

STOCHASTIC THERMODYNAMICS

## Integral fluctuation relations for entropy production at stopping times

To cite this article: Izaak Neri *et al* *J. Stat. Mech.* (2019) 104006

View the [article online](#) for updates and enhancements.



**IOP** | ebooks™

Bringing you innovative digital publishing with leading voices  
to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of  
every title for free.

# Integral fluctuation relations for entropy production at stopping times

Izaak Neri<sup>1,5</sup>, Édgar Roldán<sup>2</sup>, Simone Pigolotti<sup>3</sup>  
and Frank Jülicher<sup>4</sup>

<sup>1</sup> Department of Mathematics, King's College London, Strand, London, WC2R 2LS, United Kingdom

<sup>2</sup> ICTP—The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, 34151 Trieste, Italy

<sup>3</sup> Biological Complexity Unit, Okinawa Institute for Science and Technology and Graduate University, Onna, Okinawa 904-0495, Japan

<sup>4</sup> Max Planck Institute for the Physics of Complex Systems, Nöthnitzerstrasse 38, 01187 Dresden, Germany

E-mail: [izaak.neri@kcl.ac.uk](mailto:izaak.neri@kcl.ac.uk), [edgar@ictp.it](mailto:edgar@ictp.it), [simone.pigolotti@oist.jp](mailto:simone.pigolotti@oist.jp) and [julicher@pks.mpg.de](mailto:julicher@pks.mpg.de)

Received 20 March 2019

Accepted for publication 29 August 2019

Published 16 October 2019



Online at [stacks.iop.org/JSTAT/2019/104006](https://stacks.iop.org/JSTAT/2019/104006)

<https://doi.org/10.1088/1742-5468/ab40a0>

**Abstract.** A stopping time  $T$  is the first time when a trajectory of a stochastic process satisfies a specific criterion. In this paper, we use martingale theory to derive the integral fluctuation relation  $\langle e^{-S_{\text{tot}}(T)} \rangle = 1$  for the stochastic entropy production  $S_{\text{tot}}$  in a stationary physical system at stochastic stopping times  $T$ . This fluctuation relation implies the law  $\langle S_{\text{tot}}(T) \rangle \geq 0$ , which states that it is not possible to reduce entropy on average, even by stopping a stochastic process at a stopping time, and which we call the second law of thermodynamics at stopping times. This law bounds the average amount of heat and work a system can extract from its environment when stopped at a random time. Furthermore, the integral fluctuation relation implies that certain fluctuations of entropy production are universal or are bounded by universal functions. These universal properties descend from the integral fluctuation relation by selecting appropriate stopping times: for example, when  $T$  is a first-passage time for entropy production, then we obtain a bound on the statistics of negative records of entropy production. We illustrate these results on simple models of nonequilibrium systems described by Langevin equations and reveal two interesting phenomena. First, we demonstrate

<sup>5</sup> Author to whom any correspondence should be addressed.

Integral fluctuation relations for entropy production at stopping times

that isothermal mesoscopic systems can extract on average heat from their environment when stopped at a cleverly chosen moment and that the second law at stopping times bounds the average extracted heat. Second, we demonstrate that the efficiency at stopping times of an autonomous stochastic heat engine, such as Feynmann's ratchet, can be larger than the Carnot efficiency and that the second law of thermodynamics at stopping times bounds its efficiency at stopping times.

**Keywords:** fluctuation theorems, stochastic processes, stochastic thermodynamics

## Contents

<b>1. Introduction and statement of the main results</b>	<b>3</b>
<b>2. Preliminaries and notation</b>	<b>8</b>
<b>3. Martingales</b>	<b>8</b>
3.1. Gambling and martingales.....	9
3.2. Martingales and the optional stopping theorem .....	11
<b>4. Stochastic thermodynamics for stationary processes</b>	<b>12</b>
<b>5. Integral fluctuation relations at stopping times</b>	<b>14</b>
5.1. The martingale structure of exponential entropy production .....	14
5.2. Fluctuation relation at stopping times within a finite-time window.....	14
5.3. Fluctuation relation at stopping times within an infinite-time window.....	15
5.4. Conditional integral fluctuation relation at stopping times.....	15
<b>6. Second law of thermodynamics at stopping times</b>	<b>15</b>
6.1. Bounds on heat absorption in isothermal processes.....	16
6.2. Efficiency of heat engines at stopping times: the case of Feynman's ratchet .....	16
<b>7. Universal properties of stochastic entropy production</b>	<b>18</b>
7.1. Fluctuation properties at fixed time $T = t$ .....	18
7.2. Splitting probabilities of entropy production.....	19
7.3. Infima of entropy production.....	20
7.4. Survival probability of entropy production .....	20
7.5. Number of crossings.....	21
<b>8. The influence of finite statistics on the integral fluctuation relation at stopping times</b>	<b>22</b>
<b>9. Examples</b>	<b>22</b>
9.1. Heat extraction in an isothermal system .....	23
9.2. Illustration of an empirical test of the integral fluctuation relation at stopping times.....	24

9.3. Super Carnot efficiency for heat engines at stopping times .....	24
<b>10. Discussion</b>	<b>30</b>
<b>Acknowledgments</b> .....	<b>32</b>
<b>Appendix A. The exponential of the negative entropy production is a (uniformly integrable) martingale</b> .....	<b>32</b>
A.1. Reversible processes in discrete time .....	32
A.2. Reversible stationary processes that are right-continuous .....	32
A.3. Uniform integrability .....	33
<b>Appendix B. Integral fluctuation relation for entropy production within infinite-time windows</b> .....	<b>33</b>
<b>Appendix C. Derivation of the bound (71) on the statistics of <math>N_x</math></b> .....	<b>35</b>
<b>References</b>	<b>36</b>

---

## 1. Introduction and statement of the main results

Stochastic thermodynamics is a thermodynamics theory for the slow degrees of freedom  $\vec{X}(t)$  of a mesoscopic system that is weakly coupled to an environment in equilibrium [1–5]. Examples of systems to which stochastic thermodynamics applies are molecular motors [6, 7], biopolymers [8], self-propelled Brownian particles [9], micro-manipulation experiments on colloidal particles [10–12], and electronic circuits [13–15].

The stochastic entropy production  $S_{\text{tot}}(t)$  is a key variable in stochastic thermodynamics. It is defined as the sum of the entropy change of the environment  $S_{\text{env}}(t)$  and a system entropy change  $\Delta S_{\text{sys}}(t)$  [16]. In stochastic thermodynamics entropy production is a functional of the trajectories of the slow degrees of freedom in the system. If time is discrete and  $\vec{X}(t)$  is a variable of even parity with respect to time reversal, then the entropy production  $S_{\text{tot}}(t)$  associated with a trajectory of a nonequilibrium stationary process  $\vec{X}(t)$  is the logarithm of the ratio between the stationary probability density of that trajectory  $p(\vec{X}(1), \vec{X}(2), \dots, \vec{X}(t))$  and the probability density of the same trajectory but in time-reversed order [4, 17–19],

$$S_{\text{tot}}(t) = \log \frac{p(\vec{X}(1), \vec{X}(2), \dots, \vec{X}(t))}{p(\vec{X}(t), \vec{X}(t-1), \dots, \vec{X}(1))}, \quad (1)$$

where  $\log$  denotes natural logarithm. Here and throughout the paper we use dimensionless units for which Boltzmann's constant  $k_B = 1$ . Equation (1) is a particular case of the general expression of stochastic entropy production in terms of probability measures that we will discuss below in equation (33). The functional  $S_{\text{tot}}(t)$  is exactly equal to zero at all times for systems in equilibrium. For nonequilibrium systems, entropy production fluctuates with expected value larger than zero,  $\langle S_{\text{tot}}(t) \rangle \geq 0$ .

An interesting consequence of definition (1) is that the exponential of the negative entropy production  $e^{-S_{\text{tot}}(t)}$  is a *martingale* associated with the process  $\vec{X}(t)$  [20–22]. Historically the concept of martingales has been introduced to understand fundamental

questions in betting strategies and gambling [23]. Martingale theory [24–26], and in particular Doob’s optional stopping theorem, provides an elegant resolution to the question whether it is possible to make profit in a fair game of chance by leaving the game at a cleverly chosen moment. We can distinguish unfair games of chances, where the expected values of the gambler’s fortune decreases (or increases) with time, from fair ones, where such expected values are constant in time on average. In probability theory, these categories correspond to supermartingales (or submartingales) and martingales, respectively. In a nutshell, the optional stopping theorem for martingales states that a gambler cannot make profit on average in a fair game of chance by quitting the game at an intelligently chosen moment. The optional stopping theorem holds as long as the total amount of available money is finite. A gambler with access to an infinite budget of money could indeed devise a betting strategy that makes profit out of a fair game; the St. Petersburg game provides an example of a such a strategy, see [27] and chapter 6 of [28]. Nowadays martingales have various applications, for example, they model stock prices in efficient capital markets [29].

In this paper, we study universal properties of entropy production in nonequilibrium stationary states using martingale theory and Doob’s optional stopping theorem. In an analogy with gambling, the negative entropy production  $-S_{\text{tot}}(t)$  of a stationary process  $\vec{X}(t)$  is equivalent to a gambler’s fortune in an unfair game of chance  $\vec{X}(t)$  and the exponentiated negative entropy production  $e^{-S_{\text{tot}}(t)}$  is a martingale associated with the gambler’s fortune. In stochastic thermodynamics, Doob’s optional stopping theorem implies

$$\langle e^{-S_{\text{tot}}(T)} \rangle = 1, \quad (2)$$

where the expected value  $\langle \cdot \rangle$  is over many realizations of the physical process  $\vec{X}(t)$ , and where  $T$  is a stopping time. A stopping time  $T$  is the first time when a trajectory of  $\vec{X}$  satisfies a specific criterium; it is thus a stochastic time. This criterium must obey causality and cannot depend on the future. The relation (2) holds under the condition that either  $T$  acts in a finite-time window, i.e.  $T \in [0, \tau]$  with  $\tau$  a positive number, or that  $S_{\text{tot}}(t)$  is bounded for all times  $t \in [0, T]$ . We call (2) the *integral fluctuation relation for entropy production at stopping times* because it is an integral relation,  $\langle e^{-S_{\text{tot}}(T)} \rangle = \int d\mathbb{P} e^{-S_{\text{tot}}(T)} = 1$ , that characterises the fluctuations of entropy production. Here,  $\mathbb{P}$  is the probability measure associated with  $\vec{X}(t)$ .

The fluctuation relation at stopping times (2) can be extended into a fluctuation relation conditioned on trajectories  $\vec{X}_0^{T'}$  of random duration  $[0, T']$ , namely,

$$\left\langle e^{-S_{\text{tot}}(T)} \middle| \vec{X}_0^{T'} \right\rangle = e^{-S_{\text{tot}}(T')}, \quad (3)$$

with  $T'$  a stopping time for which  $T' \leq T$ , and where  $\vec{X}_0^s = \left\{ \vec{X}(t') \right\}_{t' \in [0, s]}$  denotes a trajectory in a finite-time window. Notice that for  $T' = 0$  we obtain  $\langle e^{-S_{\text{tot}}(T)} \mid \vec{X}(0) \rangle = e^{-S_{\text{tot}}(0)} = 1$ , since in our definition  $S_{\text{tot}}(0) = 0$ . The fluctuation relation (3) implies thus (2).

There are two important implications of the integral fluctuation relations (2) and (3). First, it holds that

$$\langle S_{\text{tot}}(T) \rangle \geq 0, \quad (4)$$

or in other words, it is not possible to reduce the average entropy production by stopping a stochastic process  $\vec{X}(t)$  at a cleverly chosen moment  $T$  that can be different for each realisation. The relation (4) is reminiscent of the second law of thermodynamics, and therefore we call it the *second law of thermodynamics at stopping times*. A second implication of the integral fluctuation relations (2) and (3) is that certain fluctuation properties of entropy production are *universal*. In what follows, we discuss in more detail these two consequences of the integral fluctuation relations.

We first discuss the second law at stopping times equation (4). Remarkably, this second law holds for *any* stopping time  $T$  defined by a (arbitrary) stopping criterium that obeys causality and does not use information about the future of the physical process; first-passage times are canonical examples of stopping times. Interestingly, the inequality (4) bounds the average amount of heat and work a system can perform on or extract from its surroundings at a stopping time  $T$ . For isothermal systems, equation (4) implies

$$\langle Q(T) \rangle \leq T_{\text{env}} \langle \Delta S_{\text{sys}}(T) \rangle. \quad (5)$$

Relation (5) states that the average amount of heat  $\langle Q(T) \rangle$  a system can on average extract at a stopping time  $T$  from a thermal reservoir at temperature  $T_{\text{env}}$  is smaller or equal than the average system entropy difference  $\langle \Delta S_{\text{sys}}(T) \rangle$  between the initial state and the final state at the stopping time. Similar considerations allow us to derive bounds on the average amount of work that a stationary heat engine, e.g. Feynman's ratchet, can extract from its surrounding when stopped at a cleverly chosen moment. Consider a system in contact with two thermal reservoirs at temperatures  $T_h$  and  $T_c$  with  $T_h \geq T_c$ . We define the *stopping-time efficiency*  $\eta_T$  associated with the stopping time  $T$  as

$$\eta_T := -\frac{\langle W(T) \rangle}{\langle Q_h(T) \rangle}, \quad (6)$$

where  $\langle W(T) \rangle$  is the average work exerted on the system in the time interval  $[0, T]$ , and  $\langle Q_h(T) \rangle$  is the average heat absorbed by the system from the hot reservoir within the same time interval. If  $\langle Q_h(T) \rangle > 0$ , then the second law at stopping times (4) implies that

$$\eta_T \leq \eta_C - \frac{\langle \Delta \mathcal{F}_c(T) \rangle}{\langle Q_h(T) \rangle}, \quad (7)$$

where  $\eta_C = 1 - (T_c/T_h)$  is the Carnot efficiency,  $\langle \Delta \mathcal{F}_c(t) \rangle = \langle \Delta v(t) \rangle - T_c \langle \Delta S_{\text{sys}}(t) \rangle$  is the generalised free energy change of the system at the stopping time  $T$ , and  $\Delta v(t) = v(X(t)) - v(X(0))$  is change of the internal energy of the system. Note that the second term in the right-hand side of (7) can be positive, and thus efficiencies at stopping times of stationary heat engines can be greater than the Carnot efficiency. This is because using a stopping time the system is, in general, no longer cyclic.

We now discuss *universal* properties of the fluctuations of entropy production. By applying the integral fluctuation relations (2) and (3) to different examples of stopping times  $T$ , we will derive the following generic relations for the fluctuations of entropy production:

Integral fluctuation relations for entropy production at stopping times

- In the simple case  $T = t$ , where  $t$  is a deterministic fixed time, relation (4) reads  $\langle S_{\text{tot}}(t) \rangle \geq 0$ , which is a well-known second-law-like relation derived in stochastic thermodynamics [4, 16], and the relation (2) is the stationary integral fluctuation relation  $\langle e^{-S_{\text{tot}}(t)} \rangle = 1$  [4, 16]. The integral fluctuation relation at fixed times implies that in a nonequilibrium process events of negative entropy production must exist and their likelihood is bounded by [4]

$$\mathbb{P}(S_{\text{tot}}(t) \leq -s) \leq e^{-s}, \quad s \geq 0, \quad (8)$$

where  $\mathbb{P}(\cdot)$  denotes the probability of an event.

- Our second choice of stopping times are first-passage times  $T_{\text{fp}} = \inf\{t : S_{\text{tot}}(t) \notin (-s_-, s_+)\}$  for entropy production with two absorbing boundaries at  $-s_- \leq 0$  and  $s_+ \geq 0$ . As we show in this paper, the integral fluctuation relation equation (2) implies that the splitting probabilities  $p_- = \mathbb{P}[S_{\text{tot}}(T) \leq -s_-]$  and  $p_+ = \mathbb{P}[S_{\text{tot}}(T) \geq s_+]$  are bounded by

$$p_+ \geq 1 - \frac{1}{e^{s_-} - e^{-s_+}}, \quad p_- \leq \frac{1}{e^{s_-} - e^{-s_+}}. \quad (9)$$

If the trajectories of entropy production are continuous, then [21]

$$p_+ = \frac{e^{s_-} - 1}{e^{s_-} - e^{-s_+}}, \quad p_- = \frac{1 - e^{-s_+}}{e^{s_-} - e^{-s_+}}. \quad (10)$$

- Global infima of entropy production,  $S_{\text{inf}} = \inf_{t \geq 0} S_{\text{tot}}(t)$ , quantify fluctuations of negative entropy production. The cumulative distribution  $\mathbb{P}[S_{\text{inf}} \leq -s]$  is equal to the splitting probability  $p_-$  in the limit  $s_+ \rightarrow \infty$ . Using (9) we obtain [21]

$$\mathbb{P}[S_{\text{inf}} \leq -s] \leq e^{-s}, \quad s \geq 0, \quad (11)$$

which implies the *infimum law*  $\langle S_{\text{inf}} \rangle \geq -1$  [21]. It is insightful to compare the two relations (8) and (11). Since  $S_{\text{inf}} \leq S_{\text{tot}}(t)$ , the inequality (11) implies the inequality (8), and (11) is thus a stronger result. Remarkably, the bound (11) is tight for continuous stochastic processes. Indeed, using (10) we obtain the probability density function for global infima of the entropy production in continuous processes [21],

$$p_{S_{\text{inf}}}(-s) = e^{-s}, \quad s \geq 0. \quad (12)$$

The mean global infimum is thus  $\langle S_{\text{inf}} \rangle = -1$ .

- The survival probability  $p_{\text{sur}}(t)$  of the entropy production is the probability that entropy production has not reached a value  $s_0$  in the time interval  $[0, t]$ . For continuous stochastic processes we obtain the generic expression

$$p_{\text{sur}}(t) = \frac{e^{-s_0} - 1}{e^{-s_0} - \langle e^{-S_{\text{tot}}(t)} \rangle_{\text{sur}}}, \quad (13)$$

where  $\langle \dots \rangle_{\text{sur}}$  is an average over trajectories that have not reached the absorbing state in the interval  $[0, t]$ .

- We consider the statistics of the number of times,  $N_x$ , that entropy production crosses the interval  $[-\Delta, \Delta]$  from  $-\Delta$  towards  $\Delta$  in one realisation of the process  $\vec{X}(t)$ . The probability of  $N_x$  is bounded by

$$\begin{aligned}\mathbb{P}(N_x = 0; \Delta) &\geq 1 - e^{-\Delta}, \\ \mathbb{P}(N_x > n) &\leq e^{-\Delta(2n+1)}.\end{aligned}\tag{14}$$

In other words, the probability of observing a large number of crossings decays at least exponentially in  $N_x$ . For continuous stochastic processes we obtain a generic expression for the probability of  $N_x$  [22], given by

$$\mathbb{P}(N_x = n_x) = \begin{cases} 1 - e^{-\Delta} & n_x = 0, \\ 2 \sinh(\Delta) e^{-2n_x \Delta} & n_x \geq 1. \end{cases}\tag{15}$$

Remarkably, all these results on universal fluctuation properties are direct consequences of the integral fluctuation relation for entropy production at stopping times, equation (2), and its conditional version, equation (3).

Some of the results in this paper have already appeared before in the literature or are closely related to existing results. A fluctuation relation analogous to (2) has been derived for the exponential of the negative housekeeping heat, see equation (6) in [30]. Since for stationary processes the housekeeping heat is equal to the entropy production, the relation (6) in [30] implies the relation (2) in this paper. The relations (10)–(12) and (15) have been derived before in [21] and [22]. Instead, the relations (3), (5), (7), (9), (13), and (14) are, to the best of our knowledge, derived here for the first time. Moreover, we demonstrate that all the results (3), (5), (7), (9)–(13) and (15) descend from the integral fluctuation relation fluctuation relations (2) and (3) in a few simple steps, and we discuss the physical meaning of the results derived in this paper on examples of simple nonequilibrium systems.

The paper is organised as follows. Section 2 introduces the notation used in the paper. In section 3, we revisit the theory of martingales in the context of gambling. In section 4, we briefly recall the theory of stochastic thermodynamics, focusing on the aspects we will use in this paper. These two sections only contain review material, and can be skipped by readers who want to directly read the new results of this paper. In section 5, we derive the first important results of this paper: the integral fluctuation relations at stopping times (2) and (3). In section 6, we derive the second law of thermodynamics at stopping times (4), and we discuss the physical implications of this law. In section 7, we use the integral fluctuation relation at stopping times to derive universal properties for the fluctuations of entropy production in nonequilibrium stationary states, including the relations (9)–(15). In section 8, we discuss the effect of finite statistics on the integral fluctuation relation at stopping times, which is relevant for experimental validation. In section 9, we illustrate the second law at stopping times and the integral fluctuation relation at stopping times in paradigmatic examples of nonequilibrium stationary states. We conclude the paper with a discussion in section 10. In the appendices, we provide details on important proofs and derivations.

## 2. Preliminaries and notation

In this paper, we will consider stochastic processes described by  $d$  degrees of freedom  $\vec{X}(t) = (X_1(t), X_2(t), \dots, X_d(t))$ . The time index can be discrete,  $t \in \mathbb{Z}$ , or continuous,  $t \in \mathbb{R}$ . We denote the full trajectory of  $\vec{X}(t)$  by  $\omega = \left\{ \vec{X}(t) \right\}_{t \in (-\infty, \infty)}$  and the set of all trajectories by  $\Omega$ . We call a subset  $\Phi$  of  $\Omega$  a measurable set, or an event, if we can measure the probability  $\mathbb{P}[\Phi]$  to observe a trajectory  $\omega$  in  $\Phi$ . The  $\sigma$ -algebra  $\mathcal{F}$  is the set of all subsets of  $\Omega$  that are measurable.

The triple  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space. We denote random variables on this probability space in upper case, e.g.  $X, Y, Z$ , whereas deterministic variables are written in lower case letters, e.g.  $x, y, z$ . An exception is the temperature  $T_{\text{env}}$ , which is a deterministic variable. Random variables are functions defined on  $\Omega$ , i.e.  $X : \omega \rightarrow X(\omega)$ . For simplicity we often omit the  $\omega$ -dependency in our notation of random variables, i.e. we write  $X = X(\omega)$ ,  $Y = Y(\omega)$ , etc. For stochastic processes,  $\vec{X}(t) = \vec{X}(t, \omega)$  is a function on  $\Omega$  that returns the value of  $\vec{X}$  at time  $t$  in the trajectory  $\omega$ . The expected value of a random variable  $X$  is denoted by  $\langle X \rangle$  or  $E[X]$  and is defined as the integral  $\langle X \rangle = E[X] = \int_{\Omega} d\mathbb{P} X(\omega)$  in the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . We write  $p_X(x)$  for the probability density function or probability mass function of  $X$ , if it exists. We denote vectors by  $\vec{x} = (x_1, x_2, \dots, x_d)$  and we use the notation  $t \wedge \tau = \min \{t, \tau\}$ .

We will consider situations where an experimental observer does not have instantaneously access to the complete trajectory  $\omega$  but rather tracks a trajectory  $\left\{ \vec{X}(s) \right\}_{s \in [0, t]} = \vec{X}_0^t$  in a finite time interval. In this case, the set of measurable events gradually expands as time progresses and a larger part of the trajectory  $\omega$  becomes visible. Mathematically this situation is described by an increasing sequence of  $\sigma$ -algebras  $\{\mathcal{F}_t\}_{t \geq 0}$  where  $\mathcal{F}_t$  contains all the measurable events  $\Phi$  associated with finite-time trajectories  $\vec{X}_0^t$ . The sequence of sub  $\sigma$ -algebras  $\{\mathcal{F}_t\}_{t \geq 0}$  of  $\mathcal{F}$  is called the *filtration* generated by the stochastic process  $X(t)$  and  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  is a filtered probability space. If time is continuous, then we assume that  $\{\mathcal{F}_t\}_{t \geq 0}$  is right-continuous, i.e.  $\mathcal{F}_t = \cap_{s > t} \mathcal{F}_s$ ; this implies that the process  $\vec{X}(t)$  consists of continuous trajectories intersected by a discrete number of jumps.

We denote by  $E[M(t)|\mathcal{F}_s](\omega)$  the conditional expectation of a random variable  $M(t)$  given a sub- $\sigma$ -algebra  $\mathcal{F}_s$  of  $\mathcal{F}$  [25, 26]. Note that conditional expectations  $E[M(t)|\mathcal{F}_s](\omega)$  are random variables on the measurable space  $(\Omega, \mathcal{F}_s)$ . Since  $E[M(t)|\mathcal{F}_s](\omega)$  is an expectation value we also use the physics notations  $E[M(t)|\mathcal{F}_s] = \langle M(t)|\mathcal{F}_s \rangle$  and  $E[M(t)|\mathcal{F}_s] = \langle M(t)|\vec{X}_0^s \rangle$ .

## 3. Martingales

In a first subsection, we introduce martingales within the example of games of chance, to illustrate how fluctuations of a stochastic process can be studied with martingale theory. The calculations in this subsection are similar to those for the stochastic entropy

production presented in sections 5 and 7, with the difference that games of chance are simpler to analyse, since they consist of sequences of independent and identically distributed random variables. In a second subsection, we present a definition of martingales that applies to stochastic processes in continuous and discrete time, and we discuss the optional stopping theorem.

### 3.1. Gambling and martingales

Games of chance have inspired mathematicians as far back as the 17th century and have laid the foundation for probability theory [31]. A question that has often been studied is the gambler's ruin problem: Consider a gambler that enters a casino and tries his/her luck at the roulette. The gambler plays until he/she has either won a pre-determined amount of money or until the initial stake is lost. We are interested in the probability of success, or equivalently in the ruin probability of the gambler.

The roulette is a game of chance that consists of a ball that rolls at the edge of a spinning wheel and falls in one of 37 coloured pockets on the wheel: 18 pockets are coloured in red, 18 pockets are coloured in black, and one is coloured in green. Before each round of the game, the gambler puts his/her bet on whether the ball will fall in either a red or a black pocket. If the gambler's call is correct, then he/she wins an amount of chips equal to the bet size, otherwise he/she loses the betted chips. The gambler cannot bet for the green pocket. The presence of the green pocket biases the game in favour of the casino: if the ball falls in the green pocket, then the casino wins all the bets. We are interested in the gambler's ruin problem: what is the probability that the gambler loses his/her initial stake before reaching a certain amount of profit?

A gambling sequence at the roulette can be formalised as a stochastic process  $X(t)$  in discrete time,  $t = 1, 2, 3, \dots$ . We define  $X(t) = 1$  if the ball falls in a red pocket and  $X(t) = -1$  if the ball falls in a black pocket in the  $t$ -th round of the game. If the ball falls in a green pocket we set  $X(t) = 0$ . We denote the bets of the gambler by the process  $Y(t)$ : if the gambler calls for red we set  $Y(t) = 1$  and if the gambler calls for black we set  $Y(t) = -1$ . The gambler does not bet on green. Finally, we assume the bet size of the gambler  $b$  is constant.

For an ideal roulette, the random variables  $X(t)$  are independently drawn from the distribution

$$p_{X(t)}(x) = \frac{18}{37}\delta_{x,1} + \frac{18}{37}\delta_{x,-1} + \frac{1}{37}\delta_{x,0}, \quad (16)$$

where  $\delta_{x,y}$  is the Kronecker's delta. The gambler's bet  $Y(t) = y(X(0), X(1), \dots, X(t-1))$  with  $y$  a function that defines the gambler's betting system. The gambler's fortune at time  $t$  is the process

$$F(t) = n + b \sum_{s=1}^t (Y(s)X(s) + |X(s)| - 1), \quad (17)$$

where  $F(0) = n$  is the initial stake.

The duration of the game is random. The gambler plays until a time  $T_{\text{play}}$  when the gambler is either ruined, i.e.  $F(T_{\text{play}}) \leq 0$ , or the gambler's fortune has surpassed for the

first time a certain amount  $m$ , i.e.  $F(T_{\text{play}}) \geq m$ . Clearly we require that  $m > n = F(0)$ , since otherwise  $T_{\text{play}} = 0$ . The ruin probability

$$p_{\text{ruin}}(n) = \mathbb{P}[F(T_{\text{play}}) \leq 0] \quad (18)$$

is the probability that the gambler loses the game.

The gambler's fortune is a *supermartingale* because it is a bounded stochastic process satisfying

$$E[F(t)|X(0), X(1), \dots, X(s)] \leq F(s), \quad (19)$$

for all  $s \leq t$  and all  $t \geq 0$ . Relation (19) means that on average the gambler will inevitably lose money, irrespective of the betting system he/she adopts.

A gambler whose fortune is expected to decrease may still be tempted to play if the probability of winning is high enough. The probabilities to win or loose the game depend on the fluctuations of  $X(t)$ . In the roulette game, the gambler's fortune  $F(t)$  can be represented as a biased random walk on the real line  $[0, m]$ , which starts at the position  $n$ , and moves each time step either a distance  $-b$  to the left or a distance  $b$  to the right. The probability to make a step to the left is  $q = 19/37 \approx 0.51$  and the probability to make a step to the right is  $1 - q = 18/37 \approx 0.49$ . Hence, the gambler's fortune is slightly biased to move towards the left where  $F(t) < 0$ . The ruin probability  $p_{\text{ruin}}(n) = \mathbb{P}[F(T_{\text{play}}) \leq 0]$  solves the recursive equation

$$p_{\text{ruin}}(n) = q p_{\text{ruin}}(n - b) + (1 - q) p_{\text{ruin}}(n + b), \quad n \in [0, m], \quad (20)$$

with boundary conditions  $p_{\text{ruin}}(0) = 1$  and  $p_{\text{ruin}}(m) = 0$ . Instead of solving the relations (20) we bound the ruin probability  $p_{\text{ruin}}(n)$  using the theory of martingales [32]. We define the process

$$M(t) = \left(\frac{q}{1-q}\right)^{F(t)/b}. \quad (21)$$

The processes  $M(t)$  is a *martingale* relative to the process  $X(t)$  [25, 26]. Indeed, we say that a bounded process  $M$  is a martingale if

$$E[M(t)|X(0), X(1), \dots, X(s)] = M(s), \quad (22)$$

for all  $s < t$ .

An important property of martingales is that their expected value evaluated at a stopping time  $T$  of the process  $X$  equals their expected value at the initial time [24–26],

$$E[M(T)] = E[M(0)]. \quad (23)$$

Equation (23) is known as *Doob's optional stopping theorem* and will constitute the main tool in this paper to derive fluctuation properties of stochastic processes. In the present example, since

$$E[M(0)] = \left(\frac{q}{1-q}\right)^{n/b}, \quad (24)$$

and since for  $q \geq 0.5$

$$p_{\text{ruin}}(n) + (1 - p_{\text{ruin}}(n)) \left( \frac{q}{1-q} \right)^{m/b+1} \geq E[M(T_{\text{play}})] \geq p_{\text{ruin}}(n) + (1 - p_{\text{ruin}}(n)) \left( \frac{q}{1-q} \right)^{m/b}, \quad (25)$$

Doob's optional stopping theorem (23) implies that

$$\frac{\left( \frac{q}{1-q} \right)^{m/b+1} - \left( \frac{q}{1-q} \right)^{n/b}}{\left( \frac{q}{1-q} \right)^{m/b+1} - 1} \geq p_{\text{ruin}}(n) \geq \frac{\left( \frac{q}{1-q} \right)^{m/b} - \left( \frac{q}{1-q} \right)^{n/b}}{\left( \frac{q}{1-q} \right)^{m/b} - 1}. \quad (26)$$

Hence, Doob's optional stopping theorem bounds the gambler's ruin probability.

The formula (26) provides useful information for the gambler. If we start the game with an initial fortune of 10£, if we play until our fortune reaches 50£ and if we bet each game 1£, then the chance of loosing our initial stake is in between 94.8 and 95.2 percent. If, on the other hand, we bet each game 10£, then the ruin probability is in between 82 and 85.6 percent. Hence, the probability of winning increases as a function of the betting size. Indeed, since the game is biased in favour of the casino, the best strategy is to reduce the number of betting rounds to the minimal possible and hope that luck plays in our favour. After all, the outcome of a single game is almost fair, since the odds of winning a single game are  $q = 19/37 \approx 0.51$ .

### 3.2. Martingales and the optional stopping theorem

We now discuss martingales and the optional stopping theorem for generic stochastic processes  $\vec{X}(t)$  in discrete or continuous time. A *martingale* process  $M(t)$  with respect to another process  $\vec{X}(t)$  is a real-valued stochastic process that satisfies the following three properties [25, 26]:

- $M(t)$  is  $\mathcal{F}_t$ -adapted, which means, in a loose sense, that  $M(t)$  is a function on trajectories  $\vec{X}(0, \dots, t)$ ;
- $M(t)$  is integrable,

$$E[|M(t)|] < \infty, \quad \forall t \geq 0; \quad (27)$$

- the conditional expectation of  $M(t)$  given the  $\sigma$ -algebra  $\mathcal{F}_s$  satisfies the property

$$E[M(t)|\mathcal{F}_s] = M(s), \quad \forall s \leq t, \quad \text{and} \quad \forall t \geq 0. \quad (28)$$

The conditional expectation of a random variable  $M(t)$  given a sub- $\sigma$ -algebra  $\mathcal{F}_s$  of  $\mathcal{F}$  is defined as a  $\mathcal{F}_s$ -measurable random variable  $E[M(t)|\mathcal{F}_s]$  for which  $\int_{\omega \in \Phi} E[M(t)|\mathcal{F}_s] d\mathbb{P} = \int_{\omega \in \Phi} M(t) d\mathbb{P}$  for all  $\Phi \in \mathcal{F}_s$  [26].

If instead of the equality (28) we have an inequality  $E[M(t)|\mathcal{F}_s] \geq M(s)$ , then we call the process a *submartingale*. If  $E[M(t)|\mathcal{F}_s] \leq M(s)$ , then we call the process a *supermartingale*.

Fluctuations of a martingale  $M(t)$  can be studied with stopping times. Stopping times are the random times when the stochastic process  $\vec{X}(t)$  satisfies for the first time

a given criterion. Stopping times do not allow for clairvoyance (the stopping criterion cannot anticipate the future) and do not allow for cheating (the criterion does not have access to side information). Aside these constraints, the stopping rule can be an arbitrary function of the trajectories of the stochastic process  $\vec{X}(t)$ .

Formally, a stopping time  $T(\omega) \in [0, \infty]$  of a stochastic process  $\vec{X}$  is defined as a random variable for which  $\{\omega : T(\omega) \leq t\} \in \mathcal{F}_t$  for all  $t \geq 0$ . Alternatively, we can also define stopping times as functions on trajectories  $\omega = \vec{X}_0^\infty$  with the property that the function  $T(\omega)$  does not depend on what happens after the stopping time  $T$ .

An important result in martingale theory is *Doob's optional stopping theorem*. There exist different versions of the optional stopping theorem, which differ in the conditions assumed for the martingale process  $M(t)$  and the stopping time  $T$ . We discuss the version of the theorem presented as theorem 3.6 in the book of Liptser and Shirayev [26].

Let  $M(t)$  be a martingale relative to the process  $\vec{X}$  and let  $T$  be a stopping time on the process  $\vec{X}$ . If  $M$  is uniformly integrable, then

$$E[M(T)] = E[M(0)]. \quad (29)$$

If time is continuous we also require that  $M(t)$  is right-continuous. A stochastic process  $M(t)$  is called uniformly integrable if

$$\lim_{m \rightarrow \infty} \sup_{t \geq 0} \int |M(t)| I_{|M(t)| \geq m} d\mathbb{P} = 0, \quad (30)$$

where  $I_\Phi(\omega)$  is the indicator function defined by

$$I_\Phi(\omega) = \begin{cases} 1 & \text{if } \omega \in \Phi, \\ 0 & \text{if } \omega \notin \Phi, \end{cases} \quad (31)$$

for all  $\omega \in \Omega$  and  $\Phi \in \mathcal{F}$ . If  $M(t)$  is not uniformly integrable, then (29) may not hold. For example if  $M(t) = B(t)$  with  $B(t)$  a Wiener process on  $[0, \infty)$  and  $T = \inf\{t \geq 0 : M(t) = m\}$ , then  $E[B(T)] = m \neq E[B(0)] = 0$ , where we have used the convention that  $0 \cdot \infty = 0$  [33].

An extended version of the optional stopping theorem holds for two stopping times  $T_1$  and  $T_2$  with the property  $\mathbb{P}[T_2 \leq T_1] = 1$ ,

$$E[M(T_1) | \mathcal{F}_{T_2}](\omega) = M(T_2(\omega), \omega), \quad (32)$$

where the  $\sigma$ -algebra  $\mathcal{F}_{T_2}$  consists of all sets  $\Phi \in \mathcal{F}$  such that  $\Phi \cap \{\omega : T_2(\omega) \leq t\} \in \mathcal{F}_t$ .

#### 4. Stochastic thermodynamics for stationary processes

In this section, we briefly introduce the formalism of stochastic thermodynamics in nonequilibrium stationary states; for reviews see [1–4]. We use a probability-theoretic approach [21, 34, 35], which has the advantage of dealing with Markov chains, Markov jump processes, and Langevin process in one unified framework. It is moreover the natural language to deal with martingales.

The stochastic entropy production  $S_{\text{tot}}(t)$  is defined in terms of a probability measure  $\mathbb{P}$  of a stationary stochastic process and its time-reversed measure  $\mathbb{P} \circ \Theta$ . The time-reversal

map  $\Theta$ , with respect to the origin  $t = 0$ , is a measurable involution on trajectories  $\omega$  with the property that  $X_i(t, \Theta(\omega)) = X_i(-t, \omega)$  for variables of even parity with respect to time reversal and  $X_i(t, \Theta(\omega)) = -X_i(-t, \omega)$  for variables of odd parity with respect to time reversal. We say that the measure  $\mathbb{P}$  is stationary if  $\mathbb{P} = \mathbb{P} \circ \mathcal{T}_t$  for all  $t \in \mathbb{R}$ , with  $\mathcal{T}_t$  the time-translation map, i.e.  $\vec{X}(t', \mathcal{T}_t(\omega)) = \vec{X}(t' + t, \omega)$  for all  $t' \in \mathbb{R}$ . In order to define an entropy production we require that the process  $\vec{X}(t, \omega)$  is reversible. This means that for all finite  $t \geq 0$  and for all  $\Phi \in \mathcal{F}_t$  it holds that  $\mathbb{P}[\Phi] = 0$  if and only if  $(\mathbb{P} \circ \Theta)[\Phi] = 0$ . In other words, if an event happens with zero probability in the forward dynamics, then this event also occurs with zero probability in the time-reversed dynamics. In probability theory, one says that  $\mathbb{P}$  and  $\mathbb{P} \circ \Theta$  are locally mutually absolute continuous.

Given the above assumptions, we define the entropy production in a stationary process  $\vec{X}$  by [17, 21, 34, 35]

$$S_{\text{tot}}(t, \omega) := \log \frac{d\mathbb{P}|_{\mathcal{F}_t}}{d(\mathbb{P} \circ \Theta)|_{\mathcal{F}_t}}(\omega), \quad t \geq 0, \quad \omega \in \Omega, \quad (33)$$

where we have used the Radon–Nikodym derivative of the restricted measures  $\mathbb{P}|_{\mathcal{F}_t}$  and  $(\mathbb{P} \circ \Theta)|_{\mathcal{F}_t}$  on  $\mathcal{F}$ . The restriction of a measure  $\mathbb{P}$  on a sub- $\sigma$ -algebra  $\mathcal{F}_t$  of  $\mathcal{F}$  is defined by  $\mathbb{P}|_{\mathcal{F}_t}[\Phi] = \mathbb{P}[\Phi]$  for all  $\Phi \in \mathcal{F}_t$ . If  $t$  is continuous, then  $S_{\text{tot}}(t, \omega)$  is right-continuous, since we have assumed the rightcontinuity of the filtration  $\{\mathcal{F}_t\}_{t \geq 0}$ . Local mutual absolute continuity of the two measures  $\mathbb{P}$  and  $\mathbb{P} \circ \Theta$  implies that the Radon–Nikodym derivative in (33) exists and is almost everywhere uniquely defined. The definition (33) states that entropy production is the log of the probability density of the measure  $\mathbb{P}$  with respect to the time-reversed measure  $\mathbb{P} \circ \Theta$ ; it is a functional of trajectories  $\omega$  of the stochastic process  $X$  and characterises their time-irreversibility.

The definition (33) of the stochastic entropy production is general. It applies to Markov chains, Markov jump processes, diffusion processes, etc. For Markov chains, the relation (33) is equivalent to the expression (1) for entropy production in terms of probability density functions of trajectories. Consider for example the case of  $\vec{X}(t) \in \mathbb{R}^d$  and  $t \in \mathbb{Z}$  and let us assume for simplicity that all degrees of freedom are of even parity with respect to time reversal. Using  $d\mathbb{P}|_{\mathcal{F}_t} = p(\vec{X}(1), \vec{X}(2), \dots, \vec{X}(t)) d\lambda|_{\mathcal{F}_t}$ , with  $\lambda|_{\mathcal{F}_t}$  the Lebesgue measure on  $\mathbb{R}^{td}$ , the entropy production is indeed of the form given by relation (1). However, formula (33) is more general than (1) because it also applies to cases where the path probability density with respect to a Lebesgue measure does not exist, as is the case with stochastic processes in continuous time.

For systems that are weakly coupled to one or more environments in equilibrium, the entropy production (33) is equal to [4]

$$S_{\text{tot}}(t) = \Delta S_{\text{sys}}(t) + S_{\text{env}}(t), \quad (34)$$

where  $S_{\text{env}}(t)$  is the entropy change of the environment, and where

$$\Delta S_{\text{sys}}(t) = -\log \frac{p_{\text{ss}}(\vec{X}(t))}{p_{\text{ss}}(\vec{X}(0))} \quad (35)$$

is the system entropy change associated with the stationary probability density function  $p_{\text{ss}}(\vec{X}(t))$  of  $\vec{X}(t)$ .

## 5. Integral fluctuation relations at stopping times

In this section we initiate the study of the fluctuations of the entropy production in stationary processes. We follow an approach similar to the one presented in section 3 for the fluctuations of a gambler's fortune, namely, we first identify a martingale process related to the entropy production, which is the exponentiated negative entropy  $e^{-S_{\text{tot}}(t)}$ , and we then apply Doob's optional stopping theorem (29) to this martingale process. Since  $e^{-S_{\text{tot}}(t)}$  is unbounded, we require uniform integrability of  $e^{-S_{\text{tot}}(t)}$  in order to apply Doob's optional stopping theorem (29). Therefore, we obtain two versions of the integral fluctuation relation at stopping times: a first version holds within finite-time windows and a second version holds when  $S_{\text{tot}}(t)$  is bounded for all times  $t \leq T$ . These two versions represent two different ways to ensure that the total available entropy in the system and environment is finite. Note that when we applied Doob's optional stopping theorem in the gambling problem, we also required that the gambler's fortune is finite. Finally, we obtain conditional integral fluctuation relations by applying the conditional version (32) of Doob's optional stopping theorem to  $e^{-S_{\text{tot}}(t)}$ .

### 5.1. The martingale structure of the exponential entropy production

The exponentiated negative entropy production  $e^{-S_{\text{tot}}(t)}$  associated with a stationary stochastic process  $\vec{X}(t)$  is a martingale process relative to  $\vec{X}(t)$ . Indeed,  $S_{\text{tot}}(t)$  is a  $\mathcal{F}(t)$ -adapted process,  $E[e^{-S_{\text{tot}}(t)}] = 1$ , and in appendix A we show that [20, 21]

$$E[e^{-S_{\text{tot}}(t)}|\mathcal{F}_s] = e^{-S_{\text{tot}}(s)}, \quad s \leq t. \quad (36)$$

As a consequence, entropy production is a submartingale:

$$E[S_{\text{tot}}(t)|\mathcal{F}_s] \geq S_{\text{tot}}(s), \quad s \leq t. \quad (37)$$

Notice that we can draw an analogy between thermodynamics and gambling by identifying the negative entropy production with a gambler's fortune and by identifying the exponential  $e^{-S_{\text{tot}}(t)}$  with the martingale (21).

### 5.2. Fluctuation relation at stopping times within a finite-time window

We apply Doob's optional stopping theorem (29) to the martingale  $e^{-S_{\text{tot}}(t)}$ . We consider first the case when an experimental observer measures a stationary stochastic processes  $\vec{X}(t)$  within a finite-time window  $t \in [0, \tau]$ . In this case, the experimental observer measures in fact the process  $e^{-S_{\text{tot}}(t \wedge \tau)}$ , where we have used the notation

$$t \wedge \tau = \min\{t, \tau\}. \quad (38)$$

The process  $e^{-S_{\text{tot}}(t \wedge \tau)}$  is uniformly integrable, as we show in the appendix A, and therefore

$$\langle e^{-S_{\text{tot}}(T \wedge \tau)} \rangle = 1, \quad (39)$$

holds for all stopping times  $T$  of  $\vec{X}(t)$ .

### 5.3. Fluctuation relation at stopping times within an infinite-time window

We discuss an integral fluctuation relation for stopping times within an infinite-time window, i.e.  $T \in [0, \infty]$ . In the appendix B we prove that if the conditions

- (i)  $e^{-S_{\text{tot}}}$  converges  $\mathbb{P}$ -almost surely to 0 in the limit  $t \rightarrow \infty$
- (ii)  $S_{\text{tot}}(t)$  is bounded for all  $t \leq T$

are met, then

$$\langle e^{-S_{\text{tot}}(T)} \rangle = 1. \quad (40)$$

The condition (i) is a reasonable assumption for nonequilibrium stationary states, since  $\langle S_{\text{tot}}(t) \rangle$  grows extensive in time as  $\langle S_{\text{tot}}(t) \rangle = \sigma t$  with  $\sigma$  a positive number. The condition (ii) can be imposed on  $T$  by considering a stopping time  $T \wedge T_{\text{fp}}$  where  $T_{\text{fp}} = \inf \{t : S_{\text{tot}}(t) \notin (-s_-, s_+)\}$  is a first-passage time with two thresholds  $s_-, s_+ \gg 1$ , which can be considered large compared to the typical values of entropy production at the stopping time  $T$ .

### 5.4. Conditional integral fluctuation relation at stopping times

We can also apply the conditional optional stopping-time theorem (32) to  $e^{-S_{\text{tot}}(t)}$ . We then obtain the conditional integral fluctuation relation (3) for stopping times  $T_2 \leq T_1$ , viz.,

$$\langle e^{-S_{\text{tot}}(T_1)} | \mathcal{F}_{T_2} \rangle = e^{-S_{\text{tot}}(T_2)}. \quad (41)$$

The relation (41) is valid either for finite stopping times  $T_1 \in [0, \tau]$  or for stopping times  $T_1$  and  $T_2$  for which  $S_{\text{tot}}(t)$  is bounded for all  $t \in [0, T_1]$ .

## 6. Second law of thermodynamics at stopping times

Jensen's inequality  $\langle e^{-S_{\text{tot}}(T \wedge \tau)} \rangle \geq e^{-\langle S_{\text{tot}}(T \wedge \tau) \rangle}$  together with the integral fluctuation relation (39) imply that

$$\langle S_{\text{tot}}(T \wedge \tau) \rangle \geq 0. \quad (42)$$

The relation (42) states that on average entropy production always increases, even when we stop the process at a random time  $T$  chosen according to a given protocol. This law is akin to the relation (19) describing that a gambler cannot make profit out of a fair game of chance, even when he/she quits the game in an intelligent manner. Analogously, (40) implies the law  $\langle S_{\text{tot}}(T) \rangle \geq 0$  for unbounded stopping times  $T$ . We have thus derived the second law (4) of thermodynamics at stopping times.

When applying this second law to examples of physical processes one obtains interesting bounds on heat and work in nonequilibrium stationary states. Below we first

discuss bounds on the average dissipated heat in isothermal processes and then bounds on the average work in stationary stochastic heat engines.

### 6.1. Bounds on heat absorption in isothermal processes

For systems that are in contact with one thermal reservoir at temperature  $T_{\text{env}}$  and for which the entropy of hidden internal degrees of freedom is negligible, the entropy production is given by (34) and (35), and the environment entropy takes the form [4, 36]

$$S_{\text{env}}(t) = -\frac{Q(t)}{T_{\text{env}}}, \quad (43)$$

where  $Q$  is the heat transferred from the environment to the system. Relation (43) relates the stochastic entropy production (34) to the stochastic heat that enters into the first law of thermodynamics. Therefore, for isothermal systems the second law at stopping times (42) reads

$$\langle Q(T \wedge \tau) \rangle \leq T_{\text{env}} \langle \Delta S_{\text{sys}}(T \wedge \tau) \rangle = T_{\text{env}} \left\langle \log \frac{p_{\text{ss}}(X(0))}{p_{\text{ss}}(X(T \wedge \tau))} \right\rangle. \quad (44)$$

Analogously, we obtain the relation (5) for unbounded stopping times  $T$ .

The relation (44) implies that it is not possible to extract on average heat from a thermal reservoir when the system state is invariant. Indeed, when  $X(T \wedge \tau) = X(0)$ , then  $\langle Q(T \wedge \tau) \rangle \leq 0$ . However, if the system entropy at the stopping time  $T$  is different than the entropy in the stationary state, then it is possible to extract on average at most an amount  $T_{\text{env}} \left\langle \log \frac{p_{\text{ss}}(X(0))}{p_{\text{ss}}(X(T \wedge \tau))} \right\rangle$  of heat from the thermal reservoir.

For systems in equilibrium  $p_{\text{ss}}(x) \sim e^{-v(x)/T_{\text{env}}}$ , such that the bound (44) reads  $\langle Q(T \wedge \tau) \rangle \leq \langle \Delta v(T \wedge \tau) \rangle$ . Moreover, according to the first law of thermodynamics  $\langle Q(T \wedge \tau) \rangle = \langle \Delta v(T \wedge \tau) \rangle$ , such that for systems in equilibrium the bound (44) is tight.

Notice that the bound on the right hand side of (44) is maximal for stopping times of the form

$$T^\dagger = \inf \{t \geq 0 : p_{\text{ss}}(X(t)) = \min_{x \in \mathcal{X}} p_{\text{ss}}(x)\}. \quad (45)$$

If  $X(t)$  is a recurrent process, i.e.  $\mathbb{P}(T^\dagger < \infty) = 1$ , and if  $\tau \rightarrow \infty$ , then

$$\langle Q(T^\dagger) \rangle \leq T_{\text{env}} \sum_{x \in \mathcal{X}} p_{\text{ss}}(x) \log \frac{p_{\text{ss}}(x)}{\min_{x' \in \mathcal{X}} p_{\text{ss}}(x')}. \quad (46)$$

### 6.2. Efficiency of heat engines at stopping times: the case of Feynman's ratchet

We consider stochastic heat engines in contact with two thermal baths at temperatures  $T_c$  and  $T_h$  with  $T_h \geq T_c$ . A paradigmatic example is Feynman's ratchet [37–41], which is composed of a ratchet wheel with a pawl that is mechanically linked by an axle to a vane. The ratchet wheel and the pawl are immersed in a hot thermal reservoir, and the vane is immersed in a cold thermal reservoir. An external mass is connected to the axle of the Feynman ratchet and follows the movement of the ratchet wheel. If the wheel turns in the clockwise direction then the axle performs work on the mass, whereas if the wheel turns in the counterclockwise direction then the mass performs work on the axle.

We now perform an analysis of the Feynman ratchet at stopping times. For example, we ask the question what is the efficiency of the ratchet when the system is stopped right before or after the ‘main event’, i.e. the passage of the pawl over the peak of the ratchet wheel. The first law of thermodynamics implies that

$$\langle Q_h(T) \rangle + \langle Q_c(T) \rangle + \langle W(T) \rangle = \langle \Delta v(T) \rangle, \quad (47)$$

where  $Q_h$  is the heat absorbed by the ratchet from the hot reservoir,  $Q_c$  is heat absorbed by the ratchet from the cold reservoir,  $v$  is the mechanical energy stored in the pawl, and  $W$  is the work performed on the external mass. For the Feynman’s ratchet, the second law of thermodynamics at stopping times  $T$  reads

$$\frac{\langle Q_h(T) \rangle}{T_h} + \frac{\langle Q_c(T) \rangle}{T_c} \leq \langle \Delta S_{\text{sys}}(T) \rangle, \quad (48)$$

where  $S_{\text{sys}}$  is the entropy of the ratchet. If  $\langle Q_h(T) \rangle > 0$ , the first and second law of thermodynamics at stopping times imply the inequality (7), i.e.

$$\eta_T \leq \eta_C - \frac{\langle \Delta F_c(T) \rangle}{\langle Q_h(T) \rangle},$$

where we have introduced the efficiency at stopping times

$$\eta_T := -\frac{\langle W(T) \rangle}{\langle Q_h(T) \rangle},$$

the Carnot efficiency

$$\eta_C = 1 - \frac{T_c}{T_h}, \quad (49)$$

and the system free energy

$$\langle \Delta F_c(t) \rangle = \langle \Delta v(t) \rangle - T_c \langle \Delta S_{\text{sys}}(t) \rangle. \quad (50)$$

For  $T = t$ ,  $\langle \Delta F_c(T) \rangle = 0$  and we obtain the classical Carnot bound  $\eta_t \leq \eta_C$ . Moreover, if  $X(T) = X(0)$ , which implies that the process stops when it returns to its original state, then  $\langle \Delta F_c(T) \rangle = 0$  and we obtain again the classical Carnot bound  $\eta_t \leq \eta_C$ . Hence, it is not possible to exceed on the Carnot efficiency when the final state equals the initial state and thus when the heat engine is a cyclic process in phase space.

However, for general stopping times  $T$ ,  $\langle \Delta F_c(T) \rangle$  is different than zero. Interestingly, for stopping times  $T$  for which  $\langle \Delta F_c(T) \rangle / \langle Q_h(T) \rangle$  is negative, the second law of thermodynamics at stopping times implies that  $\eta_T$  is bounded by a constant that is larger than the Carnot efficiency. Note that the stopping-time efficiency  $\eta_T$  is defined as the ratio of averages, and not as the average of the ratios. In general  $\langle W(T) \rangle / \langle Q_h(T) \rangle \neq \langle W(T) / Q_h(T) \rangle$ , and the latter corresponds to the average of an unbounded random variable whose value at fixed times  $T = t$  has been previously studied in [12, 42–45].

Another interesting property of thermodynamic observables at stopping times is that they can take a different sign with respect to their stationary averages. For example, it is possible that at the stopping time  $T$  the fluxes of the Feynman ratchet have the same sign as those in a refrigerator, namely,  $\langle W(T) \rangle > 0$ ,  $\langle Q_c(T) \rangle > 0$  and

$\langle Q_h(T) \rangle < 0$ . To evaluate the performance of this process, we introduce the coefficient  $\epsilon_T := -\langle Q_c(T) \rangle / \langle W(T) \rangle > 0$ , for which the second law of thermodynamics at stopping times reads  $\epsilon_T \leq \left[ \left( \frac{T_h}{T_c} - 1 \right) + \frac{\langle \Delta \mathcal{F}_h(T) \rangle}{\langle Q_c(T) \rangle} \right]^{-1}$ , with  $\langle \Delta \mathcal{F}_h(T) \rangle := \langle \Delta v(T) \rangle - T_h \langle \Delta S_{\text{sys}}(T) \rangle$ . For fixed times  $T = t$  and for stopping times  $T$  with  $X(T) = X(0)$ , we recover the classical bound  $\epsilon_T \leq T_c / (T_h - T_c)$  [46–48].

In section 9, we illustrate the bounds (5) and (7) on simple physical models.

## 7. Universal properties of stochastic entropy production

We use the fluctuation relations (39)–(41) to derive universal relations about the stochastic properties of entropy production in stationary processes.

All our results hold for nonequilibrium stationary states for which  $\lim_{t \rightarrow \infty} e^{-S_{\text{tot}}(t)} = 0$  ( $\mathbb{P}$ -almost surely).

If entropy production is bounded for  $t < T$ , then we will use the optional stopping theorem (40) for stopping times within infinite-time windows, and if entropy production is unbounded for  $t < T$ , then we will use the optional stopping theorem (39) for stopping times within finite-time windows.

### 7.1. Fluctuation properties at fixed time $T = t$

We first consider the case where the stopping time  $T$  is a fixed non-fluctuating time  $t$ , i.e.  $T = t$ . In this case the fluctuation relation (39) is the integral fluctuation theorem derived in [16],

$$\langle e^{-S_{\text{tot}}(t)} \rangle = 1, \quad (51)$$

and the second law inequality (42) yields the second law of stochastic thermodynamics [4]

$$\langle S_{\text{tot}}(t) \rangle \geq 0. \quad (52)$$

The relation (51) provides a bound on negative fluctuations of entropy production, and for isothermal systems bounds on the fluctuations of work [3]. Since  $e^{-S_{\text{tot}}(t)}$  is a positive random variable, we can use Markov's inequality, see e.g. chapter 1 of [49], to bound events of large  $e^{-S_{\text{tot}}(t)}$ , namely,

$$\mathbb{P}(e^{-S_{\text{tot}}(t)} \geq \lambda) \leq \frac{\langle e^{-S_{\text{tot}}(t)} \rangle}{\lambda}, \quad \lambda \geq 0. \quad (53)$$

Using the integral fluctuation relation (51) together with (53) we obtain

$$\mathbb{P}(S_{\text{tot}}(t) \leq -s) \leq e^{-s}, \quad s \geq 0, \quad (54)$$

which is a well-known bound on the probability of negative entropy production, see equation (54) in [4].

The relations (39) and (42) are more general than the relations (51) and (52), since the former concern an average over an ensemble of trajectories of variable length  $x_0^T$  whereas the latter concern an average over an ensemble of trajectories of fixed length  $x_0^t$ .

Integral fluctuation relations for entropy production at stopping times

Therefore, we expect that evaluating (39) at fluctuating stopping times  $T$  it is possible to derive stronger constraints on the probability of negative entropy production than (54). This is the program we will pursue in the following sections.

## 7.2. Splitting probabilities of entropy production

We consider the first-passage time

$$T_{\text{fp}} = \inf \{t \geq 0 : S_{\text{tot}}(t) \notin (-s_-, s_+)\} \quad (55)$$

for entropy production with two absorbing thresholds at  $-s_- < 0$  and  $s_+ > 0$ . If the set  $\{t : S_{\text{tot}}(t) \notin (-s_-, s_+)\}$  is empty, then  $T_{\text{fp}} = \infty$ . Since entropy production is bounded for all values  $t < T_{\text{fp}}$  we can use the optional stopping theorem (40).

We split the ensemble of trajectories  $\Omega$  into two sets  $\Omega_- = \{S_{\text{tot}}(T_{\text{fp}}) \leq -s_-\}$  and  $\Omega_+ = \{S_{\text{tot}}(T_{\text{fp}}) \geq s_+\}$ . Since  $S_{\text{tot}}(T_{\text{fp}}) \in (\infty, -s_-] \cup [s_+, \infty)$  and  $\lim_{t \rightarrow \infty} e^{-S_{\text{tot}}(t)} = 0$  ( $\mathbb{P}$ -almost surely), the splitting probabilities  $p_- = \mathbb{P}[\Omega_-]$  and  $p_+ = \mathbb{P}[\Omega_+]$  have a total probability of one,

$$p_+ + p_- = 1. \quad (56)$$

We apply the integral fluctuation relation (40) to the stopping time  $T_{\text{fp}}$ :

$$1 = p_- \langle e^{-S_{\text{tot}}(T_{\text{fp}})} \rangle_- + p_+ \langle e^{-S_{\text{tot}}(T_{\text{fp}})} \rangle_+ \quad (57)$$

where

$$\langle e^{-S_{\text{tot}}(T_{\text{fp}})} \rangle_+ = \frac{\int_{\omega \in \Omega_+} d\mathbb{P} e^{-S_{\text{tot}}(T_{\text{fp}}(\omega), \omega)}}{\int_{\omega \in \Omega_+} d\mathbb{P}}, \quad (58)$$

$$\langle e^{-S_{\text{tot}}(T_{\text{fp}})} \rangle_- = \frac{\int_{\omega \in \Omega_-} d\mathbb{P} e^{-S_{\text{tot}}(T_{\text{fp}}(\omega), \omega)}}{\int_{\omega \in \Omega_-} d\mathbb{P}}. \quad (59)$$

The relations (56) and (57) imply that

$$p_+ = \frac{\langle e^{-S_{\text{tot}}(T)} \rangle_- - 1}{\langle e^{-S_{\text{tot}}(T)} \rangle_- - \langle e^{-S_{\text{tot}}(T)} \rangle_+}, \quad (60)$$

$$p_- = \frac{1 - \langle e^{-S_{\text{tot}}(T)} \rangle_+}{\langle e^{-S_{\text{tot}}(T)} \rangle_- - \langle e^{-S_{\text{tot}}(T)} \rangle_+}. \quad (61)$$

Moreover, since  $S_{\text{tot}}(T_{\text{fp}}) \in (\infty, -s_-] \cup [s_+, \infty)$  we have that  $\langle e^{-S_{\text{tot}}(T)} \rangle_- \geq e^{s_-}$  and  $\langle e^{-S_{\text{tot}}(T)} \rangle_+ \leq e^{-s_+}$ . Using these two inequalities in (60) and (61) we obtain the universal inequalities (9) for the splitting probabilities, viz.,

$$p_+ \geq 1 - \frac{1}{e^{s_-} - e^{-s_+}}, \quad p_- \leq \frac{1}{e^{s_-} - e^{-s_+}},$$

which hold for first-passage times with  $s_-, s_+ > 0$ .

In the case where  $\vec{X}(t)$  is a continuous stochastic process  $S_{\text{tot}}(T_{\text{fp}}) \in \{-s_-, s_+\}$  holds with probability one. Using in (60) and (61) that  $\langle e^{-S_{\text{tot}}(T)} \rangle_- = e^{s_-}$  and  $\langle e^{-S_{\text{tot}}(T)} \rangle_+ = e^{-s_+}$ , we obtain

$$p_+ = \frac{e^{s_-} - 1}{e^{s_-} - e^{-s_+}}, \quad p_- = \frac{1 - e^{-s_+}}{e^{s_-} - e^{-s_+}},$$

which are the relations (10). Hence, the splitting probabilities of entropy production are universal for continuous processes.

The bounds (9) apply not only to first-passage times but hold more generally for stopping times  $T$  of  $\vec{X}(t)$  for which  $S_{\text{tot}}(T) \in (\infty, -s_-] \cup [s_+, \infty)$  holds with probability one and for which  $S_{\text{tot}}(t)$  is bounded for all  $t \in [0, T]$ . Analogously, the equalities (10) apply not only to first-passage times but hold more generally for stopping times  $T$  of  $\vec{X}(t)$  for which  $S_{\text{tot}}(T) \in \{-s_-, s_+\}$  holds with probability one and for which  $S_{\text{tot}}(t)$  is bounded for all  $t \in [0, T]$ .

### 7.3. Infima of entropy production

We can use the results of the previous subsection to derive universal bounds and equalities for the statistics of infima of entropy production. The global infimum of entropy production is defined by

$$S_{\text{inf}} = \inf_{t \geq 0} S_{\text{tot}}(t) \tag{62}$$

and denotes the most negative value of entropy production.

Consider the first-passage time  $T_{\text{fp}}$ , denoting the first time when entropy production either goes below  $-s_-$  or goes above  $s_+$ , with  $s_-, s_+ \geq 0$  and its associated splitting probability  $p_-$ . The cumulative distribution of  $S_{\text{inf}}$  is given by

$$\mathbb{P}[S_{\text{inf}} \leq -s_-] = \lim_{s_+ \rightarrow \infty} p_- . \tag{63}$$

Using the inequalities (9) we obtain the bound (11) for the cumulative distribution of infima of entropy production, i.e.

$$\mathbb{P}[S_{\text{inf}} \leq -s_-] \leq e^{-s_-} .$$

The inequality (11) bears a strong similarity with the inequality (54). However, since by definition  $S_{\text{inf}} \leq S_{\text{tot}}(t)$  for any value of  $t$ , we also have that  $\mathbb{P}[S_{\text{tot}} \leq -s_-] \leq \mathbb{P}[S_{\text{inf}} \leq -s_-]$  and therefore the inequality (11) is stronger.

For continuous processes  $\vec{X}$ , the inequality (11) becomes an equality. Indeed, if the stochastic process  $\vec{X}$  is continuous, then with probability one  $S_{\text{tot}}(T_{\text{fp}}) \in \{-s_-, s_+\}$ , and therefore  $p_-$  is given by (10). Using the relation (63), we obtain

$$\mathbb{P}[S_{\text{inf}} \leq -s_-] = e^{-s_-} . \tag{64}$$

As a consequence, for continuous stochastic processes the global infimum  $S_{\text{inf}}$  of entropy production is characterised by an exponential probability density (12) with mean value  $\langle S_{\text{inf}} \rangle = -1$ .

### 7.4. Survival probability of entropy production

We analyse the survival probability

$$p_{\text{sur}}(t) = \mathbb{P}\left[\tilde{T} > t\right], \tag{65}$$

of the first-passage time with one absorbing boundary,

$$\tilde{T} = \inf \{t : S_{\text{tot}}(t) = s_0\}, \quad (66)$$

for continuous stochastic processes  $\vec{X}(t)$ . If the set  $\{t : S_{\text{tot}}(t) = s_0\}$  is empty, then  $\tilde{T} = \infty$ .

We use the fluctuation relation (39) since for first-passage times with one absorbing boundary  $S_{\text{tot}}(t)$  is unbounded for  $t \in [0, T]$ . Applying (39) to  $\tilde{T}$  we obtain

$$1 = p_{\text{sur}}(t) \langle e^{-S_{\text{tot}}(t)} \rangle_{\text{sur}} + (1 - p_{\text{sur}}(t)) e^{-s_0} \quad (67)$$

and thus also relation (13), i.e.

$$p_{\text{sur}}(t) = \frac{e^{-s_0} - 1}{e^{-s_0} - \langle e^{-S_{\text{tot}}(t)} \rangle_{\text{sur}}}, \quad (68)$$

where

$$\langle e^{-S_{\text{tot}}(t)} \rangle_{\text{sur}} = \langle e^{-S_{\text{tot}}(t)} I_{\tilde{T}>t} \rangle / \langle I_{\tilde{T}>t} \rangle. \quad (69)$$

For positive values of  $s_0$  we expect that  $\lim_{t \rightarrow \infty} p_{\text{sur}}(t) = 0$ . This implies that  $\lim_{t \rightarrow \infty} \langle e^{-S_{\text{tot}}(t)} \rangle_{\text{sur}} = \infty$ . For negative values of  $s_0$  we expect that  $\lim_{t \rightarrow \infty} p_{\text{sur}}(t) = \mathbb{P}[S_{\text{inf}} \geq s_0]$  which holds if  $\lim_{t \rightarrow \infty} \langle e^{-S_{\text{tot}}(t)} I_{\tilde{T}>t} / \langle I_{\tilde{T}>t} \rangle \rangle = 0$ .

## 7.5. Number of crossings

We consider the number of times  $N_x$  entropy production crosses an interval  $[-\Delta, \Delta]$  from the negative side to the positive side, i.e. in the direction  $-\Delta \rightarrow \Delta$ . We can bound the distribution of  $N_x$  using a sequence of stopping times.

The probability that  $N_x > 0$  is equal to the probability that the infimum is smaller or equal than  $-\Delta$ , and therefore using (11) we obtain

$$\mathbb{P}(N_x > 0; \Delta) = \mathbb{P}[S_{\text{inf}} \leq -s_-] \leq e^{-\Delta}. \quad (70)$$

Applying the conditional fluctuation relation (41) on two sequences of stopping times, we derive in the appendix C the inequality

$$\mathbb{P}[N_x \geq n+1 | N_x \geq n] \leq e^{-2\Delta} \quad \text{with } n > 0, \quad (71)$$

and therefore

$$\mathbb{P}(N_x > n) \leq e^{-\Delta(2n+1)}, \quad (72)$$

which is the inequality (14). The probability of observing a large number of crossing decays thus at least exponentially in  $N_x$ . For continuous processes  $\vec{X}(t)$  the probability mass function  $\mathbb{P}(N_x = n_x)$  is a universal statistic given by

$$\mathbb{P}(N_x = n_x) = \begin{cases} 1 - e^{-\Delta} & n_x = 0, \\ 2 \sinh(\Delta) e^{-2n_x \Delta} & n_x \geq 1, \end{cases} \quad (73)$$

which is the same relation as derived in [22] for overdamped Langevin processes.

## 8. The influence of finite statistics on the integral fluctuation relation at stopping times

In empirical situations we may want to use the integral fluctuation relation at stopping times to test whether a given process is the stochastic entropy production [50–52]. In these cases we have to deal with finite statistics. It is useful to know how many experiments are required to verify the integral fluctuation relation at stopping times with a certain accuracy. Therefore, in this section we discuss the influence of finite statistics on tests of the integral fluctuation relation at stopping times.

We consider the case where  $T$  is a two-boundary first-passage time for the entropy production of a continuous stationary process, i.e.  $T = T_{\text{fp}} = \inf \{t \geq 0 : S_{\text{tot}}(t) \notin (-s_-, s_+)\}$ . We imagine to estimate in an experiment the average  $\langle e^{-S_{\text{tot}}(T_{\text{fp}})} \rangle$  using an empirical average over  $m_s$  realisations,

$$A = \frac{1}{m_s} \sum_{j=1}^{m_s} e^{-S_{\text{tot}}(T_j)} = e^{-s_+} + \frac{N_-}{m_s} (e^{s_-} - e^{-s_+}), \quad (74)$$

where the  $T_j$  are the different outcomes of the first passage time  $T_{\text{fp}}$  and where  $N_-$  is the number of trajectories that have terminated at the negative boundary, i.e. for which  $S_{\text{tot}}(T_j) = -s_-$ . The expected value of the sample mean is thus

$$\langle A \rangle = 1, \quad (75)$$

and the variance of the sample mean is

$$\sigma_A^2 = \langle A^2 \rangle - \langle A \rangle^2 = \frac{(1 - e^{-s_+})(e^{s_-} - 1)}{m_s}. \quad (76)$$

Hence, for small enough values of  $s_-$  a few samples  $m_s$  will be enough to test the stopping-time fluctuation relation. For large enough  $s_-$ , we obtain  $\langle A^2 \rangle - \langle A \rangle^2 \sim e^{s_-}/m_s$ . The number of required samples  $m_s$  scales exponentially in the value of the negative threshold  $s_-$ . The full distribution of the empirical estimate  $A$  of  $\langle e^{-S_{\text{tot}}(T_{\text{fp}})} \rangle$  is given by

$$p_A(a) = \frac{1}{(e^{s_-} - e^{-s_+})^{m_s}} \sum_{n=0}^{m_s} \delta_{a,n e^{s_-} + (m_s-n)e^{-s_+}} \binom{m_s}{n} (1 - e^{-s_+})^n (e^{s_-} - 1)^{m_s-n}, \quad (77)$$

where we have again used  $\delta$  for the Kronecker delta.

Since we know the full distribution of the empirical average  $A$ , the integral fluctuation relation  $\langle e^{-S_{\text{tot}}(T_{\text{fp}})} \rangle = 1$  can be tested in experiments: given a certain observed value of  $A \neq 1$  we can use the distribution (77) to compute its  $p$ -value, i.e. the probability to observe a deviation from 1 larger or equal than the empirically observed  $|A - 1|$ .

## 9. Examples

We illustrate the bounds (5) and (7) on two simple examples of systems described by Langevin equations. We demonstrate that a randomly stopped process can extract work from a stationary isothermal process and we show that heat engines can surpass the

Carnot efficiency at stopping times. Moreover, the mean of the extracted heat and the performed work are bounded by the second law of thermodynamics at stopping times.

### 9.1. Heat extraction in an isothermal system

We illustrate the second law (5), for the average heat at stopping times, in the case of a colloidal particle on a ring that is driven by an external force  $f$  and moves in a potential  $v(y)$ , where  $y \in [0, 2\pi\ell]$  denotes the position on the ring and where  $\ell$  is the radius of the ring. We ask how much heat  $\langle Q(T^*) \rangle$  a colloidal particle can extract on average from its environment at the time  $T^*$  when the particle reaches for the first time the highest peak of the landscape.

We assume that the dynamics of the colloidal particle is governed by an over-damped Langevin equation of the form

$$\frac{dX}{dt} = -\mu \frac{\partial v(X(t))}{\partial x} + \mu f + \sqrt{2d}\zeta(t) \quad (78)$$

where  $X(t) \in \mathbb{R}$ ,  $\mu$  is the mobility coefficient,  $d$  is the diffusion coefficient, and  $\zeta(t)$  is a Gaussian white noise with  $\langle \zeta(t)\zeta(t') \rangle = \delta(t' - t)$ . We assume that the environment surrounding the particle is in equilibrium at a temperature  $T_{\text{env}}$ , so that Einstein relation  $d = T_{\text{env}}\mu$  holds. The potential  $v(x)$  is a periodic function with period  $2\pi\ell$ , i.e.  $v(x) = v(x + 2\pi\ell)$ . The actual position of the particle is given by the variable  $Y(t) = X(t) - N(t)2\pi\ell \in [0, 2\pi\ell]$ , where  $N(t) \in \mathbb{Z}$  is the winding number, i.e. the net number of times the particle has traversed the ring. The heat  $Q$  can be expressed as [2, 53]

$$Q(t) = v(X(t)) - v(X(0)) + f \int_0^t dX(t'), \quad (79)$$

and the stationary distribution is [7]

$$p_{\text{ss}}(y) \sim e^{-\frac{v(y)-fy}{T_{\text{env}}}} \int_y^{y+2\pi\ell} dy' e^{\frac{v(y')-fy'}{T_{\text{env}}}}. \quad (80)$$

We simulate the model (78) for a periodic potential of the form [54],

$$v(x) = T_{\text{env}} \ln(\cos(x/\ell) + 2), \quad (81)$$

which is illustrated in figure 1(a) for  $\ell = 1$  and  $T_{\text{env}} = 1$ . In this case, the time  $T^*$  when the particle reaches for the first time the highest peak of the landscape is

$$T^* = \inf \{t \geq 0 : X(t) = 0\}, \quad (82)$$

and the stationary distribution of  $X(t)$  is given by [54]

$$p_{\text{ss}}(x) = \frac{3(g^2(2 + \cos(x/\ell)) - g \sin(x/\ell) + 2)}{2\pi\ell(3g^2 + 2\sqrt{3})(\cos(x/\ell) + 2)}. \quad (83)$$

Hence, the bound (78) reads

$$\langle Q(T^*) \rangle \geq T_{\text{env}} \int dx p_{\text{ss}}(x) \log p_{\text{ss}}(x) - T_{\text{env}} \log \frac{3g^2 + 2}{2\pi\ell(3g^2 + 2\sqrt{3})}. \quad (84)$$

In figure 1(b) we illustrate the bound (84) for  $\langle Q(T^*) \rangle$  as a function of the nonequilibrium driving  $f$ . We find that heat absorption at stopping times is significant at small values of  $f$  and in the linear-response limit of small  $f$  the bound (84) is tight; the tightness of the bound (78) holds in general for recurrent Markov processes in the linear response limit. For comparison we also plot the mean heat dissipation  $\langle Q(t) \rangle$  at a fixed time  $t = 1$ . While  $\langle Q(T^*) \rangle$  is positive at small values of  $f$ , the dissipation  $\langle Q(t) \rangle$  at fixed times is always negative. Note that at intermediate times  $\langle Q(1) \rangle > \langle Q(T^*) \rangle$ , but if we increase  $f$  furthermore then eventually  $\langle Q(1) \rangle < \langle Q(T^*) \rangle$  (not shown in the figure).

## 9.2. Illustration of an empirical test of the integral fluctuation relation at stopping times

We use the model described in section 9.1 to illustrate how the integral fluctuation relation at stopping times (2) with  $T = T_{\text{fp}} = \inf \{t \geq 0 : S_{\text{tot}}(t) \notin (-s_-, s_+)\}$  can be tested in an experimental setup. To this aim, we will verify whether the quantity  $A$  defined in (74), which is the empirical average of  $e^{-S_{\text{tot}}(T)}$ , converges for  $m_s \rightarrow \infty$  to one. We use the results on the statistics of the empirical average of  $A$ , described in section 8, to validate the statistical significance of the experimental results.

In figure 2 we plot the empirical average (74) as a function of the number of samples  $m_s$  for ten simulation runs. We also plot the theoretical curves  $1 + \sigma_A$  and  $1 - \sigma_A$ , with  $\sigma_A$  the standard deviation of  $A$ , see (76). We observe in figure 2 that all test runs lie within the  $1 \pm \sigma_A$  confidence intervals, and we can thus conclude that numerical experiments are in agreement with the integral fluctuation relation at stopping times and thus also with the fact that  $e^{-S_{\text{tot}}(t)}$  is martingale.

## 9.3. Super Carnot efficiency for heat engines at stopping times

We illustrate the bound (7) for the efficiency of stationary stochastic heat engines at stopping times with a Brownian gyrator [55], which is arguably one of the simplest models of a Feynman ratchet. This system is described by two degrees of freedom  $x_1, x_2$  that are driven by an external force field  $f(x_1, x_2)$  and interact via a potential  $v(x_1, x_2)$ . Two thermal reservoirs at temperatures  $T_h$  and  $T_c$ , with  $T_h > T_c$ , interact independently with the coordinates  $x_1$  and  $x_2$  of the system, respectively. We are specifically interested in the efficiency  $\eta_T = -\langle W(T) \rangle / \langle Q_h(T) \rangle$ , which is the ratio between the work  $-W(T)$  the gyrator performs on its surroundings in a time interval  $[0, T]$ , and the heat  $Q_h(T)$  absorbed by the gyrator in the same time interval, with  $T$  the stopping time at which a specific criterion is first satisfied. The efficiency  $\eta_T$  is a measure of the average amount of work the gyrator performs on its environment.

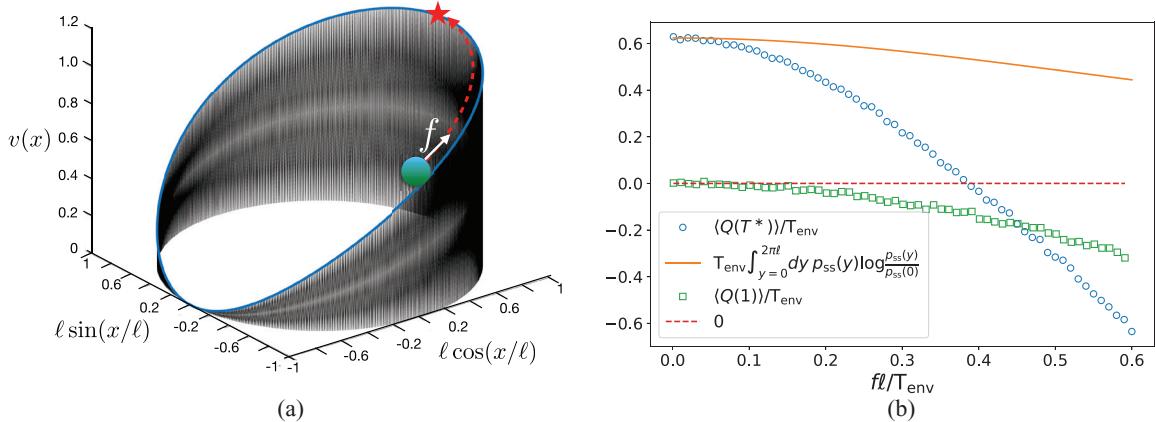
We consider a Brownian gyrator described by the two coupled stochastic differential equations [55–59]

$$\frac{dX_1}{dt} = -\mu \frac{\partial v(X_1(t), X_2(t))}{\partial x_1} + \mu f_1(X_1(t), X_2(t)) + \sqrt{2d_1} \zeta_1(t), \quad (85)$$

$$\frac{dX_2}{dt} = -\mu \frac{\partial v(X_1(t), X_2(t))}{\partial x_2} + \mu f_2(X_1(t), X_2(t)) + \sqrt{2d_2} \zeta_2(t). \quad (86)$$

Here  $\mu$  is the mobility coefficient,  $d_1 = \mu T_h$  and  $d_2 = \mu T_c$  are the diffusion coefficients of the two degrees of freedom,  $v$  is a potential, and  $f_1$  and  $f_2$  are two external nonconservative

## Integral fluctuation relations for entropy production at stopping times



**Figure 1.** Heat extraction of a colloidal particle in a nonequilibrium stationary state. Panel (a): Illustration of a colloidal particle that is driven by a nonconservative  $f$  until the time  $T^*$  when it reaches the highest point of the ‘hill’ denoted by the star, at which the process is stopped. The plotted energy function  $v(x)$  is given by (81) with parameters  $T_{\text{env}} = 1$  and  $\ell = 1$ . Panel (b): Illustration of the bound (5) for the absorbed heat  $\langle Q(T^*) \rangle$  at stopping times  $T^*$  in the model (78) with the energy function plotted in panel (a). The parameters used are  $\mu = 1$ ,  $T_{\text{env}} = 1$ ,  $\ell = 1$ . Markers denote empirical averages that estimate  $\langle Q(T^*) \rangle$  (blue circles) and  $\langle Q(1) \rangle$  (green squares) using 10000 simulated trajectories. The solid orange line denotes the bound on the right-hand side of (84) which follows from the second law of thermodynamics (4) at stopping times, and the red dashed line is simply equal to zero.

forces, whose functional form we specify below. We use the model from [57], for which the potential is

$$v(x_1, x_2) = \frac{1}{2} (u_1 x_1^2 + u_2 x_2^2 + c x_1 x_2), \quad (87)$$

with  $u_1, u_2, c > 0$  and  $c < \sqrt{u_1 u_2}$  and for which the two components of the external nonconservative force are

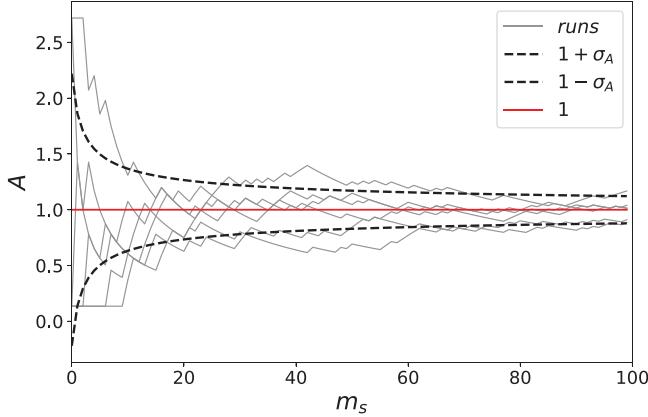
$$f_1(x_1, x_2) = kx_2, \quad f_2(x_1, x_2) = -kx_1. \quad (88)$$

The two thermal reservoirs induce two stochastic forces with amplitudes  $d_1$  and  $d_2$ , which appear in equations (85) and (86) as two independent Gaussian white noises  $\zeta_1$  and  $\zeta_2$  with zero mean and autocorrelation ( $i, j = 1, 2$ )

$$\langle \zeta_i(t) \rangle = \langle \zeta_j(t) \rangle = 0, \quad \langle \zeta_i(t) \zeta_j(t') \rangle = \delta_{i,j} \delta(t - t'). \quad (89)$$

Because of the external driving forces  $f_1$  and  $f_2$  and the presence of two thermal reservoirs at different temperatures, the gyrator develops a nonequilibrium stationary state characterised by a current in the clockwise direction, see figure 3, and a non-zero entropy production. At stationarity, we measure the work  $W$  that the external driving force exerts on the gyrator and the net heat  $Q_h$  and  $Q_c$  that the system absorbs from the hot and cold reservoirs, respectively. Following Sekimoto [39, 53], these quantities are, respectively,

## Integral fluctuation relations for entropy production at stopping times



**Figure 2.** Empirical test of the integral fluctuation relation at stopping times for  $T = T_{\text{fp}} = \inf \{t \geq 0 : S_{\text{tot}}(t) \notin (-s_-, s_+)\}$ . We plot the empirical average  $A$ , see (74), for ten simulation runs as a function of the number of samples  $m_s$ . The simulation runs are for the model defined in section 9.1 with parameters  $\mu = T_{\text{env}} = \ell = 1$  and  $f = 0.1$ , as in figure 1. The threshold parameters that define the stopping time are  $s_+ = 2$  and  $s_- = 1$ . Dashed lines denote the  $1 \pm \sigma_A$  confidence intervals using formula (76). The red line denotes  $A = 1$ .

$$W(t) = \int_0^t f_1(X_1(t'), X_2(t')) \circ dX_1(t') + \int_0^t f_2(X_1(t'), X_2(t')) \circ dX_2(t'), \quad (90)$$

$$Q_h(t) = \int_0^t \frac{\partial v(X_1(t'), X_2(t'))}{\partial x_1} \circ dX_1(t') - \int_0^t f_1(X_1(t'), X_2(t')) \circ dX_1(t'), \quad (91)$$

$$Q_c(t) = \int_0^t \frac{\partial v(X_1(t'), X_2(t'))}{\partial x_2} \circ dX_2(t') - \int_0^t f_2(X_1(t'), X_2(t')) \circ dX_2(t'), \quad (92)$$

where  $\circ$  denotes that the stochastic integrals are interpreted in the Stratonovich sense. When  $k < k_s$  this system operates as an engine [57], i.e.  $\langle W(t) \rangle < 0$ ,  $\langle Q_h(t) \rangle > 0$  and  $\langle Q_c(t) \rangle > 0$ , with

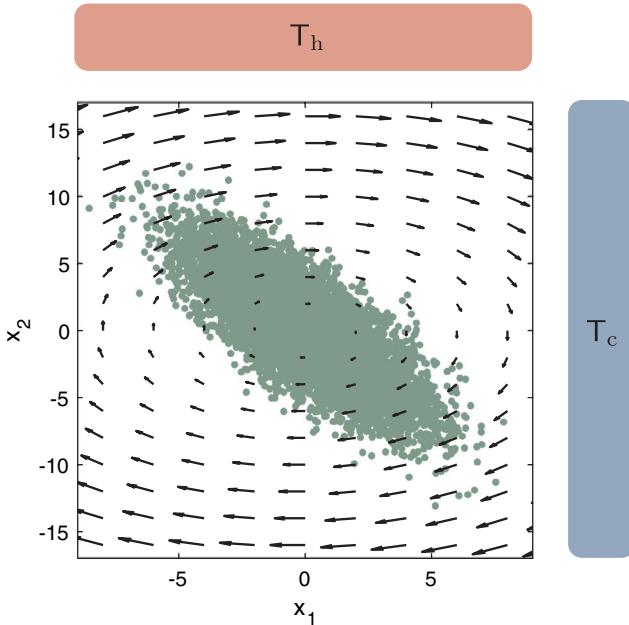
$$k_s := c \frac{\eta_C}{2 - \eta_C}, \quad (93)$$

the stall parameter and  $\eta_C$  Carnot's efficiency (49). The efficiency of the engine in the nonequilibrium stationary state satisfies

$$\eta = -\frac{\langle \frac{dW}{dt} \rangle}{\langle \frac{dQ_h}{dt} \rangle} = \frac{2k}{c+k} \leq \eta_C. \quad (94)$$

Note that when  $k \rightarrow k_s$ , then  $\eta \rightarrow \eta_C$ .

We now investigate the efficiency of the Brownian gyrator at stopping times,  $\eta_T = -\langle W(T) \rangle / \langle Q_h(T) \rangle$ . The simplest example of stopping times are the trivial stopping times  $T = t$ , with  $t$  a fixed time, for which  $\eta_t \leq \eta_C$ . A more interesting example is the time  $T_{\text{me}}$  of the *main event*, i.e. the gyrator crosses the positive  $x_2$  axis while moving in the clockwise direction, occurs for the first time. Mathematically,  $T_{\text{me}}$  can be defined



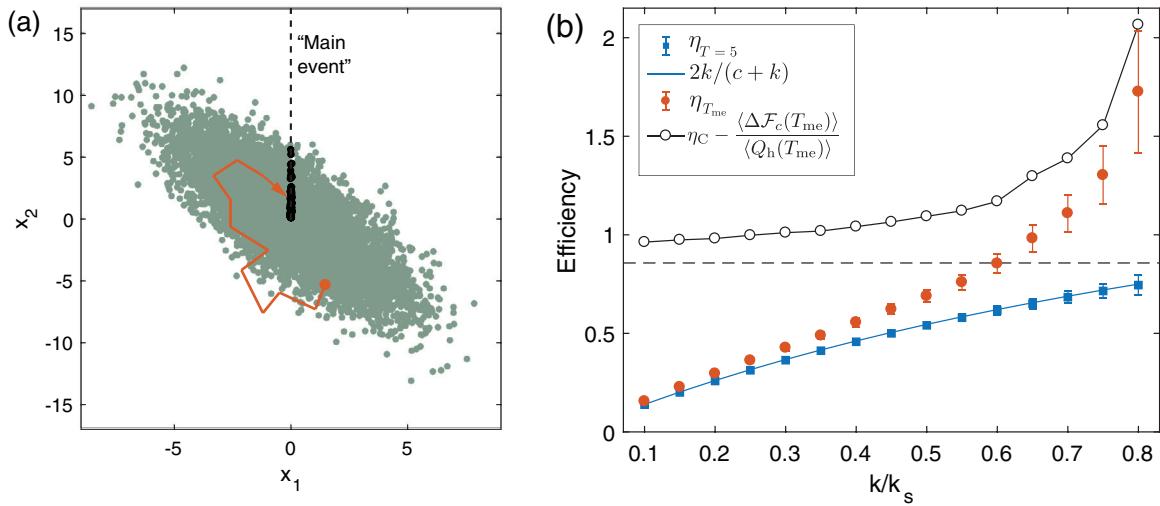
**Figure 3.** Snapshots of a Brownian gyrator (green circles), described by equations (85) and (86), sampled from the nonequilibrium stationary distribution. The black arrows illustrate the non-conservative forces given by equation (88). The system is trapped with a potential given by equation (87) and put simultaneously in contact with a hot (red box) and a cold (blue box) reservoir that act on the  $x_1$  and  $x_2$  coordinates, respectively. Values for the parameters are:  $\mu = 1$ ,  $u_1 = 1$ ,  $u_2 = 1.2$ ,  $T_c = 1$ ,  $T_h = 7$ ,  $c = 0.9$ ,  $k = 0.23$ . The markers (green circles) are obtained from the  $(x_1(t), x_2(t))$  coordinates of  $10^4$  numerical simulations of equations (85) and (86) evaluated at  $t = 5$ .

as follows: let  $z(t) = x_1(t) + ix_2(t)$  be a complex number whose real and imaginary parts are  $x_1(t)$  and  $x_2(t)$  respectively; we define  $z(t) = r(t)e^{i\varphi(t)}$ , with  $r(t) = \sqrt{x_1^2(t) + x_2^2(t)}$  its modulus and with  $\varphi(t) = \tan^{-1}(x_2(t)/x_1(t)) \in [-\pi, \pi]$  its phase; the stopping time at the main event is defined as

$$T_{me} = \inf \left\{ t > 0 : \lim_{\epsilon \downarrow 0} \varphi(t - \epsilon) > \pi/2, \lim_{\epsilon \downarrow 0} \varphi(t + \epsilon) \leq \pi/2, \varphi(t) > 0 \right\}. \quad (95)$$

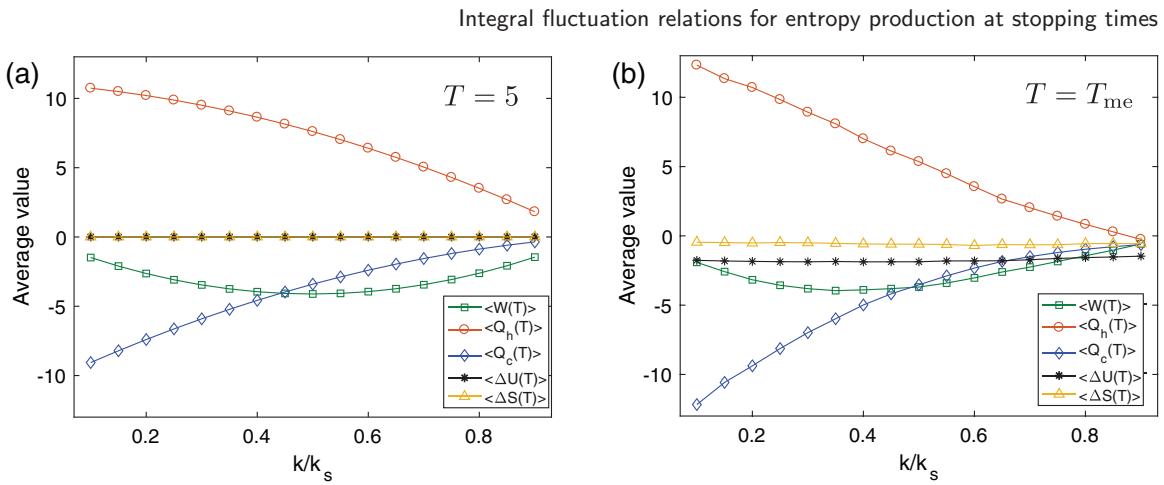
Figure 4(a) illustrates the stopping strategy defined by equation (95) with numerical simulations. In figure 4(b), we compare the values of the efficiency  $\eta_t$  at a fixed time  $t = 5$  with the efficiency  $\eta_{T_{me}}$  at the main event, both as a function of the driving parameter  $k$ . We observe that  $\eta_t$  is well described by equation (94) and thus is smaller than the Carnot efficiency, whereas  $\eta_{T_{me}}$  can surpass the Carnot efficiency if the strength of the driving force  $k$  is large enough (figure 4(b) red circles). Interestingly, the observed super Carnot efficiencies at stopping times are in agreement with the bound (7),  $\eta_{T_{me}} \leq \eta_C + \langle \Delta \mathcal{F}_c(T_{me}) \rangle / \langle Q_h(T_{me}) \rangle$  and thus compatible with the second law at stopping times (4). Moreover, the bound (7) becomes tight when  $k$  is large, which corresponds to a close-to-equilibrium limit (see figure 1(b)). Hence, efficiencies of stopped engines can surpass the Carnot bound if  $\langle Q_h(T) \rangle > 0$  and  $\langle \Delta S_{sys}(T) \rangle > T_c \langle \Delta v(T) \rangle$ , which is consistent with the bound (7).

## Integral fluctuation relations for entropy production at stopping times

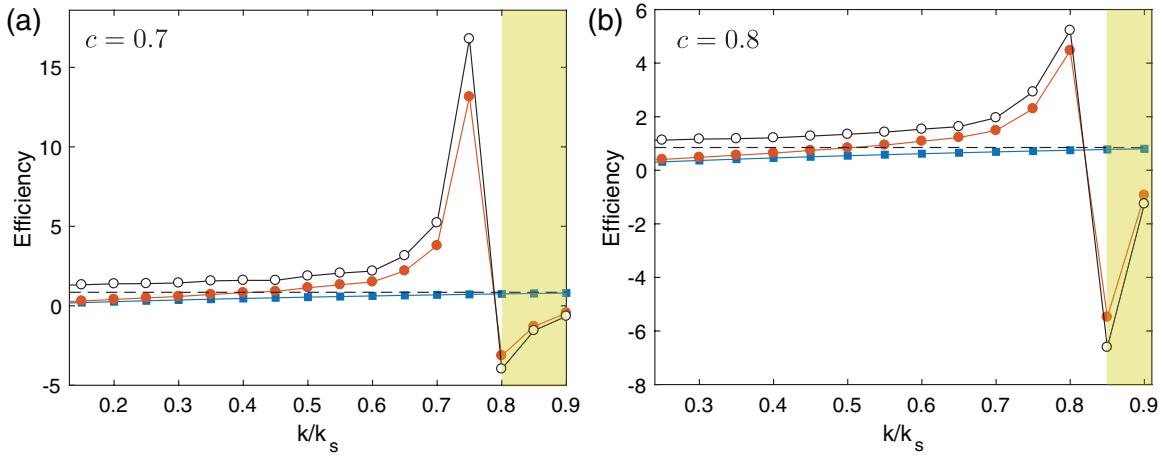


**Figure 4.** Efficiency at stopping times for the Brownian gyrator. (a) Illustration of the stopping time at the main event with stopping time  $T_{\text{me}}$ . A gyrator drawn initially from its stationary distribution (red circle) is monitored until its location crosses the black dashed line in the clockwise direction (red arrow). The location of 100 gyrators at the stopping time are shown with black filled circles whereas the green circles denote the location of the gyrator at time  $t = 5$ . (b) Simulation results for the efficiency  $\eta_t = -\langle W(t) \rangle / \langle Q_h(t) \rangle$  at fixed time  $t = 5$  (blue squares) and the efficiency  $\eta_{T_{\text{me}}} = -\langle W(T_{\text{me}}) \rangle / \langle Q_h(T_{\text{me}}) \rangle$  at the main event  $T_{\text{me}}$  (red circles), defined by equation (95), as a function of the parameter  $k/k_s$  that quantifies the strength of the nonequilibrium driving, see (88). The blue line is the theoretical value of the efficiency at large times given by equation (94), and the black circles correspond to the bound (7); the black line is a guide to the eye. The horizontal black dashed line is set at Carnot efficiency. Values of the parameters used in simulations are  $\mu = 1$ ,  $u_1 = 1$ ,  $u_2 = 1.2$ ,  $T_c = 1$ ,  $T_h = 7$ , and  $c = 0.9$ ; markers denote average values estimated from  $10^4$  independent realisations initially sampled from the stationary state. Numerical simulations are performed with the Heun's numerical integration scheme with a time step  $\Delta t = 10^{-3}$  [2]. Error bars denote the standard errors of the empirical mean.

Notice that efficiencies of stopped heat engines can be larger than one because internal system energy can be converted into useful work. To better understand this feature, we plot in figure 5 the average energetic fluxes at stopping times, namely,  $\langle W(T) \rangle$ ,  $\langle Q_h(T) \rangle$  and  $\langle Q_c(T) \rangle$ , together with the change of the internal system energy  $\langle \Delta v(T) \rangle$  and the internal system entropy change  $\langle \Delta S_{\text{sys}}(T) \rangle$ . At fixed times  $T = 5$ , we observe the well-known features of a cyclic heat engine for which  $\langle \Delta v(t) \rangle \simeq \langle \Delta S_{\text{sys}}(t) \rangle \simeq 0$  (figure 5(a)). Although the heat and work fluxes of stopped engines have the same sign as those of cyclic heat engines, the  $\langle \Delta v(T_{\text{me}}) \rangle < 0$  and  $\langle \Delta S_{\text{sys}}(T_{\text{me}}) \rangle < 0$ , which indicates that on average the energy and entropy of the gyrator are at  $T_{\text{me}}$  smaller than their initial values (figure 5(b)). This result suggests a recipe in the quest of super Carnot efficiencies at stopping times, namely by designing stopping strategies that lead to a reduction of the energy of the system. In our example,  $\langle \Delta S_{\text{sys}}(T_{\text{me}}) \rangle > \langle \Delta v(T_{\text{me}}) \rangle$ , which enables the appearance of super-Carnot stopping-time efficiencies which are nevertheless compatible with the second law at stopping times. This result motivates further



**Figure 5.** Average values of thermodynamic observables (see legend) in the Brownian gyrator are compared for a fixed time  $T = 5$  (a) with those for the stopping time  $T = T_{me}$  (b), defined in equation (95). The values of the parameters used in the numerical simulations are the same as in figure 3.



**Figure 6.** Simulation results for the stopping time efficiency  $\eta_T$ . Results for the efficiency at a fixed time  $T = 5$  (blue circles) are compared with those at the stopping time  $T = T_{me}$  of the main event. (a) and (b) are obtained for two different values of the parameter  $c$  (see legend), with the other simulation parameters set to the same values as in figure 3. The black circles correspond to the right-hand side of equation (7) and the black dashed line is set to Carnot efficiency; the lines between symbols are a guide to the eye. The yellow shaded area illustrates the range of parameters at which the system behaves as a ‘type-III heater’ at stopping times [48, 62], for which the right-hand side of (7) becomes a lower bound to the stopping-time efficiency.

research on so-called type-II efficiencies at stopping times defined as the ratio between the average input and output fluxes of entropy production [44, 60, 61].

In figure 6 we illustrate the bound provided by equation (7) for a wider range of parameters. We observe that when  $k$  exceeds the stall parameter  $k_s$  the thermodynamic fluxes obey  $\langle W(T_{me}) \rangle < 0$ ,  $\langle Q_h(T_{me}) \rangle < 0$  and  $\langle Q_c(T_{me}) \rangle < 0$ . This behaviour is still compatible with the second law of thermodynamics at stopping times. Indeed, for

this range of parameters  $\eta_{T_{\text{me}}} < 0$  and the bound (7) becomes  $\eta_{T_{\text{me}}} \geq \eta_C - \frac{\langle \Delta \mathcal{F}_c(T) \rangle}{\langle Q_h(T) \rangle}$ . This bound is corroborated by numerical simulations in figures 6(a) and (b). We define this type of operation as a ‘type III heater’, which is different than the type I and type II heaters [48, 62], that appear in stochastic thermodynamics at fixed times.

## 10. Discussion

We have derived fluctuation relations at stopping times for the entropy production of stationary processes. These fluctuation relations imply a second law of thermodynamics at stopping times—which states that on average entropy production in a nonequilibrium stationary state always increases, even when the experimental observer measures the entropy production at a stopping time—and imply that certain fluctuation properties of entropy production are universal.

We have shown that the second law of thermodynamics at stopping times has important consequences for the nonequilibrium thermodynamics of small systems. For instance, the second law at stopping times implies that it is possible to extract on average heat from an isothermal environment by applying stopping strategies to a physical system in a nonequilibrium stationary state; heat is thus extracted from the environment without using a feedback control, as is the case with Maxwell demons [63–65]. Furthermore, we have demonstrated, using numerical simulations with a Brownian gyrator, that the efficiency of a stationary stochastic heat engine can surpass Carnot’s efficiency when the engine is stopped at a cleverly chosen moment. This result is compatible with the second law at stopping times, which provides a bound on the efficiency of stochastically stopped engines. Note that the heat engines described in this paper are non cyclic devices since they are stopped at the stopping time  $T$ . It would be interesting to explore how stochastically stopped engines can be implemented in experimental systems such as, electrical circuits [66], autonomous single-electron heat engines [67], feedback traps [68], colloidal heat engines [11], and Brownian gyrators [56].

Integral fluctuation relations at stopping times imply bounds on the probability of events of negative entropy production that are stronger than those obtained with the integral fluctuation relation at fixed times. For example, the integral fluctuation relation at stopping times implies that the cumulative distribution of infima of entropy production is bounded by an exponential distribution with mean  $-1$ . Moreover, for continuous processes the integral fluctuation relation at stopping times implies that the cumulative distribution of the global infimum of entropy production is equal to an exponential distribution. A reason why the integral fluctuation relation at stopping times is more powerful than the integral fluctuation relation at fixed times—in the sense of bounding the likelihood of events of negative entropy production—is because with stopping times we can describe fluctuations of entropy production that are not accessible with fixed-time arguments, such as fluctuations of infima of entropy production.

The integral fluctuation relation at stopping times implies also bounds on other fluctuation properties of entropy production, not necessarily related to events of negative entropy production. For example, we have used the integral fluctuation relation

to derive bounds on splitting probabilities of entropy production and on the number of times entropy production crosses a certain interval. For continuous processes these fluctuation properties of entropy production are universal, and we have obtained generic expressions for these fluctuation properties of entropy production.

Since martingale theory has proven to be very useful to derive generic results about the fluctuations of entropy production, the question arises what other physical processes are martingales. In this context, the exponential of the housekeeping heat has been demonstrated to be a martingale [30, 69]. The housekeeping heat is an extension of entropy production to the case of non-stationary processes. We expect that all formulas presented in this paper extend in a straightforward manner to the case of housekeeping heat. Recently also a martingale related to the quenched dynamics of spin models [70] and a martingale in quantum systems [71] have been discovered.

There exist universal fluctuation properties that are implied by the martingale property of  $e^{-S_{\text{tot}}(t)}$ , but are not discussed in this paper. For example, martingale theory implies a symmetry relation for the distribution of conditional stopping times [21, 72–74]. These symmetry relations are also proved using the optional stopping theorem, but they cannot be seen as a straightforward consequence of the integral fluctuation relation at stopping times.

Since the integral fluctuation relation at stopping times (39) is a direct consequence of the martingale property  $e^{-S_{\text{tot}}(t)}$ , testing the fluctuation relation (39) in experiments could serve as a method to demonstrate that  $e^{-S_{\text{tot}}(t)}$  is a martingale. It is not so easy to show in an experiment that a stochastic process is a martingale: it is a herculean task to verify the condition (28). A recent experiment [75] shows that the entropy production of biased transport of single charges in a double electronic dot behaves as a martingale. The inequality (64) for the infima of entropy production was shown to be valid in this experiment. The integral fluctuation relation at stopping times (39) provides an interesting alternative to test martingality of  $e^{-S_{\text{tot}}(t)}$ , because the integral fluctuation relation at stopping times is an equality. Hence, the integral fluctuation relation at stopping times could serve as a proxy for the martingale structure of  $e^{-S_{\text{tot}}(t)}$  in experiments.

Testing the integral fluctuation relation at stopping times in experimental setting may also be advantageous with respect to testing the standard fluctuation relation at fixed times. The number of samples required to test the standard fluctuation relation increases exponentially with time, since events of negative entropy are rare. This makes it difficult to test the conditions of stochastic thermodynamics at large time scales with the standard integral fluctuation relation. Moreover, at fixed times the distribution of the empirical mean of the exponentiated negative entropy production is not known. The integral fluctuation relation at stopping times does not have these issues, since negative fluctuations of entropy can be capped at a fixed value  $-s_-$ , which is independent of time  $t$ , and these negative values can be reached at any time. Moreover, we have derived an exact universal expression for the distribution of the sample mean of the exponentiated negative entropy production at stopping times, which can be used to determine the statistical significance of empirical tests of the integral fluctuation relation at stopping times. The integral fluctuation relation at stopping times is thus a useful relation to test the conditions of stochastic thermodynamics in a certain experimental setup.

## Acknowledgments

We thank A C Barato, R Belousov, S Bo, R Chétrite, R Fazio, S Gupta, D Hartich, C Jarzynski, R Johal, C Maes, G Manzano, K Netocný, P Pietzonka, M Polettini, K Proesmans, S Sasa, F Severino, and K Sekimoto for fruitful discussions.

## Appendix A. The exponential of the negative entropy production is a (uniformly integrable) martingale

We prove that  $e^{-S_{\text{tot}}(t)}$  is a martingale. To this aim, we have to verify the two conditions (27) and (28) for  $M(t) = e^{-S_{\text{tot}}(t)}$ . We present two proofs, one for processes in discrete time using the expression (1) for the entropy production, and a general proof for reversible right-continuous processes  $\vec{X}(t)$  using the expression (33) for the entropy production.

### A.1. Reversible processes in discrete time

In discrete time

$$e^{-S_{\text{tot}}(t,\omega)} = \frac{\tilde{p}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_t)}{p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_t)} \quad (\text{A.1})$$

with  $\tilde{p}$  the probability density function associated with the time-reversed dynamics. We have simplified the notation a bit and used that  $\vec{x}(t) = \vec{x}_t$ .

The condition (27) follows from  $e^{-S_{\text{tot}}(t,\omega)} \geq 0$  and  $E[e^{-S_{\text{tot}}(t,\omega)}] = 1$  for all  $t \geq 0$ . The martingale condition (28) also holds,

$$\begin{aligned} E[e^{-S_{\text{tot}}(t,\omega)} | \mathcal{F}_s] &= \int \left( \prod_{n=s+1}^t d\vec{x}_n \right) p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_t | \vec{x}_1, \vec{x}_2, \dots, \vec{x}_s) \\ &\quad \times \frac{\tilde{p}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_t)}{p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_t)} \\ &= \int \left( \prod_{n=s+1}^t d\vec{x}_n \right) \frac{p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_t)}{p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_s)} \frac{\tilde{p}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_t)}{p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_t)} \\ &= \int \left( \prod_{n=s+1}^t d\vec{x}_n \right) \frac{\tilde{p}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_t)}{p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_s)} \\ &= \frac{\tilde{p}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_s)}{p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_s)} \\ &= e^{-S_{\text{tot}}(s,\omega)}. \end{aligned} \quad (\text{A.2})$$

### A.2. Reversible stationary processes that are right-continuous

We assume that  $\vec{X}(t)$  is right-continuous and that the two measures  $\mathbb{P}$  and  $\mathbb{P} \circ \Theta$  are locally mutually absolutely continuous, such that the entropy production (33) can be defined. Because of the definition of entropy production,

$$e^{-S_{\text{tot}}(t,\omega)} = \frac{d(\mathbb{P} \circ \Theta)|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}}. \quad (\text{A.3})$$

The condition (27) follows from  $e^{-S_{\text{tot}}(t,\omega)} \geq 0$  and  $E[e^{-S_{\text{tot}}(t,\omega)}] = 1$  for all  $t \geq 0$ .

The martingale condition (28) follows readily from (A.3):

$$E[e^{-S_{\text{tot}}(t,\omega)}|\mathcal{F}_s] = e^{-S_{\text{tot}}(s,\omega)}, \quad s \leq t. \quad (\text{A.4})$$

The relation (A.4) is a direct consequence of (A.3) and the definition of the Radon–Nikodym derivative: the Radon–Nikodym derivative  $e^{-S_{\text{tot}}(s,\omega)} = \frac{d(\mathbb{P} \circ \Theta)|_{\mathcal{F}_s}}{d\mathbb{P}|_{\mathcal{F}_s}}$  is by definition a  $\mathcal{F}_s$ -measurable random variable for which

$$E_{\mathbb{P}} \left[ I_{\Phi} \frac{d(\mathbb{P} \circ \Theta)|_{\mathcal{F}_s}}{d\mathbb{P}|_{\mathcal{F}_s}} \right] = E_{\mathbb{P} \circ \Theta} [I_{\Phi}] \quad (\text{A.5})$$

for all  $\Phi \in \mathcal{F}_s$ . We show that  $E[e^{-S_{\text{tot}}(t,\omega)}|\mathcal{F}_s]$  is a random variable with this property, and therefore (A.4) is valid. Indeed, for  $\Phi \in \mathcal{F}_s$  it holds that:

$$\begin{aligned} E_{\mathbb{P}} [I_{\Phi} E_{\mathbb{P}} [e^{-S_{\text{tot}}(t,\omega)}|\mathcal{F}_s]] &= E_{\mathbb{P}} [E_{\mathbb{P}} [I_{\Phi} e^{-S_{\text{tot}}(t,\omega)}|\mathcal{F}_s]] \\ &= E_{\mathbb{P}} [I_{\Phi} e^{-S_{\text{tot}}(t,\omega)}] \\ &= E_{\mathbb{P} \circ \Theta} [I_{\Phi}]. \end{aligned} \quad (\text{A.6})$$

Note that the martingale condition (28) is consistent with the tower property of conditional expectations:  $E[E[X|\mathcal{F}_t]|\mathcal{F}_s] = E[X|\mathcal{F}_s]$ .

### A.3. Uniform integrability

We prove that the stochastic process  $e^{-S_{\text{tot}}(t \wedge \tau)}$ , with  $t \in [0, \infty)$  and  $\tau$  a fixed positive number, is uniformly integrable. The process  $e^{-S_{\text{tot}}(t \wedge \tau)}$  can be written as

$$e^{-S_{\text{tot}}(t \wedge \tau)} = E[e^{-S_{\text{tot}}(\tau,\omega)}|\mathcal{F}_t]. \quad (\text{A.7})$$

Since a stochastic process  $Y(t)$  of the form  $Y(t) = E[Z|\mathcal{F}_t]$  is uniformly integrable [26], we obtain that  $e^{-S_{\text{tot}}(t \wedge \tau)}$  is uniformly integrable.

## Appendix B. Integral fluctuation relation for entropy production within infinite-time windows

We derive two corollaries of the optional stopping theorem, which is theorem 3.6 in [26] and equation (29) in this paper.

**Corollary 1.** *Let  $T$  be a stopping time of a stationary process  $\vec{X}(t)$  and let  $S_{\text{tot}}(t)$  be the stochastic entropy production of  $\vec{X}(t)$  as defined in (33), with the two measures  $\mathbb{P}$  and  $\mathbb{P} \circ \Theta$  locally mutually absolutely continuous. If  $t$  is continuous, then  $S_{\text{tot}}(t)$  is assumed to be right-continuous. If the two conditions*

- (i)  $\mathbb{P}(T < \infty) = 1,$
- (ii)  $\lim_{t \rightarrow \infty} E[e^{-S_{\text{tot}}(t)} I_{T>t}] = 0,$

are met, then

$$\langle e^{-S_{\text{tot}}(T)} \rangle = 1. \quad (\text{B.1})$$

Recall that  $I_\Phi(\omega)$  is the indicator function defined in (31). We use a proof analogous to the discrete time proof of theorem 8.3.5 on page 222 of [76].

**Proof.** We decompose  $\langle e^{-S_{\text{tot}}(T)} \rangle$  into three terms,

$$\langle e^{-S_{\text{tot}}(T)} \rangle = \langle e^{-S_{\text{tot}}(T \wedge t)} \rangle - \langle e^{-S_{\text{tot}}(t)} I_{T>t} \rangle + \langle e^{-S_{\text{tot}}(T)} I_{T>t} \rangle.$$

Since the right-hand side holds for arbitrary values of  $t$  we can take  $\lim_{t \rightarrow \infty}$ . Using (39) and the condition (ii) we obtain

$$\langle e^{-S_{\text{tot}}(T)} \rangle = 1 + \lim_{t \rightarrow \infty} \langle e^{-S_{\text{tot}}(T)} I_{T>t} \rangle. \quad (\text{B.2})$$

Because of condition (i), it holds that  $\lim_{t \rightarrow \infty} e^{-S_{\text{tot}}(T)} I_{T>t} = 0$  in the  $\mathbb{P}$ -almost sure sense. Because  $e^{-S_{\text{tot}}(T)} I_{T>t}$  is a nonnegative monotonic decreasing sequence we can apply the monotone convergence theorem, see e.g. [33], and we obtain

$$\lim_{t \rightarrow \infty} \langle e^{-S_{\text{tot}}(T)} I_{T>t} \rangle = \langle \lim_{t \rightarrow \infty} e^{-S_{\text{tot}}(T)} I_{T>t} \rangle = 0. \quad (\text{B.3}) \quad \square$$

**Corollary 2.** Let  $T$  be a stopping time of a stationary process  $\vec{X}(t)$  and let  $S_{\text{tot}}(t)$  be the stochastic entropy production of  $\vec{X}(t)$  as defined in (33), with the two measures  $\mathbb{P}$  and  $\mathbb{P} \circ \Theta$  locally mutually absolutely continuous. If  $t$  is continuous, then  $S_{\text{tot}}(t)$  is assumed to be right-continuous. If the two conditions

- (i)  $\lim_{t \rightarrow \infty} e^{-S_{\text{tot}}(t)} = 0$  in the  $\mathbb{P}$ -almost sure sense,
- (ii) there exist two positive numbers  $s_-$  and  $s_+$  such that  $S_{\text{tot}}(t) \in (-s_-, s_+)$  for all  $t \leq T$ ,

are met, then

$$\langle e^{-S_{\text{tot}}(T)} \rangle = 1. \quad (\text{B.4})$$

**Proof.** We show that if the conditions of the present corollary are met, then also the conditions of corollary 1 are met, and therefore (B.4) holds.

Because of condition (i), the stopping time  $T$  is almost surely finite:  $1 = \lim_{t \rightarrow \infty} \mathbb{P}[S_{\text{tot}}(t) > s_+] \leq \mathbb{P}[T < \infty]$  and  $\mathbb{P}[T < \infty] \leq 1$ , therefore  $\mathbb{P}[T < \infty] = 1$ .

Because of condition (i) and (ii) we obtain  $\lim_{t \rightarrow \infty} E[e^{-S_{\text{tot}}(t)} I_{T>t}] = 0$ :  $e^{-S_{\text{tot}}(t)} I(T > t) \leq e^{s_-} I(T > t)$  is a positive variable that is bounded from above. Hence, the dominated convergence theorem applies and  $\lim_{t \rightarrow \infty} E[e^{-S_{\text{tot}}(t)} I_{T>t}] = E[\lim_{t \rightarrow \infty} e^{-S_{\text{tot}}(t)} I_{T>t}] = 0$ .  $\square$

## Appendix C. Derivation of the bound (71) on the statistics of $N_x$

We denote by  $M_x(t)$  the number of times entropy production has crossed the interval  $[-\Delta, \Delta]$  in the direction  $-\Delta \rightarrow \Delta$  within the time interval  $[0, t]$ , and therefore  $N_x = \lim_{t \rightarrow \infty} M_x(t)$ .

We define two sequences of stopping times  $T_n$  and  $\tilde{T}_n$ , with  $n \in [0, N_x] \cap \mathbb{Z}$ , namely,

$$T_n = \inf \{t : M_x(t) \geq n\} \quad (\text{C.1})$$

and

$$\tilde{T}_n = \inf \{t : t \geq T_n, S_{\text{tot}}(t) \notin (-\Delta, s_+)\}, \quad (\text{C.2})$$

where  $s_+$  is considered to be a very large positive number.

We apply the fluctuation relation (41) to the two stopping times  $\tilde{T}_n$  and  $T_n$ . The integral fluctuation relation (41) implies that

$$E \left[ e^{-S_{\text{tot}}(\tilde{T}_n)} | N_x \geq n \right] = E \left[ e^{-S_{\text{tot}}(T_n)} | N_x \geq n \right]. \quad (\text{C.3})$$

Since  $e^{-S_{\text{tot}}(T_n)} \leq e^{-\Delta}$ , the right-hand side of (C.3) is bounded by

$$E \left[ e^{-S_{\text{tot}}(T_n)} | N_x \geq n \right] \leq e^{-\Delta}. \quad (\text{C.4})$$

The left-hand side of (C.3) can be decomposed into two terms,

$$\begin{aligned} E \left[ e^{-S_{\text{tot}}(\tilde{T}_n)} | N_x \geq n \right] &= E \left[ e^{-S_{\text{tot}}(\tilde{T}_n)} I_{S_{\text{tot}}(\tilde{T}_n) \geq s_+} | N_x \geq n \right] \\ &\quad + E \left[ e^{-S_{\text{tot}}(\tilde{T}_n)} I_{S_{\text{tot}}(\tilde{T}_n) \leq -\Delta} | N_x \geq n \right], \end{aligned} \quad (\text{C.5})$$

where  $I$  is the indicator function defined in (31). We take the limit  $s_+ \rightarrow \infty$  and obtain

$$\begin{aligned} \lim_{s_+ \rightarrow \infty} E \left[ e^{-S_{\text{tot}}(\tilde{T}_n)} | N_x \geq n \right] &= \lim_{s_+ \rightarrow \infty} E \left[ e^{-S_{\text{tot}}(\tilde{T}_n)} I_{S_{\text{tot}}(\tilde{T}_n) \leq -\Delta} | N_x \geq n \right] \\ &\geq e^{\Delta} \lim_{s_+ \rightarrow \infty} E \left[ I_{S_{\text{tot}}(\tilde{T}_n) \leq -\Delta} | N_x \geq n \right]. \end{aligned}$$

Since

$$\mathbb{P}[N_x \geq n + 1 | N_x \geq n] = \lim_{s_+ \rightarrow \infty} E \left[ I_{S_{\text{tot}}(\tilde{T}_n) \leq -\Delta} | N_x \geq n \right] \quad (\text{C.6})$$

we obtain the inequality

$$\lim_{s_+ \rightarrow \infty} E \left[ e^{-S_{\text{tot}}(\tilde{T}_n)} | N_x \geq n \right] \geq e^{\Delta} \mathbb{P}[N_x \geq n + 1 | N_x \geq n]. \quad (\text{C.7})$$

The relation (C.3) together with the two inequalities (C.4) and (C.7) imply

$$\mathbb{P}[N_x \geq n + 1 | N_x \geq n] \leq e^{-2\Delta} \quad \text{with } n > 0. \quad (\text{C.8})$$

This is the formula (71) which we were meant to prove.

## References

- [1] Maes C 2003 On the origin and the use of fluctuation relations for the entropy *Séminaire Poincaré* **2** 29–62
- [2] Sekimoto K 2010 *Stochastic Energetics* vol 799 (Berlin: Springer)
- [3] Jarzynski C 2011 Equalities and inequalities: irreversibility and the second law of thermodynamics at the nanoscale *Annu. Rev. Condens. Matter Phys.* **2** 329–51
- [4] Seifert U 2012 Stochastic thermodynamics, fluctuation theorems and molecular machines *Rep. Prog. Phys.* **75** 126001
- [5] Van den Broeck C and Esposito M 2015 Ensemble and trajectory thermodynamics: a brief introduction *Phys. A: Stat. Mech. Appl.* **418** 6–16
- [6] Jülicher F, Ajdari A and Prost J 1997 Modeling molecular motors *Rev. Mod. Phys.* **69** 1269
- [7] Reimann P 2002 Brownian motors: noisy transport far from equilibrium *Phys. Rep.* **361** 57–265
- [8] Bustamante C, Liphardt J and Ritort F 2005 The nonequilibrium thermodynamics of small systems *Phys. Today* **58** 43–8
- [9] Bechinger C, Di Leonardo R, Löwen H, Reichhardt C, Volpe G and Volpe G 2016 Active particles in complex and crowded environments *Rev. Mod. Phys.* **88** 045006
- [10] Gavrilov M, Jun Y and Bechhoefer J 2014 Real-time calibration of a feedback trap *Rev. Sci. Instrum.* **85** 095102
- [11] Martínez I A, Roldán É, Dinis L and Rica R A 2017 Colloidal heat engines: a review *Soft Matter* **13** 22–36
- [12] Martínez I A, Roldán É, Dinis L, Petrov D, Parrondo J M and Rica R A 2016 Brownian carnot engine *Nat. Phys.* **12** 67
- [13] Ciliberto S 2017 Experiments in stochastic thermodynamics: short history and perspectives *Phys. Rev. X* **7** 021051
- [14] Pekola J P 2015 Towards quantum thermodynamics in electronic circuits *Nat. Phys.* **11** 118
- [15] Pekola J and Khaymovich I 2019 Thermodynamics in single-electron circuits and superconducting qubits *Ann. Rev. Condens. Matter Phys.* **10** 193–212
- [16] Seifert U 2005 Entropy production along a stochastic trajectory and an integral fluctuation theorem *Phys. Rev. Lett.* **95** 040602
- [17] Lebowitz J L and Spohn H 1999 A gallavotti-cohen-type symmetry in the large deviation functional for stochastic dynamics *J. Stat. Phys.* **95** 333–65
- [18] Crooks G E 1998 Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems *J. Stat. Phys.* **90** 1481–7
- [19] Gaspard P 2004 Time-reversed dynamical entropy and irreversibility in markovian random processes *J. Stat. Phys.* **117** 599–615
- [20] Chetrite R and Gupta S 2011 Two refreshing views of fluctuation theorems through kinematics elements and exponential martingale *J. Stat. Phys.* **143** 543
- [21] Neri I, Roldán É and Jülicher F 2017 Statistics of infima and stopping times of entropy production and applications to active molecular processes *Phys. Rev. X* **7** 011019
- [22] Pigolotti S, Neri I, Roldán É and Jülicher F 2017 Generic properties of stochastic entropy production *Phys. Rev. Lett.* **119** 140604
- [23] Maitra A P and Sudderth W D 2012 *Discrete Gambling and Stochastic Games* vol 32 (Berlin: Springer)
- [24] Doob J L 1953 *Stochastic Processes* vol 7 (New York: Wiley)
- [25] Williams D 1991 *Probability with Martingales* (Cambridge: Cambridge University Press)
- [26] Liptser R and Shirayev A N 1977 *Statistics of Random Processes: I. General Theory (Applications of Mathematics: Stochastic Modelling and Applied Probability)* vol 5, 1st edn (Berlin: Springer) (translated from Russian by A B Aries)
- [27] Bernouilli D 1954 Exposition of a new theory on the measurement of risk *Econometrica* **22** 23–36 (Translated from Latin by Louise Sommer. Original title: Specimen Theoriae Novae de Mensura Sortis. Original year 1738)
- [28] Bernstein P L 1996 *Against the Gods: the Remarkable Story of Risk* (New York: Wiley)
- [29] LeRoy S F 1989 Efficient capital markets and martingales *J. Econ. Literature* **27** 1583–621
- [30] Chétrite R, Gupta S, Neri I and Roldán É 2018 Martingale theory for housekeeping heat *EPL* **124** 60006
- [31] Hald A 2003 *A History of Probability and Statistics and their Applications Before 1750* vol 501 (New York: Wiley)
- [32] Doyle P G and Snell J L 1984 *Random Walks and Electric Networks* vol 22 (Washington DC: Mathematical Association of America)
- [33] Tao T 2011 *An Introduction to Measure Theory* vol 126 (Providence, RI: American Mathematical Society)

## Integral fluctuation relations for entropy production at stopping times

- [34] Maes C, Redig F and Moffaert A V 2000 On the definition of entropy production, via examples *J. Math. Phys.* **41** 1528–54
- [35] Jiang D-Q, Qian M and Qian M-P 2003 *Mathematical Theory of Nonequilibrium Steady States: on the Frontier of Probability and Dynamical Systems* (Berlin: Springer)
- [36] Seifert U 2008 Stochastic thermodynamics: principles and perspectives *Eur. Phys. J. B* **64** 423–31
- [37] Feynman R P, Leighton R B and Sands M 1963 *The Feynman Lectures on Physics* vol 1 (Reading, MA: Addison-Wesley)
- [38] Parrondo J M and Español P 1996 Criticism of Feynman’s analysis of the ratchet as an engine *Am. J. Phys.* **64** 1125–30
- [39] Sekimoto K 1997 Kinetic characterization of heat bath and the energetics of thermal ratchet models *J. Phys. Soc. Japan* **66** 1234–7
- [40] Magnasco M O and Stolovitzky G 1998 Feynman’s ratchet and pawl *J. Stat. Phys.* **93** 615–32
- [41] Singh V and Johal R 2017 Feynman’s ratchet and pawl with ecological criterion: optimal performance versus estimation with prior information *Entropy* **19** 576
- [42] Sekimoto K, Takagi F and Hondou T 2000 Carnot’s cycle for small systems: irreversibility and cost of operations *Phys. Rev. E* **62** 7759
- [43] Verley G, Esposito M, Willaert T and Van den Broeck C 2014 The unlikely carnot efficiency *Nat. Commun.* **5** 4721
- [44] Verley G, Willaert T, Van den Broeck C and Esposito M 2014 Universal theory of efficiency fluctuations *Phys. Rev. E* **90** 052145
- [45] Polettini M, Verley G and Esposito M 2015 Efficiency statistics at all times: Carnot limit at finite power *Phys. Rev. Lett.* **114** 050601
- [46] Velasco S, Roco J, Medina A and Hernández A C 1997 New performance bounds for a finite-time carnot refrigerator *Phys. Rev. Lett.* **78** 3241
- [47] De Tomás C, Hernández A C and Roco J 2012 Optimal low symmetric dissipation carnot engines and refrigerators *Phys. Rev. E* **85** 010104
- [48] Rana S, Pal P, Saha A and Jayannavar A 2016 Anomalous brownian refrigerator *Phys. A: Stat. Mech. Appl.* **444** 783–98
- [49] Tao T 2012 *Topics in Random Matrix Theory* vol 132 (Providence, RI: American Mathematical Society)
- [50] Speck T, Bickle V, Bechinger C and Seifert U 2007 Distribution of entropy production for a colloidal particle in a nonequilibrium steady state *Europhys. Lett.* **79** 30002
- [51] Koski J, Sagawa T, Saira O, Yoon Y, Kutvonen A, Solinas P, Möttönen M, Ala-Nissila T and Pekola J 2013 Distribution of entropy production in a single-electron box *Nat. Phys.* **9** 644
- [52] Ciliberto S, Imparato A, Naert A and Tanase M 2013 Heat flux and entropy produced by thermal fluctuations *Phys. Rev. Lett.* **110** 180601
- [53] Sekimoto K 1998 Langevin equation and thermodynamics *Prog. Theor. Phys. Suppl.* **130** 17–27
- [54] Gomez-Marin A and Pagonabarraga I 2006 Test of the fluctuation theorem for stochastic entropy production in a nonequilibrium steady state *Phys. Rev. E* **74** 061113
- [55] Filliger R and Reimann P 2007 Brownian gyrator: a minimal heat engine on the nanoscale *Phys. Rev. Lett.* **99** 230602
- [56] Argun A, Soni J, Dabelow L, Bo S, Pesce G, Eichhorn R and Volpe G 2017 Experimental realization of a minimal microscopic heat engine *Phys. Rev. E* **96** 052106
- [57] Pietzonka P and Seifert U 2018 Universal trade-off between power, efficiency, and constancy in steady-state heat engines *Phys. Rev. Lett.* **120** 190602
- [58] Cerasoli S, Dotsenko V, Oshanin G and Rondoni L 2018 Asymmetry relations and effective temperatures for biased brownian gyrators *Phys. Rev. E* **98** 042149
- [59] Manikandan S K, Dabelow L, Eichhorn R and Krishnamurthy S 2019 Efficiency fluctuations in microscopic machines *Phys. Rev. Lett.* **122** 140601
- [60] Bejan A 2016 *Advanced Engineering Thermodynamics* (New York: Wiley)
- [61] Raz O, Subaşı Y and Pugatch R 2016 Geometric heat engines featuring power that grows with efficiency *Phys. Rev. Lett.* **116** 160601
- [62] Campisi M, Pekola J and Fazio R 2015 Nonequilibrium fluctuations in quantum heat engines: theory, example, and possible solid state experiments *New J. Phys.* **17** 035012
- [63] Toyabe S, Sagawa T, Ueda M, Muneyuki E and Sano M 2010 Experimental demonstration of information-to-energy conversion and validation of the generalized jarzynski equality *Nat. Phys.* **6** 988
- [64] Koski J V, Maisi V F, Pekola J P and Averin D V 2014 Experimental realization of a szilard engine with a single electron *Proc. Natl Acad. Sci.* **111** 13786–9
- [65] Parrondo J M, Horowitz J M and Sagawa T 2015 Thermodynamics of information *Nat. Phys.* **11** 131

## Integral fluctuation relations for entropy production at stopping times

- [66] Ciliberto S, Imparato A, Naert A and Tanase M 2013 Statistical properties of the energy exchanged between two heat baths coupled by thermal fluctuations *J. Stat. Mech.* **P12014**
- [67] Koski J V, Kutvonen A, Khaymovich I M, Ala-Nissila T and Pekola J P 2015 On-chip maxwell's demon as an information-powered refrigerator *Phys. Rev. Lett.* **115** 260602
- [68] Jun Y and Bechhoefer J 2012 Virtual potentials for feedback traps *Phys. Rev. E* **86** 061106
- [69] Chun H-M and Noh J D 2019 Universal property of the housekeeping entropy production *Phys. Rev. E* **99** 012136
- [70] Ventéjou B and Sekimoto K 2018 Progressive quenching: globally coupled model *Phys. Rev. E* **97** 062150
- [71] Manzano G, Fazio R and Roldán E 2019 Quantum martingale theory and entropy production *Phys. Rev. Lett.* **122** 220602
- [72] Roldán É, Neri I, Dörpinghaus M, Meyr H and Jülicher F 2015 Decision making in the arrow of time *Phys. Rev. Lett.* **115** 250602
- [73] Saito K and Dhar A 2016 Waiting for rare entropic fluctuations *Europhys. Lett.* **114** 50004
- [74] Dörpinghaus M, Neri I, Roldán É, Meyr H and Jülicher F 2018 Testing optimality of sequential decision-making (arXiv:[1801.01574](https://arxiv.org/abs/1801.01574))
- [75] Singh S, Roldán E, Neri I, Khaymovich I M, Golubev D S, Maisi V F, Peltonen J T, Jülicher F and Pekola J P 2019 Extreme reductions of entropy in an electronic double dot *Phys. Rev. B* **99** 115422
- [76] Kannan D 1979 *An Introduction to Stochastic Processes* (Amsterdam: North Holland)