Infimum Law and First-Passage-Time Fluctuation Theorem
for Entropy Production

Izaak Neri,1,2 Édgar Roldán,1 and Frank Jülicher1

1Max Planck Institute for the Physics of Complex Systems,
Nöthnitzer Str. 38, 01187 Dresden, Germany.
2Max Planck Institute of Molecular Cell Biology and Genetics,
Pfotenhauerstraße 108, 01307 Dresden, Germany.

We derive an infimum law and a first-passage-time fluctuation theorem for entropy production of stochastic processes at steady state. We show that the ratio between the probability densities of the first-passage time to produce $S_{\text{tot}}$ of entropy and of the first-passage time to produce $-S_{\text{tot}}$ of entropy equals $e^{S_{\text{tot}}/k}$, with $k$ Boltzmann’s constant. This first-passage-time fluctuation relation is valid for processes with one or higher order passages. In addition, we derive universal bounds for the infimum statistics of entropy production using the fact that at steady state $e^{-S_{\text{tot}}/k}$ is a martingale process. We show that the mean value of the entropy-production infimum obeys $\langle \inf S_{\text{tot}} \rangle \geq -k$ and derive a bound on the cumulative distribution of entropy-production infima. Our results are derived using a measure-theoretic formalism of stochastic thermodynamics. We illustrate our results with a drift-diffusion process and with numerical simulations of a Smoluchowski–Feynman ratchet.

I. INTRODUCTION AND STATEMENT OF THE MAIN RESULTS

The second law of thermodynamics states that the total entropy in a thermodynamic process increases with time. If such entropy production $S_{\text{tot}}$ occurs in a mesoscopic system that is in contact with a large environment, $S_{\text{tot}}$ is a stochastic quantity and the second law should be considered as a statement on the average entropy production. This idea was already formulated by J C Maxwell in the 19th century [1], but only in the last decades it has been formalized using the language of stochastic processes [2]. For stochastic processes it is possible to define the entropy produced along a single realization of the process consistent with thermodynamics [3–6]. The stochastic definition of entropy production implies a set of relations for the entropy-production fluctuations that are called fluctuation theorems [7].

An example is the detailed fluctuation theorem, which can be written as

$$\frac{p_{S}(S_{\text{tot}}; t)}{p_{S}(-S_{\text{tot}}; t)} = e^{S_{\text{tot}}/k} ,$$

where $k$ is Boltzmann’s constant. Here $p_{S}(S_{\text{tot}}; t)$ is the probability density for the entropy production $S_{\text{tot}}$ at a given time $t$. The detailed fluctuation theorem Eq. (1) is universal and holds for a broad class of physical processes, such as, deterministic nonequilibrium processes coupled to thermostats [8–10], stochastic Markovian processes [11–13] and stochastic non-Markovian processes [14–15]. Moreover, the steady-state fluctuation theorem given by Eq. (1) has been tested in several experiments, inter alia, a Rayleigh-Bénard convection cell [16], a dragged colloidal particle in an optical trap [17], a fluidized steady-state of a granular medium [18], a photochromic defect center in a diamond [19], biological systems such as the F1-ATPase rotary motor [20] and a single-electron box [21].

Until now, most work has focused on entropy-production fluctuations in a fixed time interval [6, 7, 22–23]. However, little is known about fluctuations of other statistical properties of entropy production [24–26]. Do first-passage-time fluctuations and extreme-value statistics of entropy production satisfy universal equalities? In this paper we give some novel insights on this question.

Here we derive universal fluctuation relations for the first-passage-time distributions of entropy production of classical and continuous stochastic processes in steady state. We call these relations first-passage-time fluctuation theorems. The simplest example is

$$\frac{p_{T}(t; S_{\text{tot}})}{p_{T}(t; -S_{\text{tot}})} = e^{S_{\text{tot}}/k} .$$

Here $p_{T}(t; S_{\text{tot}})$ is the first-passage-time distribution of entropy production; more precisely $p_{T}(t; S_{\text{tot}})$ is the probability density for the first-passage time $t$ at which the process has produced for the first time the entropy $S_{\text{tot}}$. The first-passage-time fluctuation theorem is illustrated in Fig. 1A.

Very interesting are also the extreme-value statistics of entropy production. We show that the average of the infimum of the stochastic entropy production in a given time interval is bounded from below by $-k$:

$$\langle \inf S_{\text{tot}} \rangle \geq -k .$$

This infimum law is illustrated in Fig. 1B and bears similarity with the second law of thermodynamics $\langle S_{\text{tot}} \rangle \geq 0$. We show that the infimum law follows from a universal lower bound on the cumulative distribution of the infimum of entropy production:

$$\Pr \left( \inf \frac{S_{\text{tot}}}{k} \geq -s \right) \geq 1 - e^{-s} ,$$

PACS numbers: 05.40.-a, 05.70.-a, 02.50.-r, 05.70.Ln
for $s \geq 0$. Here $\Pr(\cdot)$ denotes the probability of an event. Remarkably, as shown here the infimum law given by Eq. (3) is generally valid for classical and stationary stochastic processes.

The paper is structured as follows: In Sec. II we briefly review the formalism of stochastic thermodynamics based on path-probability densities. In Sec. III we formulate stochastic thermodynamics using measure theory, define entropy production for a stationary stochastic process and discuss its connection to martingale processes. The first-passage-time fluctuation theorem and the infimum law are derived in Sec. IV and Sec. V respectively. In Sec. VI we study the drift-diffusion process for which we present exact expressions of first-passage-time distributions and infimum distributions of entropy production. In Sec. VII we study a Smoluchowski-Feynman ratchet for which we obtain numerical estimates of the first-passage time and infimum statistics of entropy production. For both examples the first-passage-time fluctuation theorems and the infimum law are illustrated. We end with a discussion and an outlook in Sec. VIII.

II. STOCHASTIC THERMODYNAMICS AND ENTROPY PRODUCTION

We briefly review the basic formalism of stochastic thermodynamics and the definition of entropy production based on path-probability densities [7, 27, 28]. We consider the dynamics of a mesoscopic system in a non-equilibrium steady state described by the coarse-grained state variables $\omega(t) = (q(t), \qbar(t))$ consisting of the mesoscopic degrees of freedom $q(t)$ and $\qbar(t)$ at time $t$. The set $q(t)$ represents $n$ degrees of freedom that are even under time reversal and the set $\qbar(t)$ contains $n$ degrees of freedom that are odd under time reversal [29]. The variables $q(t)$ and $\qbar(t)$ could represent the dynamics of collective modes in a system of interacting particles, for instance, $q(t)$ could be associated with each trajectory $\omega(t)$ and $\qbar(t)$ the effective momentum of a colloidal particle in a fluid.

In a given time window $[0, t]$ the coordinates $\omega(t)$ trace a path in phase space $\omega_0 = \{q(\tau), \qbar(\tau)\}_{0 \leq \tau \leq t}$. We associate with each trajectory $\omega_0$ a probability density $P(\omega_0)$ that captures the limited information provided by the coarse-grained variables $\omega$ and the fact that the exact microstate is not known. The entropy production (or total entropy change) associated with a path $\omega_0$ in a stationary process is given by [30, 31]

$$S_{\text{tot}}(t) = k \ln \frac{P(\omega_0)}{P(\Theta\omega_0)} ,$$

where $\Theta\omega_0 = \{q(t - \tau), -\qbar(t - \tau)\}_{\tau=0}$ is the time-reversed trajectory. The steady-state average entropy production during a time interval $t$ can be expressed as a path integral [27, 32]:

$$\langle S_{\text{tot}}(t) \rangle = k \int D\omega_0 P(\omega_0) \ln \frac{P(\omega_0)}{P(\Theta\omega_0)} .$$

Entropy production is therefore the observable that quantities time irreversibility of mesoscopic trajectories [33]. In fact by measuring entropy production an observer can determine within a minimal time whether a movie of a stochastic process is run forwards or backwards [24].

Microscopic reversibility implies local detailed balance of the conditional probability densities which reads [34]:

$$\frac{P(\omega_0|\omega(0))}{P(\Theta\omega_0|\omega(t))} = e^{S_{\text{env}}(t)/k} ,$$

where $S_{\text{env}}(t)$ is the entropy change in the environment. The total entropy change $S_{\text{tot}}(t)$ can be written as

$$S_{\text{tot}}(t) = \Delta S_{\text{sys}}(t) + S_{\text{env}}(t) ,$$

FIG. 1: Illustration of the two theorems derived in the paper. A. First-passage-time fluctuation theorem for entropy production for a process with two absorbing boundaries. Examples of trajectories of the stochastic entropy production as a function of time that first reach a positive threshold $S_{\text{tot}}$ (blue thick curves) and first reach a negative threshold $-S_{\text{tot}}$ (red thin curves). The probability distributions to first reach the positive $p_T(t; S_{\text{tot}})$ and the negative $p_T(t; -S_{\text{tot}})$ thresholds at time $t$ are related by Eq. (2). B. Infimum law for entropy production. The infima of trajectory samples of total entropy production are shown in dashed lines. The mean of the infimum of the entropy production in a time interval $t$ (teal solid line) is greater than $-k$ (yellow thick solid line).
with
\[ \Delta S_{\text{sys}}(t) = -k \ln \frac{\mathcal{P}(\omega(t))}{\mathcal{P}(\omega(0))} . \]  

We call \( \Delta S_{\text{sys}}(t) \) the change in system’s entropy \([4]\). For a system in contact with one or several thermal baths the environment entropy change is related to the heat exchanged between system and environment \([38]\).

An important property of entropy production is that its exponential \( e^{-S_{\text{tot}}(t)/k} \) is a martingale process for stationary processes. A process is martingale when its expected value at any time \( t \) equals its value at a previous time \( \tau \), when the expected value is conditioned on observations up to the time \( \tau \). In other words, the expectation value of the exponential of minus the entropy change at time \( t \) equals its value at the time \( \tau \), when the expectation value is conditioned on the trajectory \( \omega_0 \) \([37]\):
\[ \langle e^{-S_{\text{tot}}(t)/k} | \omega_0 \rangle = e^{-S_{\text{tot}}(\tau)/k} . \]  

In addition, as discussed in Appendix B, a martingale process has to be integrable. Note that the martingale property of \( e^{-S_{\text{tot}}(t)/k} \) implies Jarzynski’s equality \( \langle e^{-S_{\text{tot}}(t)/k} \rangle = 1 \) \([38,39]\). The martingale property of \( e^{-S_{\text{tot}}(t)/k} \) is key to derive our two main results in sections IV and V.

III. MEASURE-THEORETIC FORMALISM OF STOCHASTIC THERMODYNAMICS

In discrete time, the expressions Eqs. (4) and (6) for entropy production are well-defined \([10]\). In continuous time, these expressions are problematic since the path-probability densities \( \mathcal{P} \) are not well defined in the continuous-time limit \([41]\). We therefore use a more general mathematical formalism based on measure theory. This measure-theoretic formalism of stochastic thermodynamics is presented in this section. We first review fundamental concepts of measure theory \([12,40]\). We then use these concepts to define the entropy production of a continuous stochastic process and discuss its properties.

A. Probability space of a stochastic process

We introduce the concept of a probability measure \( \mathbb{P} \) as a generalization of the path probability densities \( \mathcal{P} \) commonly used in stochastic thermodynamics. The difference between probability densities and probability measures is that a density associates a weight to one trajectory \( \omega \) while a measure associates a weight to sets of trajectories \( \Phi \). The value \( \mathbb{P}(\Phi) \) denotes the probability to observe a trajectory \( \omega \) in the set \( \Phi \), in other words \( \mathbb{P}(\Phi) = \mathbb{P}(\omega \in \Phi) \). In measure theory the set \( \Omega \) is called the sample space and the set \( \mathcal{F} \) of all measurable sets \( \Phi \) is a \( \sigma \)-algebra. The triple \( (\Omega, \mathcal{F}, \mathbb{P}) \) is called a probability space. Note that here and below \( \omega = \{q(\tau), \dot{q}(\tau)\}_{\tau \in (-\infty, \infty)} \) refers to a full trajectory of the state variables.

An important example of a measure \( \mathbb{P} \) on a sample space in \( \mathbb{R}^n \) is given by
\[ \mathbb{P}(\Phi) = \int_{\Phi} dx \, p(x) , \]  

where we have used the Lebesgue integral and \( p(x) \) is the probability density of the measure \( \mathbb{P} \). It is also possible to define a probability density of a measure \( \mathbb{P}(\Phi) = \int_{\omega \in \Phi} d\mathbb{P} \) with respect to a suitable reference measure \( \mathbb{Q} \) using the Radon-Nikodým theorem:
\[ \mathbb{P}(\Phi) = \int_{\omega \in \Phi} d\mathbb{Q} \, \mathcal{R}(\omega) , \]  

where we have introduced an integral over a probability space as discussed in Appendix A. The function \( \mathcal{R}(\omega) \) is called the Radon-Nikodým derivative, which we usually write as \( \mathcal{R}(\omega) = \frac{d\mathbb{P}}{d\mathbb{Q}}(\omega) \). In Eq. (12) \( \mathcal{R}(\omega) \) is a generalization of the probability density \( p(x) \) to spaces for which the Lebesgue measure does not exist, e.g., the Wiener space of trajectories of a Brownian particle (see Appendix A).

A stochastic process \( X(\omega; t) \) provides the value of an observable \( X \) at time \( t \) for a given trajectory \( \omega \). The average or expectation value of the stochastic variable \( X(t) \) is denoted by \( \langle X(t) \rangle_{\mathbb{P}} = \int_{\omega \in \mathcal{F}} X(\omega; t) d\mathbb{P} \). We often write simply \( \langle X(t) \rangle = \langle X(\omega; t) \rangle_{\mathbb{P}} \) to refer to the stochastic process and its average. In Appendix B, we present formal definitions of stochastic processes.

In what follows, we focus on stationary probability measures of steady-state processes. A stationary measure is time-translation invariant, i.e., it satisfies \( \mathbb{P} = \mathbb{P} \circ \mathcal{T}_t \) for all values of \( t \). Time translation \( \mathcal{T}_t \) is a map on trajectory space with the property \( X(\mathcal{T}_t(\omega); \tau) = X(\omega; \tau + t) \) for any stochastic process \( X \).

B. General definition of stochastic entropy production

We now define the stochastic entropy production as a stochastic process on a stationary probability measure \([31]\), which is also discussed in Appendix C. We express the entropy production in a time interval \([0, t]\) using the Radon-Nikodým derivative of the measure \( \mathbb{P}|_{F(t)} \) with respect to the time-reversed measure \( (\mathbb{P} \circ \Theta)|_{F(t)} \), viz.,
\[ S_{\text{tot}}(\omega; t) = k \ln \frac{d\mathbb{P}|_{F(t)}}{d(\mathbb{P} \circ \Theta)|_{F(t)}}(\omega) . \]  

In Eq. (13) we have introduced the restricted measure \( \mathbb{P}|_{F(t)} \) over those events in the sub-\( \sigma \)-algebra \( F(t) \subset \mathcal{F} \) that correspond to finite-time trajectories \( \omega^t \) in the time window \([0, t]\) (see Fig. 2A and Appendix B). Here the time-reversed measure \( \mathbb{P} \circ \Theta \) is defined in terms of the
time-reversal map $\Theta$, which acts on trajectories and maps future events on past events (see Fig. 2A). Time reversal $\Theta$ is defined by the property $X(\Theta(\omega); t) = X(\omega; -t)$ for all $X$ that are even under time reversal and $\tilde{X}(\Theta(\omega); t) = -\tilde{X}(\omega; -t)$ for all $\tilde{X}$ that are odd under time reversal.

For stochastic processes in continuous time, such as continuous processes or processes with jumps, Eq. (13) provides the entropy production in a given time interval $[0, t]$. In discrete time, Eq. (13) becomes the ratio of the path probability densities corresponding to the forward and backward dynamics and is equal to Eq. (5). Note that for systems that are microreversible Eq. (13) always exists [47].

The martingale property of $e^{-S_{\text{tot}}(t)/k}$ (as well as $e^{S_{\text{tot}}(t)/k}$) holds for the general definition of entropy production of stationary process given by Eq. (13). As shown in Appendix [C] this follows directly from the fact that $e^{-S_{\text{tot}}(t)/k}$ and $e^{S_{\text{tot}}(t)/k}$ are Radon-Nikodým density processes. Using the theory of martingales developed by Doob [42, 43] we derive two new general relations for entropy production. In the next section we derive the first-passage-time fluctuation theorem and in the subsequent section the infimum law for entropy production.

IV. FIRST-PASSAGE-TIME FLUCTUATION THEOREMS

In this section we derive the first-passage-time fluctuation theorems. We first introduce stopping times that generalize passage times. We then derive the first-passage time and present the main implications on entropy production fluctuations. Formal definitions of stopping times, passage times and the derivation of the first-passage-time fluctuation theorem can be found in Appendix [D].

A. General form of the first-passage-time fluctuation theorem

We now consider processes that start at the reference time $t = 0$ and that are terminated at a random time $t = T$, which is called the stopping time. The stopping time can depend on a criterion that the observables of the process have to satisfy in order to terminate the process. We define a $s_{\text{tot}}$-stopping time $T$ based on entropy production such that $S_{\text{tot}}(T) = s_{\text{tot}}$ when the process

FIG. 2: A. Illustration of trajectories of a coarse-grained variable $X$ in a time window $[0, \tau]$. Here $\mathcal{F}(\tau)$ is the sub-$\sigma$-algebra generated by trajectories in the time window $[0, \tau]$ and $P_{\mathcal{F}(\tau)}$ is the restricted measure over this time window. B. Illustration of the time-reversal map $\Theta$ on trajectories $\omega_1$, $\omega_2$ and $\omega_3$ of the coarse-grained variable $X$.

FIG. 3: Illustration of the first-passage times for entropy production. A. First-passage time $T_{FP}^{(1)}$ for entropy production with one positive absorbing boundary $s_{\text{tot}}$ (top) and first-passage time $\tilde{T}_{FP}^{(1)}$ for entropy production with one negative absorbing boundary $-s_{\text{tot}}$ (bottom). B. First-passage times $T_{FP}^{(2)}$ and $\tilde{T}_{FP}^{(2)}$ for entropy production with two absorbing boundaries at $\pm s_{\text{tot}}$. At the time $T_{FP}^{(2)}$ entropy production passes for the first time the threshold $s_{\text{tot}}$ without having reached $-s_{\text{tot}}$ before. At the time $\tilde{T}_{FP}^{(2)}$ entropy production passes for the first time the threshold $-s_{\text{tot}}$ without having reached $s_{\text{tot}}$ before.
and terminates.

An example of such an \( s_{\text{tot}} \)-stopping time is the first-passage time \( T_{\text{FP}}^{(1)} \), which determines the time at which the entropy change \( S_t(t) \) reaches for the first time the value \( s_{\text{tot}} \). Another example is given by the first-passage time \( T_{\text{FP}}^{(2)} \), the time at which entropy \( S_{\text{tot}}(t) \) passes for the first time a threshold value \( s_{\text{tot}} \), given that it has not reached \(-s_{\text{tot}}\) before. This implies that if the process first reaches \(-s_{\text{tot}}\) it terminates at a time \( T_{\text{FP}}^{(2)} \). This process is therefore equivalent to a first-passage problem with two absorbing boundaries. Note that there exist many other \( s_{\text{tot}} \)-stopping times such as times of second passage. The first-passage times \( T_{\text{FP}}^{(1)} \) and \( T_{\text{FP}}^{(2)} \) are illustrated in Fig. 3.

We have derived the following general fluctuation theorem for \( s_{\text{tot}} \)-stopping times \( T \) in Appendix D:

\[
\frac{P(\Phi T \leq t)}{P(\Theta T | \Phi T \leq t)} = e^{s_{\text{tot}}/k},
\]

where \( P(\Phi T \leq t) \) is the probability to observe a trajectory in the set \( \Phi T \leq t \) of infinite-time trajectories \( \omega \) that satisfy the termination criterion at a time \( T \leq t \). Note that these sets contain trajectories that do not actually terminate. The set \( \Theta T | \Phi T \leq t \) describes the time-reversed trajectories of \( \Phi T \leq t \). It is generated by applying the time-reversal map \( \Theta T \) to all the elements of the original set. The map \( \Theta T = T \circ \Theta \) corresponds to a time reversal within a time window \([0, T]\) with the property that \( X(\Theta T(\omega); \tau) = X(\omega; T(\omega) - \tau) \) for all stochastic processes \( X \) that are even under time reversal. The fluctuation theorem for \( s_{\text{tot}} \)-stopping times Eq. (14) is valid for continuous and stationary stochastic processes. In our derivation we use the martingale property of \( e^{S_{\text{tot}}(t)/k} \).

The probability density \( p_T \) of the \( s_{\text{tot}} \)-stopping time \( T \) is given by:

\[
p_T(t; s_{\text{tot}}) = \frac{d}{dt} P(\Phi T \leq t) .
\]

Entropy production is odd under time reversal \( S_t(\Theta T(\omega); T(\omega)) = -S_t(\omega; T(\omega)) \) as shown in Appendix C. Therefore, the first-passage times \( T = T_{\text{FP}}^{(1)} \) and \( T = T_{\text{FP}}^{(2)} \) have the property

\[
\Theta T (\Phi T \leq t) = \Phi T \leq t,
\]

where \( \tilde{T} = T_{\text{FP}}^{(1)} \) and \( \tilde{T} = T_{\text{FP}}^{(2)} \) are the corresponding first-passage times for the threshold \(-s_{\text{tot}}\) illustrated in Fig. 3. From this follows

\[
p_T(t; -s_{\text{tot}}) = \frac{d}{dt} P(\Theta T | \Phi T \leq t) .
\]

For all \( s_{\text{tot}} \)-stopping times for which \( \tilde{T} \) defined in Eq. (16) is the stopping time for the threshold \(-s_{\text{tot}}\) Eq. (17) holds and we also have

\[
\frac{p_T(t; s_{\text{tot}})}{p_T(t; -s_{\text{tot}})} = e^{s_{\text{tot}}/k} .
\]

The fluctuation theorem given by Eq. (18) therefore not only holds for distributions of \( T_{\text{FP}}^{(1)} \) and \( T_{\text{FP}}^{(2)} \) but also for higher order passage processes that depend on a single threshold \( s_{\text{tot}} \). It generalizes the results derived in Refs. [24] and [25] to stationary and continuous stochastic processes. In [24] a first-passage-time fluctuation relation analogous to Eq. (18) has been derived for stochastic processes with translation-invariant transition rates. In Ref. [25] the authors find the fluctuation relation given by Eq. (18) in the asymptotic limit \( t \to \infty \). Equation (18) implies two interesting results for first-passage times of stochastic processes that are outlined below.

### B. Fluctuation theorem for passage probabilities

The first-passage-time fluctuation theorem implies a fluctuation relation between passage probabilities of entropy production:

\[
\frac{P_+}{P_-} = e^{s_{\text{tot}}/k} .
\]

Passage probabilities are the probabilities for the process to terminate at a boundary. We distinguish two processes, one which terminates at \( s_{\text{tot}} \) and one which terminates at \(-s_{\text{tot}}\), with passage probabilities

\[
P_+ = \int_0^\infty dt \, p_T(t; s_{\text{tot}}) , \quad (20)
\]

\[
P_- = \int_0^\infty dt \, p_T(t; -s_{\text{tot}}) , \quad (21)
\]

and with \( P_+ \leq 1 \). Equation (19) follows directly from integrating Eq. (18) over time.

For the case of the first-passage times \( T_{\text{FP}}^{(2)} \) with two absorbing boundaries the process terminates in a finite time either in the positive or the negative threshold with probability one and therefore \( P_+ + P_- = 1 \). In this case, we find the following exact expressions for the two passage probabilities in terms of the threshold value \( s_{\text{tot}} \):

\[
P_+ = \frac{e^{s_{\text{tot}}/k}}{1 + e^{s_{\text{tot}}/k}} , \quad (22)
\]

\[
P_- = \frac{1}{1 + e^{s_{\text{tot}}/k}} . \quad (23)
\]

### C. Symmetry of the normalized first-passage-time distributions

The first-passage-time fluctuation relation Eq. (18) also implies an equality between the normalized first-passage-time distributions \( p_T(t|s_{\text{tot}}) \) and \( p_T(t|s_{\text{tot}}) \) which reads

\[
p_T(t|s_{\text{tot}}) = p_T(t + s_{\text{tot}}) . \quad (24)
\]
The normalized distributions are defined as:

\[ p_T(t|s_{tot}) = \frac{p_T(t; s_{tot})}{\int_0^{\infty} dt \ p_T(t; s_{tot})} , \quad (25) \]

\[ p_T(t - s_{tot}) = \frac{p_T(t; -s_{tot})}{\int_0^{\infty} dt \ p_T(t; -s_{tot})} . \quad (26) \]

The symmetric relation Eq. (24) comes from the fact that the ratio of the first-passage-time distributions in Eq. (28) is time independent.

The first-passage-time fluctuation theorem thus implies that the mean first-passage time given that the process terminates at the positive boundary, equals to the mean first-passage time given that the process terminates at the negative boundary. This remarkable symmetry extends to all of the moments of the first-passage-time distributions. A similar result has been found for waiting-time distributions in chemical kinetics [48–54] for decision-time distributions in sequential hypothesis tests [56]. These results could therefore be interpreted as a consequence of the fundamental relation for the first-passage time fluctuations of entropy production given by Eq. (24).

\[ \langle \inf S_{tot}(t) \rangle \geq -k . \quad (29) \]

The infimum law given by Eq. (29) holds for stationary stochastic processes in discrete time and for stationary stochastic processes in continuous time for which \( e^{-S_{tot}(t)/k} \) is càdlàg [57]. Càdlàg processes are stochastic processes that may contain jumps as explained in Appendix B.

\[ \langle \inf S_{tot}(t) \rangle \geq -k . \quad (29) \]

V. THE INFINUM LAW

We now derive the infimum law valid for non-equilibrium steady-state processes. We only need the martingale property of \( e^{-S_{tot}(t)/k} \) with respect to \( \mathbb{P} \).

The cumulative distribution of the supremum of \( e^{-S_{tot}(t)/k} \) satisfies

\[ \text{Pr} \left( \sup_{\tau \in [0,t]} \left\{ e^{-S_{tot}(\tau)/k} \right\} \geq \lambda \right) \leq \frac{1}{\lambda} \left\langle e^{-S_{tot}(t)/k} \right\rangle , \quad (27) \]

with \( \lambda \geq 0 \) [20]. Equation (27) corresponds to Doob’s maximal inequality for martingale processes given by Eq. (B15) in Appendix B.

Using Jarzynski’s equality \( \left\langle e^{-S_{tot}(t)/k} \right\rangle = 1 \) [38], Eq. (27) implies a lower bound on the cumulative distribution of the infimum of \( S_{tot} \) in a given time interval \([0,t] \):

\[ \text{Pr} \left( \inf \frac{S_{tot}(t)}{k} \geq -s \right) \geq 1 - e^{-s} , \quad (28) \]

with \( s \geq 0 \) and \( \inf S_{tot}(t) = \inf_{\tau \in [0,t]} \{ S_{tot}(\tau) \} \). The right hand side of Eq. (28) is the cumulative distribution of an exponential random variable \( S \) with distribution function \( p_S(s) = e^{-s} \). From Eq. (28) it thus follows that the random variable \(-\inf S_{tot}(t)/k\) dominates stochastically over \( S \). Stochastic dominance implies an inequality on the mean values of the corresponding random variables as shown in Appendix B. Equation (28) therefore implies the following universal bound for the mean infimum of entropy production:

\[ \langle \inf S_{tot}(t) \rangle \geq -k . \quad (29) \]

VI. DRIFT-DIFFUSION PROCESS

As a first illustrative example we study a drift-diffusion process of a Brownian particle with diffusion coefficient \( D \), average drift velocity \( v \), and periodic boundary conditions with period \( \ell \) (see Fig. 4). For simplicity we consider here the case of a one-dimensional Brownian motion without inertia. The state of the particle at time \( t \) can be described by a variable \( \phi(t) \in [0, \ell) \). In the illustration of a ring geometry in Fig. 4 \( \phi \) is the azimuthal angle and \( \ell = 2\pi \).

Equivalently, one can consider a stochastic process \( X(t) \) given by the net distance traveled by the Brownian particle up to time \( t \): \( X(t) = \phi(t) + \ell N_t(t) \), where \( N_t(t) \) is the winding number or the net number of clockwise turns (or minus the number of counterclockwise turns) done by the particle up to time \( t \) [24]. The time evolution of \( X(t) \) is described by the following Langevin equation

\[ \frac{dX(t)}{dt} = v + \zeta(t) . \quad (30) \]

The term \( \zeta(t) \) in Eq. (30) is a Gaussian white noise with zero-mean \( \langle \zeta(t) \rangle = 0 \) and with autocorrelation...
The entropy production in a time \( t \) in steady state is given by [24]

\[
S_{\text{tot}}(t) = k \frac{v}{D} [X(t) - X(0)] ,
\]

or equivalently \( S_{\text{tot}}(t) = k \frac{v}{D} [\phi(t) - \phi(0)] + k \ell \frac{v}{D} N_\phi(t) \).

Equation (31) implies that the first-passage and extreme value statistics of entropy production in the drift-diffusion process can be obtained from the statistics of the position \( X(t) \) of a drifted Brownian particle in the real line.

The drift-diffusion process is an example for which the first-passage-time fluctuation theorem and the infimum law can be verified analytically, as we show below.

### A. First-passage-time fluctuation theorem

The first-passage-time distribution for \( X \) to pass at time \( t \) for the first time the threshold \( L > 0 \) starting from the initial condition \( X(0) = 0 \) is given by Wald’s distribution [58, 59]

\[
p_T(t; L) = \frac{|L|}{\sqrt{4\pi Dt^3}} e^{-(L-vt)^2/4Dt} .
\]

Equation (32) implies:

\[
\frac{p_T(t; L)}{p_T(t; -L)} = e^{sL/D} .
\]

Note that the argument of the exponential equals to the Péclet number, \( Pe = vL/D \).

The distribution of the entropy first-passage time \( T_{FP}^{(1)} \) equals to the first-passage-time distribution given by Eq. (32) for the position of the particle with an absorbing boundary at \( L = s_{\text{tot}} D/v \). Replacing \( L \) by \( s_{\text{tot}} D/v \) in Eq. (33) one finds the first-passage-time fluctuation theorem for entropy production

\[
\frac{p_T(t; s_{\text{tot}})}{p_T(t; -s_{\text{tot}})} = e^{s_{\text{tot}}/k} .
\]

An analogous relation holds for the two-boundary first-passage times \( T_{FP}^{(2)} \) for entropy production and can be derived using the results in Sec. 2.2.2.2 in [58] (see also [24]).

### B. Infimum law

For the drift-diffusion process in steady state we derive an exact expression for the cumulative distribution of minus the infimum of entropy production (see Appendix F):

\[
\Pr \left( \inf \frac{S_{\text{tot}}(t)}{k} \geq -s \right)
\]

\[
= \frac{1}{2} \left[ \text{erfc} \left( \frac{-s - \sigma(t)}{2\sqrt{\sigma(t)}} \right) - e^{-s} \text{erfc} \left( \frac{s - \sigma(t)}{2\sqrt{\sigma(t)}} \right) \right] ,
\]

where \( s > 0 \), \( \text{erfc} \) is the complementary error function and \( S(t) = (S_{\text{tot}}(t))/k = (v^2/D)t \) is the average entropy production in steady state at time \( t \). Equation (35) illustrates the universal bound on the infimum cumulative distribution given by Eq. (28). Indeed, in Fig. 5 we show that the infimum cumulative distribution given by Eq. (35) is bounded from below by \( \Pr \left( \inf \frac{S_{\text{tot}}(t)}{k} \geq -s \right) \geq 1 - e^{-s} \), for different values of \( \sigma(t) \). Interestingly, the lower bound is reached when \( \sigma(t) \) becomes very large, which corresponds to the asymptotic limit of long times or the very irreversible limit at finite times \( t \). In the reversible limit of small \( \sigma(t) \), the probability \( \Pr \left( \inf \frac{S_{\text{tot}}(t)}{k} \geq -s \right) \approx 1 \), which results in a trivial upper bound for the cumulative distribution.

We also derive the following analytical expression for the mean infimum of the entropy production in steady state (see Appendix F):

\[
\left\langle \inf \frac{S_{\text{tot}}(t)}{k} \right\rangle 
= -\text{erf} \left[ \frac{\sqrt{\sigma(t)}}{2} \right] + \frac{\sigma(t)}{2} \text{erfc} \left[ \frac{\sqrt{\sigma(t)}}{2} \right] - \sqrt{\frac{\sigma(t)}{\pi}} e^{-\sigma(t)/4} ,
\]

where \( \text{erf} \) is the error function. Equation (36) satisfies the infimum law \( \left\langle \inf S_{\text{tot}}(t) \right\rangle \geq -k \). Interestingly the lower bound for the mean infimum is reached in the limit of large mean entropy production \( \sigma(t) \), as shown in Fig 6. In the equilibrium limit we have \( \left\langle \inf S_{\text{tot}}(t) \right\rangle \simeq 0 \).

### VII. Smoluchowski–Feynman Ratchet

We now discuss the statistics of first-passage times and infima of entropy production in a paradigmatic ex-
ample of a nonequilibrium process for which system entropy is varies with position in steady state. We study a Smoluchowski–Feynman ratchet for which an overdamped Brownian particle moves under influence of an constant external force in a periodic potential \(\phi\) (see Fig. 7 for a graphical illustration). Experimentally this model has been realized using colloidal particles trapped with toroidal optical potentials \(\alpha, \beta\). The steady-state fluctuation theorem given by Eq. (1) was tested using this experimental technique in Ref. \(\alpha, \beta\).

Analogously to the drift-diffusion process, we describe the dynamics of the Smoluchowski–Feynman ratchet in terms of the net distance \(X(t)\) traveled by the particle. The dynamics of the stochastic process \(X(t)\) is given by an overdamped Langevin equation

\[
\gamma \frac{dX(t)}{dt} = -\frac{\partial V(X(t))}{\partial x} + F + \xi(t),
\]

where \(\gamma\) is a friction coefficient and \(\xi\) is a Gaussian white noise with zero mean \(\langle \xi(t) \rangle = 0\) and autocorrelation \(\langle \xi(t) \xi(t') \rangle = 2kT \gamma \delta(t - t')\). Here \(V(x)\) is a periodic potential of period \(\ell\), \(V(x + m\ell) = V(x)\) with \(m \in \mathbb{Z}\), and \(\frac{\partial V(X(t))}{\partial x} = \frac{\partial V(x)}{\partial x}\big|_{X(t)}\).

The system entropy change in steady state over a time \(t\) equals \(\Delta S_{\text{sys}} = \frac{\Delta U(t)}{T} + k \ln \frac{\int_{X(0)}^{X(0)+\ell} dy e^{U(y)/kT}}{\int_{X(t)}^{X(t)+\ell} dy e^{U(y)/kT}},\)

where \(\Delta U(t) = U(X(t)) - U(X(0))\) with \(U(x) = V(x) - Fx\) equal to the effective potential felt by the particle. The change in the environment entropy is given by \(S_{\text{env}}(t) = -Q(t)/T\), where \(Q(t)\) is the heat transferred from the environment to the particle in a time \(t\). Following Sekimoto \(\alpha, \beta\), \(Q(t) = \int_0^t (-\gamma dX(s)/ds + \xi(s)) \circ dX(s) = \int_0^t \partial U(X(s))/\partial x \circ dX(s) = \Delta U(t),\) where \(\circ\) denotes the Stratonovich product. Therefore, in this example,

\[
S_{\text{env}}(t) = -\frac{\Delta U(t)}{T}.
\]

The change in total entropy over a time \(t\) is given by \(S_{\text{tot}}(t) = \Delta S_{\text{sys}}(t) + S_{\text{env}}(t),\) with:

\[
S_{\text{tot}}(t) = k \ln \frac{\int_{X(0)}^{X(0)+\ell} dy e^{U(y)/kT}}{\int_{X(t)}^{X(t)+\ell} dy e^{U(y)/kT}}.
\]

For the choice

\[
V(x) = kT \ln[\cos(2\pi x/\ell) + 2]
\]

the integrals in Eq. (40) can be calculated analytically \(\alpha, \beta\). The stochastic entropy production in steady state given by Eq. (40) equals

\[
\frac{S_{\text{tot}}(t)}{k} = f[X(t) - X(0)] - \ln \frac{\psi(X(t), f)}{\psi(X(0), f)},
\]

with \(f = F\ell/2kT, \psi(x, f) = f^2[\cos(x) + 2] - f \sin(x) + 2\).

We perform numerical simulations of Smoluchowski-Feynman ratchet Eq. (37) in steady state with a potential given by Eq. (41). We then obtain numerical estimates

**FIG. 7:** Illustration of a Smoluchowski–Feynman ratchet. A Brownian particle (gray sphere) immersed in a thermal bath of temperature \(T\) moves in a periodic potential \(\phi\) (black shaded curve) with friction coefficient \(\gamma\). The coordinate \(\phi\) is the azimuthal angle of the particle. When applying an external force \(F\) in the azimuthal direction, the particle reaches a nonequilibrium steady state. In this example, \(\phi = kT \ln[\cos(\phi) + 2], \alpha = R \cos(\phi)\) and \(\beta = R \sin(\phi)\), with \(R = 2\).
of the first-passage times of entropy production and of
the infima distributions of entropy production using the
expression for entropy production \( \langle S \rangle \). These empirical
estimates are used to test our universal results.

A. First-passage-time fluctuation theorems

First we study the first-passage times \( T_{FP}^{(2)} \) for en-
tropy production with two absorbing boundaries at the
threshold values \( s_{tot} \) and \( -s_{tot} \) (with \( s_{tot} > 0 \)). Figure
\( 8 \) shows the empirical first-passage-time distribution \( p_{FP}(t; s_{tot}) \) to first reach the positive threshold (blue squares) together with the first-passage time distribution \( p_{FP}(t; -s_{tot}) \) (scaled by a factor \( e^{s_{tot}/k} \)) to first reach
the negative threshold (red circles). Since both distribu-
tions coincide we confirm validity of the first-passage-
time fluctuation theorem given by Eq. \( 18 \) (see top inset in Fig. \( 8 \)). Note that for
small values of the thresholds the passage probabilities
\( p_{FP} \) also fulfilled in this example (see bottom inset in Fig. \( 8 \)). As a result, the integral first-

![Figure 8: Empirical two-boundary first-passage-time distributions of entropy production to first reach a positive threshold \( p_{FP}(t; s_{tot}) \) (blue squares) and the rescaled distribution for the negative threshold \( p_{FP}(t; -s_{tot}) e^{s_{tot}/k} \) (red circles) for the Smoluchowski–Feynman ratchet. The distributions are obtained from 10^4 numerical simulations and the threshold values are set to \( \pm s_{tot} = \pm 2.2k \). The simulations are done using Euler numerical scheme with the following parameters: \( F = 4 \) pN, \( T = 300 \) K, \( \ell = 2\pi \) nm, \( \gamma = 8.4 \) pN s/\( nm \) and simulation time step \( \Delta t = 0.0127 \) s. Top inset: Empirical passage probabilities of entropy production to the positive (\( P_+ \), blue squares) and to the negative (\( P_- \), red circles) thresholds as a function of the threshold values. The analytical expressions for \( P_+ \) given by Eq. \( 22 \) (blue solid line) and \( P_- \) given by Eq. \( 23 \) (red dashed line) are also shown. Bottom inset: Logarithm of the ratio between the empirical passage probabilities \( P_+ \) and \( P_- \) as a function of the threshold value (magenta diamonds). The solid line is a line of slope 1.

B. Infimum law

We now study the infimum properties of entropy pro-
duction in the Smoluchowski–Feynman ratchet. First
we compute numerical estimates of the cumulative dis-
tribution of the entropy production infimum measured
over a fixed time interval for different values of the ex-
ternal force \( F \). Figure \( 10 \) shows that the cumulative
distribution of minus the entropy production infimum is
bounded from below by \( 1 - e^{-\alpha} \) which confirms the
universality of Eq. \( 25 \). Interestingly, the bound is
tighter for larger values of the average entropy produc-
tion \( \sigma(t) = \langle S_{tot}(t) \rangle/k \). Strikingly, as shown in Fig. \( 10 \)
the cumulative distribution for the infimum of entropy produ-
tion of the Smoluchowski–Feynman ratchet is nearly
identical to the corresponding cumulative distribution of
the drift-diffusion process given by Eq. \( 35 \). This equiv-

equivalence between the infimum cumulative distributions holds
even for small values of \( \sigma(t) \) where the shape of the po-
tential \( V(x) \) affects the entropy-production fluctuations.

In Fig. \( 11 \) we furthermore show that the infimum law
\( \langle \inf S_{tot}(t) \rangle \geq -k \) holds for the Smoluchowski–Feynman
ratchet using numerical simulations. Moreover, we also
find that the mean infimum of entropy production follows
the functional dependency on \( \sigma(t) \) given by Eq. \( 56 \)
obtained for the drift-diffusion process. These results point
towards a universal behaviour of the entropy produc-
tion infimum as a function of the average entropy pro-
In this paper we have derived universal relations for first-passage times and extreme-value statistics for entropy production of stationary stochastic processes. Our results do not follow from the standard fluctuation theorem but from other generic properties of stochastic entropy production in particular its martingale property. We have derived a fluctuation theorem given by Eq. (18) for first-passage times of entropy production valid for continuous stochastic processes in steady state. This theorem states that at any time instance it is exponentially more likely to have at this time first produced a positive amount of entropy as compared to the possibility to have at this time first reduced entropy by the same amount.

Secondly, we have discussed properties of the infimum of entropy production. Note that it can only be smaller or equal than zero, since at the initial time entropy is zero. We have shown that the mean infimum of entropy production in steady state is lower bounded by minus the Boltzmann constant. This bound, which we call the infimum law holds for stochastic processes in discrete time and stochastic processes in continuous time that may have discontinuous jumps.

VIII. DISCUSSION AND OUTLOOK

FIG. 9: Empirical one-boundary first-passage-time distributions of entropy production to first reach a positive threshold \( p_+ (t; s_{tot}) \) (blue squares) and the rescaled distribution for the negative threshold \( p_- (t; -s_{tot}) e^{s_{tot}/k} \) (red circles) for a Smoluchowski-Feynman ratchet in steady state. The estimate of \( p_+ (t; s_{tot}) \) (\( p_- (t; -s_{tot}) \)) is obtained measuring the time elapsed by the entropy production to first reach a single absorbing boundary in \( s_{tot} = 2.2k \) \((-s_{tot} = -2.2k)\) in \(10^4\) simulations. The simulations are done with the same parameters as in Fig. 8 and the empirical probabilities are calculated over a total simulation time of \( \tau_{max} = 20s\). Top inset: Empirical passage probabilities of entropy production in the positive-threshold simulations \( (P_+, \text{blue squares}) \) and in the negative-threshold simulations \( (P_-, \text{red circles}) \) as a function of the value of the thresholds. The expressions \( P_+ \) given by Eq. (22) (blue solid line) and \( P_- \) given by Eq. (23) (red dashed line) are also shown. Bottom inset: Logarithm of the ratio between \( P_+ \) and \( P_- \) as a function of the threshold value (magenta diamonds). The solid line is a line of slope 1.

FIG. 10: Cumulative distribution of the infimum of entropy production obtained from numerical simulations of a Smoluchowski-Feynman ratchet in steady state for different values of the mean entropy change \( \sigma (t) = \langle S_{tot} (t) \rangle/k \) compared with the bound given by \( 1 - e^{-s} \) (solid yellow line). The corresponding dashed curves are the values of the cumulative distribution of the infimum in the drift-diffusion process given by Eq. (35). Simulation parameters: \( 10^4 \) simulations with \( T = 300K, \ell = 2\pi\text{nm}, \gamma = 8.4\text{pNs/nm}, \) simulation time step \( \Delta t = 0.00127s \) and total simulation time \( t = 0.4s\).

FIG. 11: Mean value of the entropy production infimum as a function of the mean entropy change \( \sigma (t) = \langle S_{tot} (t) \rangle/k \) obtained from numerical simulations of a Smoluchowski-Feynman ratchet in steady state with different simulation time steps \( \Delta t \). The yellow thick bar is the bound of infimum law given by \( -k\). Simulation parameters: \( 10^4 \) simulations with \( T = 300K, \ell = 2\pi\text{nm}, \gamma = 8.4\text{pNs/nm}, \) and total simulation time \( t = 0.4s\).
The fundamental results derived here imply thermodynamic constraints on mesoscopic nonequilibrium processes that result from microscopic reversibility. For instance, in simple models of enzymatic reactions it has been shown that the waiting-time distributions for the forward and backward reactions are exactly the same [18-22]. Single-molecule data on kinesin motor stepping provides evidence that the waiting-time distributions of forward and backward steps are the same [23]. Interestingly, these relations are very similar to the first-passage-time fluctuation theorem on the normalized distributions given by Eq. (24). This raises the question whether reactions times are first-passage times or entropic stopping times and whether these relations found in chemical kinetics can be understood in terms of entropy-production fluctuations. Analogously, we expect that the infimum law can provide novel bounds on extreme-value statistics in enzymatic reactions and other biomolecular processes.

Acknowledgments

We thank Prof. em. Dr. sc. techn. Heinrich Meyr, Dr. -Ing. Meik Dörpinghaus and Mostafa Khalili-Marandi for enlightening discussions on decision theory and the Gichman-Skorochod method [67-68]. We also thank Dr. Andre Cardoso Barato for valuable discussions on stochastic thermodynamics, and Dr. Matteo Polettini for making us aware about the cycle fluctuation theorems in Refs. [53] and [55]. E.R. acknowledges funding from Grupo Interdisciplinar de Sistemas Complejos (GISC, Madrid, Spain) and from Spanish Government, Grant TerMic (FIS2014-52486-R).

References

Variables that are even under time reversal do not change sign under time reversal (e.g. position variables) whereas variables that are odd under time reversal change sign under time reversal (e.g. momenta or current).


For a system in contact with $\ell$ thermal reservoirs at temperatures $T_i$, $i = 1, \ldots, \ell$, the change of environment entropy $S_{\text{env}}(t)$ equals to $S_{\text{env}}(t) = -\sum_{i=1}^{\ell} Q_i(t)/T_i$, with $Q_i(t)$ the heat exchanged between the $i$-th reservoir and the system.

The martingale property follows from

\[ \int d\omega_i^T \mathcal{P}(\omega_i^T|\omega_i^0) e^{-S_{\text{tot}}(t)/k} = \int d\omega_i^T \mathcal{P}(\omega_i^T|\omega_i^0) e^{\mathcal{I}(\omega_i^0)/\mathcal{P}(\omega_i^0)} = \int d\omega_i^T \mathcal{P}(\omega_i^T|\omega_i^0) e^\mathcal{I}(\omega_i^0) = e^{-S_{\text{tot}}(\tau)/k}. \]


We have

\[ e^{-S_{\text{tot}}(t)/k} = \int d\omega(0)\mathcal{P}(\omega(0)) \int d\omega_0^T \mathcal{P}(\omega_0^T|\omega(0)) e^{-\mathcal{I}(\omega(0))/k} = e^{-S_{\text{tot}}(\tau)/k} \]

1, using Eq. (10) for $\tau = 0$.


The normalization constant of the path probability density diverges in the continuum limit.


The Radon-Nikodým derivative in Eq. (13) is well defined as long as the measure $\mathcal{P}|_{\mathcal{I}(t)}$ is absolutely continuous with respect to the time-reversed measure $\mathcal{P}|_{\Theta|_{\mathcal{I}(t)}}$. Absolute continuity means in the present context that irreversible transitions form a set of zero measure.


Cädlig is an acronym for continue à droite, limite à gauche, which means everywhere right continuous with left limits.


Appendix A: General concepts on measure theory

We introduce probability spaces, integration over a probability space and the Radon-Nikodým derivative following the treatises by Royden and Fitzpatrick [44] and Tao [46] on measure theory.

1. Measure space

A measurable space is a couple $(\Omega, \mathcal{F})$ with $\Omega$ a set and $\mathcal{F}$ a $\sigma$-algebra of subsets of $\Omega$. We say that a subset $\Phi$ of $\Omega$ is $\mathcal{F}$-measurable when it belongs to $\mathcal{F}$, i.e., $\Phi \in \mathcal{F}$. The $\sigma$-algebra contains thus all measurable subsets of $\Omega$.

We measure sets in a measurable space using a function $\mu$ that we call a measure. A measure is a function $\mu : \mathcal{F} \to [0, \infty]$, which has the property $\mu(\emptyset) = 0$ and is countably additive, i.e., for any countable disjoint collection of measurable sets $\{\Phi_i\}_{i=1}^{\infty}$ we have

$$\mu(\bigcup_{i=1}^{\infty} \Phi_i) = \sum_{i=1}^{\infty} \mu(\Phi_i).$$

The triple $(\Omega, \mathcal{F}, \mu)$ is called a measure space. A measure space $(\Omega, \mathcal{F}, \mu)$ is $\sigma$-finite when there exists a countable collection of measurable sets $\{\Phi_i\}_{i=1}^{\infty}$ such that $\Omega = \bigcup_{i=1}^{\infty} \Phi_i$ and $\mu(\Phi_i) < \infty$ for all values of $i$.

A probability measure $\mathbb{P}$ is a $\sigma$-finite measure for which $\mathbb{P}(\Omega) = 1$. The corresponding measure space is also called a probability space.

2. Integration over a measure space

A function $f : \Omega \to [0, \infty]$ is $\mathcal{F}$-measurable when $f^{-1}(U)$ is $\mathcal{F}$-measurable for every open subset $U \subset [0, \infty]$. The mean value $E_\mu(f)$ of a function $f$ over the set $\Omega$ is defined as

$$E_\mu(f) = \int_{\omega \in \Omega} f(\omega) d\mu.$$  \hspace{1cm} (A1)

In Eq. (A1) we have used the unsigned integral over a measure space $(\Omega, \mathcal{F}, \mu)$ as defined in subsection 1.4 of [44] or chapter 18 of [44]. We call a positive-valued function $f$ integrable when its expected value is finite, i.e.,

$$\int_{\omega \in \Phi} f(\omega) d\mu < \infty.$$  \hspace{1cm} (A2)

The measure of a set $\Phi$ can also be written in terms of the above unsigned integral

$$\mu(\Phi) = \int_{\omega \in \Phi} d\mu.$$  \hspace{1cm} (A3)

3. Random variables

The elements $\omega$ of the sample space $\Omega$ contain the full information on a realization or outcome of the process. Often we are only interested in partial information given by the value of a specific measurable observable. We describe measurable observables in terms of random variables. A random variable $X$ is defined as a function $X : \Omega \to E : \omega \to X(\omega)$ from a probability space

5. Radon-Nikodým theorem

The Radon-Nikodým theorem relates two integrals under a change of measure. The theorem also defines the Radon-Nikodým derivative, which generalizes the concept of a probability density over the real coordinates space $\mathbb{R}^n$.

Let $(\Omega, \mathcal{F}, \mu)$ and $(\Omega, \mathcal{F}, \nu)$ be two $\sigma$-finite measure spaces. The measure $\mu$ is absolute continuous with respect to the measure $\nu$ if for all $\Phi \in \mathcal{F}$ with $\nu(\Phi) = 0$ we have that $\mu(\Phi) = 0$. The Radon-Nikodým theorem asserts that for these two measure spaces there exists a unique $\mathcal{F}$-measurable and nonnegative function $\mathcal{R}$ on $\Omega$ for which

$$\mu(\Phi) = \int_{\omega \in \Phi} \mathcal{R}(\omega) d\nu,$$  \hspace{1cm} (A6)

for all $\Phi \in \mathcal{F}$ (see subsection 18.4 in [44]). The function $\mathcal{R}$ is uniquely determined up to sets of zero-measure, i.e., $\mu$-almost everywhere. The measurable function $\mathcal{R}(\omega)$ is called the Radon-Nikodým derivative of the measure $\mu$ with respect to the measure $\nu$. We use for $\mathcal{R}(\omega)$ the more conventional notation

$$\mathcal{R}(\omega) = \frac{d\mu}{d\nu}(\omega).$$  \hspace{1cm} (A7)

The Radon-Nikodým derivative generalizes the concept of a probability density over the Euclidean space $\mathbb{R}^n$. The probability density of a measure on $\mathbb{R}^n$ is the Radon-Nikodým derivative of the measure with respect to the Lebesgue measure. The Radon-Nikodým derivative also applies to $\sigma$-finite measures on infinite dimensional vector spaces for which there is no analogue of the Lebesgue measure. We elaborate on this point further.

A probability density $p_X$ of a real-valued random variable $X$ can be defined in terms of the Radon-Nikodým derivative. Consider a measure space $(\Omega, \mathcal{F}, \mu)$, a $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}$ and a measure space on the real line $(\mathbb{R}, \mathcal{F}(\mathbb{R}), \lambda)$ with $\mathcal{F}(\mathbb{R})$ the Borel $\sigma$-algebra and $\lambda$ the Lebesgue measure on the real line. We furthermore consider the measure space $(\mathbb{R}, \mathcal{F}(\mathbb{R}), \mu_X)$ of the random variable $X$ with the measure $\mu_X = \mu(X^{-1}(\Phi))$ for all $\Phi \in \mathcal{F}(\mathbb{R})$. The probability density $p_X$ of the random variable $X$ is then given by the Radon-Nikodým derivative of $\mu_X$ with respect to $\lambda$:

$$p_X(x) = \frac{d\mu_X}{d\lambda}(x)$$  \hspace{1cm} (A8)

for $x \in \mathbb{R}$. More generally, the Radon-Nikodým derivative of two absolutely-continuous measures $\mu_X$ and $\nu_X$ on $\mathcal{F}(\mathbb{R})$ is given by their probability-density ratio:

$$\frac{d\mu_X}{d\nu_X}(x) = \frac{d\mu_X}{d\lambda}(x) \frac{d\lambda}{d\nu_X}(x)$$

$$= \frac{d\mu_X}{d\lambda}(x) \left( \frac{d\nu_X}{d\lambda}(x) \right)^{-1} = \frac{p_X(x)}{q_X(x)}.$$  \hspace{1cm} (A9)
On infinite-dimensional vector spaces the Lebesgue measure does not exist. On these vector spaces we use Radon-Nikodým derivatives as a generalization of probability densities. Infinite-dimensional vector spaces appear naturally as the space of trajectories of a process in continuous time, see Fig. 12. The Wiener space of continuous trajectories and the Skorochod space of càdlàg trajectories are examples of infinite-dimensional vector spaces. Càdlàg trajectories contain continuous trajectories as particular cases but may also have jumps. In Fig. 12 B we give an example of a continuous trajectory over the time interval [0, t] and in Fig. 12 C we show a càdlàg trajectory over the same interval in time. Although infinite-dimensional vector spaces do not have a Lebesgue measure, Gaussian measures on infinite-dimensional vector spaces exist and are σ-finite [71]. An example is the Wiener measure defined on the space of continuous trajectories (see Fig. 12 B) [71].

**Appendix B: Measure-theoretic concepts for stochastic processes**

We introduce stochastic processes, martingale processes, stopping times, Doob’s theorems for martingales and Radon-Nikodým density processes following chapter XI on martingales in [43] and the treatise [73].

1. **Stochastic processes**

Observables of a physical process evolve over time. The corresponding phase space of trajectories evolves thus as well. In measure theory we use sequences of sub-σ-algebras to describe this dynamical evolution and a corresponding sequence of probability spaces.

We introduce time through an ordered index set I of time points t ∈ I with an order-relation symbol by ≤. An event at time t happens before the time t when τ ≤ t. For processes in continuous time the index set I is, e.g., the set of positive real numbers \( \mathbb{R}^+ \), the set of negative real numbers \( \mathbb{R}^- \) or a time interval \([0, t] \). In discrete time the index set is, e.g., the set of positive integer numbers \( \mathbb{Z}_+ \) or a sequence of natural numbers \( \{1, 2, \ldots, n\} \).

We consider a probability space \(( \Omega, \mathcal{F}, P )\) and a sequence \( \{ \mathcal{F}(t) \} \) of sub-σ-algebras \( \mathcal{F}(t) \subseteq \mathcal{F} \) with \( t \in I \). As time evolves the available information on the process increases and therefore we consider an increasing sequence of sub-σ-algebras, i.e., \( \mathcal{F}(τ) \subseteq \mathcal{F}(t) \), when τ ≤ t. We call \(( \Omega, \mathcal{F}, \{ \mathcal{F}(t) \}, P )\) a filtered probability space and \( \{ \mathcal{F}(t) \} \) a filtration. A filtered probability space is equivalent to a sequence of probability spaces \(( \Omega, \mathcal{F}(t), P |_{\mathcal{F}(t)} )\), with \( P |_{\mathcal{F}(t)} \) the restriction of \( P \) to the sub-σ-algebra \( \mathcal{F}(t) \). The restricted measure satisfies \( P |_{\mathcal{F}(t)} ( \Phi ) = P ( \Phi |_{\mathcal{F}(t)} ) \) for all \( \Phi \in \mathcal{F}(t) \).

A stochastic process \( X(ω; t) \) is a map \( X : I \times \Omega \to \mathbb{R} \) with \( X(ω; t) \) the value of a physical observable at time \( t \) for a given realization \( ω \) of the process. We say that the stochastic process \( X \) is adapted to the filtration \( \{ \mathcal{F}(t) \} \) when the functions \( X(ω; t) \) are \( \mathcal{F}(t) \)-measurable for any \( t \in I \). In continuous time we generally consider that \( X(ω; t) \) is càdlàg. Càdlàg processes are everywhere right-continuous and have left limits everywhere or in other words \( \lim_{τ→t^-} X(ω; τ) = X(ω; t) \) (for all \( τ > t \)) and \( \lim_{τ→t^-} X(ω; τ) \) exists (for all \( τ < t \)). Càdlàg processes have thus jumps as illustrated in Fig. 12 C.

The average or expectation value of a stochastic process \( X(ω; t) \) with respect to \( P \) is given by the integral \( E_P [ X(t) ] = \int_{ω ∈ \mathcal{F}(t)} X(ω; t) \, dP \). We often write more shortly \( ⟨ X(ω; t) ⟩ \) for the average \( E_P [ X(ω; t) ] \). For a \( \{ \mathcal{F}(t) \} \)-adapted process \( X \) we have:

\[
\langle X(ω; t) \rangle = \int_{ω ∈ \Phi} X(ω; t) \, dP
\]

\[
= \int_{ω ∈ \Phi} X(ω; t) \, dP |_{\mathcal{F}(t)} \tag{B1}
\]

for all sets \( \Phi ∈ \mathcal{F}(t) \).

2. **Martingales processes**

A process \( X(ω; t) \) adapted to the filtration \( \{ \mathcal{F}(t) \} \) is a martingale with respect to \( P \) when \( ⟨ X(ω; t) ⟩ < ∞ \) and

\[
\int_{ω ∈ \Phi} X(ω; τ) \, dP = \int_{ω ∈ \Phi} X(ω; t) \, dP \tag{B2}
\]

for any \( t ∈ I \), \( τ < t \) and \( \Phi ∈ \mathcal{F}(τ) \). A sub-martingale satisfies \( ⟨ X(ω; t) ⟩ < ∞ \) and

\[
\int_{ω ∈ \Phi} X(ω; τ) \, dP \leq \int_{ω ∈ \Phi} X(ω; t) \, dP \tag{B3}
\]

for any \( t ∈ I \), \( τ < t \) and \( \Phi ∈ \mathcal{F}(τ) \).

Martingale processes can also be expressed in terms of conditional expectations. The conditional expectation \( E_P (X(ω; t) | \mathcal{F}(τ)) \) of \( X(ω; t) \) given the σ-algebra \( \mathcal{F}(τ) \) is a \( \mathcal{F}(τ) \)-measurable function that satisfies

\[
\int_{ω ∈ \Phi} E_P (X(ω; t) | \mathcal{F}(τ)) \, dP = \int_{ω ∈ \Phi} X(ω; t) \, dP \tag{B4}
\]

for any \( \Phi ∈ \mathcal{F}(τ) \). The martingale condition Eq. (B2) thus reads

\[
X(ω; τ) = E_P (X(ω; t) | \mathcal{F}(τ)) \tag{B5}
\]

for \( τ < t \).

3. **Radon-Nikodým density process**

The Radon-Nikodým density process of the filtered probability space \(( \Omega, \mathcal{F}, \{ \mathcal{F}(t) \}, P )\) with respect to the filtered probability space \(( \Omega, \mathcal{F}, \{ \mathcal{F}(t) \}, Q )\) is the stochastic process

\[
R(ω; t) = \frac{dP |_{\mathcal{F}(t)}}{dQ |_{\mathcal{F}(t)}}(ω) \tag{B6}
\]
for \( t \in I \). The Radon-Nikodym density process is positive, adapted to the filtration \( \{ \mathcal{F}(t) \} \) and it is martingale process with respect to the measure \( Q \).

We now show that the Radon-Nikodym density process given by Eq. (B6) is a martingale process with respect to \( Q \). Consider two sub-\( \sigma \)-algebras \( \mathcal{F}(\tau) \) and \( \mathcal{F}(t) \) with \( \tau < t \). We first write the measure \( P(\Phi) \) of a set \( \Phi \in \mathcal{F}(\tau) \) as an integral over the probability space \((\Omega, \mathcal{F}(\tau), P|_{\mathcal{F}(\tau)})\)

\[
P(\Phi) = \int_{\omega \in \Phi} dP|_{\mathcal{F}(\tau)} = \int_{\omega \in \Phi} \mathcal{R}(\omega; \tau) dQ|_{\mathcal{F}(\tau)} = \int_{\omega \in \Phi} \mathcal{R}(\omega; t) dQ .
\]  

(B7)

Alternatively, we write the measure \( P(\Phi) \) as an integral over the probability space \((\Omega, \mathcal{F}(t), P|_{\mathcal{F}(t)})\)

\[
P(\Phi) = \int_{\omega \in \Phi} dP|_{\mathcal{F}(t)} = \int_{\omega \in \Phi} \mathcal{R}(\omega; t) dQ|_{\mathcal{F}(t)} = \int_{\omega \in \Phi} \mathcal{R}(\omega; t) dQ .
\]  

(B8)

We thus have the equality

\[
\int_{\omega \in \Phi} \mathcal{R}(\omega; \tau) dQ = \int_{\omega \in \Phi} \mathcal{R}(\omega; t) dQ ,
\]  

(B9)

for all sets \( \Phi \in \mathcal{F}(\tau) \), which is identical to the Eq. (B2) that defines a martingale process. The process \( \mathcal{R}(\omega; t) \) is therefore a martingale process with respect to the measure \( Q \).

4. Stopping times and first-passage times

A stopping time \( T \) is a random variable \( T : \Omega \rightarrow I \cup +\infty : \omega \rightarrow T(\omega) \) for which \( \{ \omega \in \Omega : T(\omega) \leq t \} \) is \( \mathcal{F}(t) \)-measurable for all times \( t \in I \).

First-passage times of a \( \{ \mathcal{F}(t) \} \)-adapted stochastic process \( X(\omega; t) \) are examples of stopping times. The first-passage-time \( T(\omega; A) \) of such a stochastic process \( X(\omega; t) \) determines the earliest time at which the process reaches the measurable set \( A \):

\[
T(\omega; A) = \inf \{ t \in I : X(\omega; t) \in A \}
\]  

(B10)

with \( T(\omega; A) = +\infty \) when the set \( \{ t \in I : X(\omega; t) \in A \} \) is empty. The set \( A \) is measurable if \( \{ \omega \in \Omega : X(\omega; t) \in A \} \in \mathcal{F} \) holds.

We define the sub-\( \sigma \)-algebra \( \mathcal{F}_T \) of \( \mathcal{F} \) corresponding to the stopping time \( T \) as

\[
\mathcal{F}_T = \{ \Phi \in \mathcal{F} : \forall t \in I, \Phi \cap \{ \omega : T(\omega) \leq t \} \in \mathcal{F}(t) \}
\]  

(B11)

We write \( \Phi_{T \leq t} \) for the set of trajectories for which \( T(\omega) \leq t \):

\[
\Phi_{T \leq t} = \{ \omega \in \Omega : T(\omega) \leq t \}
\]  

(B12)

The set \( \Phi_{T \leq t} \) is \( \mathcal{F}_T \)-measurable.

5. Doob’s optional sampling theorem

We consider a martingale process \( X(\omega; t) \) adapted to the filtration \( \{ \mathcal{F}(t) \} \) and two stopping times \( T_1(\omega) \) and \( T_2(\omega) \), with the property that for each \( \omega \in \Omega \), \( T_1(\omega) \leq T_2(\omega) \) and \( T_2(\omega) < \infty \). Doob’s optional sampling theorem states that

\[
\int_{\omega \in \Phi} X(\omega; T_2(\omega)) dP = \int_{\omega \in \Phi} X(\omega; T_1(\omega)) dP
\]  

(B13)

for each set \( \Phi \in \mathcal{F}_{T_1} \). In Doob’s treatise [14], the optional sampling theorem is derived for discrete-time processes. The relation (B13) is also valid for continuous-time processes \( X(\omega; t) \) that are càdlàg and for which the stopping times \( T_1(\omega) \) and \( T_2(\omega) \) are bounded.

To derive the first-passage-time fluctuation theorem we use in Appendix C a particular case of the optional sampling theorem. Consider the filtered probability space \((\Omega, \mathcal{F}(t), \{ \mathcal{F}(\tau) \}, P)\), a \( \{ \mathcal{F}(t) \} \)-adapted martingale process \( X(\omega; t) \) and a stopping time \( T \). We apply Doob’s optional sampling theorem given by Eq. (B13) using \( T_1(\omega) = \min \{ T(\omega), t \} \), \( T_2(\omega) = t \) and the set \( \Phi = \Phi_{T \leq t} \) (which is \( \mathcal{F}_{T_1} \)-measurable):

\[
\int_{\omega \in \Phi_{T \leq t}} X(\omega; t) dP = \int_{\omega \in \Phi_{T \leq t}} X(\omega; T(\omega)) dP
\]  

(B14)

On the right hand side of Eq. (B14) we have used that \( T_1(\omega) = T(\omega) \) for all \( \omega \in \Phi_{T \leq t} \).

6. Doob’s maximal inequality

Doob’s maximal inequality provides an upper bound on the cumulative distribution of the supremum of a positive submartingale process \( X(\omega; t) \) [12], viz.,

\[
Pr(\sup_{t \leq t} X(\omega; t) \geq \lambda) \leq 1 - \frac{\lambda}{\lambda} \mathbb{E}_P[X(\omega; t)]
\]  

(B15)

with \( \lambda \geq 0 \). Doob’s maximal inequality given by Eq. (B15) is valid for discrete-time processes and for continuous-time processes that are càdlàg [14, 15].

Appendix C: Entropy production of a stochastic process

Here we apply the above concepts on measure theory and stochastic processes to trajectories of coarse-grained
variables of a physical process. We further define entropy production of a physical process and show its martingale properties.

1. Probability space of a physical process

We define a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) for which the outcomes \(\omega = (q(\tau), \dot{q}(\tau))_{\tau \in (-\infty, \infty)}\) are the full trajectories of the coarse-grained variables (see also the discussion in Section IIIA and IIIB). The space of finite-time trajectories \(\omega_0^t = (q(\tau), \dot{q}(\tau))_{0 \leq \tau \leq t}\) generates a sub-\(\sigma\)-algebra \(\mathcal{F}^+(t)\) and a corresponding probability space \((\Omega, \mathcal{F}^+(t), \mathbb{P}|_{\mathcal{F}^+(t)})\) for all values \(t \geq 0\). Analogously, the trajectories \(\omega_{-t} = (q(\tau), \dot{q}(\tau))_{-t \leq \tau \leq 0}\) generate a sub-\(\sigma\)-algebra \(\mathcal{F}^-(t)\) and a corresponding probability space \((\Omega, \mathcal{F}^-(t), \mathbb{P}|_{\mathcal{F}^-(t)})\) for all values \(t \geq 0\). The sequences \(\{\mathcal{F}^+(t)\}_{t \geq 0}\) and \(\{\mathcal{F}^-(t)\}_{t \geq 0}\) are filtrations that describe, respectively, future and past events with respect to the reference time point \(t = 0\).

We define the map \(\Theta_t\) as \(\Theta_t = T_t \circ \Theta\). Recall that \(T_t\) is a time-translation map and \(\Theta\) a time-reversal map with respect to the reference point \(t = 0\). These maps are defined, respectively, in subsection IIIA and subsection IIIB. Note that \(\Theta\) maps the past on the future (with respect to the reference point \(t = 0\)) such that \(\Theta((\mathcal{F}^+(t))) = \mathcal{F}^-(t)\) whereas \(\Theta_t\) has the property that it inverts the part of the trajectory in the time window \([0, t]\).

2. Definition of entropy production

We define entropy production over a time interval \([0, t]\) of a stationary stochastic process using the Radon-Nikodým derivative between the forward measure \(\mathbb{P}|_{\mathcal{F}^+(t)}\) and the time-reversed measure \(\mathbb{P} \circ \Theta_t|_{\mathcal{F}^+(t)}\) \([35, 76, 77]\), viz.,

\[
S_{\text{tot}}(\omega; t) = k \ln \frac{d\mathbb{P}|_{\mathcal{F}^+(t)}}{d(\mathbb{P} \circ \Theta_t)|_{\mathcal{F}^+(t)}}(\omega) . \tag{C1}
\]

for \(t \geq 0\). Equation (C1) gives the entropy production of one realization of a physical process over a time interval \([0, t]\), and applies to trajectories of processes in discrete time and continuous time as illustrated in Fig. 12. For stationary processes the measure \(\mathbb{P}\) is time-translation invariant, i.e., \(\mathbb{P} = \mathbb{P} \circ T_t\). The time-reversed measure thus simplifies into \(\mathbb{P} \circ \Theta_t = \mathbb{P} \circ T_t \circ \Theta = \mathbb{P} \circ \Theta\) and we get the steady-state expression of entropy production given by Eq. (13) in the main text (note that in the main text for simplicity we have left the "+" index away in the sub-\(\sigma\)-algebra \(\mathcal{F}^+)\).

We can write an analogous expression for entropy production over the time interval \([-t, 0]\) as:

\[
S_{\text{tot}}(\omega; -t) = k \ln \frac{d\mathbb{P}|_{\mathcal{F}^-(t)}}{d(\mathbb{P} \circ \Theta_{-t})|_{\mathcal{F}^-(t)}}(\omega) . \tag{C2}
\]

for \(t \geq 0\).

3. Entropy production under time reversal

Entropy production is a stochastic process that is odd under time-reversal such that \(S_{\text{tot}}(\Theta(\omega); t) = -S_{\text{tot}}(\omega; -t)\). Another interesting property is the change of sign of entropy under the operation \(\Theta_t\), viz.,

\[
S_{\text{tot}}(\Theta_t(\omega); t) = k \ln \frac{d\mathbb{P}|_{\mathcal{F}^+(t)}}{d(\mathbb{P} \circ \Theta_t)|_{\mathcal{F}^+(t)}}(\Theta_t(\omega))
= k \ln \frac{d(\mathbb{P} \circ \Theta_t)|_{\mathcal{F}^+(t)}}{d(\mathbb{P} \circ \Theta_t)|_{\mathcal{F}^+(t)}}(\omega)
= k \ln \frac{d(\mathbb{P} \circ \Theta_t)|_{\mathcal{F}^+(t)}}{d(\mathbb{P} \circ \Theta_t)|_{\mathcal{F}^+(t)}}(\omega)
= -S_{\text{tot}}(\omega; t) , \tag{C3}
\]

where we have used that \(\Theta_t = \Theta_t^{-1}\).

4. The exponential of entropy production is a martingale process

The exponential of minus the entropy production \(e^{-S_{\text{tot}}(t)/k}\) is a martingale with respect to the measure \(\mathbb{P}\) and \(e^{S_{\text{tot}}(t)/k}\) is a martingale with respect to the time-reversed measure \(\mathbb{P} \circ \Theta\). The martingale property is generally valid for stationary measures \(\mathbb{P}\). We use this martingale property to derive the first-passage-time fluctuation theorem and the infimum law. In fact, the martingale property is the only property of \(S_{\text{tot}}(t)\) we use to derive the infimum law.

The exponential of minus the entropy production \(e^{-S_{\text{tot}}(t)/k}\) equals to

\[
e^{-S_{\text{tot}}(t)/k} = \frac{d\mathbb{P}|_{\mathcal{F}(t)}}{d(\mathbb{P} \circ \Theta)|_{\mathcal{F}(t)}} . \tag{C4}
\]

for \(t \geq 0\). Equation (C4) follows from the definition of entropy production given by Eq. (13). From Eq. (C4) we see that \(e^{-S_{\text{tot}}(t)/k}\) is the Radon-Nikodým density process of the filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}(t)\}, \mathbb{P})\) with respect to the filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}(t)\}, \mathbb{P} \circ \Theta)\). The process \(e^{-S_{\text{tot}}(t)/k}\) is thus a martingale process because Radon-Nikodým density processes are martingales, as shown in Appendix B.3 For analogous reasons \(e^{S_{\text{tot}}(t)/k}\) is a martingale with respect to the time-reversed measure \(\mathbb{P} \circ \Theta\).

Appendix D: Derivation of the first-passage-time fluctuation theorem

In this section we derive the cumulative first-passage-time fluctuation theorem given by Eq. (14) in Section IV.
i.e.,

\[ \frac{\mathbb{P}(\Phi_{T(\omega_{\text{tot}})} \leq t)}{\mathbb{P}(\Theta_{T(\omega_{\text{tot}})} (\Phi_{T(\omega_{\text{tot}})} \leq t))} = e^{s_{\text{tot}}/k}, \quad (D1) \]

for \( s_{\text{tot}} \)-stopping times \( T(\omega_{\text{tot}}) \). The map \( \Theta_{T(\omega_{\text{tot}})} \) is the time-reversal map \( \Theta_t = T_t \circ \Theta \) applied at the stopping time \( t = T(\omega_{\text{tot}}) \). Nota the in the main text we write simply \( T \) for \( T(\omega_{\text{tot}}) \). In the next subsection we define the \( s_{\text{tot}} \)-stopping times \( T(\omega_{\text{tot}}) \) and we show that first-passage times for entropy production are also \( s_{\text{tot}} \)-stopping times. We then derive the first-passage-time fluctuation theorem in the subsequent subsection.

1. First-passage times as \( s_{\text{tot}} \)-stopping times

We call \( T(\omega; s_{\text{tot}}) \) an \( s_{\text{tot}} \)-stopping time when entropy production reaches a value \( s_{\text{tot}} \) when the process terminates:

\[ S_{\text{tot}}(\omega; T(\omega; s_{\text{tot}})) = s_{\text{tot}} \quad . \tag{D2} \]

For continuous processes, an example of an \( s_{\text{tot}} \)-stopping time is a first-passage time \( T_{FP}^{(1)}(\omega; s_{\text{tot}}) \) with one absorbing boundary for which the process terminates when entropy production reaches or exceeds for the first time a positive threshold \( s_{\text{tot}} > 0 \):

\[ T_{FP}^{(1)}(\omega; s_{\text{tot}}) = \inf \{ t \in \mathbb{R}_+ \cup \{ \infty \} : S_{\text{tot}}(\omega; t) \geq s_{\text{tot}} \} \quad . \tag{D3} \]

When the process does not terminate the first-passage time equals to infinity \( T_{FP}^{(1)}(\omega; s_{\text{tot}}) = +\infty \). Analogously, we define a first-passage time for the entropy production to reach or fall below a negative threshold \(-s_{\text{tot}}\):

\[ T_{FP}^{(1)}(\omega; -s_{\text{tot}}) = \inf \{ t \in \mathbb{R}_+ \cup \{ \infty \} : S_{\text{tot}}(\omega; t) \leq -s_{\text{tot}} \} \quad . \tag{D4} \]

As a second example we consider a first-passage time \( T_{FP}^{(2)}(\omega; s_{\text{tot}}) \) of entropy production with two absorbing boundaries at \( s_{\text{tot}} \) and \(-s_{\text{tot}}\) (with \( s_{\text{tot}} > 0 \)). We set \( T_{FP}^{(2)}(\omega; s_{\text{tot}}) = T_{FP}^{(1)}(\omega; s_{\text{tot}}) \) for \( T_{FP}^{(1)}(\omega; s_{\text{tot}}) < T_{FP}^{(2)}(\omega; s_{\text{tot}}) \) and \( T_{FP}^{(2)}(\omega; s_{\text{tot}}) = +\infty \) for \( T_{FP}^{(1)}(\omega; s_{\text{tot}}) > T_{FP}^{(2)}(\omega; s_{\text{tot}}) \). Analogously we define a \(-s_{\text{tot}}\)-stopping time \( T_{FP}^{(2)}(\omega; -s_{\text{tot}}) \). We set \( T_{FP}^{(2)}(\omega; -s_{\text{tot}}) = T_{FP}^{(1)}(\omega; -s_{\text{tot}}) \) for \( T_{FP}^{(1)}(\omega; -s_{\text{tot}}) < T_{FP}^{(2)}(\omega; -s_{\text{tot}}) \) and \( T_{FP}^{(2)}(\omega; -s_{\text{tot}}) = +\infty \) for \( T_{FP}^{(1)}(\omega; -s_{\text{tot}}) > T_{FP}^{(2)}(\omega; -s_{\text{tot}}) \). Note that other \( s_{\text{tot}} \)-stopping times can be defined such as the times of second passage.

Recall that the set \( \Phi_{T(s_{\text{tot}}) \leq t} \) is the \( \mathcal{F}(t) \)-measurable set of all trajectories \( \omega \) that terminate before time \( t \), i.e., \( \Phi_{T(s_{\text{tot}}) \leq t} = \{ \omega \in \Omega : T(\omega; s_{\text{tot}}) \leq t \} \). We define a corresponding conjugate set as \( \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}}) \leq t}) = \{ \omega \in \Omega : T(\Theta_{T(s_{\text{tot}})}(\omega); s_{\text{tot}}) \leq t \} \) using the time-reversal map \( \Theta_{T(s_{\text{tot}})} \) which reverses in time the trajectory \( \omega \) and translates this trajectory over its stopping time \( T(\omega; s_{\text{tot}}) \). The map \( \Theta_{T(s_{\text{tot}})} \) is thus the concatenation \( \Theta_{T(s_{\text{tot}})} = T_{s_{\text{tot}}} \circ \Theta \). For the first-passage times \( T_{FP}^{(1)}(\omega; s_{\text{tot}}) \) and \( T_{FP}^{(2)}(\omega; s_{\text{tot}}) \) we find using Eq. (C3) : \( \Theta_{T_{FP}^{(1)}(s_{\text{tot}})} (\Phi_{T_{FP}^{(1)}(s_{\text{tot}}) \leq t}) = \Phi_{T_{FP}^{(1)}(-s_{\text{tot}}) \leq t} \) and \( \Theta_{T_{FP}^{(2)}(s_{\text{tot}})} (\Phi_{T_{FP}^{(2)}(s_{\text{tot}}) \leq t}) = \Phi_{T_{FP}^{(2)}(-s_{\text{tot}}) \leq t} \). In other words, the time-reversal operation \( \Theta_{T_{FP}^{(1)}(s_{\text{tot}})} \) on the set \( \Phi_{T_{FP}^{(1)}(s_{\text{tot}})} \) and the time-reversal operation \( \Theta_{T_{FP}^{(2)}(s_{\text{tot}})} \) on the set \( \Phi_{T_{FP}^{(2)}(s_{\text{tot}})} \) are equivalent to a change of the sign in the threshold value from \( s_{\text{tot}} \) to \(-s_{\text{tot}}\).
2. First-passage-time fluctuation relation

The first-passage-time fluctuation theorem given by Eq. (D1) is derived from the following identities:

\[
\frac{\mathbb{P} (\Phi_{T(s_{\text{tot}})} \leq t)}{\mathbb{P} (\Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t))} = \int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} d\mathbb{P}|\mathcal{F}(t) \tag{D5}
\]

\[
\int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) \tag{D7}
\]

\[
\int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) = \int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) \tag{D8}
\]

\[
\int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) = \int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) \tag{D9}
\]

\[
\int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) = \int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) \tag{D10}
\]

\[
\int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) = \int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) \tag{D11}
\]

\[
\int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) = \int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) \tag{D12}
\]

\[
\int_{\omega \in \Theta_{T(s_{\text{tot}})} (\Phi_{T(s_{\text{tot}})} \leq t)} e^{S_{\text{tot}}(\omega; t)/k} d(\mathbb{P} \circ \Theta)|\mathcal{F}(t) = e^{S_{\text{tot}}(\omega; t)/k} \tag{D13}
\]

In Eq. (D6) we write the measures of the \(\mathcal{F}(t)\) measurable sets \(\Theta_{T(s_{\text{tot}})} \leq t\) and \(\Theta_{T(s_{\text{tot}})} \leq t\) in terms of an integral over a probability space (see Eq. (A1) in appendix A). In Eq. (D7) we first use the Radon-Nikodým theorem Eq. (A6) to change the integral over the measure \(\mathbb{P}\) to an integral over the measure \(\mathbb{P} \circ \Theta\) and then use our measure-theoretic definition of entropy production Eq. (13) to write the Radon-Nikodým derivative in terms of entropy production \(S_{\text{tot}}(t)\). Since \(e^{S_{\text{tot}}(t)/k}\) is a \(\mathbb{P} \circ \Theta\)-martingale we use in Eq. (D8) Doob’s optional sampling theorem Eq. (B15). In order to apply this theorem we additionally need that \(e^{S_{\text{tot}}(t)}\) is càdlàg. In Eq. (D9) we use the fact that \(T(\omega; s_{\text{tot}})\) is an \(s_{\text{tot}}\)-stopping time such that \(s_{\text{tot}} = S_{\text{tot}}(\omega; T(\omega; s_{\text{tot}}))\). In Eq. (D10) we have changed the variables in the integral using Eq. (A5) with the measurable morphism \(\Theta_{T(s_{\text{tot}})}\). In Eq. (D11) we use the definition of \(\Theta_{T(s_{\text{tot}})} = \Theta \circ T(s_{\text{tot}})\) as defined in the previous subsection. In Eq. (D12) we use that \(\mathbb{P}\) is a stationary measure and thus \(\mathbb{P} = \mathbb{P} \circ T(s_{\text{tot}})\). In the last Eq. (D13) we use that \(\Phi_{T(s_{\text{tot}})}\) has non-zero measure and \(da/a = 1\) for \(a \neq 0\).

Appendix E: Bounds on the expectation value using stochastic dominance

Consider two positive valued random variables \(X \geq 0\) and \(Y \geq 0\). We define the cumulative distributions as:

\[
F_X(x) = \Pr(X \leq x) \tag{E1}
\]

\[
F_Y(y) = \Pr(Y \leq x) \tag{E2}
\]

We say that \(X\) dominates \(Y\) stochastically when the cumulative distribution functions of \(X\) and \(Y\) satisfy the relation \(F_X(x) \geq F_Y(x)\). When \(X\) dominates \(Y\) stochastically we have that the mean value of \(X\) is smaller than the mean value of \(Y\): \(\langle X \rangle \leq \langle Y \rangle\). This follows directly from the relation \(\langle X \rangle = \int_0^\infty d\tau (1 - F_X(x))\) between the mean and the cumulative distribution.

Appendix F: Infimum of the position of a drift-diffusion process

The probability of the supremum value of a stochastic process at time \(t\) to be less or equal than \(L > 0\) equals to the survival probability of the process to lie in the interval \((-\infty, L)\) at time \(t\) [73, 79]:

\[
\Pr (\sup X(t) \leq L) = \Pr(X(s) \leq L; s \leq t) = Q_X(L, t). \tag{F1}
\]

For a general stochastic process, the survival probability in an interval can be calculated from the first-passage-time distribution as follows:

\[
Q_X(L, t) = 1 - \int_0^t d\tau p_T(\tau; L). \tag{F2}
\]

We now consider two conjugate drift-diffusion processes:

1. \(X_+(t)\) with velocity \(v\), diffusion \(D\), and initial condition \(X_+(0) = 0\)

2. \(X_-(t)\) with velocity \(-v\) and diffusion \(D\), and initial condition \(X_-(0) = 0\)

with \(v > 0\). The infimum value of \(X_+(t)\) equals to minus the supremum of the conjugate process \(X_-(t)\). In the following, we derive analytical expressions for the statistics of the supremum of \(X_-(t)\) that can be then used to obtain the statistics of the infimum of \(X_+(t)\).
The survival probability of \( X_-(t) \) can be obtained from its first-passage-time distribution to reach \( L \), which is given by

\[
pr(t; L) = \frac{L}{\sqrt{4\pi D t^3}} e^{-(L+vt)^2/4D t} .
\]  

Replacing Eq. (F3) in (F2) results in the following expression for the survival probability of \( X_- \):

\[
Q_{X_-}(L, t) = 1 - \frac{1}{2} \left[ \text{erfc} \left( \frac{L + vt}{\sqrt{4Dt}} \right) + e^{-vL/D} \text{erfc} \left( \frac{L - vt}{\sqrt{4Dt}} \right) \right] .
\]  

where \( \text{erfc} \) is the complementary error function. Equation (F4) yields the cumulative density function of the supremum of \( X_- \), as follows from Eq. (F1). From the relation between the conjugate processes, we find that

\[
\text{Pr}(\inf X_+(t) \leq L) = \text{Pr}(\inf X_+(t) \geq -L) = \text{Pr}(\sup X_-(t) \leq L) = Q_{X_-}(L, t) .
\]  

Using Eq. (F4) and the property \( \text{erfc}(x) + \text{erfc}(-x) = 2 \), we obtain an analytical expression for the cumulative distribution of the infimum of the position of a drift-diffusion process with positive velocity:

\[
\text{Pr}(\inf X_+(t) \leq L) = 1 - \frac{1}{2} \left[ \text{erfc} \left( \frac{-L - vt}{\sqrt{4Dt}} \right) - e^{-vL/D} \text{erfc} \left( \frac{L - vt}{\sqrt{4Dt}} \right) \right] .
\]  

From the definition of the entropy production in the drift-diffusion process given by Eq. (31) and using Eq. (30), we find the following expression for the time evolution of the entropy production:

\[
\frac{1}{k} \frac{dS_{\text{tot}}(t)}{dt} = \frac{v^2}{D} + \eta(t) ,
\]  

with \( \langle \eta(t) \rangle = 0 \) and \( \langle \eta(t)\eta(t') \rangle = 2(v^2/D)\delta(t-t') \). Equation (F7) can be seen as a Langevin equation with effective drift velocity \( v_{\text{eff}} = v^2/D \) and effective diffusion coefficient \( D_{\text{eff}} = v^2/D \).

For the stochastic process \( S_{\text{tot}}(t)/k = (v/D)X_+(t) \) the infimum distribution can be obtained by replacing in (F6) \( v \) and \( D \) by its effective values for the process \( S_{\text{tot}}(t)/k \), given by \( v_{\text{eff}} = v^2/D \) and \( D_{\text{eff}} = v^2/D \). Defining \( \sigma(t) = (S_{\text{tot}}(t)/k) = (v^2/D)t \) we obtain:

\[
\text{Pr} \left( - \frac{\inf S_{\text{tot}}(t)}{k} \leq s \right) = 1 - \frac{1}{2} \left[ \text{erfc} \left( \frac{-s - \sigma(t)}{\sqrt{2\sigma(t)}} \right) - e^{-s} \text{erfc} \left( \frac{s - \sigma(t)}{\sqrt{2\sigma(t)}} \right) \right] ,
\]  

which equals to Eq. (F5).

We now derive an analytical expression for the mean infimum of \( X_+ \), which equals to minus the supremum of the conjugate process \( X_- \):

\[
\langle \inf X_+(t) \rangle = - \langle \sup X_-(t) \rangle .
\]  

The cumulative distribution of the supremum of \( X_- \) is given by

\[
\text{Pr}(\sup X_-(t) \leq L) = Q_{X_-}(L, t) = 1 - \frac{1}{2} \left[ \text{erfc} \left( \frac{-L - vt}{\sqrt{4Dt}} \right) - e^{-vL/D} \text{erfc} \left( \frac{L - vt}{\sqrt{4Dt}} \right) \right] ,
\]  

where we have used Eq. (F4) and the property \( \text{erfc}(x) + \text{erfc}(-x) = 2 \). The distribution of the supremum of \( X_- \) can be found deriving Eq. (F10) with respect to \( L \), which yields:

\[
\text{Pr}(\sup X_-(t) = L) = \frac{1}{\sqrt{\pi Dt}} e^{-v^2 t/4D} + \frac{v}{2D} e^{-vL/D} \text{erfc} \left( \frac{v\sqrt{t}}{2\sqrt{D}} \right) .
\]  

The mean of the supremum of \( X_- \) can be calculated integrating its probability distribution

\[
\langle \sup X_-(t) \rangle = \int_0^\infty dL \text{Pr}(\sup X_-(t) = L) L ,
\]  

which after some algebra yields

\[
\langle \sup X_-(t) \rangle = \frac{D}{v} \text{erf} \left( \frac{v\sqrt{t}}{2D} \right) - \frac{v}{2D} \text{erfc} \left( \frac{v\sqrt{t}}{2D} \right) + \sqrt{\frac{Dt}{\pi}} e^{-v^2 t/4D} .
\]  

From Eqs. (F9) and (F14) we find an exact expression for the mean infimum of a drift-diffusion process with positive velocity:

\[
\langle \inf X_+(t) \rangle = -\frac{D}{v} \text{erf} \left( \frac{v\sqrt{t}}{2D} \right) + \frac{v}{2D} \text{erfc} \left( \frac{v\sqrt{t}}{2D} \right) - \sqrt{\frac{Dt}{\pi}} e^{-v^2 t/4D} .
\]  

Replacing \( v \) by \( v_{\text{eff}} = v^2/D \) and \( D \) by \( D_{\text{eff}} = v^2/D \) and using \( \sigma(t) = (S_{\text{tot}}(t)/k) = (v^2/D)t \), we obtain an analytical expression for the mean infimum of entropy production in a drift-diffusion process at time \( t \):

\[
\langle \inf S_{\text{tot}}(t)/k \rangle = -\text{erf} \left( \frac{\sqrt{\sigma(t)}}{2} \right) + \frac{\sigma(t)}{2D} \text{erfc} \left( \frac{\sqrt{\sigma(t)}}{2D} \right) - \sqrt{\frac{\sigma(t)}{\pi}} e^{-\sigma(t)/4D} ,
\]  

which equals to Eq. (40).