

Information-theoretical limit on the estimates of dissipation by molecular machines using single-molecule fluorescence resonance energy transfer experiments

Kevin Song¹, Dmitrii E. Makarov^{2,3,*} and Etienne Vouga¹

¹Department of Computer Science, University of Texas at Austin, Austin, Texas 78712, USA

² Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, USA

³ Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas 78712, USA

Abstract.

Single-molecule fluorescence resonance energy transfer (FRET) experiments are commonly used to study the dynamics of molecular machines. While *in vivo* molecular processes often break time-reversal symmetry, the temporal directionality of cyclically operating molecular machines is often not evident from single-molecule FRET trajectories, especially in the most common 2-color FRET studies. Solving a more quantitative problem of estimating the energy dissipation/entropy production by a molecular machine from single-molecule data is even more challenging. Here we present a critical assessment of several practical methods of doing so, including Markov-model-based methods and a model-free approach based on an information-theoretical measure of entropy production that quantifies how (statistically) dissimilar observed photon sequences are from their time reverses. The Markov model approach is computationally feasible and may outperform model free approaches, but its performance strongly depends on how well the assumed model approximates the true microscopic dynamics. Markov models are also not guaranteed to give a lower bound on dissipation. On the other hand, model-free, information-theoretical methods systematically underestimate entropy production at low photoemission rates, and long memory effects in the photon sequences make these methods demanding computationally. There is no clear winner among the approaches studied here, and all methods deserve to belong to a comprehensive data analysis toolkit.

* makarov@cm.utexas.edu

1. Introduction

Single-molecule experiments, particularly single-molecule fluorescence resonance energy transfer, or FRET (see, e.g., refs.¹⁻⁹), and single-molecule force spectroscopy (see, e.g., refs.¹⁰⁻¹³) are some of the most promising tools for probing mechanistic details about the function of molecular machines. Yet the temporal directionality, which is a key property of life, is rarely evident when life is probed at molecular scales¹⁴ where the randomness induced by thermal fluctuations dominates over directional motion. In some cases, such as direct observation of a molecular motor walking along its track, the directionality of motion is evident, but in many other cases where machines operate cyclically (e.g. enzymatic turnover) this has proven to be a tremendous experimental challenge¹⁵⁻¹⁷.

In parallel, the same question has been raised in the field of stochastic thermodynamics(see, e.g., refs.¹⁸⁻³⁰). The root of the difficulty is that true microscopic information is rarely accessible to measurements. Rather, one observes low-dimensional projections of high-dimensional dynamics, and irreversibility is often hidden by such projections³¹⁻³³. In response, numerous new methods have been proposed for inferring thermodynamic signatures of irreversibility from “partial” observations^{18, 19, 22, 23, 34-37}. More specifically, such studies usually seek to estimate the entropy production rate: unlike an equilibrium system, a “machine” of any kind dissipates heat into its surroundings thereby increasing the surroundings’ entropy. Thus, the entropy production rate offers a *physical* measure of irreversibility. At the same time, this quantity is also an information-theoretical measure of irreversibility, as it is related to the relative probabilities of the forward and time-reversed paths taken by the machine^{38, 39}. In particular, a forward and time-reversed path are equally probable for an equilibrium system.

Here we ask how well one can differentiate between time-symmetric and time-asymmetric dynamics of a molecular system based on a stream of photons emitted by it, as in FRET studies. We focus on the case of “fast” molecular processes, where the lack of timescale separation between the photoemission process and the observed dynamics necessitates photon-by-photon analysis^{4, 40-52}. We compare two types of approaches to this problem. One is based on (hidden) Markov modeling of the system combined with maximum-likelihood inference of model parameters. The other is a “model free approach” based on estimating relative probabilities of the forward and time-reversed photon sequences. In comparing the two methods, we are mostly concerned with questions of principle rather than with statistical errors of the methods. In particular, we wish to compare the relative performance of the two methods in the hypothetical limit of infinite amount of data: in this case the Markov-model-based method will recover the exact entropy production if the topology of the assumed dynamic model coincides with that of the “true” one, even in the limit where the photons emitted by the system are sparse; in contrast, its performance can be considerably poorer if a “wrong” Markov model is assumed. On the other hand, the second method does not rely on any knowledge about the true dynamics of the system, but its performance depends significantly, and non-trivially, on the photocount rate. For the Markov-model-based method, statistical errors may further lead to overestimation of the entropy production or even to erroneous inference of directionality for a system that, in reality, obeys detail balance. Overall, there is no clear winner: we argue that both types of approaches should be included in a comprehensive toolkit.

2. Model.

We consider an N -state molecular system (illustrated in Figure 1, left), with its dynamics governed by the master equation of the form (see, e.g., refs.⁵³⁻⁵⁶)

$$\frac{d\mathbf{p}}{dt} = \mathbf{K}\mathbf{p}. \quad (1)$$

Here $\mathbf{p} = (p(1), \dots, p(N))^T$ is the vector of occupation probabilities for each of the states and $\mathbf{K}$ is the matrix of transition rate coefficients, with K_{ij} , $i \neq j$, being the transition rate from j to i , and with $K_{ii} = -\sum_{j \neq i} K_{ji}$ being the escape rate from the state i .

The system is assumed to be in a steady state, with the vector of steady-state probabilities $\mathbf{p}_{ss}$ satisfying

$$\mathbf{K}\mathbf{p}_{ss} = 0. \quad (2)$$

When the system satisfies the detailed balance condition,

$$K_{ij}p_{ss}(j) = K_{ji}p_{ss}(i), \quad (3)$$

its dynamics are time-reversible, and the steady-state probabilities are related to the free energies associated with each state. We are, however, particularly interested in molecular systems that violate detailed balance, such that the vector $\mathbf{p}_{ss}$ describes a *nonequilibrium* steady state satisfying Eq. 2 but not Eq. 3.

For such a nonequilibrium system, our goal is to quantify the degree to which it deviates from equilibrium; from a thermodynamic standpoint, this is usually accomplished by considering the average entropy production per unit time, which is related to the heat dissipated by the system, and which, for the system at hand, is given by the known expression^{38, 39}:

$$\sigma \equiv \langle dS/dt \rangle = \sum_{i < j} [K_{ij}p_{ss}(j) - K_{ji}p_{ss}(i)] \log \frac{K_{ij}}{K_{ji}}. \quad (4)$$

From an information-theoretical standpoint, the entropy production is given by the Kullback-Leibler (KL) divergence between the probability of the forward path (i.e. the state of the system $i(t)$ as a function of time) and its time reverse:

$$\langle dS/dt \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \langle \log \frac{p[i(t)]}{p[i(\tau-t)]} \rangle_{p[i(t)]}, \quad (5)$$

where the average is performed over the ensemble of forward paths. Eq. 4 can be derived from Eq. 5. Moreover, the elements of the kinetic matrix $\mathbf{K}$ can be evaluated, and thus the entropy production can be computed, from a long trajectory $i(t)$ (or from many short trajectories) if such trajectories can be observed.

The difficulty in estimating the entropy production arises from the fact that the instantaneous microscopic state i of the system is not measurable at every time, or, sometimes, is not measurable precisely at all^{14, 22, 23, 34}. Here we envisage a single-molecule fluorescence resonance energy transfer (FRET) experiment, a common technique of probing molecular states of the system, which are encoded in the colors of photons emitted by the molecule. We start with a simplified – and usually unrealistic – model of N -color FRET, where there is a one-to-one correspondence between the color of the photon emitted and the state of the molecule. In this case, the uncertainty in the knowledge of the state comes entirely from the fact that the state of the system is unknown during the dark time periods in between the photons. We then

consider more realistic scenarios, where there are fewer photon colors than microscopic states and/or where, in each, a mix of photons of each color is emitted. It turns out that both of these cases can be described using the same formalism. In all of these cases we do not consider, explicitly, the physics of photoexcitation^{57, 58}; instead we make the following two assumptions: (i) the emitted photons obey Poisson statistics; this assumption is reasonable when the photoexcitation rate is much higher than the emission rate. (ii) The emission rate μ of the photons of all possible colors is independent of the state of the molecule; this assumption is reasonable for FRET experiments, in which this emission rate is determined by the excitation rate of the donor probe.

3. Estimating entropy production when photon colors unambiguously report on the molecular states.

FRET experiments differentiate between molecular states according to their FRET efficiencies. For example, in a two-color FRET experiment, a mix of photons of two colors is observed in each molecular state, with the state of the system determining the fractions of photons of each color. That is, the photon color itself does not identify the molecular state, and multiple photons must be seen before the molecular state can be even guessed. This means that if the molecule leaves the state before it emits a sufficient number of photons, identification of that state is challenging. Ideally, one wants the FRET efficiencies (equal to the probabilities of detecting the acceptor photon in FRET) to be as different as possible for different states. Here we consider the extreme (and usually unrealistic) case where the color of the photons emitted in each state is unique to that state. For a two-color FRET experiment performed on a two-state molecular system that would imply that the FRET efficiencies in the two states are 0 and 1. For a molecule with N states we assume that N colors are available to identify, unambiguously, each of the states. In this case we label both molecular states and photon colors with the same discrete index i .

Of course, in practice the complexity of the experiment increases drastically with the number of FRET dyes and thus colors. Moreover, each state may emit photons of “wrong colors”, and experimental effects⁵⁹ such as cross-talk add to the complexity. This situation, considered in the next section, can be described by introducing a matrix of probabilities R_{ji} that a photon of color j is observed when the molecule is in state i , as illustrated in Fig. 3, right. This description also covers the most common case where the number of colors available is smaller than the number of states. Section 4 provides a general theory for such a case; the present case corresponds to a diagonal matrix R_{ji} – ideally, of course, one should strive to keep R_{ji} as close to diagonal as possible.

Observation of photons will yield a sequence of photon colors $i_1, i_2, \dots, i_N$ detected at times $t_1, t_2, \dots, t_n$, where the time sequence $t_1, t_2, t_3, \dots$ is drawn from a Poisson process, with the average number of photons emitted per unit time equal to μ . Given such a sequence, one may attempt estimating the entropy production of the process using equation Eq. 5, but with photon sequences replacing the true microscopic states of the system. In particular, we consider two estimates:

$$\left\langle \frac{dS}{dt} \right\rangle_{ti} = \lim_{n \rightarrow \infty} \sigma_{ti}^{(n)}, \sigma_{ti}^{(n-1)} = \frac{\mu}{(n-1)} \left\langle \log \frac{p[i_1, t_1, i_2, t_2, \dots, i_n, t_n]}{p[i_n, 0, i_{n-1}, t_n - t_{n-1}, \dots, i_1, t_n]} \right\rangle_{p[i_1, 0, i_2, t_2, \dots, i_n, t_n]}, \quad (6)$$

$$\left\langle \frac{dS}{dt} \right\rangle_i = \lim_{n \rightarrow \infty} \sigma_i^{(n)}, \sigma_i^{(n-1)} = \frac{\mu}{(n-1)} \left\langle \log \frac{p[i_1, i_2, \dots, i_n]}{p[i_n, i_{n-1}, \dots, i_1]} \right\rangle_{p[i_1, i_2, \dots, i_n]}. \quad (7)$$

Here $p[i_1, t_1, i_2, t_2, \dots, i_n, t_n]$ is the joint probability of observing a photon of color i_1 at time t_1 , etc. (more precisely, this is a probability density with respect to the continuous variables $t_1 \dots t_n$). Note that, because of translational invariance, it only depends on the differences between photon arrival times, and so one can arbitrarily set $t_1 = 0$ to indicate that the clock starts with the observation of the first photon. Similarly, $p[i_1, i_2, \dots, i_n]$ is the joint probability to observe the sequence of photon colors appearing in the argument of this function.

Since for a long photon sequence its total time, $\tau = t_n - t_1 \approx (n - 1)/\mu$ is essentially a deterministic function of n , the inverse of this time appears instead of the factor $1/\tau$ in the above two equations. Finally, the quantities $\sigma_{ti}^{(n)}$ and $\sigma_i^{(n)}$ can be regarded as the n -th Markov order estimates of the entropy production. In particular, we have $\sigma_i^{(m)} = \sigma_i^{(n)}$ for any $m \geq n$ if the photon sequence is a true n -th order Markov process (that is, if the probability of observing a photon of a certain color only depends on the colors of the previous n photons). If, however, the photon sequence is an infinite-order Markov process then we expect $\sigma_i^{(n)}$ to be n -dependent for any n .

We emphasize that, since a photon sequence emitted by the system does not provide complete microscopic information about the system's true trajectory, Eqs. 6 and 7 are not expected, in general, to be equivalent to Eq. 5; instead one hopes to use them to recover a lower bound on the true entropy production¹⁹. Eq. 7, as compared to Eq. 6, further discards temporal information about the photons and simply considers the KL divergence between photon sequences. We note that the sequence of emission times $t_1, t_2, \dots, t_n$ in our model obeys a Poisson process and is therefore time reversible, but the molecular states $i_1, i_2, \dots$ (and thus the observed photon colors) are, in general, not statistically independent of the emission times. We then expect, based on the log sum inequality⁶⁰, that $\left\langle \frac{dS}{dt} \right\rangle_i \leq \left\langle \frac{dS}{dt} \right\rangle_{ti}$. In fact, it turns out that the two estimates are nearly identical except when the inter-photon time μ^{-1} is comparable to or is longer than the inherent timescale of the system's dynamics (Appendix A), a regime where, clearly, the number of photons emitted is insufficient to report on the inherent dynamics in question. Because most of the data presented here are far away from this low-emission-rate regime, we will ignore the differences between Eqs. 6 and 7 and will use the simpler Eq. 7 to estimate the entropy production using this method.

In practice, estimating Eq. 6 or 7 from the data is a daunting task because the sequence of photons emitted by the system, unlike the dynamics of the system itself, is not generally described by a Markov process. But in the particular scenario considered here, with the photon color faithfully reporting on the molecular state, the sequence of photon colors $i_1, i_2, \dots, i_n, \dots$ is in fact a discrete-time Markov process. Indeed, suppose that, at $t = 0$, a photon of color i has been observed. Then the state of the system is also known to be i with certainty, and, since the system is Markovian, this state completely determines the probability of any future observation performed on the system. The (conditional) probability that the next photon will be of color j and will be observed between times t and $t + dt$ is the product of the probability $\mu dt \exp(-\mu t)$ to see the next photon within that time window, regardless of its color, times the

conditional probability $p(j, t|i, 0) = (e^{\mathbf{K}t})_{ji}$ that the system is in state j at the moment the photon is emitted. For the purpose of evaluating Eq. 7, in particular, we need the probability $p(j|i)$ that the photon of color j is emitted after a photon of color i , which is given by

$$p(j|i) = \int_0^\infty (e^{\mathbf{K}t})_{ji} e^{-\mu t} \mu dt \quad (8)$$

The conditional probabilities $p(j|i)$ now completely determine the Markov chain of the photon colors. As a result, Eq. 7 is easily evaluated to give

$$\left\langle \frac{dS}{dt} \right\rangle_i = \sigma_i^{(1)} = \mu \sum_{j,i} p(j|i) p_{ss}(i) \log \frac{p(j|i)}{p(i|j)}, \quad (9)$$

where the steady-state probabilities of observation of photons of different colors are the solutions of the eigenvalue problem $\sum_i p(j|i) p_{ss}(i) = p_{ss}(j)$. Note that, without the factor μ , equal to the inverse mean inter-photon time, Eq. 9 would give the estimated average entropy production per observed photon.

As an example, let us give an explicit expression for the entropy production thus estimated for the case where the molecular system in question can be modeled as a 3-state cycle (Fig. 1, right). The system undergoes transitions in the clockwise direction with a rate k_F and in the counterclockwise direction with a rate k_B . It emits photons at a constant rate μ ; the probability of emitting the photon of color j while in state i is generally specified by a parameter R_{ji} , as discussed below, but in the present case only the photons of “true color” are emitted, and thus we have $R_{ij} = \delta_{ij}$, where δ_{ij} is Kronecker’s delta.

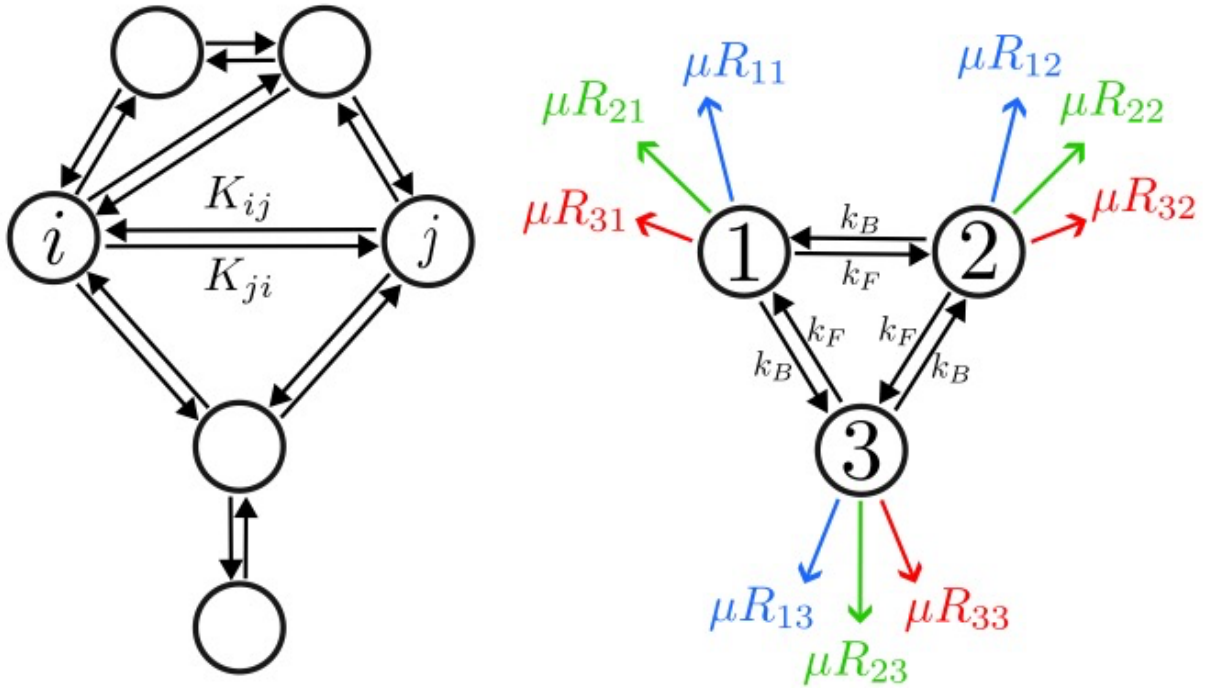


Figure 1. Left: the general model of dynamics assumed here is described by the kinetic coefficients K_{ji} representing transition probabilities (per unit time) from a discrete i state to a state j . Right: an example of an energy dissipating cycle when the forward rate k_F is different from the backward rate. The dynamics of the cycle are monitored by observing the photons emitted by the system. In each state, the arrival times of the photons obey a Poisson process with an emission rate μ (that is, quantum effects such as photon antibunching are neglected^{57, 61}). The probability that a photon has color j when emitted while the molecule is in the state i is given by R_{ji} , and so the total emission rate for color 1 in state 1 is given by μR_{11} , for color 2 in state 1 is given by μR_{21} , etc.

In this case the true entropy production (Eq. 4) is given by

$$\langle dS/dt \rangle = (k_F - k_B) \log \frac{k_F}{k_B}, \quad (10)$$

whereas Eqs. 7-9 give the following estimate:

$$\begin{aligned} & \left\langle \frac{dS}{dt} \right\rangle_i \\ &= \frac{\mu}{\mu^2 + 3(k_B^2 + k_F^2 + k_F\mu + k_B\mu + k_Bk_F)} \left[(k_B^2 + k_F^2 + k_B\mu + k_Bk_F) \ln \frac{k_B^2 + k_F^2 + k_B\mu + k_Bk_F}{k_B^2 + k_F^2 + k_F\mu + k_Bk_F} \right. \\ & \quad \left. + (k_B^2 + k_F^2 + k_F\mu + k_Bk_F) \ln \frac{k_B^2 + k_F^2 + k_F\mu + k_Bk_F}{k_B^2 + k_F^2 + k_B\mu + k_Bk_F} \right] \\ & (11) \end{aligned}$$

Inspecting Eq. 11 we find that:

$$\lim_{\mu \rightarrow 0} \left\langle \frac{dS}{dt} \right\rangle_i = 0. \quad (12)$$

Thus the estimated entropy production vanishes in the limit of low emission rate, an intuitively expected result, as the photons emitted by the molecule become statistically independent in this case. On the other hand, in the limit $\mu \gg k_B, k_F$ many photons will be emitted while the molecule resides in one state (with the average residence time being $(k_B + k_F)^{-1}$), and thus the state of the molecule is directly measurable with a high temporal resolution ($\sim \mu^{-1}$), much higher than the time associated with the molecular dynamics. In this case, we expect the estimate of Eq. 11 to approach the true entropy production, and indeed, we find from Eq. 11 that

$$\lim_{\mu \rightarrow \infty} \left\langle \frac{dS}{dt} \right\rangle_i = \left\langle \frac{dS}{dt} \right\rangle = (k_F - k_B) \log \frac{k_F}{k_B}. \quad (13)$$

The estimated entropy production increases monotonically with increasing emission rate (Fig. 2).

4. More realistic scenarios: No one-to-one correspondence between photon colors and states.

We now discuss the more realistic case, where there is no one-to-one correspondence between the emitted colors and the molecular states, but rather each state can be characterized by a set of probabilities of emitting photons of different colors. Specifically, we assume that while in state i the molecule will emit a color j with a probability R_{ji} , with R_{ii} being the probability of emitting the “true” (or dominant) color associated with the state (Fig. 1, right). The matrix $\mathbf{R}$ has the property $R_{1i} + R_{2i} + R_{3i} = 1$ assuming no dark molecular states. This matrix $\mathbf{R}$ may be viewed as a “recoloring matrix”: if we first imagine the idealized scenario of the previous section, in which each molecular state i produces photons of the true color associated with i , then the action of $\mathbf{R}$ is to recolor those, randomly, into colors 1, 2, and 3 according to the probabilities R_{1i}, R_{2i}, R_{3i} .

The (common) scenario where there are fewer FRET colors than states, such as the two-color FRET, can also be described by this formalism. If, for example, color 1 corresponds to the donor photon color and color 2 to the acceptor photon color then we have $R_{33} = R_{32} = R_{31} = 0$ (i.e. no 3rd color photons are present). Moreover, the quantity

$$\epsilon_i = \frac{R_{2i}}{R_{2i} + R_{1i}} = R_{2i} \quad (14)$$

is the usual FRET efficiency⁶¹ associated with the state i .

A key (and somewhat counterintuitive) observation that will be critical for the following discussion is that, unlike the original photon stream before recoloring, *the sequence of observed “recolored” photons is generally non-Markov*. Intuitively, this occurs because the color of the photon no longer faithfully reports on the state, and so accounting for photons beyond the most recently emitted one can give more information on the current state of the system. Because of the non-Markovianity, Eq. 7 can no longer be reduced to the much simpler Eq. 9. Nevertheless, we can view Eq. 9 as a Markov *approximation* to Eq. 7. It is potentially a useful approximation in practice, since the conditional probabilities $p(j|i) = p(j, i)/p(i)$ are easily estimated from photon sequences; here $p(j, i)$ is the fraction of pairs where photon j is preceded by photon i in the stream of photons, and $p(i)$ is the fraction of i -colored photons in the stream. In contrast, estimation of the probabilities of long photon sequences, as in Eq. 7, is demanding computationally and requires much larger amounts of data⁶². To examine the quality of this approximation, we note that these joint probabilities can be written in the form

$$p'(j) = \sum_{\alpha} R_{j\alpha} p_{SS}(\alpha) \quad (15)$$

and

$$p'(j, i) = \sum_{\beta, \alpha} R_{j\beta} R_{i\alpha} p(\beta, \alpha) = \sum_{\beta, \alpha} R_{j\beta} R_{i\alpha} p(\beta|\alpha) p_{SS}(\alpha). \quad (16)$$

Or, written in matrix form, $\boldsymbol{\pi}' = \mathbf{R}\boldsymbol{\pi}\mathbf{R}^T$, where $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ are the matrices whose elements are the joint probabilities $p(j, i)$ and $p'(j, i)$. The primes in these expressions are used to emphasize that these probabilities are associated with the *recolored* photon stream and are different from those evaluated for the photon stream described in Section 3 where each state emits a single unique color. Using Eq. 8, Eq. 9 can now be adapted to describe the Markovian estimate (indicated by the subscript “ m ”):

$$\left\langle \frac{dS}{dt} \right\rangle_m = \sigma_i^{(1)} = \mu \sum_{i,j} p'(i, j) \log \frac{p'(i|j)}{p'(j|i)}. \quad (17)$$

We now consider the cases of 3-color and 2-color FRET separately.

Three color FRET: the number of colors is still the same as the number of observed states, but in each state the molecule can emit photons of “wrong” colors, with specified probabilities. As an example, consider the recoloring matrix of the form:

$$\mathbf{R} = \begin{pmatrix} 1-p & p/2 & p/2 \\ p/2 & 1-p & p/2 \\ p/2 & p/2 & 1-p \end{pmatrix}, \quad (18)$$

where the photon of the “true color” is emitted with a probability $1 - p$ while photons of the two wrong colors are emitted with probabilities equal to $p/2$. For $p \rightarrow 0$ this model is identical to the one considered in the previous section, with color completely identifying a molecular state. For the model of Fig. 1, the Markovian estimate for the entropy production, Eq. 17, is then evaluated to give:

$$\left\langle \frac{dS}{dt} \right\rangle_m = \frac{\mu}{3[\mu^2 + 3(k_B^2 + k_F^2 + k_F\mu + k_B\mu + k_Bk_F)]} \times$$

$$\left\{ [-((-2+p)p\mu^2) + 3(k_B^2 + k_F(\mu + k_F) + k_B(-((-2+p)p\mu) + k_F))] \log \frac{(-2+p)p\mu^2 - 3k_B^2 + 3k_B((-2+p)p\mu - k_F) - 3k_F(\mu + k_F)}{(-2+p)p\mu^2 - 3k_B^2 + 3(-2+p)p\mu k_F - 3k_F^2 - 3k_B(\mu + k_F)} + [-((-2+p)p\mu^2) + 3(k_B^2 + k_B(\mu + k_F) + k_F(-((-2+p)p\mu) + k_F))] \log \frac{(-2+p)p\mu^2 - 3k_B^2 + 3(-2+p)p\mu k_F - 3k_F^2 - 3k_B(\mu + k_F)}{(-2+p)p\mu^2 - 3k_B^2 + 3k_B((-2+p)p\mu - k_F) - 3k_F(\mu + k_F)} \right\} \quad (19)$$

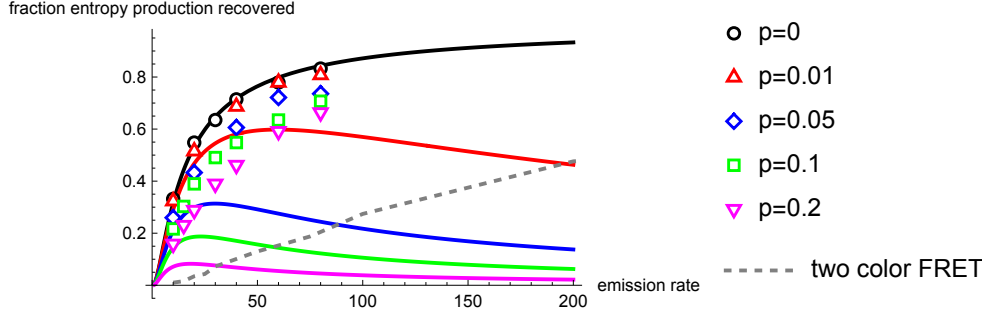


Figure 2. Entropy production estimates (divided by the true entropy production, Eq. 10) for a 3-color FRET experiment measuring the dynamics of a 3-state cycle depicted in Fig. 1, right. The model parameters are $k_F = 2k_B$, and the emission rate μ is measured in dimensionless units set by k_B . The parameter p quantifies the probability that a photon of a “wrong” color is emitted (see Eq. 18). For $p = 0$ the photon color unambiguously reports on the state of the molecule at the moment when the photon is emitted. Solid lines show the estimates $\sigma_i^{(1)}$ obtained from the Markovian estimator (Eq. 19), while open markers show the estimates $\sigma_i^{(\infty)}$ using the sequence probabilities, Eq. 27 (that is, non-Markov effects are fully taken into account). The dashed line shows the results for a two-color photon FRET experiment performed on the same system. In this experiment, state 1 emits photons of color 1 with probability $R_{11} = 0.9$ and color 2 with probability $R_{21} = 0.1$; state 2 emits photons of color 2 with probability $R_{22} = 0.9$ and color 1 with probability $R_{12} = 0.1$; state 3 emits photons of color 1 and 2 with equal probabilities, $R_{13} = R_{23} = 0.5$. See text for further details. The Markov estimator $\sigma_i^{(1)}$ gives zero entropy production in this case; the estimate shown is the result of using Eq. 27.

An interesting feature of Eq. 19 is that the estimated entropy production is a nonmonotonic function of the emission rate μ . In fact, it can be verified that $\lim_{\mu \rightarrow \infty} \left\langle \frac{dS}{dt} \right\rangle_m = 0$ for any finite value of p , while $\lim_{\mu \rightarrow \infty} \left\langle \frac{dS}{dt} \right\rangle_m = (k_F - k_B) \log k_F/k_B$ (cf. Eq. 13) for $p = 0$. Thus an optimal value of the emission rate exists, where the Markov approximation gives the most accurate estimate of the entropy production. This behavior has to do with the fact that, as the inter-photon time μ^{-1} decreases, longer strings of photons are correlated, requiring longer memory to be considered to improve entropy production estimates. Note that this finding is general, as we observe similar behavior for more complicated models exhibiting non-Markovian dynamics, see Section 8.

Two-color FRET. To use the formalism described above, we will assume that the molecule can only emit photons of colors 1 (call it donor) and 2 (acceptor). Then the recoloring matrix has a row of zeros and can be written in the form:

$$\mathbf{R} = \begin{pmatrix} 1 - R_{21} & 1 - R_{22} & 1 - R_{23} \\ R_{21} & R_{22} & R_{23} \\ 0 & 0 & 0 \end{pmatrix}, \quad (20)$$

where the second row contains the FRET efficiencies for each state. The Markovian estimate of the entropy production, $\langle \frac{dS}{dt} \rangle_m$, can then be obtained using Eqs. 15-17 and 20, as before.

Without doing explicit calculations, however, it is easy to see¹⁴ that $\langle \frac{dS}{dt} \rangle_m = 0$. Indeed, a sufficiently long Markovian sequence of two colors, (e.g. 1222121...) is time-reversible as it always obeys detailed balance (i.e., there is the same number of 12 and 21 pairs in the sequence). Does this result imply that two-color FRET is incapable to discern the time-arrow of a molecular machine operating in nonequilibrium steady state? No, but it implies that non-Markov effects in the photon color sequence must be considered when evaluating Eq. 7, and that Eq. 9 is not useful as an entropy-production estimate for two-color FRET experiments.

5. Evaluating entropy production from sequence probabilities.

It is clear from the above discussion that unless – unrealistically – the colors of photons emitted by the molecule in each microscopic state are different, non-Markov effects in the photon sequences are important when evaluating the entropy production using Eqs. 6 or 7. Evaluation of probabilities of long (large n) sequences of photons in Eqs. 6-7 is demanding computationally and requires very long trajectories. If the underlying dynamics (i.e. the values of k_F and k_B) are known, however, computing Eqs. 6 and 7 becomes a much easier task. Of course, this is not useful in practice, as once k_F and k_B are known, the exact value of the entropy production is also known (Eq. 10), and Eqs. 6-7 are unnecessary. But such estimates are useful for learning what kind of information about molecular dynamics is accessible, *in principle*, from the sequences of photons emitted by molecules.

To evaluate Eq. 6 (or Eq. 7) numerically, consider a sequence of n photons emitted by the system in its steady state. The probability of observing a sequence of photons with colors $i_1, i_2, \dots, i_n$ detected within infinitesimal time intervals $(t_1, t_1 + dt_1), (t_2, t_2 + dt_2), \dots, (t_n, t_n + dt_n)$, is given by

$$p[i_1, t_1, i_2, t_2, \dots, i_n, t_n] dt_1 dt_2 \dots dt_n = \mu^n R_{i_n j_n} (e^{\mathbf{K}(t_n - t_{n-1})})_{j_n j_{n-1}} e^{-\mu(t_n - t_{n-1})} dt_n \dots R_{i_2 j_2} (e^{\mathbf{K}(t_2 - t_1)})_{j_2 j_1} e^{-\mu(t_2 - t_1)} dt_2 R_{i_1 j_1} p_{SS}(j_1) dt_1, \quad (21)$$

where summation is assumed over all repeated indices (i.e., $j_1, \dots, j_n$). Note that this quantity is the likelihood function, which plays a key role in FRET-based hidden-Markov models^{41, 63} further discussed below. Note also that, in view of time-translational invariance of photon sequences, $p[i_1, t_1, i_2, t_2, \dots, i_n, t_n] = p[i_1, 0, i_2, t_2 - t_1, \dots, i_n, t_n - t_1]$ only depends on the time differences $t_2 - t_1, t_3 - t_2, \dots$. Introducing new diagonal matrices $\rho^{(i)}$

$$\rho_{j\alpha}^{(i)} = R_{ij} \delta_{j\alpha}, \quad (22)$$

where δ is Kronecker's delta, we can rewrite the sequence probability density in a compact form as

$$p[i_1, t_1, i_2, t_2, \dots, i_n, t_n] = \mu^n e^{-\mu(t_n - t_1)} \mathbf{1} \boldsymbol{\rho}^{(i_n)} e^{\mathbf{K}(t_n - t_{n-1})} \dots \boldsymbol{\rho}^{(i_2)} e^{\mathbf{K}(t_2 - t_1)} \boldsymbol{\rho}^{(i_1)} \mathbf{p}_{ss}, \quad (23)$$

where $\mathbf{1} = (1 \dots 1)$ is the N -dimensional vector with all of its components equal to 1. The probability density of the time-reversed sequence is, similarly,

$$p[i_n, t_n, i_{n-1}, t_{n-1}, \dots, i_1, t_1] = \mu^n e^{-\mu(t_n - t_1)} \mathbf{1} \boldsymbol{\rho}^{(i_1)} e^{\mathbf{K}t_1} \dots \boldsymbol{\rho}^{(i_{n-1})} e^{\mathbf{K}(t_n - t_{n-1})} \boldsymbol{\rho}^{(i_n)} \mathbf{p}_{ss}, \quad (24)$$

and, therefore,

$$\log \frac{p[i_1, t_1, i_2, t_2, \dots, i_n, t_n]}{p[i_n, t_n, i_{n-1}, t_{n-1}, \dots, i_1, t_1]} = \log \frac{\mathbf{1} \boldsymbol{\rho}^{(i_n)} e^{\mathbf{K}(t_n - t_{n-1})} \dots \boldsymbol{\rho}^{(i_2)} e^{\mathbf{K}(t_2 - t_1)} \boldsymbol{\rho}^{(i_1)} \mathbf{p}_{ss}}{\mathbf{1} \boldsymbol{\rho}^{(i_1)} e^{\mathbf{K}(t_2 - t_1)} \dots \boldsymbol{\rho}^{(i_{n-1})} e^{\mathbf{K}(t_n - t_{n-1})} \boldsymbol{\rho}^{(i_n)} \mathbf{p}_{ss}} \quad (25)$$

We assume that a sufficiently long sequence of photons has a “self-averaging” property^{18, 19}, and thus $\langle \frac{dS}{dt} \rangle_{ti}$ can be estimated as

$$t_n \langle \frac{dS}{dt} \rangle_{ti} \approx \log \frac{\mathbf{1} \boldsymbol{\rho}^{(i_n)} e^{\mathbf{K}(t_n - t_{n-1})} \dots \boldsymbol{\rho}^{(i_2)} e^{\mathbf{K}(t_2 - t_1)} \boldsymbol{\rho}^{(i_1)} \mathbf{p}_{ss}}{\mathbf{1} \boldsymbol{\rho}^{(i_1)} e^{\mathbf{K}(t_2 - t_1)} \dots \boldsymbol{\rho}^{(i_{n-1})} e^{\mathbf{K}(t_n - t_{n-1})} \boldsymbol{\rho}^{(i_n)} \mathbf{p}_{ss}} \quad (26)$$

for a sufficiently long sequence of photons, or as an average of Eq. 26 performed over multiple sufficiently-long sampled n -photon sequences. This procedure is numerically efficient since it only involves matrix multiplication.

Eq. 7 can be estimated similarly to give

$$t_n \langle \frac{dS}{dt} \rangle_i \approx \log \frac{\mathbf{1} \boldsymbol{\rho}^{(i_n)} \boldsymbol{\pi} \boldsymbol{\rho}^{(i_{n-1})} \dots \boldsymbol{\pi} \boldsymbol{\rho}^{(i_1)} \mathbf{p}_{ss}}{\mathbf{1} \boldsymbol{\rho}^{(1)} \boldsymbol{\pi} \boldsymbol{\rho}^{(i_2)} \dots \boldsymbol{\pi} \boldsymbol{\rho}^{(i_n)} \mathbf{p}_{ss}}, \quad (27)$$

where $\boldsymbol{\pi}$ is the matrix with elements $\pi_{ji} = p(j|i)$.

As seen from Fig. 2, the entropy production estimates that include non-Markov effects are significantly greater than those evaluated using the Markov approximation. Moreover, such estimates also show nonzero entropy production for the two-color FRET case, where the Markov approximation fails to differentiate between reversible and irreversible dynamics. Unfortunately, Eqs. 26 and 27 cannot be used as practical entropy production estimators, as they require knowledge of the underlying dynamics of the system. They do, however, show the theoretical limit of what fraction of true entropy production can be recovered in *principle* from Eqs. 6-7 if a practical algorithm is available for this task. In the next section, we discuss one proposal for such an algorithm.

6. Estimating entropy production using compression algorithms.

Compression algorithms can be used to estimate the “information” content in a long string s , representing, e.g., an English text or a time series⁶⁴. This approach, in the context of single-molecule dynamics, has been used to evaluate the predictability (i.e., Markovianity) of single-molecule trajectories⁶⁵⁻⁶⁸. A compression algorithm uses the string to build a “dictionary” that is then used to construct a shorter representation of the string. The basic idea of detecting time asymmetry using such an algorithm is that time asymmetry causes the string s to be less compressible when attempting to compress it using a dictionary built from the time reverse of s .

More specifically, entropy production is proportional to the KL divergence between the forward path of the system (here, sequence of photons) and its time reverse, see Eqs. 5-7. To estimate this quantity using a compression algorithm, we use the method developed by Ro et. al⁶⁹, which in turn builds on the Ziv-Merhav algorithm⁷⁰. This method is based on computing the cross-parsing complexity between two strings (i.e. sequences of photon colors in our case). The cross-parsing complexity of a sample string s using a string r as a dictionary, denoted $C(s || r)$, is the least number of phrases from r needed to entirely encode s . For example, if $s = 123133$ and $r = 1233$, then s can be expressed as three phrases from r : 123, 1, and 33. Thus, $C(s || r) = 3$. Given two long, independent sequences s_1 and s_2 of photon colors, each of length N , the KL divergence between sequences of photon colors (Eq. 7) is estimated as⁶⁹

$$D_{KL}(s|s^R) = \frac{\log N}{N} [C(s_1||s_2^R) - C(s_1||s_2)],$$

where s_2^R represents the time reversal of s_2 .

7. Maximum likelihood estimators (MLE) of entropy production.

The entropy production estimators discussed so far did not require, in principle, any knowledge of the underlying dynamics of the molecular system, as they attempt to evaluate the photon sequence probabilities required by Eqs. 6-7 directly. The method described in Section 5, of course, did require the knowledge of the rate matrix $\mathbf{K}$, but only as a *practical* shortcut to evaluating the probabilities of very long photon sequences: in principle, given sufficient statistics and computer power Eqs. 6-7 could be evaluated directly without any knowledge of $\mathbf{K}$. Hidden Markov models (HMMs), which currently belong to the standard toolkit of FRET experiments, offer an entirely different approach^{17, 40-42, 44, 45, 63, 71-73}. These approaches fall into two broad classes. The first uses binning to obtain quantities such as FRET efficiencies as a function of time. This requires that the inter-photon time μ^{-1} is far shorter than the timescales of the dynamics of interest. See refs.^{17, 74, 75} for a recent survey and discussion of such methods. The second approach uses individual photon arrival times and colors to infer the kinetic parameters of the system and is particularly suitable for studying fast molecular phenomena such as transition paths^{4, 7, 49}. Examples of this approach include the Schröder-Grubmüller⁵² and the Gopich-Szabo methods^{40, 63}, H2MM^{1, 41, 42}, and Bayesian inference methods from the Presse group^{44, 45}. A central quantity in these single-photon approaches is a likelihood function evaluated from photon sequences (see below). The Gopich-Szabo method and the H2MM maximize this function to determine the optimal model parameters. Bayesian non-parametric approaches^{44, 45} go beyond this and sample the posterior probabilities in the model and parameter space, albeit at increased computational expense. Here we limit our discussion to the maximum likelihood method applied to individual photon sequences, resulting in an entropy estimator that we will refer to as the “maximum likelihood estimator” or MLE.

In the maximum likelihood method, for a given photon sequence the probability of Eq. 21,

$$p[i_1, t_1, i_2, t_2, \dots, i_n, t_n] \equiv L(\mathbf{K}, i_1, t_1, i_2, t_2, \dots, i_n, t_n), \quad (28)$$

is the likelihood function that is to be maximized with respect to the matrix $\mathbf{K}$ of the rate coefficients (and possibly photon count rates and other model parameters). Once the optimal $\mathbf{K}$ is found, the entropy production can be estimated directly from Eq. 4. Importantly, the structure of the matrix $\mathbf{K}$ is determined by the topology of the assumed kinetic scheme (Fig. 1), which is usually postulated beforehand. In general, there is no guarantee that the assumed model is correct; the situation is further complicated by the fact that it may be impossible, in principle, to differentiate among different equivalent models (possibly of comparable complexity) based on the available data⁷⁶.

In principle, for an infinitely long sequence of photons and for any given topology of the assumed Markov model (cf. Fig. 1, left) the likelihood function is expected to be the delta function (or possibly a sum of delta functions) with respect to the model parameters^{71, 77, 78}. Thus if the topology of the kinetic scheme is known ahead of time, the maximum likelihood method can, in principle, give the exact entropy production, regardless of the photon count rate μ (i.e. even in the limit where many molecular transitions take place in between consecutive photoemission events). Therefore, unlike all other methods considered here so far, MLE's performance is not limited by the photon count rate. In practice, of course, we expect that statistical errors will strongly affect the performance of the method at low photon counts⁷¹.

One important consequence of such statistical errors is that, for an equilibrium system obeying detailed balance (i.e. with zero entropy production), they will generally bias *the MLE toward predicting positive entropy production*, simply because Eq. 4 always yields a nonnegative number. That is, if $\mathbf{K}$ satisfies detailed balance exactly, Eq. 4 will give zero entropy production, but any statistical errors in the entries of the matrix $\mathbf{K}$ can only increase this value (an exception is the case of linear model topology, where detailed balance is always satisfied regardless of the rate coefficients). This complicates differentiating equilibrium from nonequilibrium systems using MLE approaches.

Another fundamental limitation of the MLE method is that its performance depends (in an unpredictable way) on the assumed (and generally unknown) topology of the Markov kinetic scheme. For example, consider two 3-state kinetic schemes shown in Fig. 3. If the linear scheme (Fig. 3, right) is assumed as the underlying Markov model, it will automatically satisfy detailed balance, and the MLE method will always predict zero entropy production, even if the true dynamics involves a dissipative cycle (Fig. 3, left).

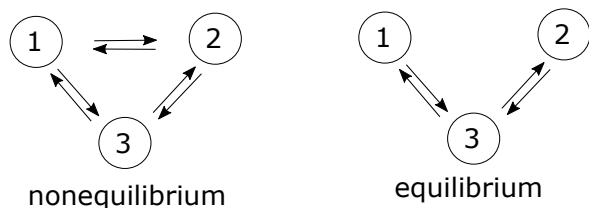


Figure 3. Two 3-state kinetic models: the one on the left may represent a dissipative cycle with nonzero entropy production, while the one on the right always satisfies detailed balance and has zero entropy production.

The consequences of assuming an incorrect Markov model for MLE estimation of entropy production will be further explored in the next Section. Bayesian non-parametric approaches, which simultaneously sample both in the parameter space and the model space (see, e.g.,

refs.^{44, 45, 79}) are a potentially promising way to overcome this problem, but they come at high computational cost and will not be considered here. In practice, physical, chemical, or structural insight, along with Bayesian criteria for model selection that take model complexity into account^{80, 81}, may guide one toward better hidden Markov models of the observable process.

8. Maximum likelihood vs. “histogram” estimators vs compression algorithms: a case study.

Here we compare the performance of different entropy production estimators using the toy model of a 3-state single-file random walk shown in Figure 4. The model involves two particles that can occupy 3 possible sites, with no more than one particle occupying each site. Therefore, a particle can hop onto another site only when that site is unoccupied. Of the two particles (gray and red in Fig. 4) only one (red) is observed. Thus the observer sees the red particle hopping between 3 sites, similarly to Fig. 1, right, and Fig. 3, left. More precisely, we will assume that the system is probed via 3-color FRET that can differentiate between the red particle occupying each of the 3 sites (Fig. 1, right). Moreover, we assume the 3-color FRET setup of Section 3, where the color of the photon uniquely identifies the corresponding site. Unlike the example of Section 3, however, the hopping dynamics of the observed particle are inherently non-Markov (in fact, infinite-order Markov)^{37, 65}.

We consider two scenarios: in the first (Fig. 4, left) both random walkers are driven, stepping clockwise (counterclockwise) with rates k_+ (k_-) when the transition leads to an unoccupied site. In the second (Fig. 4, right), only the observed particle is driven (with the rates k_+ and k_- for clockwise and counterclockwise steps), while the unobserved gray particle has equal rates k_0 for stepping in either direction.

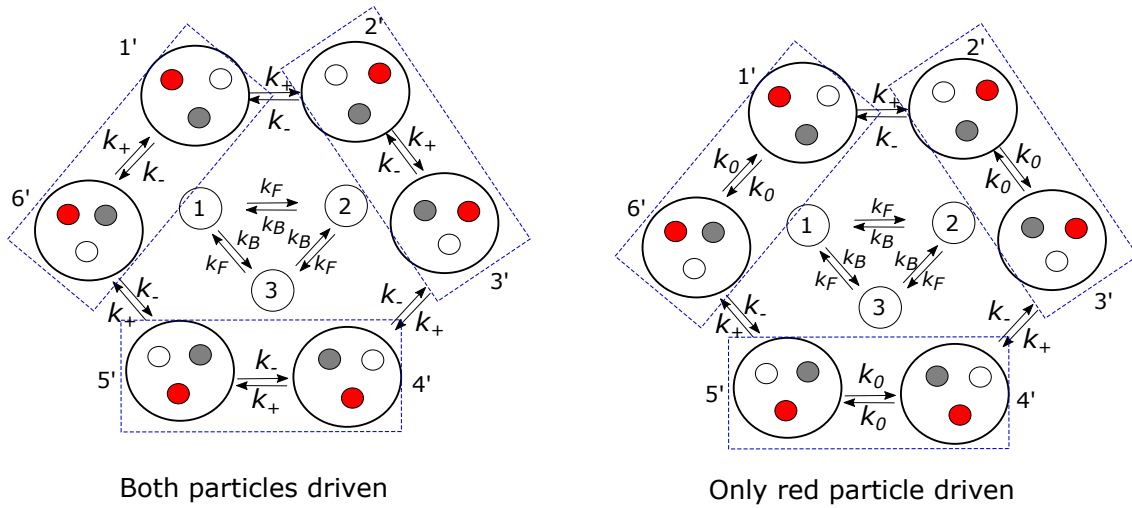


Figure 4. Continuous-time single-file random walk, in which two random walkers can each occupy one of the 3 sites, with at most one walker occupying each site. Only one of the two walkers (red particle) is observed. The “microscopic” Markov process describing the system contains 6 states arranged on a ring. Left: when both walkers are driven, each makes transitions in clockwise and counterclockwise directions with rates k_+ and k_- , whenever such transitions lead to an unoccupied site. Right: the observed walker is driven (with rates k_+ and k_- for clockwise and counterclockwise transitions), but the unobserved walker has no preference as to go clockwise or counterclockwise (transition rate k_0 in either direction). The 6 microscopic states (enumerated 1', 2', ..., 6') are not observable directly. Rather, coarse states including pairs of microscopic states (e.g. coarse state 1 consists of states

2' and 3') are observed, as indicated by rectangles. A Markovian approximation to this random walk assumes that the transitions between states 1-3 happen with rates k_F and k_B in the clockwise and counterclockwise directions, as in Fig. 1, right.

As illustrated in Fig. 4, each of the scenarios can be described using an underlying microscopic Markov model with 6 states^{37, 65, 82, 83}, where a coarse-grained state with a known location of the observed particle consists of two microscopic states corresponding to the two possible positions of the unobserved particle. The microscopic states form a ring, with the transition rates indicated in Fig. 4. The exact entropy production for both models is then readily estimated using Eq. 4:

$$\sigma_A = (k_+ - k_-) \log(k_+/k_-) \quad (29)$$

for the case where both particles are driven, and

$$\sigma_B = \frac{k_0(k_+ - k_-)}{2k_0 + k_+ + k_-} \log(k_+/k_-) \quad (30)$$

for the case of a single driven particle. As discussed in Section 7, MLE with a hidden Markov model that assumes the correct kinetic scheme with 6 states shown in Fig. 4 will, in principle, recover the correct kinetic parameters and thus give the exact entropy production in each case, given infinite amount of data. But as the exact number of states and the topology of the Markov model is unknown, here we investigate how choosing the “wrong” Markov model⁴⁰ affects entropy production estimates. The most natural choice of such a model here would be the one identifying the observable states of the walker with the true microscopic states of the system. This means that the system’s dynamics would be interpreted using a 3-state Markov model (Figure 1, right and Figure 3, left). The MLE will then predict the *effective* transition rates k_F and k_B for hopping in the clockwise and counterclockwise directions (Fig. 4).

Since in the limit of infinite photocount rate μ the site of the observed particle is known exactly at any moment of time, it seems plausible that k_F and k_B will be such that the average frequencies of clockwise and counterclockwise transitions will be the same as the ones in the original non-Markov model (in what follows, we will refer to this model as the “optimal Markov model”). More specifically, as the pairs of the microscopic states (2',3'), (4'5'), and (6',1') are mapped onto states 1, 2, and 3 of the 3-state model (we use primes to denote microscopic states, as in Fig. 4), we expect that the effective 3-state model will preserve the fluxes between the coarse states 1,2,3, e.g.

$$K_{4'3'}p_{3'} - K_{3'4'}p_{4'} = k_F p_1 - k_B p_2 = (k_F - k_B)/3. \quad (31)$$

Here $p_i = 1/3$ is the steady-state population of a coarse state, $p_{i'}$ is the steady-state population of a microscopic state, and $K_{j'i'}$ is the transition rate from a microscopic state i' to a microscopic state j' . For example, $K_{4'3'} = k_-$ and $K_{3'4'} = k_+$. Furthermore, the 3-state model should also preserve the (splitting) probabilities to go in the clockwise and counterclockwise directions, e.g.:

$$\frac{k_F}{k_B} = \frac{K_{4'3'}p_{3'}}{K_{1'2'}p_{2'}} = \frac{k_+p_{3'}}{k_-p_{2'}} \quad (32)$$

Solving Eq. 2 for the steady-state probabilities of the microscopic kinetic scheme and using Eqs. 31-32 one finds

$$k_F = \frac{k_+}{2}, k_B = \frac{k_-}{2} \quad (33)$$

for the case where both particles are driven. This result is intuitively appealing: the observed particle can hop to an adjacent site half of the time (i.e., when this site is unoccupied), so that the effective 3-state kinetic scheme, indeed, preserves the transition frequencies.

For the case where only one particle is driven we find:

$$k_F = \frac{k_-(k_++k_0)}{2k_0+k_++k_-}, k_B = \frac{k_+(k_-+k_0)}{2k_0+k_++k_-} \quad (34)$$

Using these rate coefficients in the Markov entropy production estimates (Eq. 13), $\sigma = (k_F - k_B) \log(k_F/k_B)$, we obtain the following optimal Markov entropy production estimates:

$$\sigma_A^{eff} = \frac{(k_+-k_-)}{2} \log(k_+/k_-) \quad (35)$$

$$\sigma_B^{eff} = \frac{k_0(k_+-k_-)}{2k_0+k_++k_-} \left\{ \log \frac{k_+}{k_-} + \log \frac{k_0+k_-}{k_0+k_+} \right\} = \sigma_B + \frac{k_0(k_+-k_-)}{2k_0+k_++k_-} \log \frac{k_0+k_-}{k_0+k_+} \quad (36)$$

Here the superscript “eff” indicates that those are not the true entropy production values but ones obtained from a Markov model with effective rates. For the case where both particles are driven, the effective 3-state model thus constructed underestimates the entropy production by a factor of two. In the case where only a single particle is driven, the accuracy of the estimate depends on the rate k_0 of transitions of the undriven particle. For $k_0 \rightarrow \infty$, the 3-state model recovers the exact entropy production, a result that is easily understood by examining Figure 4, right. Indeed, in this limit, the interconversion timescale k_0^{-1} within coarse states is much shorter than the mean dwell time $(k_B + k_F)^{-1}$ within coarse states, resulting in effectively Markovian transitions between coarse states⁶⁵.

The limit $k_0 \rightarrow 0$ is more interesting. In this limit, both the true entropy production σ_B and its effective Markov estimate σ_B^{eff} go to zero; however, we also have

$$\lim_{k_0 \rightarrow 0} (\sigma_B^{eff} / \sigma_B) = 0, \quad (37)$$

and thus the *relative* accuracy of the Markov estimate becomes increasingly poor as the rate k_0 decreases.

In our numerical MLE estimates, we generated photon sequences from the kinetic scheme in Figure 4 using the standard Kinetic Monte Carlo method (see, e.g., refs.^{53, 54, 57, 84}). We

then used the Fretica software^{85, 86} (version 9.11.2023) developed by D. Nettels, B. Schuler and coworkers. The log-likelihood of sequences consisting of 200,000 photons, was jointly optimized with respect to k_F , k_B , and μ . This causes photon trajectories simulated at low values of μ to represent a longer time period than those simulated at high μ . The entropy production was then estimated using $\sigma^{MLE} = (k_F - k_B) \log(k_F/k_B)$ (cf. Eq. 13). In general, we find that the estimated parameters k_F and k_B are not equal to their “optimal” values, Eq. 33-34, except in the limit $\mu \rightarrow \infty$ (Figs. 5 and 6).

We now turn to the model-free entropy production estimators that are based on Eq. 7. We call them “histogram” estimators, as they are based on computing multidimensional histograms for the probabilities of different photon sequences. As noted above, the advantage of such an estimator is that it does not require any knowledge or guess about the underlying microscopic dynamics, but a disadvantage is that its performance (even if unlimited data are available) is limited by the photon count rate μ . Moreover, when photon sequences exhibit significant memory, the numerical evaluation of Eq. 7 becomes prohibitive, and further approximations assuming memory cutoffs must be introduced. That is, instead of estimating the entropy production $\sigma_i^{(\infty)}$, one estimates $\sigma_i^{(n)}$ at some finite Markov order n . Here we report data for $n = 1, 2$, and 3 , estimated using the longer of either 10^8 photons or the amount of time needed to simulate 10^5 state transitions.

The simplest (and least costly computationally) estimate, $\sigma_i^{(1)}$, assumes that the photon sequence is a 1st order Markov process, with only pair of consecutive photons being correlated. In this case, the entropy production is described by Eqs. 15-17, with a suitably chosen matrix R . For example, since states $2'$ and $3'$ lead to emission of photons of color 1, we have $R_{12} = R_{13} = 1$, but $R_{22} = R_{23} = R_{32} = R_{33} = 0$ etc. We can also use Eq. 27 to obtain the “infinite-Markov-order” (albeit unobtainable in real applications) estimate of Eq. 7, which is not hampered by computational issues and which provides an information-theoretical limit on the entropy production obtained from Eq. 27 given a finite photocount rate μ .

The results are illustrated in Figure 5 (two driven particles) and 6 (one driven particle). We find that the infinite-order-Markov estimate converges toward the true entropy production (Eqs. 29,30) as the photon count rate increases, with nearly exact result achieved when the mean inter-photon time is 1-2 orders of magnitude shorter than the mean lifetime of an observable state in the system. The performance of the 1-st order Markov estimate is comparable to that of the optimal Markov model estimate, Eqs. 35-36. Curiously, this estimate shows a non-monotonic dependence on the photocount rate μ , similar to that observed in Fig. 2. For sufficiently high values of μ the 1st-order Markov estimate is slightly better (i.e. greater) than the optimal Markov estimate, but the effect is rather small for all model parameters that we have considered.

The maximum likelihood estimate of the entropy production rate, as anticipated above, coincides with the optimal Markov model estimate, Eqs. 35-36, in the limit $\mu \rightarrow \infty$, but, somewhat surprisingly, it predicts a greater entropy production rate for small photocount rates. While this may seem like a welcome news (predicted values closer to the true one, as in Fig. 6), it is not: In general, good entropy production estimators are expected to provide a lower bound, yet the MLE method may predict *greater* values than the *true* entropy production (Fig. 5) and thus suggest that the observed process is more irreversible than it actually is.

The higher-order entropy production estimates, $\sigma_i^{(2)}$ and $\sigma_i^{(3)}$, produce modest improvements, and the compression-based method leads to more significant improvement at intermediate photocount rates μ , but it results in spurious non-monotonic μ dependence at high photocount rates. We note that higher-order estimates $\sigma_i^{(n)}$ quickly become computationally prohibitive and entail large statistical errors⁶² thus requiring more data (thereby making them even more demanding). For example, when estimating $\sigma_i^{(2)}$ or $\sigma_i^{(3)}$ at higher values of μ , we find that there is insufficient data to compute probabilities of all reverse sequences, leading to a division by zero in Equation 7. In Figures 5 and 6, we do not plot the values for $\sigma_i^{(2)}$ or $\sigma_i^{(3)}$ when this happens, leading to their lines disappearing in some parts of the plot.

As the compression-based estimates do not rely on Markovianity of photon sequences (or their being close to Markov), we expect that the compression approach will be even more advantageous when another source of non-Markovianity – namely, the emission of photons of different colors from each state, as in Section 4 – is present. On the other hand, a drawback of the compression approach is the lack of existing methods for controlling its errors in a systematic way: although one hopes that it will converge to the exact result in the limit of infinite trajectory length, in practice this convergence is extremely slow⁶².

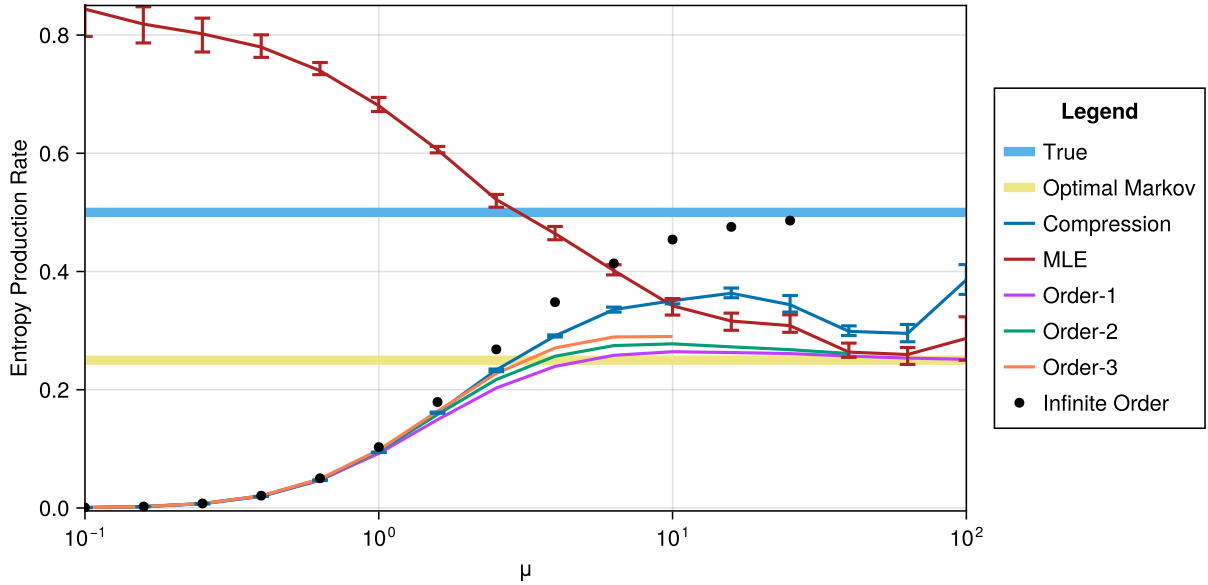


Figure 5. Entropy production rate estimates (in units of $k_B T k_+$) for single-file diffusion (Fig. 4, left) with $k_- = 0.5 k_+$ plotted as a function of the photocount rate μ measured in units of k_+ . The “infinite order estimate” is obtained from Eq. 27, and an “order- n ” estimate $\sigma_i^{(n)}$ assumes that the sequence of photons is n^{th} – order Markov. The “MLE” estimate assumes that the dynamics of the system can be described as a Markov process with 3 interconnected states. The optimal Markov estimate (Eq. 35) assumes a 3-state Markov model with the rates preserving the transition frequencies of the original model and underestimates the true entropy production (Eq. 29) by a factor of 2. The “infinite-order” Markov estimate is based on Eq. 27 and represents the $\sigma_i^{(\infty)}$ limit. Error bars represent the maximum range observed over 5 trials.

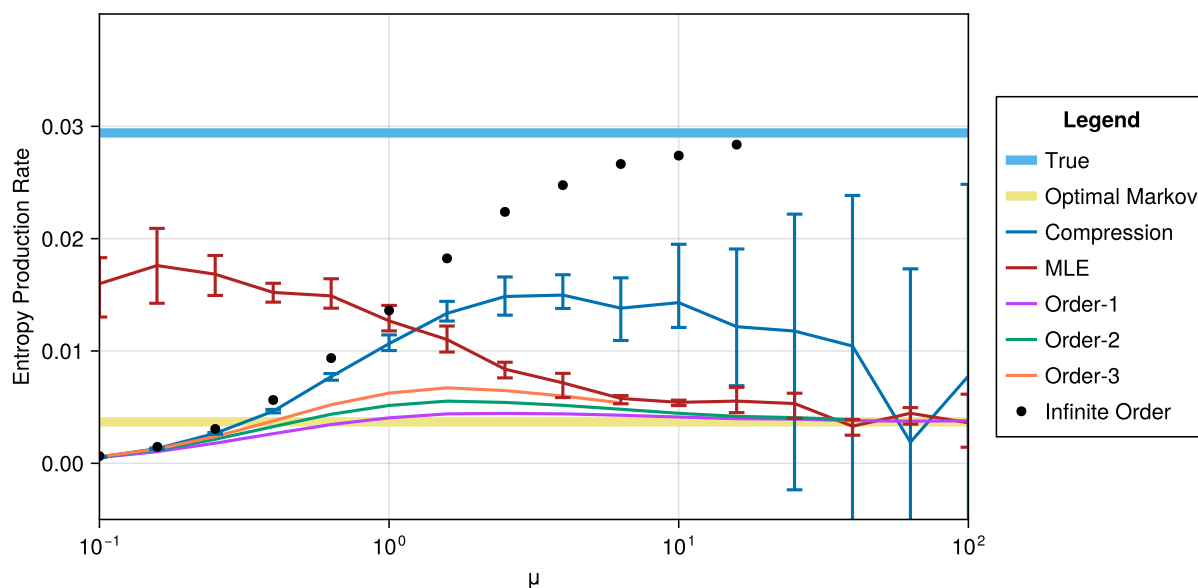


Figure 6. Entropy production rate estimates (in units $k_B T k_+$) for single-file diffusion (Fig. 4, right) with $k_- = 0.5 k_+$ and $k_0 = 0.1 k_+$, plotted as a function of the photocount rate μ measured in units of k_+ . The “infinite order estimate” is obtained from Eq. 27, and an “order- n ” estimate $\sigma_i^{(n)}$ assumes that the sequence of photons is n^{th} – order Markov. The “MLE” estimate assumes that the dynamics of the system can be described as a Markov process with 3 interconnected states. The optimal Markov estimate (Eq. 36) assumes a 3-state Markov model with the rates preserving the transition frequencies of the original model and significantly underestimates the true entropy production (Eq. 30). The “infinite-order” Markov estimate is based on Eq. 27 and represents the $\sigma_i^{(\infty)}$ limit. Error bars represent the maximum range observed over 5 trials.

9. Beyond Poissonian photoemission statistics: more realistic models for FRET photophysics.

So far, the discussion was limited to the case where the photon arrival times were described by Poisson statistics. For a single-molecule source of light, this is only an approximation (see, e.g., ref.^{61, 87}). Photons emitted by a single-molecule source tend to “repel” one another or “anti-bunch”, as detection of one photon implies that the molecule is in the ground state and is unlikely to immediately emit another photon. Even more importantly, the cycle of laser excitation and photon emission is a nonequilibrium process, and one must ask whether the nonequilibrium nature of photophysics will contribute to the entropy production estimates obtained using Eqs. 6 and 7. Here we address this question by considering time reversibility of single-molecule emitters that do not undergo any conformational dynamics. Specifically, Fig. 7, left, shows a simplified scheme of a FRET, involving 3 states. After a photon has been emitted by either the donor or acceptor molecule, the system is found in its ground state G. The FRET donor can then be excited to the state D* by a laser, which is followed by either reemission of a photon by the donor (D* to G transition) or by energy transfer to the A* state where the acceptor is excited. This is eventually followed by a transition from A* to G where an acceptor photon is emitted. This scheme contains a nonequilibrium cycle (cf. Fig 3,

left), and the question is whether its nonequilibrium nature can be deduced from the sequence of the donor and acceptor photons.

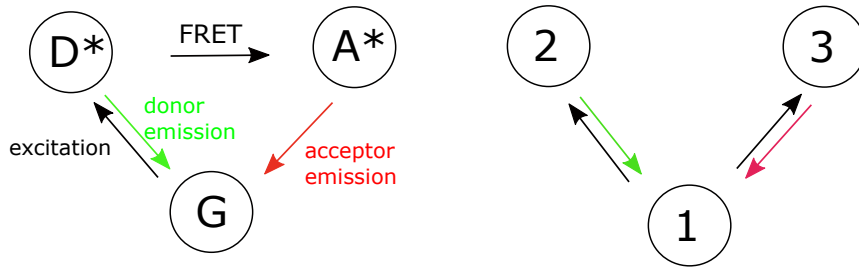


Figure 7. Left: Simplified scheme of FRET, where excitation of the donor from the ground state of the system results either in photon emission from the donor or energy transfer to the acceptor, which subsequently emits a lower-frequency photon. Colored arrows indicate the transitions that are observable as photons. The system contains a nonequilibrium cycle $G \rightarrow D^* \rightarrow A^* \rightarrow G$. Right: a hypothetical equilibrium 3-state system, where only transitions $2 \rightarrow 1$ and $3 \rightarrow 1$ (indicated, by analogy with the scheme on the left, with colored arrows) are observable.

For the purpose of evaluating Eq. 7, one needs to consider the sequence of photon colors $i_1, i_2, \dots, i_n$ (where $i \in \{D, A\}$) and discard the information about their arrival times $t_1, t_2, \dots, t_n$. Since we know the state (i.e., G) of the system with certainty after detecting the photon of either color, there is no memory in the color; in fact, the colors are independent random numbers drawn with probabilities $p_D = 1 - \epsilon$, and $p_A = \epsilon$, where ϵ is the FRET efficiency. Such a random sequence of two possible colors is time-reversible¹⁴. Therefore, at the level of Eq. 7, the photon color sequence does not reflect the irreversibility of the excitation/emission process.

When the temporal information about photon arrival times is included, the situation is different. We can record the full information about the photons as a sequence of pairs $(\tau_1, i_1), (\tau_2, i_2) \dots, (\tau_n, i_n)$, where $\tau_m = t_m - t_{m-1}$ is the dark time preceding the detection of the m -th photon. Consider, for instance, the following 3-photon sequence, $(\tau_{m-1}, A), (\tau_m, A) \dots, (\tau_{m+1}, D)$. The time τ_m is drawn from the distribution $\psi_A(\tau)$ conditional upon observing the acceptor photon (and thus going through the sequence of states $G \rightarrow D^* \rightarrow A^* \rightarrow G$), which is reflected in the subscript “A”. The time τ_{m+1} is drawn from the distribution $\psi_D(\tau)$ conditional upon going through the sequence $G \rightarrow D^* \rightarrow G$ resulting in the emission of a donor photon. According to the statistical properties of the above sequence, its time reversal results in the sequence $(\dots, D), (\tau_{m+1}, A), (\tau_m, A)$, with the lag time preceding the second photon (A) drawn from $\psi_D(\tau)$. But this is different from the distribution $\psi_A(\tau)$ that this time would have in the forward-in-time photon sequence. Therefore, the statistics of the original and time-reversed photon sequences are different when photon arrival times are taken into account, and thus Eq. 6 would predict irreversible dynamics in this case. Stated more briefly, the lag-time distribution conditioned on the first photon being emitted by the donor and the next by the acceptor is given by $p_{D \rightarrow A}(\tau) = \psi_A(\tau)$, which is different from $p_{A \rightarrow D}(t) = \psi_D(\tau)$, and thus the photon sequence is time-asymmetric. We note that such asymmetry has been previously observed in correlation functions of donor and acceptor emission intensity^{57, 58}.

Unfortunately, this time irreversibility is spurious rather than indicative of nonequilibrium photophysics, as it could lead one to erroneously conclude that the kinetic scheme shown in Fig. 7, right, which obeys detailed balance (cf. Fig. 3, right), is time reversible as well! Indeed, this result inevitably follows from the difference between the distributions $\psi_A(\tau)$ and $\psi_D(\tau)$, which generally exists even for the equilibrium scheme on the right.

The spurious irreversibility found here is akin to the finding of ref.³⁷, and our conclusion is, similarly, that (1) Eq. 6 does not correctly uncover the driven nature of the FRET photophysical processes and (2) the temporal information in photon sequences should be discarded and Eq. 7 should be used, in which case photophysical effects will not contribute to the estimated entropy production. We note that the failure of Eq. 7 to quantify irreversibility of the photoemission process is not surprising: In order to do so one generally expects^{14, 22} both forward and backward transitions (e.g. both the photoexcitation $G \rightarrow D^*$ and the photoemission $D^* \rightarrow G$ steps in Fig. 7, left) to be observable, which is not the case.

10. Conclusions.

Inferring dissipation by molecular machines, as studied by FRET, is a difficult task, and all of the methods considered here have advantages and disadvantages. Generally, we have considered two classes of approaches, model-based ones (such as hidden Markov maximum likelihood estimators) and “model-free” ones (which are based on direct estimation of the KL divergence between the forward and time-reversed photon sequences). The maximum likelihood estimates depend on how close the assumed model is to the true microscopic dynamics. If the two coincide, the MLE method will, in principle, give the exact entropy production and exact transition rates given enough data^{71, 78}. In practice, of course, low photocount rates may result in significant statistical errors⁷⁸. Moreover, statistical errors will bias MLE toward positive estimated values of entropy production (and thus toward erroneous characterization of the dynamics as irreversible) when the system satisfies detailed balance and thus when the true entropy production is exactly zero.

When the assumed Markov model is not an accurate description of true dynamics then the entropy production estimate depends on both the photon count rate μ and on the model itself. For example, a Markov model assuming a linear kinetic scheme will always predict zero entropy production by construction. In contrast, a non-Markov model involving discrete states arranged linearly may still show time asymmetry¹⁴. A Markov model that assumes wrong dynamics may further overestimate the entropy production, making the dynamics appear more irreversible than the true dynamics at low photocount rates.

Model free estimators that approximate Eqs. 6-7 are limited by the photon count rate and can recover the true entropy production only in the limit $\mu \rightarrow \infty$. This is not surprising, since the state of the system is unknown during the dark periods between photons. The examples studied here suggest that the true entropy production can be recovered when both the infinite Markov order and infinite emission rate limits are taken, and that finite-Markov order estimates $\sigma_i^{(n)}$ at finite emission rates provide a lower bound on entropy production. Except for an analytically solvable model of Section 3, however, we do not have a strict proof of the above two statements.

In practice, we find that Eq. 7, which discards information about photon arrival times and only considers their sequences, does not result in any loss of accuracy and, moreover, avoids artifacts associated with spurious contribution of photophysical effects to the entropy production (Section 9). Discarding temporal information is of significant computational advantage, as the “photon color states” usually form a small discrete set, while photon arrival times are continuous and require further discretization⁶⁶.

Still, in practice, high-Markov-order estimates are computationally challenging, but compression-based approaches (particularly the method of ref. ⁶⁹) hold some promise. On the other hand, the compression approach suffers from the systematic errors of the compression method introduced by the finite amount of data⁶² (i.e. compression of limited amounts of data is never perfect) and from their rather weak dependence on the amount of data. Overall, there seems to be no clear winner among the methods considered here – it seems advisable to have a toolkit containing both model-based and model-free methods.

Acknowledgments.

Comments from and discussions with Kristian Blom, Aljaž Godec, Irina Gopich, Helmut Grubmüller, Gilad Haran, Hagen Hofmann, Thorsten Hugel, Steve Pressé, and Benjamin Schuler are gratefully acknowledged. This work was supported by the Robert A. Welch Foundation (Grant No. F- 1514 to DEM), the National Science Foundation (Grant Nos. CHE 1955552 to DEM and IIS 2212048 to EV), and by the Adobe Inc.

Appendix A. Entropy production from photon sequences with and without temporal information.

In most of this paper, we used Eq. 7 to estimate the entropy production from photon sequences. This equation discards the information about the arrival times of the photons, which is included in Eq. 6. In practice, we find the two estimates to be nearly identical when the inter-photon times μ^{-1} are shorter than the typical dwell times in each state of the system. To illustrate this, here we consider the Markov entropy production estimators based on Eqs. 6 and 7, which neglect correlations between photons beyond those between consecutive photon pairs.

The first order Markov estimate of the entropy production including temporal information is given by

$$\sigma_{ti}^{(1)} = \mu^2 \int_0^\infty dt \sum_{i,j} p(j, t|i, 0) p_{ss}(i) e^{-\mu t} \log \frac{p(j, t|i, 0)}{p(i, t|j, 0)}, \quad p(j, t|i, 0) = (e^{Kt})_{ji}, \quad (\text{A1})$$

where the summation accounts for averaging over states and integration accounts for averaging over emission times in Eq. 6. The estimate $\sigma_i^{(1)}$ without the temporal information is given by Eqs. 8-9. For the 3-state model of Fig. 1, right, this result is further given by Eq. 11. In Figure A1 we plot $\sigma_{ti}^{(1)}$ and $\sigma_i^{(1)}$, as a function of the photon count rate μ , for this three state system using the same parameters as in Fig. 2. Both estimates give nearly identical results for the emission rates in the range $0.5 < \mu k_B < 100$.

fraction entropy production recovered

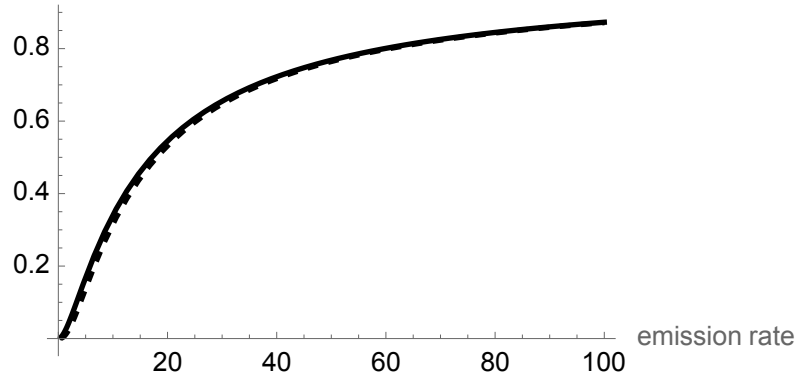


Figure 8. First order estimates of entropy production including temporal information ($\sigma_{ti}^{(1)}$, solid line) and not including it ($\sigma_i^{(1)}$ dashed line) as a function of the photon count rate μ for the 3-state model of Figure 1, right, using 3-color FRET. Photon colors directly correspond to the molecular states. The model parameters are $k_F = 2k_B$, and the emission rate μ is measured in dimensionless units set by k_B .

References

1. H. Mazal and G. Haran, *Curr Opin Biomed Eng* **12**, 8-17 (2019).
2. R. Kellner, H. Hofmann, A. Barducci, B. Wunderlich, D. Nettels and B. Schuler, *Proc Natl Acad Sci U S A* **111** (37), 13355-13360 (2014).
3. N. R. Marzano, B. P. Paudel, A. M. van Oijen and H. Ercroyd, *Sci Adv* **8** (50), eadd0922 (2022).
4. H. S. Chung and W. A. Eaton, *Curr Opin Struct Biol* **48**, 30-39 (2018).
5. B. Schuler and W. A. Eaton, *Curr Opin Struct Biol* **18** (1), 16-26 (2008).
6. B. Schuler and H. Hofmann, *Curr Opin Struct Biol* **23** (1), 36-47 (2013).
7. F. Sturzenegger, F. Zosel, E. D. Holmstrom, K. J. Buholzer, D. E. Makarov, D. Nettels and B. Schuler, *Nat Commun* **9** (1), 4708 (2018).
8. R. Roy, S. Hohng and T. Ha, *Nat Methods* **5** (6), 507-516 (2008).
9. O. Opanasyuk, A. Barth, T. O. Peulen, S. Felekyan, S. Kalinin, H. Sanabria and C. A. M. Seidel, *J Chem Phys* **157** (3), 031501 (2022).
10. M. T. J. Halma, D. B. Ritchie, T. R. Cappellano, K. Neupane and M. T. Woodside, *Proc Natl Acad Sci U S A* **116** (39), 19500-19505 (2019).
11. C. J. Bustamante, Y. R. Chemla, S. Liu and M. D. Wang, *Nature Reviews Methods Primers* **1** (1), 25 (2021).
12. H. Yu, M. G. Siewny, D. T. Edwards, A. W. Sanders and T. T. Perkins, *Science* **355** (6328), 945-950 (2017).
13. J. O. Andreasson, S. Shastry, W. O. Hancock and S. M. Block, *Curr Biol* **25** (9), 1166-1175 (2015).
14. A. Godec and D. E. Makarov, *J Phys Chem Lett* **14** (1), 49-56 (2023).
15. C. Ratzke, B. Hellenkamp and T. Hugel, *Nat Commun* **5**, 4192 (2014).
16. L. Vollmar, J. Schimpf, B. Hermann and T. Hugel, *Nat Commun* **15** (1), 569 (2024).
17. M. Götz, A. Barth, S. S. R. Bohr, R. Börner, J. Chen, T. Cordes, D. A. Erie, C. Gebhardt, M. C. A. S. Hadzic, G. L. Hamilton, N. S. Hatzakis, T. Hugel, L. Kisley, D. C. Lamb, C. de Lannoy, C. Mahn, D. Dunukara, D. de Ridder, H. Sanabria, J. Schimpf, C. A. M. Seidel, R. K. O. Sigel, M. B. Sletfjerding, J. Thomsen, L. Vollmar, S. Wanninger, K. R. Weninger, P. Xu and S. Schmid, *Nature Communications* **13** (1), 5402 (2022).
18. E. Roldan and J. M. Parrondo, *Phys Rev Lett* **105** (15), 150607 (2010).
19. E. Roldan and J. M. Parrondo, *Phys Rev E Stat Nonlin Soft Matter Phys* **85** (3 Pt 1), 031129 (2012).
20. B. Ertel, J. van der Meer and U. Seifert, *Phys Rev E* **105** (4-1), 044113 (2022).
21. B. Ertel, J. van der Meer and U. Seifert, *Int J Mol Sci* **24** (8) (2023).
22. J. v. d. Meer, B. Ertel and U. Seifert, *Phys. Rev. X* **12**, 031025 (2022).
23. J. van der Meer, J. Degunther and U. Seifert, *Phys Rev Lett* **130** (25), 257101 (2023).
24. T. R. Gingrich, J. M. Horowitz, N. Perunov and J. L. England, *Phys Rev Lett* **116** (12), 120601 (2016).
25. S. K. Manikandan, D. Gupta and S. Krishnamurthy, *Phys Rev Lett* **124** (12), 120603 (2020).
26. S. Otsubo, S. Ito, A. Dechant and T. Sagawa, *Phys Rev E* **101** (6-1), 062106 (2020).
27. Q. Yu, D. Zhang and Y. Tu, *Phys Rev Lett* **126** (8), 080601 (2021).
28. D. J. Skinner and J. Dunkel, *Proc Natl Acad Sci U S A* **118** (18) (2021).

29. I. Di Terlizzi, M. Gironella, D. Herraiez-Aguilar, T. Betz, F. Monroy, M. Baiesi and F. Ritort, *Science* **383** (6686), 971-976 (2024).
30. G. Knotz, T. M. Muenker, T. Betz and M. Kruger, *Frontiers in Physics* **11**, 11:1331835 (2024).
31. D. Hartich and A. Godec, *Phys. Rev. X* **11**, 041047 (2021).
32. D. Hartich and A. Godec, *Phys. Rev. Research* **5**, L032017 (2023).
33. D. Hartich and A. Godec, arXiv:2112.08978, 2112.08978 (2022).
34. P. E. Harunari, A. Dutta, M. Polettini and É. Roldán, *Physical Review X* **12**, 041026 (2022).
35. G. Bisker, M. Polettini, T. R. Gingrich and J. M. Horowitz, *J. Stat. Mech*, 093210 (2017).
36. J. Li, J. M. Horowitz, T. R. Gingrich and N. Fakhri, *Nat Commun* **10** (1), 1666 (2019).
37. K. Blom, K. Song, E. Vouga, A. Godec and D. E. Makarov, *Proc Natl Acad Sci U S A* **121** (17) (2024).
38. L. Peliti and S. Pigolotti, *Stochastic Thermodynamics*. (Princeton University Press, 2021).
39. U. Seifert, *Rep Prog Phys* **75** (12), 126001 (2012).
40. H. S. Chung and I. V. Gopich, *Phys Chem Chem Phys* **16** (35), 18644-18657 (2014).
41. H. Y. Aviram, M. Pirchi, Y. Barak, I. Riven and G. Haran, *J Chem Phys* **148** (12), 123303 (2018).
42. H. Y. Aviram, M. Pirchi, H. Mazal, Y. Barak, I. Riven and G. Haran, *Proc Natl Acad Sci U S A* **115** (13), 3243-3248 (2018).
43. L. Grabenhorst, F. Sturzenegger, M. Hasler, B. Schuler and P. Tinnefeld, *J Am Chem Soc* **146** (5), 3539-3544 (2024).
44. A. Saurabh, M. Fazel, M. Safar, I. Sgouralis and S. Presse, *Biophys Rep (N Y)* **3** (1), 100089 (2023).
45. A. Saurabh, M. Safar, M. Fazel, I. Sgouralis and S. Presse, *Biophys Rep (N Y)* **3** (1), 100087 (2023).
46. I. Terterov, D. Nettels, D. E. Makarov and H. Hofmann, *Biophys Rep (N Y)* **3** (3), 100116 (2023).
47. H. S. Chung, T. Cellmer, J. M. Louis and W. A. Eaton, *Chem Phys* **422**, 229-237 (2013).
48. H. S. Chung and W. A. Eaton, *Nature* **502** (7473), 685-688 (2013).
49. H. S. Chung, J. M. Louis and W. A. Eaton, *Science* **335**, 981-984 (2012).
50. L. P. Watkins and H. Yang, *Biophys J* **86** (6), 4015-4029 (2004).
51. H. Yang and X. S. Xie, *The Journal of Chemical Physics* **117** (24), 10965-10979 (2002).
52. G. F. Schröder and H. Grubmüller, *The Journal of Chemical Physics* **119** (18), 9920-9924 (2003).
53. R. Elber, D. E. Makarov and H. Orland, *Molecular Kinetics in Condense Phases: Theory, Simulation, and Analysis*. (Wiley and Sons, 2020).
54. D. E. Makarov, in *Reviews in Computational Chemistry*, edited by A. L. Parrill and K. B. Lipkowitz (John Wiley & Sons, 2017), Vol. 30.
55. N. G. v. Kampen, *Stochastic processes in physics and chemistry*, 3rd ed. (Elsevier, Amsterdam ; Boston, 2007).
56. R. Zwanzig, *Nonequilibrium Statistical Mechanics*. (Oxford University Press, 2001).
57. Z. Wang and D. E. Makarov, *J. Phys. Chem. B* **107**, 5617 (2003).
58. D. Nettels, I. V. Gopich, A. Hoffmann and B. Schuler, *Proc Natl Acad Sci U S A* **104** (8), 2655-2660 (2007).

59. B. Hellenkamp, S. Schmid, O. Doroshenko, O. Opanasyuk, R. Kühnemuth, S. Rezaei Adariani, B. Ambrose, M. Aznauryan, A. Barth, V. Birkedal, M. E. Bowen, H. Chen, T. Cordes, T. Eilert, C. Fijen, C. Gebhardt, M. Götz, G. Gouridis, E. Gratton, T. Ha, P. Hao, C. A. Hanke, A. Hartmann, J. Hendrix, L. L. Hildebrandt, V. Hirschfeld, J. Hohlbein, B. Hua, C. G. Hübner, E. Kallis, A. N. Kapanidis, J.-Y. Kim, G. Krainer, D. C. Lamb, N. K. Lee, E. A. Lemke, B. Levesque, M. Levitus, J. J. McCann, N. Naredi-Rainer, D. Nettels, T. Ngo, R. Qiu, N. C. Robb, C. Röcker, H. Sanabria, M. Schlierf, T. Schröder, B. Schuler, H. Seidel, L. Streit, J. Thurn, P. Tinnefeld, S. Tyagi, N. Vandenberg, A. M. Vera, K. R. Weninger, B. Wünsch, I. S. Yanez-Orozco, J. Michaelis, C. A. M. Seidel, T. D. Craggs and T. Hugel, *Nature Methods* **15** (9), 669-676 (2018).
60. J. V. Stone, *Information theory: A Tutorial Introduction*. (Sebtel Press, 2015).
61. D. E. Makarov, *Single Molecule Science: Physical Principles and Models*. (CRC Press, Taylor & Francis Group, Boca Raton, 2015).
62. T. Schurmann and P. Grassberger, *Chaos* **6** (3), 414-427 (1996).
63. I. V. Gopich and A. Szabo, *J Phys Chem B* **113** (31), 10965-10973 (2009).
64. C. E. Shannon, *Bell System Technical Journal* **January**, 57-64 (1951).
65. K. Song, D. E. Makarov and E. Vouga, *Physical Review Research* **5** (1), L012026 (2023).
66. K. Song, R. Park, A. Das, D. E. Makarov and E. Vouga, *J Chem Phys* **159** (6) (2023).
67. C. B. Li, H. Yang and T. Komatsuzaki, *Proc Natl Acad Sci U S A* **105** (2), 536-541 (2008).
68. C. B. Li, H. Yang and T. Komatsuzaki, *J Phys Chem B* **113** (44), 14732-14741 (2009).
69. S. Ro, B. Guo, A. Shih, T. V. Phan, R. H. Austin, D. Levine, P. M. Chaikin and S. Martiniani, *Phys Rev Lett* **129** (22), 220601 (2022).
70. J. Ziv and N. Merhav, *IEEE T. Inform. Theory* **39** (1993).
71. I. V. Gopich, *J Chem Phys* **142** (3), 034110 (2015).
72. I. V. Gopich, D. Nettels, B. Schuler and A. Szabo, *J Chem Phys* **131** (9), 095102 (2009).
73. Z. Kilic, I. Sgouralis and S. Presse, *Biophys J* **120** (3), 409-423 (2021).
74. A. Saurabh, L. W. Q. Xu and S. Presse, *Nat Commun* **15** (1), 3627 (2024).
75. M. Götz, A. Barth, S. S. R. Bohr, R. Börner, J. Chen, T. Cordes, D. A. Erie, C. Gebhardt, M. C. A. S. Hadzic, G. L. Hamilton, N. S. Hatzakis, T. Hugel, L. Kisley, D. C. Lamb, C. de Lannoy, C. Mahn, D. Dunukara, D. de Ridder, H. Sanabria, J. Schimpf, C. A. M. Seidel, R. K. O. Sigel, M. B. Sletfjerding, J. Thomsen, L. Vollmar, S. Wanninger, K. R. Weninger, P. Xu and S. Schmid, *Nature Communications* **15** (1), 3626 (2024).
76. O. Flomenbom, J. Klafter and A. Szabo, *Biophys J* **88** (6), 3780-3783 (2005).
77. S. Pressé and I. Sgouralis, *Data Modeling for the Sciences: Applications, Basics, Computations*. (Cambridge, 2023).
78. M. Pirchi, R. Tsukanov, R. Khamis, T. E. Tomov, Y. Berger, D. C. Khara, H. Volkov, G. Haran and E. Nir, *J Phys Chem B* **120** (51), 13065-13075 (2016).
79. M. Safar, A. Saurabh, B. Sarkar, M. Fazel, K. Ishii, T. Tahara, I. Sgouralis and S. Presse, *Biophys Rep (N Y)* **2** (4), 100088 (2022).
80. G. Schwarz, *The Annals of Statistics* **6** (2), 461-464, 464 (1978).
81. D. J. C. MacKay, *Information theory, inference, and learning algorithms*. (Cambridge University Press, Cambridge, UK ; New York, 2003).
82. J. Shin, A. M. Berezhkovskii and A. B. Kolomeisky, *J Chem Phys* **154** (20), 204104 (2021).
83. J. Shin and A. B. Kolomeisky, *J Chem Phys* **153** (12), 124103 (2020).
84. D. T. Gillespie, *J. Comput. Phys.* **22**, 403 (1976).

85. D. Nettels and B. Schuler, (<https://schuler.bioc.uzh.ch/programs/>).
86. A. Chowdhury, D. Nettels and B. Schuler, *Annu Rev Biophys* **52**, 433-462 (2023).
87. H. Carmichael, *An Open Systems Approach to Quantum Optics*. (Springer, Berlin, New York, 1993).