# Time-Resolved Statistics of Snippets as General Framework for Model-Free Entropy Estimators

Jann van der Meer,* Julius Degünther,* and Udo Seifert
*II. Institut für Theoretische Physik, Universität Stuttgart, 70550 Stuttgart, Germany*

(Received 30 November 2022; revised 20 February 2023; accepted 7 June 2023; published 23 June 2023)

Irreversibility is commonly quantified by entropy production. An external observer can estimate it through measuring an observable that is antisymmetric under time reversal like a current. We introduce a general framework that allows us to infer a lower bound on entropy production through measuring the time-resolved statistics of events with any symmetry under time reversal, in particular, time-symmetric instantaneous events. We emphasize Markovianity as a property of certain events rather than of the full system and introduce an operationally accessible criterion for this weakened Markov property. Conceptually, the approach is based on snippets as particular sections of trajectories between two Markovian events, for which a generalized detailed balance relation is discussed.

DOI: 10.1103/PhysRevLett.130.257101

*Introduction and illustration of main result.*—Thermodynamic equilibrium is characterized by the absence of dissipative and irreversible processes. While dissipation is observed as heat production in the environment, the related concept of irreversibility is often summarized under the notion of an "arrow of time." Therefore, thermodynamic features like heat or entropy production can be revealed by applying inference and estimation methods to time series of appropriate observables.

The most prominent class of entropy estimators is based on simply measuring a time-asymmetric quantity like a steady-state current. This concept has been developed into methods as varied as the thermodynamic uncertainty relation (TUR) [1–3], apparent entropy production rates [4], and, in the paradigmatic case of Markov networks, fluctuation theorems for incomplete, effective descriptions [5–7]. These methods are closely related to a second category of estimators that are based on first identifying a coarse-grained model, on which time-asymmetric currents are then identified in a second step. Whereas older work mainly focuses on lumping Markov states together into coarse-grained ones, see, e.g., [8,9], much richer and more accurate descriptions are obtained by including waiting times and milestoning in semi-Markov models [10–14]. Thermodynamic consistency of the resulting entropy estimators is ensured through information-theoretic [15] reasoning, which has already found its way into stochastic thermodynamics [16] and, in particular, into entropy estimation [17–19].

Further estimation techniques are applicable by presuming a particular underlying model or equation of motion. If, for example, a master equation describes the system on some underlying level, this insight can be utilized in, e.g., minimization [20,21] and decimation [22] methods or by analyzing the communication between subsystems [23]. More generally, not only the mean but also the distribution of fluctuating entropy production is estimated and studied through application of a variety of mathematical tools including, e.g., large deviation theory [24–26], fluctuation-response relations [27–29], and martingale and decision theory [30–32].

From a broader perspective, "thermodynamic inference" [33] is not limited to the estimation of a single quantity like entropy production. Ranging from linear systems [34] over active particles [35–41] to living systems [42,43], even qualitatively distinguishing nonequilibrium from equilibrium can be challenging. In contrast, concepts like the TUR or, more recently, waiting and first passage time distributions are not only able to infer entropy production, but also driving affinities of thermodynamic cycles [13,44,45] or topological features of transition paths [46,47].

The methods described above fall short of sufficient generality and versatility to give a nontrivial result in a model system that is as simple as the three-state network shown in Fig. 1(a) with its coarse-grained version in Fig. 1(b). The coarse-grained model cannot sustain any steady-state current, as it consists of two objects, 1 and $G$, of which only the former, 1, is a state. Since $G$ is not Markovian, we cannot employ any of the methods derived for Markov networks. Moreover, merely distinguishing between residing in state 1 or outside is not sufficient to infer irreversibility, even if the respective waiting times are included [34].

In this Letter, we introduce a framework to obtain a more general and flexible estimator for the mean entropy production rate $\langle\sigma\rangle$. It remains agnostic of the underlying model and can be applied to any conceivable coarse graining of some dynamics described by a path weight,

 

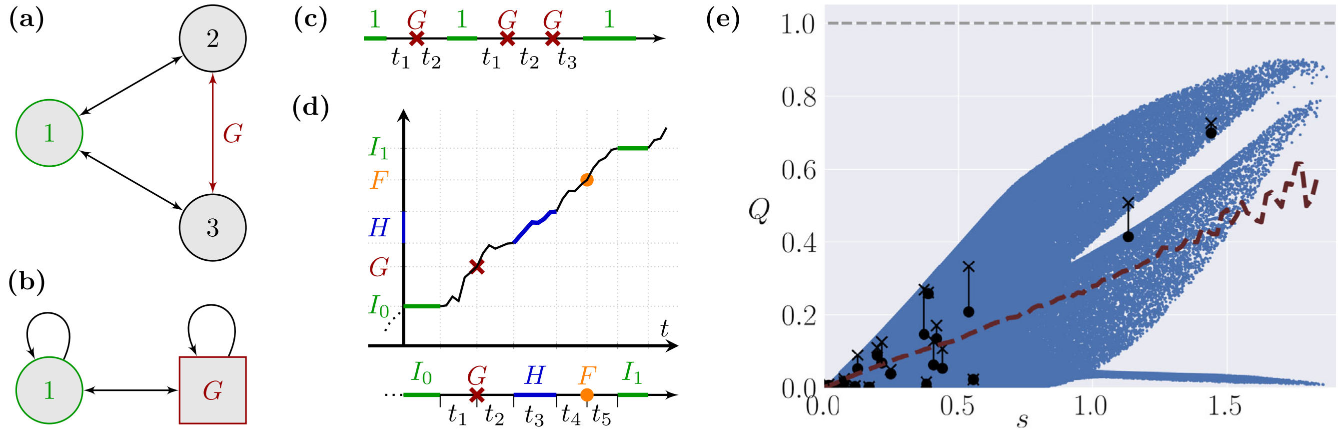


FIG. 1. Entropy estimation based on time-resolved statistics. (a) Paradigmatic three-state Markov network. We assume that only state 1 and transitions along the edge $G$, but not their direction, can be observed, i.e., transitions $2 \to 3$ and $3 \to 2$ are observable but indistinguishable. (b) Effective network of the coarse-grained description in which $G$ is not a state. (c) Time series of observed events showing two possible trajectory snippets with corresponding waiting times. (d) A trajectory snippet from a network with multiple types of measurements, which can be instantaneous ($G$, $F$) or extended in time ($H$), and corresponding time series. (e) Scatter plot of the quality of the estimator $Q = \langle\hat{\sigma}\rangle/\langle\sigma\rangle$ for the network from (a). The rates were parametrized as $k_{ij} = \kappa_{ij} e^{A_{ij}/2}$ with $\kappa_{ij} = \kappa_{ji}$ and $A_{ij} = -A_{ji}$. A sample of $10^6$ sets of rate amplitudes $\kappa_{ij}$ is drawn from a uniform distribution on [0.01, 10]; the affinity is fixed as $A_{ii+1} = 1$. A measure of the asymmetry of the network is $s = \langle(\kappa_{ij} - \langle\kappa_{ij}\rangle)^2\rangle/\langle\kappa_{ij}\rangle^2$ where the averages are taken over the three links. For the dots, only the first two terms in the sum in Eq. (1) are considered. The X markers show the improvement if the first four terms in the sum in Eq. (1) are considered. The dashed line shows the average quality factor as a function of $s$.

e.g., a Langevin or a master equation dynamics. In particular, the estimator is able to exploit data consisting of time-symmetric instantaneous non-Markovian events like the observation of a (nondirected) transition. For the model shown in Fig. 1, the general entropy estimator derived below becomes

$$\langle\hat{\sigma}\rangle = \frac{1}{\langle t\rangle}\sum_{k\geq 1}\left(\prod_{j=1}^{k}\int_0^\infty dt_j\right)\psi(t_1,\ldots,t_k)\ln\frac{\psi(t_1,\ldots,t_k)}{\psi(t_k,\ldots,t_1)}, \tag{1}$$

where the time-resolved statistics is expressed via the waiting time distribution $\psi(t_1,\ldots,t_k)$ associated with observing $k-1$ transitions along edge $G$ with waiting times

$$1\xrightarrow{t_1}G\xrightarrow{t_2}\cdots\xrightarrow{t_{k-1}}G\xrightarrow{t_k}1 \tag{2}$$

between two visits of state 1. A nontrivial estimator $\langle\hat{\sigma}\rangle$ is equivalent to $\psi(t_1,\ldots,t_k) - \psi(t_k,\ldots,t_1) \neq 0$ for at least some $t_1,\ldots,t_k$, which could be understood as an analog of a steady-state current. The time conversion factor $1/\langle t\rangle$ measures the rate of visits to state 1 during a long trajectory. We will prove that the estimator $\langle\hat{\sigma}\rangle$ provides a lower bound on the total entropy production,

$$\langle\sigma\rangle \geq \langle\hat{\sigma}\rangle \geq 0. \tag{3}$$

*General setup.*—In the general setup, we consider coarse-grained trajectories $\Gamma$ that may contain measurements of any kind, which we will refer to as events. Figure 1(d) shows an example. These events can be instantaneous, like $G$ and $F$, or last for a certain duration, as for $I_0$, $H$, and $I_1$. Possible events include observations whose microscopic origin cannot be identified, like a transition whose direction is not resolved as in Fig. 1(a) or, more generally, "lumped transitions," where different pathways cannot be distinguished. This class comprises situations in which, e.g., the passage of a particle (or molecule) through a certain point (or volume) in space can be detected, whereas no further information about its directionality or other internal mechanisms is available.

Moreover, we distinguish between two classes of measurements, which we denote by Markovian and non-Markovian events. If registering the event determines the state of the underlying fundamental system completely, data prior to the measurement do not contain any significant information anymore and can be disregarded. In this sense, we understand these events as Markovian. Two simple examples are the observation of a directed transition [13,14] or a state in a Markov network. Markovian events allow us to cut the trajectory into smaller sections without loss of information. A section that starts and ends at a Markovian event contains the same information regardless of the remaining trajectory it is embedded in. We denote the sections that result from cutting a trajectory at every such Markovian event as "trajectory snippets" $\Gamma^s$. These snippets are a crucial concept of our approach and can be identified if at least one recurring Markovian event is observed.

The time-series data obtained by observing such a coarse-grained trajectory consist of snippets, of which we illustrate one in the lower part of Fig. 1(d). Along its way from the initial Markovian event $I_0$ to the final one $I_1$, the snippet contains the events $G$, $H$, and $F$ with corresponding waiting times $t_1, \ldots, t_5$. We use the term "waiting times" in a more general way to include both genuine waiting times between consecutive events as well as residence times for events of finite duration.

*Derivation of main result.*—For a stationary process, the entropy production rate in the long time limit takes the form of a Kullback-Leibler divergence after averaging [17,19],

$$\langle \sigma \rangle = \lim_{T\to\infty} \langle S \rangle / T = \frac{1}{T} \sum_{\zeta} \mathcal{P}[\zeta] \ln \left( \mathcal{P}[\zeta] / \mathcal{P}[\tilde{\zeta}] \right), \quad (4)$$

where $\langle \cdot \rangle$ denotes the average over many realizations $\zeta$. The time-reversal operation $\zeta \mapsto \tilde{\zeta}$ is an involution whose exact form must be justified by the underlying physical mechanisms. A coarse-grained trajectory $\Gamma$ is the result of a many-to-one mapping $\zeta \mapsto \Gamma(\zeta)$. This mapping determines the time-reversal operation on the coarse-grained level in terms of the underlying one, since coarse graining respects $\tilde{\zeta} \mapsto \tilde{\Gamma}$. Thus, a coarse-grained entropy production rate $\langle \hat{\sigma} \rangle$ can be defined, which provides an estimator for the actual entropy production in the sense that

$$\langle \sigma \rangle \geq \langle \hat{\sigma} \rangle \equiv \frac{1}{T} \sum_{\Gamma} \mathcal{P}[\Gamma] \ln \frac{\mathcal{P}[\Gamma]}{\mathcal{P}[\tilde{\Gamma}]} \geq 0. \quad (5)$$

This result can be derived from the log-sum inequality, which is well known in information theory [15,18,19,33]. In this abstract form, the estimator $\langle \hat{\sigma} \rangle$ is both simple and universal, but often not practical, since a statistically significant amount of long trajectories with negative entropy production is needed to determine $\mathcal{P}[\Gamma]$ and $\mathcal{P}[\tilde{\Gamma}]$.

Instead, it is more feasible to cut $\Gamma$ into trajectory snippets of shorter length and collect the statistics of these smaller snippets. If the first (second,...) part of the trajectory $\Gamma$ is denoted by $\Gamma_1$ ($\Gamma_2, \ldots$), the path weight $\mathcal{P}[\Gamma_k]$ of each individual section can be conditioned on its past in the form

$$\mathcal{P}[\Gamma] = \mathcal{P}[\Gamma_1] \mathcal{P}[\Gamma_2 | \Gamma_1] \cdots \mathcal{P}[\Gamma_n | \Gamma_{n-1}, \Gamma_{n-2}, \ldots]. \quad (6)$$

If we can cut the full trajectory in such a way that the coarse-grained initial state $I_{k-1}$ of a trajectory section $\Gamma_k$ suffices to determine the state of the system on the fundamental level, the path weight factorizes into the contributions from the individual snippets in the form

$$\mathcal{P}[\Gamma] = \mathcal{P}[\Gamma_1^s] \mathcal{P}[\Gamma_2^s | I_1] \cdots \mathcal{P}[\Gamma_n^s | I_{n-1}]. \quad (7)$$

This condition is satisfied if we can observe at least one recurring Markovian event, say $I$, which we define via the Markov property

$$\mathcal{P}[\Gamma(t > s) | I, \Gamma(t < s)] = \mathcal{P}[\Gamma(t > s) | I], \quad (8)$$

formulated for a trajectory $\Gamma(t)$ where the Markovian event $I$ is observed at time $s$. This time instant $s$ can then be used as a cutting locus to study individual trajectory snippets $\Gamma_k^s$. Being able to consider shorter parts of the trajectories individually also emphasizes the major practical advantages that the concept of snippets offers. Note that, technically, continued residence within a time-symmetric Markovian event, e.g., within a Markov state, gives rise to infinitely many trivial snippets inside the state that can be discarded, as they do not contain information about the irreversibility of the process.

In general, a snippet $\Gamma^s$ is a section of the full coarse-grained trajectory $\Gamma$. As a small trajectory on its own, a snippet is characterized by an initial Markovian event $I$, final Markovian event $J$, duration $t$, and possibly additional observations including events and waiting times, summarized under the symbol $\mathcal{O}$. The probability distribution $\psi_{I\to J}(t; \mathcal{O})$ to observe a coarse-grained trajectory $\Gamma^s$ of this form is given by the sum over all contributing microscopic trajectories $\gamma$ denoted by $\gamma_{I\to J}(t; \mathcal{O})$, i.e.,

$$\psi_{I\to J}(t; \mathcal{O}) \equiv \mathcal{P}[\Gamma^s | I] = \sum_{\gamma_{I\to J}(t;\mathcal{O})} \mathcal{P}[\gamma_{I\to J}(t; \mathcal{O}) | I]. \quad (9)$$

These probability distributions can be interpreted as generalized waiting time distributions, whose normalization is written as $\sum_{J,\mathcal{O}} \int_0^\infty dt \psi_{I\to J}(t; \mathcal{O}) = 1$. If the additional observations $\mathcal{O}$ contain continuous degrees of freedom, e.g., position in continuous space or further waiting times, the sum over $\mathcal{O}$ has to be replaced by an appropriate integral.

Up to boundary terms, using (7), the coarse-grained entropy production rate (5) becomes a steady-state average

$$T \langle \hat{\sigma} \rangle = \left\langle \ln \frac{\mathcal{P}[\Gamma_1^s | I_0]}{\mathcal{P}[\tilde{\Gamma}_1^s | \tilde{I}_1]} + \cdots + \ln \frac{\mathcal{P}[\Gamma_n^s | I_{n-1}]}{\mathcal{P}[\tilde{\Gamma}_n^s | \tilde{I}_n]} \right\rangle, \quad (10)$$

where $n$ is the number of snippets. Denoting the probability that a random snippet begins with $I$ by $P(I)$, we use $\mathcal{P}[\Gamma_k^s] = P(I_{k-1}) \mathcal{P}[\Gamma_k^s | I_{k-1}]$ to calculate (10) as

$$\langle \hat{\sigma} \rangle = \sum_{I} \frac{n P(I)}{T} \left( \sum_{\Gamma^s} \mathcal{P}[\Gamma^s | I] \ln \frac{\mathcal{P}[\Gamma^s | I]}{\mathcal{P}[\tilde{\Gamma}^s | \tilde{J}]} \right) \quad (11)$$

for a long, stationary trajectory. To simplify, we define $\langle t \rangle \equiv T/n$, which measures the average waiting time between two events that initialize a snippet. After using Eq. (9) to express the path weights in terms of the waiting time distribution, we obtain our main result

$$\langle \hat{\sigma} \rangle = \frac{1}{\langle t \rangle} \sum_{IJ,\mathcal{O}} \int_0^\infty dt \pi_I \psi_{I\to J}(t;\mathcal{O}) \ln \frac{\psi_{I\to J}(t;\mathcal{O})}{\psi_{\tilde{J}\to\tilde{I}}(t;\tilde{\mathcal{O}})}, \quad (12)$$

with $\pi_I \equiv P(I)$ and the transformations $I \mapsto \tilde{I}$, $J \mapsto \tilde{J}$, and $\mathcal{O} \mapsto \tilde{\mathcal{O}}$ under time reversal. As long as this behavior is known, $\mathcal{O}$ may contain any kind of events, even time-symmetric ones. In this sense, Eq. (12) provides a model-free estimator without relying on particular classes of dynamics, events, or model systems. We first illustrate how to apply the general Eq. (12) to the paradigmatic example of Fig. 1, which exclusively contains time-symmetric events.

*Paradigmatic example.*—In the model of Fig. 1, the coarse-grained description includes transitions along the edge $G$, whose direction is not resolved, as instantaneous non-Markovian events and a single observed Markov state $I = 1$. Thus, each snippet starts as soon as the system exits state 1 and is terminated as soon the system revisits this state. The observations $\mathcal{O}$ along a generic trajectory snippet consist of $k-1$ transitions along the observed edge $G$ and the waiting times $t_1, \ldots, t_k$ as defined by the trajectory snippet (2). The associated waiting time distribution is given by $\psi_{1\to 1}(t;\mathcal{O}) = \psi(t_1, \ldots, t_k)$, with $t$ implicitly fixed via $t = \sum_{i=1}^k t_i$. Since both the transition $G$ and the state 1 are even under time reversal, the reversed trajectory is obtained by simply reading the trajectory snippet (2) backward and hence is associated with the waiting time distribution $\psi(t_k, \ldots, t_1)$. Therefore, the general result (12) reduces to Eq. (1).

From a practical point of view, generating sufficient statistics for all waiting time distributions might not be feasible. Since each term in the sum in Eq. (1) has a non-negative contribution to $\langle \hat{\sigma} \rangle$, considering only snippets that contain a maximum of $k-1$ transitions along the observed edge $G$ already results in a nontrivial estimator. The scatter plot in Fig. 1(e) illustrates this procedure. For the blue dots, only snippets with $k \leq 2$, i.e., $\psi(t)$ and $\psi(t_1, t_2)$, were considered. The black dots denote a random selection from these, for which we have also calculated the improvements of $\langle \hat{\sigma} \rangle$ when going to $k \leq 4$, shown by the corresponding black X markers. As Fig. 1(e) also shows, the estimator inherently benefits from asymmetries in the network, which tend to produce time series with more distinct forward and backward directions.

*Snippets as dressed transitions.*—The time-resolved statistics is condensed in a generalized waiting time distribution $\psi_{I\to J}(t;\mathcal{O})$ that can be interpreted on the level of individual transitions $I \to J$ dressed with $t$ and $\mathcal{O}$ as additional information. We introduce

$$a_{I\to J}(t;\mathcal{O}) \equiv \ln \frac{\pi_I \psi_{I\to J}(t;\mathcal{O})}{\pi_{\tilde{J}} \psi_{\tilde{J}\to\tilde{I}}(t;\tilde{\mathcal{O}})}, \quad (13)$$

which vanishes identically under the condition of apparent detailed balance on the coarse-grained level, defined as $\pi_I \psi_{I\to J}(t;\mathcal{O}) = \pi_{\tilde{J}} \psi_{\tilde{J}\to\tilde{I}}(t;\tilde{\mathcal{O}})$. This suggests to use $a_{I\to J}(t;\mathcal{O})$ as a thermodynamic measure of irreversibility inherent to the dressed transition $I \to J$. Indeed, if at least one $a_{I\to J}(t;\mathcal{O})$ does not vanish, the system cannot be in equilibrium. Moreover, the events $I$, $J$ and those included in $\mathcal{O}$ must be part of a thermodynamic cycle with nonvanishing affinity if $a_{I\to J}(t;\mathcal{O}) \neq 0$. Note that the condition of apparent detailed balance does not imply that the underlying system is in equilibrium, since, depending on the observables, it is always possible to miss hidden cycles with nonvanishing affinity.

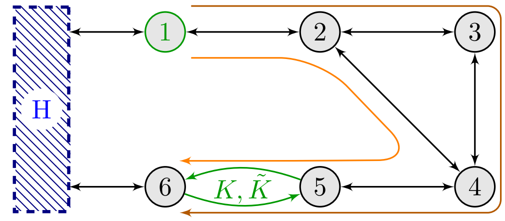


FIG. 2. Illustrative example with a hidden cycle. We assume that the Markov state 1, the directed transitions $K = (56)$ and $\tilde{K} = (65)$, and the non-Markovian event $H$ can be observed. Generically, $a_{1\to K}(t, \mathcal{O} = \emptyset)$ depends on time, which allows us to infer the existence of a hidden cycle, here formed by 2, 3, and 4. Similarly, we detect a hidden cycle within $H$ by studying the time dependence of $a_{1\to\tilde{K}}(t;\mathcal{O})$ for snippets containing $H$.

Surprisingly, we are even able to infer hidden cycles that do not necessarily include the observed events. We illustrate this method in Fig. 2. We observe an explicit time dependence of $a_{1\to K}(t, \mathcal{O} = \emptyset)$, i.e., broken local detailed balance [33] in the trajectory snippets from 1 to $K$, if and only if the affinity of the small cycle containing the states 2,3,4 does not vanish. This broken symmetry between forward and backward transitions

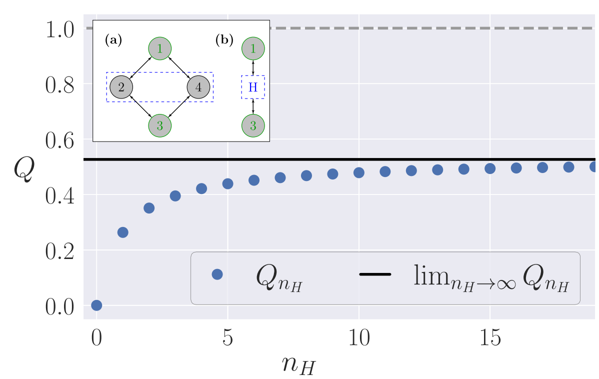


FIG. 3. Demonstration of the systematic bias in the entropy estimator if the trajectory is cut at a non-Markovian state $H$. The inset shows the full (a) and coarse-grained (b) network. $Q_{n_H} \equiv \langle \hat{\sigma}_{n_H} \rangle / \langle \sigma \rangle$ results from cutting the trajectory at every $n_H$th occurrence of $H$. For $n_H \to \infty$ it converges to the estimator obtained by cutting the trajectory at the Markov states 1 and 3. The rates are $k_{12} = 3$, $k_{21} = 1$, $k_{23} = 9$, $k_{32} = 3$, $k_{34} = 27$, $k_{43} = 9$, $k_{41} = 81$, and $k_{14} = 27$.

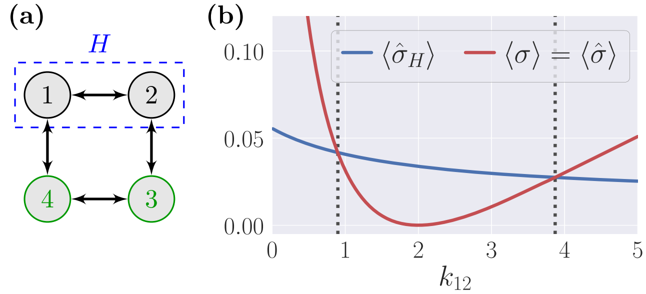


FIG. 4. Illustration of the fact that cutting snippets at non-Markovian events can lead to an overestimation of entropy production. (a) Four-state network, where all rates are set to 1 except for $k_{43} = 2$ and $k_{12}$, which remains variable. (b) Entropy production $\langle\sigma\rangle$, entropy estimator $\langle\hat{\sigma}\rangle$, which in this case fully recovers $\langle\sigma\rangle$, and false entropy estimator $\langle\hat{\sigma}_H\rangle$. For the latter, the compound state $H$ is erroneously treated as a Markov state, leading to $\langle\hat{\sigma}_H\rangle \geq \langle\sigma\rangle$ for values of $k_{12}$ around its equilibrium value $k_{12}^{\mathrm{eq}} = 2$.

generalizes extant discussions for Markov networks [48] and Langevin dynamics [49].

*"Non-Markovian" snippets.*—So far, a Markov property in the form of Eq. (8) has turned out to be crucial. What happens if we cut a trajectory into snippets at, say, a hidden compound state where the factorization (7) is not valid? For such snippets with non-Markovian cut loci, here dubbed non-Markovian snippets, our main result cannot be applied. Nevertheless, the limit $T \to \infty$ in (4) suggests asymptotic consistency of an entropy estimator of the form (12) if the length $t$ of such non-Markovian snippets becomes sufficiently large on average. Conditioning the waiting time distribution $\psi_{H\to J}(t;\mathcal{O})$ on a non-Markovian event $H$ while still disregarding the past of the trajectories introduces a systematic bias. This bias can be used as an operationally verifiable, model-independent criterion for Markovianity of some measured event $H$. If the observed trajectory $\Gamma$ is cut at, say, every $n_H$th occurrence of $H$ rather than every occurrence of $H$, the implied entropy estimator $\langle\hat{\sigma}_{n_H}\rangle$ is independent of $n_H$ if and only if the factorization (7) can be applied, i.e., if $H$ qualifies as a cut locus for snippets. As an example, Fig. 3 shows a four-state Markov network where states 2 and 4 form a compound state $H$. Non-Markovian snippets obtained by cutting at every $n_H$th occurrence of $H$ lead to an estimator that depends on $n_H$ and improves for $n_H \to \infty$.

Thus, shorter non-Markovian snippets reduce the statistical error of finite sample sizes at the cost of introducing systematic bias. As the example in Fig. 4 shows, this systematic bias may even overestimate $\langle\sigma\rangle$, i.e., the error is qualitatively different from merely disregarding information, which always decreases entropy production.

*Concluding perspective.*—This Letter has established a framework for constructing entropy estimators based on coarse-grained data, which may include any kind of measurement whose behavior under time reversal is known and the associated time-resolved statistics. While we have assumed a stationary process, i.e., a nonequilibrium steady state, the underlying concepts are sufficiently general so that future work can adapt this approach to time-dependent situations like periodic driving.

Moreover, we want to emphasize the shift in focus regarding the Markov property. While Markovianity is usually understood as a system property, e.g., in the case of overdamped Langevin dynamics or Markov networks, the present formalism is based on identifying particular observable events at which the Markov property is satisfied as "Markovian." In accordance with Occam's razor, we do not make any assumptions about unobservable parts of the system, which renders this approach more applicable to model complex real-world scenarios.

Finally, our Letter also provides a starting point for thermodynamic inference beyond the estimation of a single quantity like entropy production. We have pointed out how broken local detailed balance can be detected qualitatively. If more details about the system are known, these concepts can be quantified into estimators for driving affinity and topology of the thermodynamic cycles, as Ref. [13] has shown for trajectories cut at observed directed transitions. The broader framework developed here invites further advances in this direction of more refined thermodynamic inference schemes.

We thank Benjamin Ertel for valuable discussions.

*These authors contributed equally to this work.

[1] A. C. Barato and U. Seifert, Thermodynamic Uncertainty Relation for Biomolecular Processes, Phys. Rev. Lett. **114**, 158101 (2015).
[2] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Dissipation Bounds All Steady-State Current Fluctuations, Phys. Rev. Lett. **116**, 120601 (2016).
[3] J. M. Horowitz and T. R. Gingrich, Thermodynamic uncertainty relations constrain non-equilibrium fluctuations, Nat. Phys. **16**, 15 (2020).
[4] M. Esposito, Stochastic thermodynamics under coarse-graining, Phys. Rev. E **85**, 041125 (2012).
[5] N. Shiraishi and T. Sagawa, Fluctuation theorem for partially masked nonequilibrium dynamics, Phys. Rev. E **91**, 012130 (2015).
[6] M. Polettini and M. Esposito, Effective Thermodynamics for a Marginal Observer, Phys. Rev. Lett. **119**, 240601 (2017).
[7] G. Bisker, M. Polettini, T. R. Gingrich, and J. M. Horowitz, Hierarchical bounds on entropy production inferred from partial information, J. Stat. Mech. (2017) 093210.
[8] S. Rahav and C. Jarzynski, Fluctuation relations and coarse-graining, J. Stat. Mech. (2007) P09012.
[9] S. Bo and A. Celani, Entropy production in stochastic systems with fast and slow time-scales, J. Stat. Phys. **154**, 1325 (2014).
[10] I. A. Martinez, G. Bisker, J. M. Horowitz, and J. M. R. Parrondo, Inferring broken detailed balance in the absence of observable currents, Nat. Commun. **10**, 3542 (2019).

[11] D. Hartich and A. Godec, Emergent Memory and Kinetic Hysteresis in Strongly Driven Networks, Phys. Rev. X **11**, 041047 (2021).

[12] B. Ertel, J. van der Meer, and U. Seifert, Operationally accessible uncertainty relations for thermodynamically consistent semi-Markov processes, Phys. Rev. E **105**, 044113 (2022).

[13] J. van der Meer, B. Ertel, and U. Seifert, Thermodynamic Inference in Partially Accessible Markov Networks: A Unifying Perspective from Transition-Based Waiting Time Distributions, Phys. Rev. X **12**, 031025 (2022).

[14] P. E. Harunari, A. Dutta, M. Polettini, and E. Roldán, What to Learn from a Few Visible Transitions' Statistics?, Phys. Rev. X **12**, 041026 (2022).

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Telecommunications and Signal Processing (Wiley, Hoboken, NJ, and Canada, 2006).

[16] S. Ito and A. Dechant, Stochastic Time Evolution, Information Geometry, and the Cramér-Rao Bound, Phys. Rev. X **10**, 021056 (2020).

[17] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Dissipation: The Phase-Space Perspective, Phys. Rev. Lett. **98**, 080602 (2007).

[18] A. Gomez-Marin, J. M. R. Parrondo, and C. Van den Broeck, Lower bounds on dissipation upon coarse-graining, Phys. Rev. E **78**, 011107 (2008).

[19] E. Roldan and J. M. R. Parrondo, Estimating Dissipation from Single Stationary Trajectories, Phys. Rev. Lett. **105**, 150607 (2010).

[20] D. J. Skinner and J. Dunkel, Improved bounds on entropy production in living systems, Proc. Natl. Acad. Sci. U.S.A. **118**, e2024300118 (2021).

[21] D. J. Skinner and J. Dunkel, Estimating Entropy Production from Waiting Time Distributions, Phys. Rev. Lett. **127**, 198101 (2021).

[22] G. Teza and A. L. Stella, Exact Coarse Graining Preserves Entropy Production out of Equilibrium, Phys. Rev. Lett. **125**, 110601 (2020).

[23] D. H. Wolpert, Minimal entropy production rate of interacting systems, New J. Phys. **22**, 113013 (2020).

[24] H. Touchette, The large deviation approach to statistical mechanics, Phys. Rep. **478**, 1 (2009).

[25] P. Pietzonka, A. C. Barato, and U. Seifert, Universal bounds on current fluctuations, Phys. Rev. E **93**, 052145 (2016).

[26] C. Maes, Frenetic Bounds on the Entropy Production, Phys. Rev. Lett. **119**, 160601 (2017).

[27] T. Harada and S. I. Sasa, Equality Connecting Energy Dissipation with a Violation of the Fluctuation-Response Relation, Phys. Rev. Lett. **95**, 130602 (2005).

[28] A. Dechant and S.-i. Sasa, Fluctuation–response inequality out of equilibrium, Proc. Natl. Acad. Sci. U.S.A. **117**, 6430 (2020).

[29] A. Dechant and S. I. Sasa, Improving Thermodynamic Bounds Using Correlations, Phys. Rev. X **11**, 041061 (2021).

[30] E. Roldan, I. Neri, M. Dörpinghaus, H. Meyr, and F. Jülicher, Decision Making in the Arrow of Time, Phys. Rev. Lett. **115**, 250602 (2015).

[31] I. Neri, É. Roldán, and F. Jülicher, Statistics of Infima and Stopping Times of Entropy Production and Applications to Active Molecular Processes, Phys. Rev. X **7**, 011019 (2017).

[32] S. Pigolotti, I. Neri, É. Roldán, and F. Jülicher, Generic Properties of Stochastic Entropy Production, Phys. Rev. Lett. **119**, 140604 (2017).

[33] U. Seifert, From stochastic thermodynamics to thermodynamic inference, Annu. Rev. Condens. Matter Phys. **10**, 171 (2019).

[34] D. Lucente, A. Baldassarri, A. Puglisi, A. Vulpiani, and M. Viale, Inference of time irreversibility from incomplete information: Linear systems and its pitfalls, Phys. Rev. Res. **4**, 043103 (2022).

[35] É. Fodor, C. Nardini, M. E. Cates, J. Tailleur, P. Visco, and F. van Wijland, How Far from Equilibrium is Active Matter?, Phys. Rev. Lett. **117**, 038103 (2016).

[36] T. Speck, Stochastic thermodynamics for active matter, Europhys. Lett. **114**, 30006 (2016).

[37] C. Nardini, É. Fodor, E. Tjhung, F. van Wijland, J. Tailleur, and M. E. Cates, Entropy Production in Field Theories without Time-Reversal Symmetry: Quantifying the Non-Equilibrium Character of Active Matter, Phys. Rev. X **7**, 021007 (2017).

[38] D. Mandal, K. Klymko, and M. R. DeWeese, Entropy Production and Fluctuation Theorems for Active Matter, Phys. Rev. Lett. **119**, 258001 (2017).

[39] P. Pietzonka and U. Seifert, Entropy production of active particles and for particles in active baths, J. Phys. A **51**, 01LT01 (2018).

[40] L. Dabelow, S. Bo, and R. Eichhorn, Irreversibility in Active Matter Systems: Fluctuation Theorem and Mutual Information, Phys. Rev. X **9**, 021009 (2019).

[41] E. Flenner and G. Szamel, Active matter: Quantifying the departure from equilibrium, Phys. Rev. E **102**, 022607 (2020).

[42] F. S. Gnesotto, F. Mura, J. Gladrow, and C. P. Broedersz, Broken detailed balance and non-equilibrium dynamics in living systems: A review, Rep. Prog. Phys. **81**, 066601 (2018).

[43] E. Roldan, J. Barral, P. Martin, J. M. R. Parrondo, and F. Jülicher, Quantifying entropy production in active fluctuations of the hair-cell bundle from time irreversibility and uncertainty relations, New J. Phys. **23**, 083013 (2021).

[44] A. C. Barato and U. Seifert, Universal bound on the Fano factor in enzyme kinetics, J. Phys. Chem. B **119**, 6555 (2015).

[45] P. Pietzonka, A. C. Barato, and U. Seifert, Affinity- and topology-dependent bound on current fluctuations, J. Phys. A **49**, 34LT01 (2016).

[46] R. Satija, A. M. Berezhkovskii, and D. E. Makarov, Broad distributions of transition-path times are fingerprints of multidimensionality of the underlying free energy landscapes, Proc. Natl. Acad. Sci. U.S.A. **117**, 27116 (2020).

[47] A. M. Berezhkovskii and D. E. Makarov, On distributions of barrier crossing times as observed in single-molecule studies of biomolecules, Biophys. Rep. **1**, 100029 (2021).

[48] A. M. Berezhkovskii and D. E. Makarov, On the forward/backward symmetry of transition path time distributions in nonequilibrium systems, J. Chem. Phys. **151**, 065102 (2019).

[49] A. M. Berezhkovskii, G. Hummer, and S. M. Bezrukov, Identity of Distributions of Direct Uphill and Downhill Translocation Times for Particles Traversing Membrane Channels, Phys. Rev. Lett. **97**, 020601 (2006).