiencies experienced both in this report and in [36]. As we noted above, this plateau can be traced back to the degenerate form of the joint CGF of work and heat as well as its finite domain of convergence. The universal theory suggested in [10, 11] does simply not consider this situation, such that smoothness and non-degeneracy of the CGF (or likewise the joint LDF $I(q_1, w)$) become additional prerequisites for the universal theory to hold. Whether or not these assumptions are always fulfilled for systems with a finite state space is at least not obvious.
As a final remark, we should also remember that differentiability of the CGF was a condition for the Gärtner-Ellis theorem (cf. Section 3.3) to hold, which was the cornerstone of the relation between CGFs and LDFs. Exploiting its consequences notwithstanding may in principle be considered a mathematical negligence, though we believe that the result’s general soundness as well as the good agreement with independent numerical simulations more than justify this recklessness.

4 CONCLUSION

In this report, we investigated the thermodynamics and efficiency fluctuations of the Brownian gyrator, a microscopic heat engine consisting of a two-dimensional Brownian particle trapped by a parabolic potential and interacting with two heatbaths at different temperatures. Mathematically, the system is described by an Ornstein-Uhlenbeck process, a Markovian, Gaussian process satisfying a linear stochastic differential equation.

We derived the full dynamics of the particle in the form of its transition probabilities or propagator. A characteristic feature is the stationary state reached in the long-time limit, which exhibits a nonvanishing probability current. Physically, this is the result of the interplay of various asymmetries in the alignment of the system’s constituents. On the one hand, we have two reservoirs at different temperatures that couple to different directions in the plane and thus lead to anisotropic thermal fluctuations. On the other hand, the stiffness of the parabolic potential is anisotropic, too, and crucially its principal axes must not coincide with the irradiating directions of the heatbaths. The combination of these circumstances implies that thermal excitations and mechanical relaxations have different preferred directions, and the effect is that the particle spins around the origin on average. This way, it exerts a torque on the environment, which can in turn be used to perform work against an external force, so that the system converts heat extracted from one of the reservoirs into mechanical work.

We characterized the different operational regimes of this macroscopic heat engine and showed that, depending on the strength of the external “countertorque”, the system may work as an ordinary heat engine, a heat pump or refrigerator, or a dud engine which trivially heats the two reservoirs using the power provided from the external mechanical force. For the remainder of the report, we then focused on the ordinary heat engine regime, where the particle extracts heat from the hot reservoir to work against the external torque.

Inspired by a recently presented universal theory of the efficiency fluctuations of microscopic heat engines, we set off to study the large deviation function of the Brownian gyrator. This function characterizes the exponential decay of the probability to observe specific values for the heat engine’s efficiency in the long-time limit. The efficiency large deviation function can be obtained from the joint large deviation function of work and heat by a minimization along lines of constant efficiency. In order to compute this latter joint large deviation function, we exploited the Gärtner-Ellis theorem, a general result from large deviation theory that relates large deviation functions of observables to their cumulant generating functions.

For the calculation of the joint cumulant generating function of work and heat, we presented two different analytical methods, both of which were based on the path integral
description of Markovian Gaussian processes. These tools allowed us to express the cumulant generating function as the functional trace of a differential operator, which we then evaluated in two different ways. With the first method, we diagonalized the aforementioned operator by means of a Fourier expansion and expressed the functional trace as an infinite sum of its eigenvalues. This so-called Matsubara series was formally divergent, but we were able to resum it using a contour integral in the complex plane. The drawback of this procedure was that we could not incorporate contributions from the boundary because the Fourier expansion restricts the space of trajectories to periodic functions, i.e. closed paths. However, we were able to restore parts of these boundary terms by considering them separately and independently, which led to a reduced domain of convergence of the cumulant generating function obtained from the bulk contributions.

To understand the influence of the boundary terms better, we developed a second method to investigate the functional trace encoding the joint cumulant generating function of work and heat. Instead of the differential operator itself, we expressed the trace by the operator’s inverse or Green’s matrix. We worked out an exact procedure to construct this inverse, including all boundary contributions previously neglected. Unfortunately, this construction requires the solution of high-dimensional systems of linear equations, which we were unable to obtain in closed form. Rather, we had to fall back to numerical methods, so that we lost analytic tractability of the infinite-time limit needed to extract the large deviation function. Nevertheless, the numerical results agreed well with the closed trajectory approximation we investigated before. The only deviation, to be traced back to the boundary terms, was an expected modification of the cumulant generating function’s domain of convergence.

From the infinite time average of the cumulant generating function of work and heat, we then extracted the large deviation function of efficiency fluctuations. As mentioned above, this was achieved by minimizing the associated large deviation function of work and heat along lines of constant efficiency. The resulting efficiency large deviation functions from the two different calculation schemes were consistent and agreed with the results found from independent numerical simulations. The large deviation function exhibits similarities to the universal shape found in another work mentioned earlier, but also shows some differences. The predictions made for its values at the macroscopic efficiency as well as for the limits of infinite efficiency and the general shape held true in the Brownian gyrator case, too. However, compared to the universal shape, the efficiency large deviation function found in this report appears to be truncated, leading to a plateau of maximum values as opposed to a unique maximum predicted by the universal theory. Most importantly, this maximum region always lies around zero efficiency, whereas the universal theory claims the maximum to be reached at the Carnot efficiency for steady-state machines as the one under study.

To explain the observed deviations, we pointed out a few reasons why the universal theory does not cover systems like the Brownian gyrator and also linked our findings to other recent works. On the one hand, the universal theory was explicitly formulated for systems with a finite state space, which is definitely violated by the Brownian gyrator in the form studied here. The most important implication of this was the resulting boundedness of system energy and entropy, meaning that entropy fluctuations in the environment dominate in the long-time limit. It would therefore be interesting to see how the Brownian gyrator evolves under the presence of a confining wall that
restricts the potential energy and system entropy to finite values. On the other hand, the case of a finite domain of convergence and degenerate convexity of the cumulant generating function is not considered by the universal theory. This is, however, the situation encountered by us and others for Brownian-gyrator-like systems. To what extent this peculiarity is caused by the continuous state space of these systems is an open question.

All in all, the efficiency fluctuations of the Brownian gyrator exhibit an intriguing variety of features, some of which agree with predictions of previous works, others showing new behavior unseen before. A completely analytic derivation of the cumulant generating and efficiency large deviation functions would be worthwhile having, in particular to better understand the general structure and the influence of boundary contributions, although we do not expect any significant alteration of the final result. One of the most alluring properties of the Brownian gyrator is its comparably broad accessibility both theoretically and experimentally, which gives hope for more and exciting insights in the future.
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Ralf Eichhorn for giving me the opportunity for this internship and for his constant tremendous support during my time at NORDITA. I think I have learned a lot about stochastic processes and thermodynamics from an excellent teacher and researcher. I greatly appreciate his outstanding guidance and continuous approachability during my stay as well as the openness he would offer towards me and all other students and colleagues in discussions, talks, and breaks, which created an extraordinarily pleasant working atmosphere as well.

Besides my advisor, I would like to give my special thanks to Stefano Bo, with whom I would work on this and other projects, and who made me feel particularly welcome at NORDITA and in Stockholm both at work and afterwards. He, too, would offer a lot to learn from, and I am grateful for extensive and illuminating scientific discussions and clarifications, as well as all those less scientific chats and activities.

I also want to thank my office mates, who contributed to a working environment with a very decent mix of focus and distraction. Christoph was always open for delight- and insightful conversations on physics or the world in general, and I am also happy about the times we spent outside the institute. Raffaele provided me with some useful tips as well, and I also greatly enjoyed the shared lunch breaks. Eduardo’s and Victor’s stays here overlapped with mine for a shorter period only, but I am glad about the common time nonetheless. Moreover, I would like to thank Erik Aurell for his hospitality and some interesting ideas in the context of path integrals.

For their integration and generosity in general and the Monday fika in particular, I am also thankful to all people at NORDITA. I enjoyed my time here at the institute a lot.

Likewise, I appreciate the financial support from the Stiftung der Deutschen Wirtschaft and the Erasmus+ program of the European Union, without which this experience would not have been possible, as well as the facilities and support provided by NORDITA itself.

All the more, I would like to thank my family and friends back home, who encouraged my plans for this internship and accompanied it from a distance (and partially from close by, too). Finally, I also thank the friends I made here for entertaining evening/weekend activities and travel adventures.

Hej då osthyvlar!
A SOLUTION OF THE GENERAL ORNSTEIN-UHLENBECK SDE

In this appendix, we will present a derivation of the transition probabilities $p(t, x \mid t_0, x_0)$ and steady-state distribution $p_\infty(x)$ of an arbitrary $d$-dimensional Ornstein-Uhlenbeck process $x(t)$. It is characterized by the SDE

$$d x(t) = - A x(t) \, dt + B \, d\omega(t), \quad (A.1)$$

where $\omega(t)$ is a $d$-dimensional standard Brownian motion and $A$ and $B$ are constant $(d \times d)$-matrices.

Fokker-Planck equation. The SDE (A.1) is equivalent to a Fokker-Planck equation [7, 9] for the transition probabilities,

$$\frac{\partial}{\partial t} p(t, x \mid t_0, x_0) = \sum_{i,j} A_{ij} \frac{\partial}{\partial x_i} [x_j p(t, x \mid t_0, x_0)] + D_{ij} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} p(t, x \mid t_0, x_0), \quad (A.2)$$

where $D = \frac{1}{2} B B^T$. This partial differential equation (PDE) is of first order in time and second order in space. It has to be solved together with the initial condition

$$p(t_0, x \mid t_0, x_0) = \delta(x - x_0) \quad (A.3)$$

and boundary conditions reflecting the geometry of the state space. Here we will assume natural boundary conditions, meaning that the propagator vanishes at infinity, $p(t, x \mid t_0, x_0) \to 0$ as $\|x - x_0\| \to \infty$. To arrive at the general solution, we Fourier transform the propagator with respect to $x$:

$$p(t, x \mid t_0, x_0) = \int d^n k \, \bar{p}(t, k \mid t_0, x_0) e^{i k^T x}. \quad (A.4)$$

Substituting this into (A.2) leads to an equivalent PDE for the Fourier-transformed propagator,

$$\frac{\partial}{\partial t} \bar{p}(t, k \mid t_0, x_0) = - \sum_{i,j} A_{ij} k_i \frac{\partial}{\partial k_j} \bar{p}(t, k \mid t_0, x_0) - D_{ij} k_i k_j \bar{p}(t, k \mid t_0, x_0),$$

which is of first order in both $t$ and $k$. We may express it in matrix notation as

$$\frac{\partial}{\partial t} \bar{p} = - \left[ k^T A \partial_k + k^T D k \right] \bar{p}. \quad (A.5)$$

Gaussian ansatz. Along the lines of [9], we reduce the complexity by means of a Gaussian ansatz

$$\bar{p}(t, k \mid t_0, x_0) = e^{-\frac{1}{2} k^T \Sigma(t) k - i k^T \mu(t)} \quad (A.6)$$

with the symmetric matrix $\Sigma(t)$ and the vector $\mu(t)$ to be determined. Plugging this into (A.5), we find

$$(\text{LHS}) = - \left[ \frac{1}{2} k^T \dot{\Sigma} + i k^T \dot{\mu} \right] \bar{p},$$

$$(\text{RHS}) = - \left[ k^T A (-\Sigma k - i \mu) + k^T D k \right] \bar{p}.$$
Comparing coefficients on both sides, we obtain two ODEs for \( \mu \) and \( \Sigma \), namely
\[
\dot{\mu} = -A \mu, \quad (A.7a)
\]
\[
\dot{\Sigma} = -A \Sigma - \Sigma A^T + 2D. \quad (A.7b)
\]
These equations have to be solved for the initial conditions inferred from the initial distribution (A.3) of \( p \):
\[
\bar{p}(t_0, k | t_0, x_0) = e^{-i k^T x_0}.
\]
W.l.o.g., we may assume \( t_0 = 0 \), so that the initial conditions
\[
\mu(0) = x_0, \quad \Sigma(0) = 0 \quad (A.8)
\]
are implied. Equation (A.7a) for \( \mu \) can then be integrated immediately, resulting in
\[
\mu(t) = e^{-A t} x_0. \quad (A.9)
\]
In order to solve (A.7b) for \( \Sigma \), we first consider the homogeneous equation (\( D = 0 \)). Its general solution is given by
\[
\Sigma^{\text{hom}}(t) = e^{-t A} C e^{-t A^T}, \quad (A.10)
\]
where \( C \) is a constant matrix, as one can easily verify by differentiation. (The matrices \( E^{(\alpha,\beta)} \) with \( E_{ij}^{(\alpha,\beta)} = (e^{-t A})_{i\alpha} \delta_{\alpha\beta} (e^{-t A^T})_{\beta j} \) are linearly independent.) The full solution to (A.7b) is then obtained by adding a particular solution \( \Sigma^{\text{part}} \) of the inhomogeneous equation. Assuming \( \Sigma^{\text{part}} \) to be constant, we obtain a linear equation
\[
A \Sigma^{\text{part}} + \Sigma^{\text{part}} A^T = 2D. \quad (A.11)
\]
The solution can be found with the aid of computer algebra software: the solution for the two-dimensional overdamped Brownian gyrator is given by (2.5). The general solution to (A.7b) is then given by
\[
\Sigma(t) = \Sigma^{\text{hom}}(t) + \Sigma^{\text{part}}. \quad (A.12)
\]
\textbf{Transition probabilities.} To arrive at the propagator \( p(t, x | 0, x_0) \), we need to compute the inverse Fourier transform of our ansatz (A.6). After normalization, we finally obtain
\[
p(t, x | 0, x_0) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma(t)}} e^{-\frac{1}{2} [x - \mu(t)]^T \Sigma^{-1}(t) [x - \mu(t)]}, \quad (A.13)
\]
\footnote{Note that the antisymmetric part of \( A \Sigma \) drops out in the bilinear form \( k^T A \Sigma k \). Any matrix \( M \) can be decomposed into a symmetric component \( M_+ \) and an antisymmetric component \( M_- \) according to \( M_{\pm} = \frac{1}{2} (M \pm M^T) \). When sandwiched with the same vector \( k \), however, we find for the antisymmetric part:
\[
k^T M_- k = (k^T M_- k)^T = -k^T M_- k,
\]
thus implying that it vanishes.}
with \( \mu(t) \) given by (A.9) and \( \Sigma(t) \) given by (A.12) and (A.11).

**Steady state.** The system relaxes to a steady state in the limit \( t \to \infty \) if and only if the matrix \( A \) is positive definite. In this case, we have

\[
\mu(\infty) = 0, \quad \Sigma(\infty) = \Sigma_{\text{part}},
\]

which yields a steady state PDF

\[
p_\infty(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma(\infty)}} e^{-\frac{1}{2} x^T \Sigma(\infty)^{-1} x}.
\]

**B DERIVATION OF THE PATH WEIGHT**

The fundamental idea of the path integral is to assign a probability density \( p[x] \) on the space of trajectories \( P = \{x : x = [x(t)]_{t=0}^T, \text{ s.t. } \text{d}x(t) = -Ax(t)\text{d}t + Bd\omega(t)\} \) satisfying the equations of motion. The expected value of a stochastic function \( Y_\tau \) can then be computed as a sum over all paths weighted by the density \( p \),

\[
\langle Y_\tau \rangle = \sum_{x \in P} p[x] Y_\tau [x].
\]

However, this sum over paths is not well-defined naively, and we need a way to actually characterize all paths in a systematic manner.

**Discretization of trajectories.** The key concept to do this is discretization. Defining a set of discretized paths

\[
P_N := \{(x_0, \ldots, x_N) : x_k = x \left( k \frac{T}{N} \right), \quad x \in P \},
\]

the claim is that \( P_N \to P \) as \( N \to \infty \). If this convergence is assured, the definition of the path weight (3.1) is a logical consequence,

\[
p[x] = \lim_{N \to \infty} \prod_{k=0}^{N-1} p(t_{k+1}, x_{k+1} | t_k, x_k) p_0(x_0).
\]

Note, however, that this is only valid for Markovian processes where a joint probability of \( N \) points can be written as a chain of propagators connecting successive points. To define the second ingredient in (B.1), the sum over paths, we postulate that the sets \( P_N \) can be expressed equivalently as

\[
P_N := \{(x_0, \ldots, x_N) : x_k \in \mathbb{R}^d \},
\]

thus requiring that given an arbitrary number of points, there always exists a trajectory connecting them. The sum over paths can then be written as the limit of sums over discretized paths, \( i.e. \)

\[
\sum_{x \in P} \cdots \equiv \lim_{N \to \infty} \int \prod_{k=0}^N d^d x_k \cdots
\]
Infinitesimal transitions. To complete the picture, our remaining job is to evaluate the path weight \((B.3)\), whose core component is the propagator for infinitesimal transitions of duration \(\Delta t = \tau/N\) as \(N \to \infty\). Crucially, we do not need the actual solution \((A.13)\) to compute this quantity. The infinitesimal character of the transition probabilities allows us to infer them directly from the corresponding stochastic differential equation \((A.1)\). For this reason, the path integral is an independent and comprehensive approach to problems posed by SDEs, similarly to Itô calculus and Fokker-Planck equations.

The time-discretized version of \((A.1)\) is

\[
x_{k+1} - x_k = -A \frac{1}{2} (x_k + x_{k+1}) \Delta t + B g_k \sqrt{\Delta t}.
\]

(B.6)

Two remarks regarding the choice of discretization are in order. First, notice that we discretized \(x(t)\) in the drift term according to the mid-point rule because stochastic products in the path integral formalism are to be interpreted in the Stratonovich sense \([7, 27]\). In the given case, this is a technical peculiarity of minor importance because it will not make a difference in the end. Nevertheless, a similar discretization would have to be applied to the noise term if it depended on \(x\), and this would give a crucial contribution to the resulting path weight. Second, we recall that the variance \(\langle \omega(t)^2 \rangle\) of the Wiener process is of order \(t\) \([7, 29]\), which is the reason that the noise increments in \((B.6)\) enter at order \(\sqrt{\Delta t}\).

The Wiener increments \(g_k\) are simply independent, standard-Gaussian distributed random variables with probability density

\[
p_{g_k}(g) = \frac{e^{-g^2/2}}{(2\pi)^{d/2}}.
\]

(B.7)

Given a certain noise history \((g_0, g_1, \ldots)\) and the initial condition \(x_0\), the trajectory \((x_0, x_1, \ldots)\) is entirely determined by the equation of motion \((B.6)\). In particular, the conditioned transition probability \(p(x_{k+1} \mid x_k, g_k)\) to find the particle at \(x_{k+1}\) in the next step if it was at \(x_k\) in the previous step and experiences a random force \(g_k\) is deterministic:

\[
p(t_{k+1}, x_{k+1} \mid t_k, x_k, g_k) = \delta \left[x_{k+1} - x_k + A \frac{1}{2} (x_k + x_{k+1}) \Delta t - B g_k \sqrt{\Delta t}\right].
\]

(B.8)

This discrete conditional propagator can be rewritten in exponential form by expressing the \(\delta\)-distribution as a Fourier integral \([49]\):

\[
p(t_{k+1}, x_{k+1} \mid t_k, x_k, g_k) = \int \frac{d^d k}{(2\pi)^d} e^{-i k^T \left[x_{k+1} - x_k + A \frac{1}{2} (x_k + x_{k+1}) \Delta t - B g_k \sqrt{\Delta t}\right]}.
\]

The unconditional propagator for infinitesimal time steps is obtained by averaging over the white noise distribution \((B.7)\), so that we obtain

\[
p(t_{k+1}, x_{k+1} \mid t_k, x_k) = \int \frac{d^d g_k}{(2\pi)^d} p(x_{k+1} \mid x_k, g_k) p_{g_k}(g_k)
\]

\[
= \int \frac{d^d k}{(2\pi)^d} e^{-\frac{\Delta t}{2} k^T B B^T k - ik^T \left[x_{k+1} - x_k + A \frac{1}{2} (x_k + x_{k+1})\right]}.
\]
Evaluating the remaining integral over \( k \), this yields

\[
p(t_{k+1}, x_{k+1} | t_k, x_k) = \left[ \frac{\det(2D)}{(2\pi \Delta t)^d} \right]^{1/2} e^{-\frac{\Delta t}{2} \left[ \frac{x_{k+1}-x_k}{\Delta t} + A \frac{1}{2} (x_k+x_{k+1}) \right]^T (2D)^{-1} \left[ \frac{x_{k+1}-x_k}{\Delta t} + A \frac{1}{2} (x_k+x_{k+1}) \right]}
\]

with \( 2D = BB^T \) as usual.

**Chain of propagators.** Composing the propagators along the trajectory according to (B.3), we thus find

\[
\prod_{k=0}^{N-1} p(t_{k+1}, x_{k+1} | t_k, x_k) \propto e^{-\frac{\Delta t}{2} \sum_{k=0}^{N-1} \left[ \frac{x_{k+1}-x_k}{\Delta t} + A \frac{1}{2} (x_k+x_{k+1}) \right]^T (2D)^{-1} \left[ \frac{x_{k+1}-x_k}{\Delta t} + A \frac{1}{2} (x_k+x_{k+1}) \right]},
\]

where the omitted normalization is just the \( N \)-fold product of the prefactor in (B.9). In the limit \( N \to \infty \) and thus \( \Delta t \to 0 \), this becomes the path weight conditioned on the initial position,

\[
p(x | x(0) = x_0) \propto e^{-\frac{1}{2} \int_0^\tau dt \langle \dot{x}(t) + Ax(t) \rangle^T (2D)^{-1} \langle \dot{x}(t) + Ax(t) \rangle},
\]

as given in (3.4). Note that the normalization factor we left out admittedly diverges in this limit, but so does the sum over paths (B.5), and since normalization is assured in every time step according to (B.9), these infinities cancel when evaluating expectation values.

The path weight (B.12) was obtained for fixed initial conditions, meaning that the path integration (B.5) should not include the \( k = 0 \) integral. To obtain the path weight for arbitrary initial conditions, one simply multiplies by the initial distribution \( p_0 \) as in (B.3), i.e. \( p(x | x(0) = x_0) = p(x | x(0) = x_0) p_0(x_0) \).

**C CONSTRUCTION OF GREEN’S MATRICES FOR SYSTEMS OF SECOND ORDER ODES**

The method of Green’s functions is a well-established technique for studying differential equations in mathematics and physics [50–52]. In this appendix, we will deal with the boundary value problem (BVP)

\[
R\ddot{y} + S\dot{y} + Ty = f, \quad y(0) = y_0, \quad y(\tau) = y_\tau,
\]

where \( y, f : [0, \tau] \to \mathbb{R}^d \) are vector-valued functions, \( R, S, T \in \mathbb{R}^{d \times d} \) are constant matrices, and the dots indicate derivatives with respect to the function’s argument. We are looking for a solution \( y \) of this equation for arbitrary inhomogeneities \( f \) and require that \( y \) take the values \( y_0 \) and \( y_\tau \) on the boundary of the interval \([0, \tau]\) as specified in (C.1). The method suggested here can be extended straightforwardly to more general boundary conditions involving also derivatives, as we will see below. In its one-dimensional version, this problem is a textbook example [50, 51] and even the
more general case of non-constant coefficients is well-understood. Here we will present a generalization to higher dimensions that is much less well known in the literature.

Introducing a matrix-valued operator
\[ \hat{L}(t) := R \partial_t^2 + S \partial_t + T, \]
we may write the above system of second-order ordinary differential equations (ODEs) as
\[ \hat{L}(t) y(t) = f(t). \]

In order to solve this problem, we will construct Green’s matrix of the differential operator. This will allow us to compute the solution for arbitrary inhomogeneities as a convolution of Green’s matrix with the respective inhomogeneity. Adding a solution of the homogeneous problem to fix the correct boundary values will lead to the full solution of (C.1).

**Definition of Green’s matrix.** Green’s matrix \( \Gamma \) is defined as the inverse of the differential operator (C.2) in the sense that
\[ \hat{L}(t) \Gamma(t,t') = \delta(t - t') \quad \text{with} \quad \Gamma(0,t') = \Gamma(\tau,t') = 0, \]
where 0 and 1 denote the d-dimensional zero and unit matrices, respectively. Note that we require that Green’s matrix satisfy homogeneous boundary conditions. The defining equation (C.4) may also be expressed in components as
\[ \sum_j \hat{L}_{ij}(t) \Gamma_{jk}(t,t') = \delta_{ik} \delta(t - t'). \]

In its role as an operator inverse, Green’s matrices are not only auxiliary functions to solve systems of ODEs, but also acquire a valuable meaning on their own, e.g. in the theory of path integrals as in Sections 3.4–3.6 of this report. Therefore, the construction of Green’s matrices is much more than just a mathematical tool.

Assuming that we know Green’s matrix defined by (C.4), we can construct the solution to the problem (C.1) in two steps. First, we define a particular solution
\[ y_P(t) := \int_0^\tau dt' \Gamma(t,t')f(t'). \]
Since Green’s matrix satisfies (C.4), this function \( y_P \) fulfills equation (C.3) as can be checked by substitution. On the other hand, we immediately see that \( y_P(0) = y_P(\tau) = 0 \) because \( \Gamma \) vanishes on the boundary. Hence we obtain the full solution to the BVP (C.1) by adding a solution \( y_H \) of the homogeneous problem (\( f = 0 \)) associated with \( \hat{L} \), i.e.
\[ R\ddot{y}_H + S\dot{y}_H + Ty_H = 0, \quad y_H(0) = y_0, \quad y_H(\tau) = y_\tau, \]
This implies that \( y(t) := y_P(t) + y_H(t) \) is the solution of (C.1). The benefit of this “detour” using Green’s matrix is that we can use the same homogenous solution \( y_H \) for different inhomogeneities \( f \); all one has to do is compute a new particular solution
every time. Moreover, once the general solution of the homogeneous problem is known, Green’s matrix can be constructed from it rather directly.

Regarding a more general set of boundary conditions, we remark that one can of course impose other constraints for the values of \( y \) and its derivative at \( t = 0 \) and \( t = \tau \). The general method is still applicable in this case: We choose Green’s matrix to satisfy the associated homogeneous boundary conditions in its first argument, compute a particular solution \( y_P \) from it, and add a solution \( y_H \) of the homogeneous ODE that fulfills the desired boundary conditions.

**Definition of the fundamental matrix.** Before we delve into the construction of Green’s matrix, we will take a closer look at the homogenous problem (C.6). The construction procedure for Green’s matrix suggested here makes heavy use of the so-called fundamental matrix \( U \) of the operator \( \hat{L} \), which encodes the most general solution of the homogeneous problem associated with \( \hat{L} \). Looking at this differential operator, we note that it is linear, \( d \)-dimensional, and of second order. The general theory of linear ODEs [52, 53] thus implies that there exist \( 2d \) linearly independent solutions of the homogeneous problem

\[
\hat{L}(t) u(t) = 0.
\]  

(C.7)

Let \( u^{(1)}, \ldots, u^{(2d)} \) denote such a complete set of independent solutions. The matrix

\[
U(t) := \begin{pmatrix}
u_1^{(1)}(t) & u_1^{(2)}(t) & \cdots & u_1^{(2d)}(t) \\
\vdots & \vdots & & \vdots \\
u_d^{(1)}(t) & u_d^{(2)}(t) & \cdots & u_d^{(2d)}(t)
\end{pmatrix},
\]  

(C.8)

whose column vectors are these fundamental solutions of (C.7), is then called a fundamental matrix of \( \hat{L} \). It has dimensions \( d \times 2d \). Obviously, \( U \) itself satisfies the homogeneous problem (C.7) in its matrix form, i.e. \( \hat{L}(t) U(t) = 0 \). More importantly, the general solution of (C.7) is given by \( U(t) c \), where \( c = (c_1 \cdots c_{2d})^T \) is a constant vector, since this generates the most general linear combination of the fundamental solutions \( u^{(k)} \). Hence this provides the second ingredient of our solution to (C.1): We can write \( y_H(t) = U(t) c \) and determine the coefficients \( c_1, \ldots, c_{2d} \) using the \( 2d \) boundary conditions (C.6).

Similarly, the general solution of the associated matrix problem is \( U(t) C \) with a constant matrix \( C \in \mathbb{R}^{2d \times d} \). We will use this below because it provides a starting point for the construction of Green’s matrix, which we still need in order to compute the first ingredient of the solution to (C.1), the particular solution \( y_p \). Before we actually compute the fundamental matrix associated with \( \hat{L} \), we will lay out our method to calculate \( \Gamma \).

**Properties of Green’s matrix.** To find a construction procedure for \( \Gamma \), let us collect some of the properties following from its definition (C.4). These observations are very similar to the one-dimensional case [51]. First, we notice that the \( \delta \)-inhomogeneity effectively means that \( \Gamma \) satisfies the homogeneous equation everywhere except at \( t = t' \):

\[
\hat{L}(t) \Gamma(t,t') = 0 \quad (t \neq t').
\]  

(C.9)
Second, we remember that integration is a smoothening operation in the sense that the integral of a \(k\)-times differentiable function is itself \((k+1)\)-times differentiable. Therefore, the \(\delta\)-discontinuity on the right-hand side of (C.4) must manifest itself in the highest derivative appearing on the left-hand side, such that

\[
\sum_j R_{ij} \partial_t^2 \Gamma_j(t, t') \sim \delta_{ik} \delta(t - t') \quad \text{as} \quad t \to t'.
\]

Consequently, the components of the first derivative of \(\Gamma\) must have steps at \(t = t'\) whose sizes are determined by \(R^{-1}\):

\[
\partial_t \Gamma_{ij}(t, t') \big|_{t = t' +} - \partial_t \Gamma_{ij}(t, t') \big|_{t = t' -} = (R^{-1})_{ij}.
\]

Going one order higher we observe that the integral of a step function is continuous, meaning that

\[
\Gamma_{ij}(t, t') \big|_{t = t' +} - \Gamma_{ij}(t, t') \big|_{t = t' -} = 0.
\]

Finally, we also note down the boundary conditions again because they complete the set of defining properties of \(\Gamma\):

\[
\Gamma_{ij}(0, t') = \Gamma_{ij}(\tau, t') = 0 \quad \text{for all} \quad t' \in [0, \tau].
\]

The procedure to construct Green’s matrix, which we will work out in detail in the next paragraph, then uses the properties collected here to determine \(\Gamma\) uniquely. The basic idea is to split the interval \([0, \tau]\) at \(t'\). According to (C.9), \(\Gamma\) must obey the homogeneous equation in both subintervals. Therefore, we can make an ansatz of the form of the general solution \(U(t) C\) in both these subintervals, enforce the respective boundary conditions (C.12), and connect the two solutions via the continuity and derivative-jump conditions (C.11) and (C.10).

**Construction procedure.** Let us work out the steps sketched in the last paragraph in more detail. As we will see, the computation of Green’s matrix is thereby reduced to solving a system of \(4d^2\) linear, algebraic equations for the unknown coefficients of the general matrix ansatz.

— 1: Find the fundamental matrix (C.8) whose columns are \(2d\) linearly independent solutions of the homogeneous problem (C.7). We will explain how this is achieved in the next paragraph.

— 2: Decompose \(\Gamma\) into \(\Gamma^<(t, t')\) if \(t < t'\) and \(\Gamma^>(t, t')\) if \(t > t'\). According to (C.9), both components satisfy the homogeneous equation, so that we can write

\[
\Gamma^<(t, t') = U(t) C^< \quad \text{and} \quad \Gamma^>(t, t') = U(t) C^>.
\]

The matrices \(C^<\) and \(C^>\) will depend on \(t'\) (and \(\tau\)), but not on \(t\). Their coefficients leave us with a total of \(4d^2\) unknowns to be determined by conditions (C.10)–(C.12).

— 3: Solve for \(d^2\) components of each of the matrices \(C^<\) and \(C^>\) by imposing the
boundary conditions (C.12):

\[ U(0) C^\less = 0, \quad (C.14a) \]
\[ U(\tau) C^\greater = 0. \quad (C.14b) \]

— 4: Connect the two solutions at \( t = t' \) by means of the continuity condition (C.11),

\[ U(t') [C^\greater - C^\less] = 0, \quad (C.15) \]

and the derivative jump condition (C.11),

\[ \dot{U}(t') [C^\greater - C^\less] = R^{-1}. \quad (C.16) \]

Note that (C.15) does not imply \( C^\less = C^\greater \) because \( U(t) C \) is a \((d \times d)\)-matrix, such that (C.15) constrains only \( d^2 \) components of \( C^\less \) and \( C^\greater \) together. Similarly, (C.16) provides another \( d^2 \) constraints. Together with (C.14), we thus found \( 4d^2 \) equations fixing the \( 4d^2 \) unknowns of \( C^\less \) and \( C^\greater \).

— 5: Green’s matrix \( \Gamma \) for the differential operator \( \hat{L} \) thus reads

\[ \Gamma(t,t') = U(t) \left[ \Theta(t'-t) C^\less + \Theta(t-t') C^\greater \right], \quad (C.17) \]

where \( \Theta \) denotes the Heaviside step function. With this result, the only remaining task is to find a fundamental matrix of \( \hat{L} \).

**Computing the fundamental matrix.** The fundamental matrix consists of a complete set of linearly independent solutions of the homogeneous problem (C.7). Dividing by \( R \), we rewrite this in the form

\[ \ddot{u} + \dot{S} \dot{u} + \dot{T} u = 0 \quad \text{with} \quad \dot{S} = R^{-1} S, \quad \dot{T} = R^{-1} T. \quad (C.18) \]

As noted before, this is a system of second-order homogeneous ODEs. With the definition \( v(t) := \dot{u}(t) \) we convert it into an equivalent system of first order:

\[
\begin{pmatrix}
\dot{u} \\
\dot{v}
\end{pmatrix}
= M
\begin{pmatrix}
u \\
v
\end{pmatrix},
\]

where

\[
M = \begin{pmatrix}
0 & 1 \\
-T & -S
\end{pmatrix}.
\]

For now, let us assume that the so-defined matrix \( M \) is diagonalizable; we will discuss the more general case below. If \( M \) is diagonalizable, there exists a complete set of eigenvectors \( \omega^{(1)}, \ldots, \omega^{(2d)} \), whose corresponding eigenvalues we denote by \( \kappa^{(1)}, \ldots, \kappa^{(2d)} \). Furthermore, let \( \Omega = \begin{pmatrix} \omega^{(1)} & \cdots & \omega^{(2d)} \end{pmatrix} \) be the matrix of eigenvectors, such that

\[ \Omega^{-1} M \Omega = \begin{pmatrix}
\kappa^{(1)} & & \\
& \ddots & \\
& & \kappa^{(2d)}
\end{pmatrix}. \quad (C.20) \]
Multiplying (C.18) by $\Omega^{-1}$ from the left and inserting an identity in the form of $1 = \Omega\Omega^{-1}$ on the right-hand side leads to

$$
\Omega^{-1} \begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} \kappa^{(1)} \\ \vdots \\ \kappa^{(2d)} \end{pmatrix} \Omega^{-1} \begin{pmatrix} u \\ v \end{pmatrix}.
$$

(C.21)

This ODE system is completely decoupled and can be integrated immediately, yielding a general solution

$$
\Omega^{-1} \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} e^{\kappa^{(1)}t} \\ \vdots \\ e^{\kappa^{(2d)}t} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_{2d} \end{pmatrix}
$$

(C.22)

with constants $c_1, \ldots, c_{2d}$. The columns of the matrix on the right-hand side thus provide $2d$ linearly independent solutions of the transformed system. Multiplying by $\Omega$ from the left, we obtain the fundamental solutions of the original first-order system,

$$
\begin{pmatrix} u^{(\alpha)}(t) \\ v^{(\alpha)}(t) \end{pmatrix} = e^{\kappa^{(\alpha)}t} \omega^{(\alpha)} \quad (\alpha = 1, \ldots, 2d),
$$

(C.23)

where $\omega^{(\alpha)}$ are the eigenvectors of $M$ as before. The vectors $v^{(\alpha)}$ are in fact redundant since we know that $v^{(\alpha)} = \dot{u}^{(\alpha)}$ by definition. For the fundamental matrix $U$ of the second-order system (C.7), we only need the $u^{(\alpha)}$ vectors. Arranging them as the columns of $U$ like in (C.8) yields the desired fundamental matrix. In compact form, we can write this as

$$
\begin{pmatrix} U(t) \\ V(t) \end{pmatrix} = \Omega \begin{pmatrix} e^{\kappa^{(1)}t} \\ \vdots \\ e^{\kappa^{(2d)}t} \end{pmatrix},
$$

(C.24)

meaning that $U(t)$ consists of the first $d$ rows of the matrix on the right-hand side.

Let us return to the case that the coefficient matrix $M$ of the first-order system (C.18) is not diagonalizable. In this case, one can still find a similarity transformation that brings $M$ into Jordan normal form [54], i.e.

$$
\tilde{M} = \Omega^{-1} M \Omega = \begin{pmatrix} M_{\kappa^{(1)}} \\ \vdots \\ M_{\kappa^{(r)}} \end{pmatrix},
$$

(C.25)

where $M_{\kappa^{(\alpha)}}$ are Jordan blocks for the eigenvalue $\kappa^{(\alpha)}$. The matrix $\Omega$ here consists of the generalized eigenvectors of $M$. The rest of the above construction can be carried over immediately. Integrating the associated ODE in Jordan normal form leads to a
matrix exponential of $\tilde{M}t$:

$$\Omega^{-1} \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = e^{\tilde{M}t} \begin{pmatrix} c_1 \\ \vdots \\ c_{2d} \end{pmatrix}$$  \hspace{1cm} (C.26)$$

The fundamental matrix of the first-order system is again obtained by multiplying with $\Omega$ from the left, leading to

$$\begin{pmatrix} U(t) \\ V(t) \end{pmatrix} = \Omega e^{\tilde{M}t}.$$  \hspace{1cm} (C.27)$$

Of course, we are only interested in the first $d$ rows of this relation as before, which define the fundamental matrix $U$.

To understand the structure of $U$ in the case of non-diagonalizable $M$ better, we note that the matrix exponentials of Jordan blocks are given by

$$\exp \left[ t \begin{pmatrix} \kappa & 1 \\ 0 & \kappa \end{pmatrix} \right] = \begin{pmatrix} e^{\kappa t} & t e^{\kappa t} \\ 0 & e^{\kappa t} \end{pmatrix}, \quad \exp \left[ t \begin{pmatrix} \kappa & 1 & 0 \\ 0 & \kappa & 1 \\ 0 & 0 & \kappa \end{pmatrix} \right] = \begin{pmatrix} e^{\kappa t} & t e^{\kappa t} & \frac{t^2}{2} e^{\kappa t} \\ 0 & e^{\kappa t} & 0 \\ 0 & 0 & e^{\kappa t} \end{pmatrix}, \ldots$$  \hspace{1cm} (C.28)$$

From this we understand that multiplying by the generalized eigenvectors in (C.27) leads to terms of the form

$$\begin{pmatrix} u^{(1)}(t) \\ v^{(1)}(t) \end{pmatrix} = e^{\kappa t} \omega^{(1)},$$

$$\begin{pmatrix} u^{(2)}(t) \\ v^{(2)}(t) \end{pmatrix} = e^{\kappa t} \left[ \omega^{(2)} + t \omega^{(1)} \right],$$

$$\begin{pmatrix} u^{(3)}(t) \\ v^{(3)}(t) \end{pmatrix} = e^{\kappa t} \left[ \omega^{(3)} + t \omega^{(2)} + \frac{t^2}{2} \omega^{(1)} \right], \ldots$$

for each Jordan block, where $\omega^{(1)}$ is the actual eigenvector for $\kappa$, and $\omega^{(2)}, \omega^{(3)}, \ldots$ are the generalized eigenvectors. Thus we have constructed the most general solution of the homogeneous problem $\tilde{L}u = 0$.

**Summary.** The BVP (C.1) can be solved by decomposing $y(t) = y_H(t) + y_P(t)$. The homogeneous solution $y_H$ satisfies (C.1) with $f = 0$. To compute it, we find the fundamental matrix $U$, write $y_H(t) = U(t)c$, and solve for the $2d$ constant components of the vector $c$ by imposing the boundary conditions.

The particular solution $y_P$ is obtained as a convolution of Green’s matrix $\Upsilon$ with the inhomogeneity $f$ as in (C.5). Green’s matrix is constructed from the fundamental matrix according to the above described procedure.

**D SIMULATION PARAMETERS**
<table>
<thead>
<tr>
<th>label</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$u_1$</th>
<th>$u_2$</th>
<th>$\alpha$</th>
<th>$f_{\text{ext}}$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\tau_\infty$</th>
<th>$\bar{\eta}$</th>
<th>$\eta_C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>0.1</td>
<td>4</td>
<td>2</td>
<td>$\frac{\pi}{4}$</td>
<td>$-0.5$</td>
<td>1</td>
<td>1</td>
<td>0.348</td>
<td>0.667</td>
<td>0.95</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>0.333</td>
<td>5</td>
<td>1</td>
<td>$\frac{\pi}{4}$</td>
<td>$-0.5$</td>
<td>1</td>
<td>1</td>
<td>0.436</td>
<td>0.4</td>
<td>0.67</td>
</tr>
</tbody>
</table>

REFERENCES


[53] P. Hartman, Ordinary Differential Equations, 2nd ed. (Siam, 2002).