

# Dissipative timescales from coarse-graining irreversibility

Freddy A Cisneros<sup>1</sup>, Nikta Fakhri<sup>2</sup> and Jordan M Horowitz<sup>3,4,5</sup>

<sup>1</sup> Applied Physics Program, University of Michigan, Ann Arbor, MI 48109, USA

<sup>2</sup> Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>3</sup> Department of Biophysics, University of Michigan, Ann Arbor, MI 48109, USA

<sup>4</sup> Center for the Study of Complex Systems, University of Michigan, Ann Arbor, MI 48109, USA

<sup>5</sup> Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

E-mail: [jmhorow@umich.edu](mailto:jmhorow@umich.edu)

**Abstract.** We propose and investigate a method for identifying timescales of dissipation in nonequilibrium steady states modeled as discrete-state Markov jump processes. The method is based on how the irreversibility—measured by the statistical breaking of time-reversal symmetry—varies under temporal coarse-graining. We observe a sigmoidal-like shape of the irreversibility as a function of the coarse-graining time whose functional form we derive for systems with a fast driven transition. This theoretical prediction allows us to develop a method for estimating the dissipative time scale from time-series data by fitting estimates of the irreversibility to our predicted functional form. We further analyze the accuracy and statistical fluctuations of this estimate.

## 1. Introduction

Dissipation, the irrecoverable loss of energy by a system to its surroundings, is a defining feature of nonequilibrium steady states. Typically, it is quantified by the entropy production rate  $\Sigma$  [1]. This single number, however, elides over the details of the microscopic dynamics: different systems may have the same entropy production rate although their microscopic details may differ. This observation suggests that the development of more refined characterizations of dissipation could aid in our understanding of the nature of nonequilibrium steady states.

One way to characterize a physical process is to identify the characteristic time and length scales on which it operates [2, 3]. For dissipation, we are then aiming to probe the characteristics of the microscopic dynamics that result in energy loss to the surroundings. However, nonequilibrium steady states continually exchange energy with their environment, making it somewhat unclear how to tweeze apart, without a model, the various processes that contribute to dissipation. In this article, we introduce a method to identify and measure these timescales by exploiting the deep connection between irreversibility and dissipation.

One of the central results of stochastic thermodynamics has been the formulation of a quantitative connection between the entropy production rate  $\Sigma$  and the statistical irreversibility (time-reversal asymmetry) of the dynamics as measured by the relative entropy between the probability  $P(\gamma_t)$  of a microscopic trajectory  $\gamma_t$  of length  $t$  and the probability  $P(\tilde{\gamma}_t)$  of the time-reversed trajectory  $\tilde{\gamma}_t$  [4, 5, 6, 7, 8, 9],

$$\Sigma = \lim_{t \rightarrow \infty} \frac{k_B}{t} D[P(\gamma_t) || P(\tilde{\gamma}_t)], \quad (1)$$

where  $k_B$  is the Boltzmann constant, which we will set to one for the remainder of this paper for notational simplicity. The relative entropy between two probability distributions  $p$  and  $q$ ,  $D(p||q) = \int p(x) \ln[p(x)/q(x)] dx$ , is an information-theoretic measure of distinguishability. For this reason, we can interpret the relative entropy in Eq. (1) as a measure of the distinguishability between the forward and reverse trajectories, namely, the *irreversibility*. Furthermore, the entropy production rate itself can be linked to energy dissipation, once we have identified the thermodynamic reservoirs in the environment. For example, when a system is coupled to two heat baths at temperatures  $T_1 > T_2$ , the entropy production rate  $\Sigma = (1/T_2 - 1/T_1)\dot{Q}$  is proportional to the heat flow from reservoir 1 to 2. A useful implication of (1) is that the detection of a nonzero irreversibility from experimental time-series data allows one to deduce that a system is out of equilibrium, just from observations of dynamic fluctuations [10, 11, 12, 13, 14, 15, 16]. However, it has been observed in theoretical and experimental studies that the amount of irreversibility detected is sensitive to the types and numbers of observed variables as well as to the data acquisition rate [17, 18, 19].

Although the data-dependence of the measured irreversibility has been considered a nuisance, here we are going to attempt to exploit it to identify a dissipative timescale. To this end, imagine coarse-graining the dynamics in time by removing all information

on a timescale shorter than  $\tau$ . Examples include making stroboscopic observations at intervals of length  $\tau$  or obtaining data as local time averages over windows of width  $\tau$ . From an experimental point of view, such coarse graining is quite natural as one can never make measurements with infinite precision. No matter the procedure, we assume the result is a collection of  $N$  sequential observations that we can collect into a coarse-grained trajectory  $\gamma_N^\tau$ , from which we can determine the *observed irreversibility* above the timescale  $\tau$

$$\sigma(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N\tau} D[P(\gamma_N^\tau) || P(\tilde{\gamma}_N^\tau)]. \quad (2)$$

Under the mild assumption that coarse-graining commutes with time reversal, the observed irreversibility bounds the entropy production from below,  $\sigma(\tau) \leq \Sigma$  [20, 21, 22]. At each level of coarse-graining, we remove information about the dynamics at timescales shorter than  $\tau$ , implying that the relative entropy  $\sigma(\tau)$  remains sensitive to any irreversible processes operating on timescales longer than  $\tau$ . These arguments suggest that the variation of  $\sigma(\tau)$  with coarse-graining scale  $\tau$  is potentially sensitive to a characteristic timescale relevant to dissipation. Indeed, this was observed in an experimental study of the actomyosin network of a starfish oocyte [15]. The fluctuations of endogenous probe particles embedded in the cortex were measured and used to compute  $\sigma(\tau)$ . The observed irreversibility as a function of  $\tau$  was found to peak at a value consistent with the rate of turnover of adenosine triphosphate (ATP) by the myosin motors responsible for the nonequilibrium fluctuations in the network. In this article, we use simple models of nonequilibrium steady states with well-defined characteristic timescales to analyze this dependence of the observed irreversibility on coarse-graining timescale  $\tau$  in an effort to understand what can be inferred from measurements of  $\sigma(\tau)$ .

It is worth noting that coarse-graining irreversibility as a means to characterize dissipation is not unique to this study and represents a potentially powerful approach to assessing the properties of dissipation in nonequilibrium steady states. Indeed, similar observations were exploited in Ref. [23, 24] where the authors coarse-grained progressively larger interdependencies among a set of interacting degrees of freedom. This allowed them to identify how the complexity of interactions among the system's degrees of freedom contributes to dissipation.

Our approach to identify dissipative timescales complements other proposals in the literature that decompose the dissipation at distinct locations and length scales. Like in our present work, the authors of Ref. [25] begin their analysis with the connection between dissipation and irreversibility embodied in Eq. (1). They then identify a local irreversibility by computing the observed irreversibility within small grids of a spatially extended system. Applications of this methodology to active fluids identified regions where irreversibility was prominent, either as a consequence of mobility-induced phase separation or the inclusion of fixed asymmetric obstacles, such as chevrons. Alternative approaches aim to decompose the irreversibility into distinct frequency and wave number components. For instance, the authors of Ref. [26], suggest evaluating the relative entropy rate by replacing the true microscopic dynamics by a Gaussian diffusion process

with the same power spectrum. The resulting approximate relative entropy rate depends solely on the power spectrum and can be decomposed into frequency and wavenumber contributions. This allows one to identify time and length scales from the shape of this decomposition. A similar though distinct approach has been proposed based on the Harada-Sasa equality [27] that decomposes the dissipation into frequency and wavenumber dependent violations of the fluctuation-dissipation theorem [28, 29]. While offering a principled identification of spatiotemporal scales for diffusive dynamics, the Harada-Sasa equality is known to fail for discrete-state Markov models [30, 31].

We begin in Sec. 2 by introducing our models of interest and collecting their relevant properties. Then in Sec. 3, we numerically and theoretically analyze a model with a single dissipative timescale. We observe that for highly dissipative systems the observed irreversibility  $\sigma(\tau)$  has a simple functional form that we are able to predict using a perturbation theory. This functional form then serves as an ansatz whose fit to experimental or numerical data can be used to estimate the dissipative timescale. The performance of this estimation procedure is analyzed numerically in Sec. 4, before concluding in Sec. 5.

## 2. Coarse-graining Markov jump processes

Our focus will be systems that can be modeled as discrete-state Markov jump processes. Their dynamics are captured by how the vector of probabilities  $\mathbf{p}(t) = \{p_1(t), \dots, p_n(t)\}$  over the states  $i = 1, \dots, n$  evolves with time  $t$  [32],

$$\dot{\mathbf{p}}(t) = \hat{W}\mathbf{p}(t), \quad (3)$$

where the off-diagonal elements of the transition rate matrix  $W_{ij}$  give the transition probabilities per unit time to jump from  $j \rightarrow i$  and the diagonal elements  $W_{ii} = -\sum_{j \neq i} W_{ji}$  are fixed by probability conservation. We will further assume that  $W_{ij} \neq 0$  only if  $W_{ji} \neq 0$ , which as we will see guarantees a well-defined entropy production rate [1].

The formal solution of (3) with initial condition  $\mathbf{p}(0)$  is given by the matrix exponential

$$\mathbf{p}(t) = e^{\hat{W}t}\mathbf{p}(0). \quad (4)$$

We will assume that our dynamics are ergodic such that (4) tends to a unique steady-state distribution  $\boldsymbol{\pi} = \{\pi_1, \dots, \pi_n\}$  given as the solution of  $\hat{W}\boldsymbol{\pi} = 0$ .

Our interest lies in the thermodynamics of the steady state. For such Markov processes, the entropy production rate  $\Sigma$  can be readily obtained (say by direct evaluation of (1)) [33],

$$\Sigma = \sum_{i \neq j} W_{ij}\pi_j \ln \frac{W_{ij}}{W_{ji}}. \quad (5)$$

The link to energy dissipation can be made once we have identified the reservoirs coupled to the system that supply the energy (and particles) required for transitions between the

states. Formally, this is implemented by assuming that the transition rates satisfy a local detailed balance condition [33]: for example, if the transition from  $k \rightarrow l$  is mediated by a single thermal reservoir at inverse temperature  $\beta$ , then we require  $\ln(W_{kl}/W_{lk}) = \beta q_{kl}$ , where  $q_{kl}$  is the heat flow into the reservoir accompanying the transition.

Now, a natural and analytically tractable method for coarse-graining continuous-time dynamics are stroboscopic observations every  $\tau$ . The output is a collection of  $N$  observations  $\gamma_N^\tau = \{i_1, \dots, i_N\}$ . Such stroboscopic measurements are attractive because the resulting coarse-grained dynamics is a discrete-time Markov chain with the same steady-state distribution  $\boldsymbol{\pi}$ . The transition probabilities  $M_{ij}(\tau)$  are given as the solution of [3] with Kronecker delta initial condition  $p_q(0) = \delta_{qj}$ :

$$\hat{M}(\tau) = e^{\hat{W}\tau}. \quad (6)$$

For such observed dynamics, we can readily evaluate the relative entropy rate per unit time in Eq. (2) to obtain the observed irreversibility. Note that the probability to observe the trajectory  $\gamma_N^\tau = \{i_1, \dots, i_N\}$  beginning in the steady-state is determined by the transition probabilities via

$$P(i_1, \dots, i_N) = M_{i_N i_{N-1}} \cdots M_{i_2 i_1} \pi_{i_1}. \quad (7)$$

Thus, comparing the forward and reverse trajectory  $\tilde{\gamma}_N^\tau = \{i_N, \dots, i_1\}$ , we find that

$$\sigma(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N\tau} \sum_{i_1, \dots, i_N} M_{i_N i_{N-1}} \cdots M_{i_2 i_1} \pi_{i_1} \ln \frac{M_{i_N i_{N-1}} \cdots M_{i_2 i_1} \pi_{i_1}}{M_{i_1 i_2} \cdots M_{i_{N-1} i_N} \pi_{i_N}} \quad (8)$$

$$= \lim_{N \rightarrow \infty} \frac{1}{N\tau} \left[ (N-1) \sum_{i,j} M_{ij} \pi_j \ln \frac{M_{ij}}{M_{ji}} + \sum_{i,j} M_{ij} \pi_j \ln \frac{\pi_j}{\pi_i} \right]. \quad (9)$$

Upon taking the limit, and noting that only transitions between different states contribute, we arrive at

$$\sigma(\tau) = \frac{1}{\tau} \sum_{i \neq j} M_{ij} \pi_j \ln \frac{M_{ij}}{M_{ji}}. \quad (10)$$

Before proceeding to our analysis of the observed irreversibility, it is worth noting a few algebraic properties of the transition rate matrix that will be helpful in our theoretical calculations [32]. First,  $\hat{W}$  need not be symmetric, which implies that each eigenvalue  $\lambda^\alpha$  can have associated with it a distinct right  $\mathbf{u}^\alpha$  and left  $\mathbf{v}^\alpha$  eigenvector given as the solutions of

$$\hat{W} \mathbf{u}^\alpha = -\lambda^\alpha \mathbf{u}^\alpha \quad (11)$$

$$\mathbf{v}^{\alpha,T} \hat{W} = -\lambda^\alpha \mathbf{v}^{\alpha,T}. \quad (12)$$

We have included a minus sign for later convenience in order that all eigenvalues have a non-negative real part [32]. While the eigenvectors need not exist, for simplicity we assume that they do and that they form a complete basis: that is, the eigenvectors can be chosen to satisfy  $\mathbf{u}^\alpha \cdot \mathbf{v}^\beta = \delta^{\alpha\beta}$  such that they verify the completeness relation  $\sum_\alpha \mathbf{u}^\alpha \mathbf{v}^{\alpha,T} = \hat{I}$ , with  $\hat{I}$  being the identity matrix. Furthermore, we choose to label the eigenvalue of  $\hat{W}$  corresponding to the steady state distribution as  $\lambda^0 = 0$ . The

remaining  $n - 1$  eigenvalues can be ordered using their real part from smallest to largest  $\lambda^0 \leq \lambda^1 \leq \dots \leq \lambda^{n-1}$ . A useful consequence of these assumptions is that the transition matrix  $\hat{M}(\tau)$  can be decomposed as

$$\hat{M}(\tau) = \boldsymbol{\pi} \mathbf{1}^T + \sum_{\alpha=1}^{n-1} e^{-\lambda^\alpha \tau} \mathbf{u}^\alpha \mathbf{v}^{\alpha,T}. \quad (13)$$

### 3. Observed irreversibility with a dissipative timescale

#### 3.1. Model with single dissipative transition

To test our hypothesis that the observed irreversibility contains information about the dissipative timescales, we consider models containing a single dissipative transition with rate  $\nu$ .

With this in mind, our transition rate matrix can be divided into two parts, an undriven or equilibrium part  $\hat{W}^{\text{eq}}$  and a driven or nonequilibrium part  $\hat{W}^{\text{neq}}$ ,

$$\hat{W} = \hat{W}^{\text{eq}} + \nu \hat{W}^{\text{neq}}, \quad (14)$$

with  $\nu$  setting the relative scale. We choose the equilibrium transition rates to have an Arrhenius form,

$$W_{ij}^{\text{eq}} = e^{-(B_{ij} - E_j)} \quad (15)$$

with symmetric ‘energy barriers’  $B_{ij} = B_{ji}$  and ‘state energies’  $E_j$ . If  $\hat{W} = \hat{W}^{\text{eq}}$  ( $\nu = 0$ ), then the system would relax to a steady state characterized by an equilibrium Gibbs distribution  $\pi_j^{\text{eq}} \propto e^{-E_j}$  with weights determined by the dimensionless energies  $E_j$ . It is for this reason that we say  $\hat{W}^{\text{eq}}$  describes undriven or equilibrium-like dynamics. Every transition will be an equilibrium transition except for transitions between  $1 \leftrightarrow 2$  ( $W_{12}^{\text{eq}} = W_{21}^{\text{eq}} = 0$ ), which we will drive out of equilibrium. Thus, we choose all the off-diagonal elements of  $\hat{W}^{\text{neq}}$  to be zero, apart from those entries corresponding to the transition  $1 \leftrightarrow 2$ . Here, driving is implemented by adding an asymmetric force  $F$  to the Arrhenius form,

$$W_{12}^{\text{neq}} = e^{-(B^{\text{neq}} - E_2 - F/2)} \quad (16)$$

$$W_{21}^{\text{neq}} = e^{-(B^{\text{neq}} - E_1 + F/2)}, \quad (17)$$

with a barrier  $B^{\text{neq}}$ .

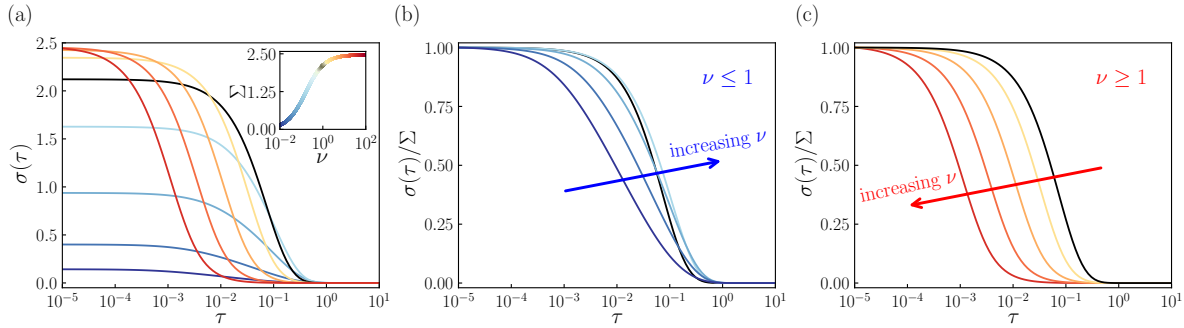
By construction, the sole source of nonequilibrium driving comes from the asymmetry of the driving force in  $\hat{W}^{\text{neq}}$ . This effect becomes explicit when calculating the entropy production rate [\(5\)](#) for this model

$$\Sigma = \nu(W_{21}^{\text{neq}}\pi_1 - W_{12}^{\text{neq}}\pi_2)F = J_{12}F, \quad (18)$$

where we have introduced the probability flux between states 1 and 2,  $J_{12} = \nu(W_{21}^{\text{neq}}\pi_1 - W_{12}^{\text{neq}}\pi_2)$ . Thus, nonzero  $F$  is required to have nonequilibrium systems in steady state.

### 3.2. Numerical analysis

For our system with a single driven transition, we can determine closed-form expressions for the transition matrix  $\hat{M}(\tau)$ , and from (10) the observed irreversibility  $\sigma(\tau)$ . To gain insight, we focus on a system with  $n = 4$  states and make the simplifying choice of setting all  $E_j = B_{ij} = B^{\text{neq}} = 0$ . This assumption sets the timescale  $W_{ij}^{\text{eq}} = 1$  against which the rates  $W_{12}^{\text{neq}} = (W_{21}^{\text{neq}})^{-1} = e^{F/2}$  are compared. The resulting expressions for  $\sigma(\tau)$  are plotted in figure 1 for a representative value of the driving force ( $F = 5$ ) as a function of timescale  $\nu$  and coarse-graining scale  $\tau$ . In figure 1(a), we see that for every value of the timescale  $\nu$ , as we vary  $\tau$  the observed irreversibility displays a cross over from the steady-state entropy production rate  $\Sigma$  (the left limit) to zero. Moreover, the shape of this crossover is sensitive to the value of  $\nu$ . For  $\nu \leq 1$  (blue tones), the dissipative transition rate limits  $\Sigma$  (inset). As  $\nu$  increases, not only does the value of  $\Sigma$  increase, but the knee of  $\sigma(\tau)$  shifts in location and steepness. However, when  $\nu \geq 1$ , we see that  $\Sigma$  is no longer limited by  $\nu$ , rather by the driving force (inset). As we continue to increase  $\nu$  in this regime (red tones), the dominant effect on  $\sigma(\tau)$  is to simply shift the knee leftward, while maintaining the same relative shape. These observations become more pronounced in figures 1(b)&(c), where we normalize the observed irreversibility by the entropy production rate,  $\sigma(\tau)/\Sigma$ . Such normalization highlights the qualitative change in the  $\nu$  dependence above and below  $\nu = 1$ .



**Figure 1.** (a) Observed irreversibility  $\sigma(\tau)$  as a function of coarse-graining scale  $\tau$  for various values of the dissipative transition rate, from  $\nu = 10^{-2}$  (blue) to  $\nu = 10^2$  (red) in increments of  $10^{0.5}$ , color coded along the temperature spectrum. Inset: steady-state entropy production rate  $\Sigma$  as a function of driven transition rate  $\nu$ . (b) Normalized observed irreversibility  $\sigma(\tau)/\Sigma$  as a function of coarse-graining scale  $\tau$  for slow driven transitions  $\nu \leq 1$ . (c) Normalized observed irreversibility  $\sigma(\tau)/\Sigma$  as a function of coarse-graining scale  $\tau$  for fast driven transitions  $\nu \geq 1$ . Parameters:  $n = 4$ ,  $E_j = 0$ ,  $B_{ij} = 0$ ,  $B^{\text{neq}} = 0$ , and  $F = 5$ .

This analysis reveals that even for this relatively simple model, the shape of  $\sigma(\tau)$  is sensitive to the interplay between the timescale  $\nu$  and the strength of thermodynamic driving  $F$ . The exception is when the driven transitions occur on a fast enough timescale (here  $\nu > 1$ ), such that the driving force is the limiting factor determining the entropy production rate. In such a limit, the shape of  $\sigma(\tau)$  does not vary with timescale  $\nu$ , only

the location of the knee.

### 3.3. Functional form of the observed irreversibility

When the dissipative transition rate is large, we have observed a rather stable shape of the observed irreversibility. To understand this simple dependence of  $\sigma(\tau)$  on  $\nu$  we will now analyze the observed irreversibility perturbatively for  $\nu \gg 1$ . An additional goal of this section will be finding a sufficiently generic functional form for  $\sigma(\tau)$  that can be leveraged in Sec. [4](#) as a data-fitting method to estimate the dissipative timescale.

To this end, we will need to approximate  $\hat{M}(\tau)$  for fast driven transitions, which is facilitated by looking at the four eigenvalues of the transition rate matrix  $\hat{W}$  in our 4-state model with  $\nu$  and  $F$  the only nonzero parameters:

$$\lambda^0 = 0 \tag{19}$$

$$\lambda^1 = \lambda^2 = 4 \tag{20}$$

$$\lambda^3 = 2 + 2\nu \cosh(F/2). \tag{21}$$

Clearly, for  $\nu \gg 1$ , one eigenvalue is much larger than the others and scales with the driven transition rate,  $\lambda^3 \sim \nu$ . Thus, at least for times up to the dissipative timescale  $\tau \sim 1/\nu$  the dynamics will be dominated by the largest eigenvalue. We can now approximate  $\hat{M}(\tau)$  and the observed irreversibility on this timescale assuming a single large dominant eigenvalue. This approximation can be done in some generality, so that our theoretical analysis will apply beyond the simple model analyzed in the previous subsection.

We begin the derivation by considering a transition rate matrix  $\hat{W}$  whose largest eigenvalue is much larger than all other eigenvalues,  $\lambda \equiv \lambda^{n-1} \gg \lambda^\alpha$ . For notational convenience, we also introduce the projector  $\hat{P} = \mathbf{u}^{n-1} \mathbf{v}^{n-1,T}$  onto the eigenspace of  $\lambda$ . Then for times  $\tau \lesssim 1/\lambda$ , we can approximate [\(13\)](#) as

$$\hat{M}(\tau) \approx \boldsymbol{\pi} \mathbf{1}^T + \sum_{\alpha=1}^{n-2} (1 - \lambda^\alpha \tau) \mathbf{u}^\alpha \mathbf{v}^{\alpha,T} + e^{-\lambda \tau} \hat{P}. \tag{22}$$

Substituting in the completeness relation and the eigendecomposition of  $\hat{W}$ , allows us to arrive at the suggestive form

$$\hat{M}(\tau) \approx \hat{I} + \tau \hat{W} + (e^{-\lambda \tau} - 1 + \lambda \tau) \hat{P}. \tag{23}$$

Using this approximation for the transition matrix, we can now approximate the observed irreversibility as,

$$\sigma(\tau) \approx \frac{1}{\tau} \sum_{i \neq j} (\tau W_{ij} + (e^{-\lambda \tau} - 1 + \lambda \tau) P_{ij}) \pi_j \ln \left( \frac{\tau W_{ij} + (e^{-\lambda \tau} - 1 + \lambda \tau) P_{ij}}{\tau W_{ji} + (e^{-\lambda \tau} - 1 + \lambda \tau) P_{ji}} \right). \tag{24}$$

While this is an accurate approximation in our region of interest, it remains a complicated function of the elements of the transition rate matrix, making [\(24\)](#) unsuitable for inference. We can simplify this expression further by noting that



$(e^{-\lambda\tau} - 1 + \lambda\tau) \ll \lambda\tau$  when  $\tau \lesssim 1/\lambda$ . This allows us to expand the logarithm to second order, giving rise to

$$\sigma(\tau) \approx \Sigma + a \left( \frac{e^{-\lambda\tau} - 1 + \lambda\tau}{\tau} \right) + b \left( \frac{e^{-\lambda\tau} - 1 + \lambda\tau}{\tau} \right)^2 \quad (25)$$

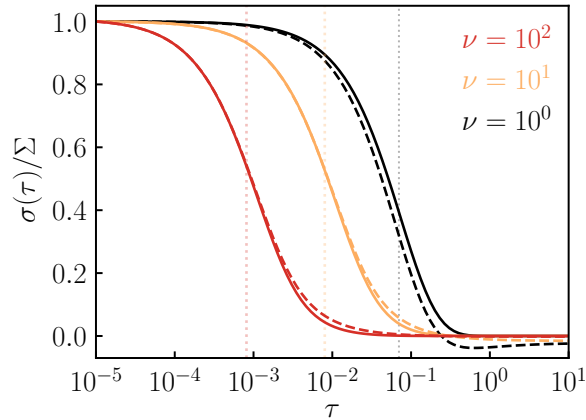
where

$$a = \sum_{i \neq j} W_{ij} \pi_j \left( \frac{P_{ij}}{W_{ij}} - \frac{P_{ji}}{W_{ji}} \right) + P_{ij} \pi_j \ln \frac{W_{ij}}{W_{ji}} \quad (26)$$

$$b = \sum_{i \neq j} \frac{1}{2} W_{ij} \pi_j \left[ \left( \frac{P_{ji}}{W_{ji}} \right)^2 - \left( \frac{P_{ij}}{W_{ij}} \right)^2 \right] + P_{ij} \pi_j \left( \frac{P_{ij}}{W_{ij}} - \frac{P_{ji}}{W_{ji}} \right). \quad (27)$$

The resulting expression (25) has the advantage of making the dependence of  $\sigma(\tau)$  on  $\lambda$  and  $\tau$  explicit, and hides all the details of the elements of the transition rate matrix in two multiplicative constants,  $a$  and  $b$ .

To test the accuracy of our approximation, we plot the normalized observed irreversibility  $\sigma(\tau)/\Sigma$  in figure 2 for  $\nu \geq 1$  (solid) alongside our prediction in Eq. (25) (dashed). As anticipated, our approximation is fairly accurate for times  $\tau \lesssim 1/\lambda$  and improves as  $\lambda \sim \nu$  increases.



**Figure 2.** Normalized observed irreversibility (solid) compared to the normalization of our approximation in Eq. (25) (dashed) for  $\nu = \{10^0, 10^1, 10^2\}$  (colors are the same as in figure 1). Vertical dashed lines are  $\tau = 1/\lambda$  for the respective  $\hat{W}$  matrices. Parameters:  $n = 4$ ,  $B_{ij} = 0$ ,  $B^{\text{neq}} = 0$ ,  $E_j = 0$ , and  $F = 5$ .

We see that when there is a strong timescale separation between the equilibrium and driven transitions a simple story emerges for the observed irreversibility: as we coarse grain past the timescale of the dissipative transition, the shape of the observed irreversibility takes on a simple form whose scale is controlled by the largest eigenvalue of the transition rate matrix. We can thus identify in these scenarios the (inverse) eigenvalue  $1/\lambda$  as the dissipative timescale: observations faster than  $1/\lambda$  appear irreversible, while observations slower appear reversible.

#### 4. Measuring the dissipative timescale

As mentioned in the introduction, part of our motivation for this study was the experiments conducted in Ref. [15] where the observed irreversibility was measured. Naturally, this raises the question of whether the analysis in the previous section can be used as a means to extract the dissipative timescale, that is the dominant eigenvalue  $\lambda$  from experimental data.

##### 4.1. Estimates using a fit ansatz

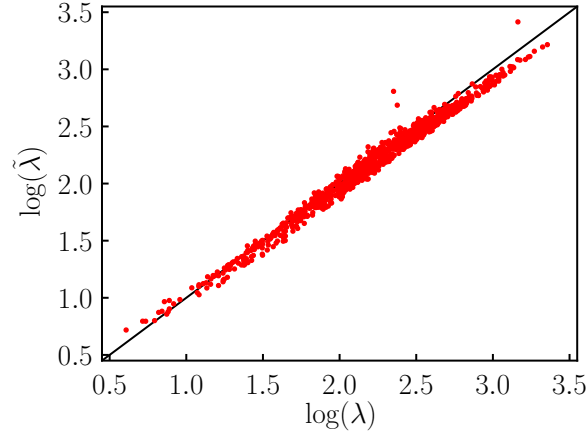
The rather simple functional form for the observed irreversibility derived in Eq. (25) suggests a method to infer  $\lambda$  from measured values of the normalized observed irreversibility  $\sigma(\tau)/\Sigma$ . Namely, we use (25) to inform a fitting ansatz

$$f(\tau) = 1 + \tilde{a} \left( \frac{e^{-\tilde{\lambda}\tau} - 1 + \tilde{\lambda}\tau}{\tilde{\lambda}\tau} \right) + \tilde{b} \left( \frac{e^{-\tilde{\lambda}\tau} - 1 + \tilde{\lambda}\tau}{\tilde{\lambda}\tau} \right)^2, \quad (28)$$

with fitting parameters  $\tilde{a}, \tilde{b}$  and  $\tilde{\lambda}$ . We can then generate an estimate for the dominant eigenvalue  $\tilde{\lambda}$  by fitting (28) to measured values of the normalized observed irreversibility.

To test the accuracy of this approach, we consider the case where measured values of the observed irreversibility are unbiased and have infinite precision. To this end, we generate  $10^3$  random transition rate matrices according to the model in Eqs. (14)-(17), where the  $E_j$ ,  $B_{ij}$  and  $B^{\text{neq}}$  are sampled uniformly in the range  $[0, 1]$ , the driving force  $F$  is sampled uniformly on  $[0, 5]$ , and the timescale  $\nu$  is sampled from a log-base 10 uniform distribution from  $[1, 10^2]$ . For each transition rate matrix we evaluate  $\sigma(\tau)$  numerically at values uniformly-spaced on a log-scale over the range  $\tau \in [\tau_m, \tau_M] = [10^{-5}, 10^1]$ . This forms the data that we use to fit our ansatz  $f(\tau)$  (28) using a nonlinear least-squares fit function. We plot the estimated  $\tilde{\lambda}$  versus the true eigenvalue  $\lambda$  for the  $10^3$  random transition rate matrices in figure 3.

Clearly our estimates correlate well with the true value as the cluster of points lie near the diagonal. However, there is a clear bias in our estimate that is uniform on a log scale, as the estimates tend to fall just below the diagonal. Indeed, we can estimate this bias by determining the (normalized) average deviation over our random samples:  $\langle (\lambda - \tilde{\lambda})/\lambda \rangle \approx 8\%$ . This bias is likely due to the fact that we are fitting a highly nonlinear function of  $\lambda$ , and such nonlinearities are known to lead to biased estimates [34]. Nevertheless, we can conclude in this highly-dissipative regime that fitting the observed irreversibility to our ansatz  $f(\tau)$  does allow one to reasonably well estimate the order of magnitude of the dissipative timescale, if not the true value.



**Figure 3.** Estimate of the dominant eigenvalue  $\tilde{\lambda}$  versus the true dominant eigenvalue  $\lambda$  (red dots) obtained by fitting  $f(\tau)$  (28) to the normalized observed irreversibility  $\sigma(\tau)/\Sigma$  for  $10^3$  randomly generated transition rate matrices as described in the text. The black line is a guide to the eye and represents perfect estimation.

#### 4.2. Implications of finite data

Numerically or experimentally generated time-series data is finite, which can result in bias and noise in the estimates of the observed irreversibility. In this section, we assess how these errors due to a finite sample size propagate to our estimate of  $\tilde{\lambda}$ .

To perform this assessment, we numerically generate continuous-time trajectories  $\gamma_T$  of a finite duration  $T$  for random transition rate matrices using the Gillespie algorithm [35]. We then coarse grained these trajectories by observing the system state every  $\tau$ , forming the coarse-grained trajectory  $\gamma_N^\tau = \{i_1, i_2, \dots, i_N\}$ . From each such coarse-grained trajectory, we estimate the observed irreversibility using a plug-in estimator [10, 11]. Namely, we directly estimate the steady-state distribution as well as the transition probabilities by counting the number of visits to each state and the number of each type of transition,

$$\tilde{\pi}_j = \frac{1}{N} \sum_{\alpha=1}^N \delta_{ji_\alpha} \quad (29)$$

$$\tilde{M}_{jk} = \frac{1}{N-1} \sum_{\alpha=1}^{N-1} \delta_{ji_{\alpha+1}} \delta_{ki_\alpha}. \quad (30)$$

These estimates are then plugged into (10) to generate our estimate of the observed irreversibility

$$\tilde{\sigma}(\tau) = \frac{1}{\tau} \sum_{i \neq j} \tilde{M}_{ij} \tilde{\pi}_j \ln \frac{\tilde{M}_{ij}}{\tilde{M}_{ji}}. \quad (31)$$

While more sophisticated methods are available to estimate the observed irreversibility [10, 16, 19, 24, 25], we chose this method for its computational simplicity. Moreover, our concern here is not with how to accurately estimate the observed irreversibility, but

how errors, which are inevitable in any method, propagate to estimates of the dissipative timescale. The estimated  $\tilde{\sigma}(\tau)$  is normalized using  $\tilde{\sigma}(\tau_m)$ , our estimate at the smallest observation time  $\tau$  considered ( $\tau_m = 10^{-5}$ ), and then fit by  $f(\tau)$  (28) to generate an estimate of the timescale  $\tilde{\lambda}$ .

The results for  $10^2$  random transition rate matrices are presented in figure 4 using the same ensemble described above in Sec. 4.1 except with fixed  $\nu = 10^2$ . For each rate matrix a trajectory of length  $T = 10^7$  was generated from which  $\tilde{\lambda}$  was estimated. We then repeated the procedure on the same trajectory, but instead taking the first  $T = 10^6$  of the trajectory, then  $T = \{10^5, 10^4\}$ , allowing us to visualize how reducing the amount of data affects our estimates. Figure 4(a) displays the normalized relative entropy rates for a representative trajectory, displaying almost no discernible disagreement on this scale. Indeed, the expected mean deviation

$$\langle \Delta \rangle = \left\langle \frac{\sum_{i=1}^N |\sigma(\tau_i) - \tilde{\sigma}(\tau_i)|}{\sum_{i=1}^N \sigma(\tau_i)} \right\rangle \quad (32)$$

of the observed irreversibilities remain below 6.3% (Table 1).

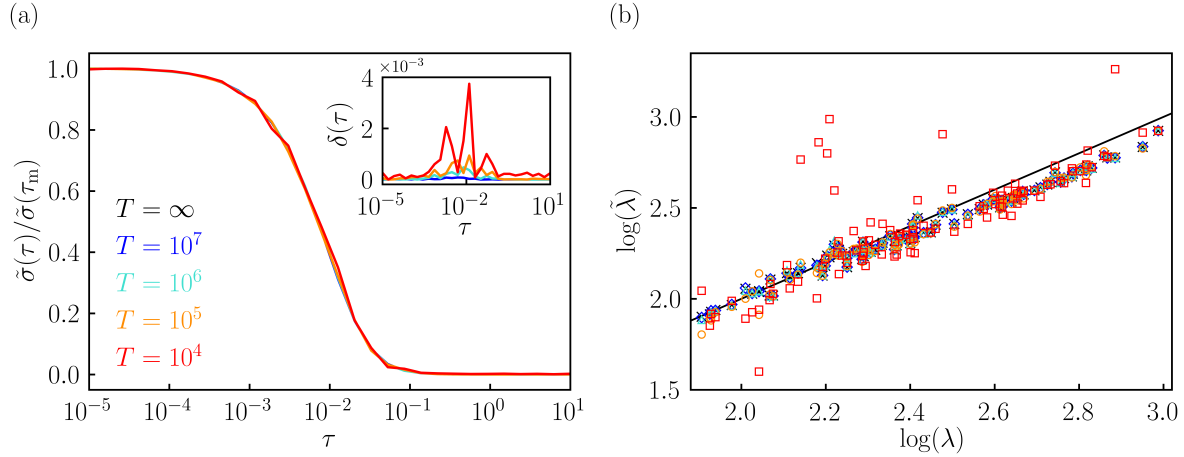
**Table 1.** Percent error in the estimates of the observed irreversibility  $\tilde{\sigma}(\tau)$  and the dissipative time scale  $\tilde{\lambda}$  for various observation times  $T$ .

$T$	$10^4$	$10^5$	$10^6$	$10^7$	$\infty$
$\langle \Delta \rangle$	6.3%	1.9%	0.6%	0.2%	0.0%
$\langle  \lambda - \tilde{\lambda}  / \lambda \rangle$	35%	13%	12%	12%	12%

In figure 4(b) estimates of  $\tilde{\lambda}$  are compared for various observation times  $T = \{10^4, 10^5, 10^6, 10^7\}$ . In addition, we include estimates from the noiseless data studied in the previous subsection Sec. 4.1, which in the present language corresponds to a trajectory of infinite length  $T = \infty$ . For observation times  $T \geq 10^5$ , the finite sample estimates are nearly as accurate as having perfectly accurate estimates of the observed irreversibility. Noise appears detrimental only for the shortest studied observation time  $T = 10^4$ , reaching 35%, about 3 times larger than our unavoidable error floor of  $\approx 12\%$  due to our nonlinear fitting method. The potential origin of this increased error can be seen in the inset of figure 4(a) where we plot the absolute error,

$$\delta(\tau) = \left| \frac{\sigma(\tau)}{\Sigma} - \frac{\tilde{\sigma}(\tau)}{\tilde{\sigma}(\tau_m)} \right| \quad (33)$$

as a function of  $\tau$  for various values of  $T$ . This error is concentrated near the cross-over where  $\tau \sim 1/\lambda$  with additional deviations for long coarse-graining times where the number of samples  $N = T/\tau$  is least. This suggests that accurate estimation of the knee region of the observed irreversibility is required when implementing our proposed estimation method, and even small deviations can adversely affect the estimate of the dissipative timescale.



**Figure 4.** (a) Normalized observed irreversibility estimate  $\tilde{\sigma}(\tau)/\tilde{\sigma}(\tau_m)$  as a function of  $\tau$  for various observation times  $T$  for a random transition rate matrix. We have included  $\sigma(\tau)/\Sigma$  (black) to illustrate the convergence of  $\tilde{\sigma}(\tau)/\tilde{\sigma}(\tau_m)$ . Inset: The corresponding absolute error  $\delta(\tau)$  as a function of  $\tau$ . (b) Estimates of  $\tilde{\lambda}$  obtained by fitting  $f(\tau)$  (28) to the normalized observed irreversibility estimates  $\tilde{\sigma}(\tau)/\tilde{\sigma}(\tau_m)$  for  $10^2$  randomly generated transition rate matrices. The diagonal black line is a guide to the eye and represents perfect estimation.

## 5. Conclusion

Under coarse-graining the observed irreversibility  $\sigma(\tau)$  displays a characteristic cross-over from its steady state value  $\Sigma$  to zero. When there is strong timescale separation between the driven transition and the undriven transitions, this profile takes on a simple shape with a single characteristic timescale. Using a perturbative approach we derived an approximate functional form and identified this dissipative scale with the largest eigenvalue  $\lambda$  of the transition rate matrix. Above this scale irreversibility is observable, whereas below it irreversibility is hidden. General theoretical analysis that goes beyond the fast driven limit remains an open question, complicated by our observation that the functional form of the observed irreversibility can vary as we change the transition speed and driving.

Building on our theoretical analysis, we proposed a fitting ansatz that allows us to estimate the dissipative timescale  $\lambda$  from time-series data. We observed that this method reasonably well reproduces the magnitude of  $\lambda$ , but remains biased due to the nonlinear fitting method. The error of our estimates is further sensitive to any error in the estimates of the observed irreversibility when data is finite. Accurate estimates in the cross-over region are necessary to most effectively use our proposed estimator. More refined fitting forms that account for a wider range of scales may be less affected by errors in the relative entropy determination, but this remains for future research.

## Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 2142466. N.F. and J.M.H. acknowledge the KITP program Active20 supported by National Science Foundation through Grant No. PHY-1748958. N. F. acknowledges National Science Foundation CAREER Grant No. PHYS-1848247.

## References

- [1] Seifert U 2012 *Rep. Prog. Phys.* **75** 126001
- [2] Chaikin P M and Lubensky T C 2000 *Principles of condensed matter physics* (Cambridge University Press, Cambridge)
- [3] Kubo R, Toda M and Hashitsume N 1985 *Statistical Physics II: Nonequilibrium Statistical Mechanics* (Springer-Verlag, Berlin)
- [4] Maes C and Netočný K 2003 *J. Stat. Phys.* **110** 269–310
- [5] Gaspard P 2004 *J. Stat. Phys.* **117** 599–615
- [6] Jarzynski C 2006 *Phys. Rev. E* **73** 046105
- [7] Kawai R, Parrondo J M R and Van den Broeck C 2007 *Phys. Rev. Lett.* **98** 080602
- [8] Horowitz J M and Jarzynski C 2009 *Phys. Rev. E* **79** 021106
- [9] Parrondo J M R, Van den Broeck C and Kawai R 2009 *New J. Phys.* **11** 073008
- [10] Roldán E and Parrondo J M R 2010 *Phys. Rev. Lett.* **105** 150607
- [11] Roldán E and Parrondo J M R 2012 *Phys. Rev. E* **85** 031129
- [12] Muy S, Kundu A and Lacoste D 2013 *J. Chem. Phys.* **139** 124109
- [13] Tusch S, Kundu A, Verley G, Blondel T, Miralles V, Démoulin D, Lacoste D and Baudry J 2014 *Phys. Rev. Lett.* **112** 180604
- [14] Martínez I A, Bisker G, Horowitz J M and Parrondo J M R 2019 *Nat. Comm.* **10** 1–10
- [15] Tan T H, Watson G A, Gingrich T R, Horowitz J M and Fakhri N Scale-dependent irreversibility in biological active matter arXiv:2107.05701
- [16] Skinner D J and Dunkel J 2021 *Proc. Natl. Acad. Sci. U.S.A.* **118** e2024300118
- [17] Gomez-Marin A, Parrondo J M R and Van den Broeck C 2008 *Europhys. Lett.* **82** 50002
- [18] Gomez-Marin A, Parrondo J M R and Van den Broeck C 2008 *Phys. Rev. E* **78** 011107
- [19] Roldán E, Barral J, Martin P, Parrondo J M R and J ulicher F 2021 *New J. Phys.* **23** 083013
- [20] Hartich D and Godec A 2022 Comment on "Inferring broken detailed balance in the absence of observable currents" arXiv:2112.08978v3
- [21] Bisker G, Martinez I A, Horowitz J M and Parrondo J M R 2022 Reply: Inferring entropy production from waiting time distributions arXiv:2202.02064
- [22] Cover T M and Thomas J A 2006 *Elements of Information Theory* 2nd ed (Wiley-Interscience, New York)
- [23] Lynn C W, Holmes C M, Bialek W and Schwab D J 2022 *Phys. Rev. Lett.* **129** 118101
- [24] Lynn C W, Holmes C M, Bialek W and Schwab D J 2022 *Phys. Rev. E* **106** 034102
- [25] Ro S, Guo B, Shih A, Phan T V, Austin R H, Levine D, Chaikin P M and Martiniani S 2022 *Phys. Rev. Lett.* **129** 220601
- [26] Seara D S, Machta B B and Murrell M P 2021 *Nat. Comm.* **12** 392
- [27] Harada T and Sasa S 2005 *Physical Review Letters* **95** 130602
- [28] Nardini C, Fodor E, Tjhung E, van Wijland F, Tailleur J and Cates M E 2017 *Phys. Rev. X* **7** 021007
- [29] Dadhichi L P, Maitra A and Ramaswaym S 2018 *J. Stat. Mech.* 123201
- [30] Lippiello E, Baiesi M and Sarracino A 2014 *Phys. Rev. Lett.* **112** 140602
- [31] Wang S W, Kawaguchi K, Sasa S i and Tang L H 2016 *Phys. Rev. Lett.* **117** 070601

- [32] Van Kampen N G 2007 *Stochastic Processes in Physics and Chemistry* 3rd ed (Elsevier Ltd., New York)
- [33] Esposito M and Van den Broeck C 2010 *Phys. Rev. E* **82** 011143
- [34] Gore J, Ritort F and Bustamante C 2003 *Proc. Natl. Acad. Sci. USA* **100** 12564–12569
- [35] Gillespie D T 1977 *The Journal of Physical Chemistry* **81** 2340–2361