Home   Search   Collections   Journals   About   Contact us   My IOPscience

Inferring dissipation from current fluctuations

You may also be interested in:

Local detailed balance: a microscopic derivation
M Bauer and F Cornu

Non-equilibrium steady states: fluctuations and large deviations of the density and of thecurrent
Bernard Derrida

First-order dynamical phase transition in models of glasses
Juan P Garrahan, Robert L Jack, Vivien Lecomte et al.

Stochastic thermodynamics, fluctuation theorems and molecular machines
Udo Seifert

BEST statistics of Markovian fluxes: a tale of Eulerian tours and Fermionic ghosts
Matteo Polettini

Dynamical fluctuations for semi-Markov processes
Christian Maes, Karel Netoný and Bram Wynants

Violation of the fluctuation–dissipation theorem in glassy systems: basic notions and the numerical evidence
A Crisanti and F Ritort

Large-deviation principles, stochastic effective actions, path entropies, and the structure and meaning of thermodynamic descriptions
Eric Smith

Joint probability distributions and fluctuation theorems
Reinaldo García-García, Vivien Lecomte, Alejandro B Kolton et al.

Inferring dissipation from current fluctuations

Todd R Gingrich[1,3]**, Grant M Rotskoff**[2] **and**
Jordan M Horowitz[1]

[1] Physics of Living Systems Group, Department of Physics, Massachusetts Institute of Technology, 400 Technology Square, Cambridge, MA 02139, United States of America
[2] Biophysics Graduate Group, University of California, Berkeley, CA 94720, United States of America

E-mail: toddging@mit.edu

Abstract

Complex physical dynamics can often be modeled as a Markov jump process between mesoscopic configurations. When jumps between mesoscopic states are mediated by thermodynamic reservoirs, the time-irreversibility of the jump process is a measure of the physical dissipation. We rederive a recently introduced inequality relating the dissipation rate to current fluctuations in jump processes. We then adapt these results to diffusion processes via a limiting procedure, reaffirming that diffusions saturate the inequality. Finally, we study the impact of spatial coarse-graining in a two-dimensional model with driven diffusion. By observing fluctuations in coarse-grained currents, it is possible to infer a lower bound on the total dissipation rate, including the dissipation associated with hidden dynamics. The tightness of this bound depends on how well the spatial coarse-graining detects dynamical events that are driven by large thermodynamic forces.

Keywords: entropy production, large deviation theory, coarse graining, nonequilibrium fluctuations

(Some figures may appear in colour only in the online journal)

1. Introduction

Experiments capable of probing molecular-scale dynamics have led to a wealth of data about the operation of nanoscale machines [1–3]. Like their macroscopic counterparts, nanomachines convert free energy from the environment into useful work. Due to molecular fluctuations, however, nanoscale machines behave stochastically. Though the stochasticity renders nanomachines less predictable, the fluctuations also offers a unique tool to deduce additional physical properties of the machine. Indeed, the fluctuation-dissipation theorem (FDT) demonstrates that near-equilibrium fluctuations allow one to deduce the nonequilibrium response to small perturbations, an idea that has become a central theme of statistical mechanics [4–9]. Analogous exploitation of far-from-equilibrium fluctuations is highly desirable [10–14], particularly for studying living systems [15–17].

In this work, we demonstrate how it is possible to extract information from nonequilibrium fluctuations under general steady-state conditions. Our central assumption is that the mesoscale dynamics, in a quantum or a classical system, can be described by a Markovian stochastic process. For example, the Markovian mesoscopic description applies to biochemical kinetics of enzymatic reactions. In this context, dynamical fluctuations have been used to constrain proposed models for enzymatic pathways, providing bounds on the number of distinct intermediate states [18, 19]. More recently, advances in the theory of Markov processes [20–23] have been used to relate the dynamical fluctuations to the physical dissipation rate of the nonequilibrium dynamics [24–27]. This paper reviews and extends that connection between fluctuations and dissipation, a connection which can be thought of as a generalization of the FDT. Where the FDT offers an equality—by measuring fluctuations we may determine the dissipation exactly [28]—the far-from-equilibrium analog guarantees an inequality: the extent of fluctuations sets a lower bound on the dissipation rate. Hence experimental observations of macroscopic, coarse-grained fluctuations provide enough information to estimate a bound on the dissipation rate of a process far from equilibrium.

In this paper, we study this generalization of the FDT bound from the perspective of coarse-graining, a viewpoint which we anticipate will complement experimental applications. After reviewing the formalism of Markov jump processes in section 2, we consider the consequence of temporal coarse-graining in section 3. By focusing on dynamical quantities which are time-averaged over long trajectories, we arrive at the so-called Level 2.5 large deviations, which quantify dynamical fluctuations in the jump process [20–23]. Mirroring our previous work [26], section 4 develops our central relationship between nonequilibrium dynamical fluctuations and dissipation. In its most general form, the relationship pertains to a fluctuating vector quantity in the space of all microscopic currents; the implication for scalar currents is detailed in section 5. In addition to the temporal coarse-graining, we investigate the role of spatial resolution by studying two limits: a diffusive limit in section 6 and the spatial coarse-graining of a particular diffusion process in section 7. A consequence of the coarse-graining is that the macrostate dynamics may not be Markovian. Remarkably, even without the Markov property, coarse-grained fluctuations can bound the dissipation rate so long as the macroscopic dynamics emerges from an underlying Markovian description, suggesting that suitable analysis of dynamical fluctuations can complement existing methods for deducing dissipation from time series data [29].

2. Markov jumps on graphs

2.1. Dissipation

Before studying the consequences of coarse-graining, we first build a detailed dynamical model. At the smallest scales, the microscopic laws of physics are Markovian and deterministic. By
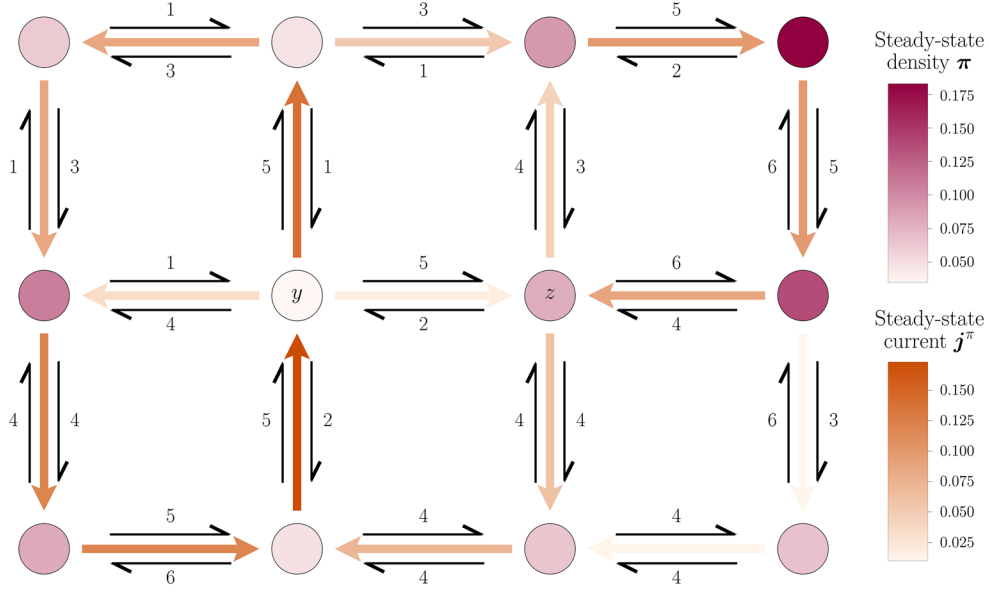
Figure 1. Example of a Markov jump process: mesostates are shaded according to the steady-state density π. Black arrows on the edges label the transition rates, and the colored edges indicate the steady-state current across each edge, j^π.

clustering microscopic configurations into sufficiently many mesostates, mesoscopic dynamics may often be described by a continuous-time Markov jump process on the set of coarse-grained states [30–32]. Such a model can be constructed, for example, from atomistic simulations, as in the case of protein conformational dynamics [33, 34]. The jump process can be viewed as a random walk on a graph, with each of the *N* vertices corresponding to one of the possible mesostates. Figure 1 depicts an example with twelve states. Any two vertices, *y* and *z*, are connected by an edge when the system can jump between *y* and *z*. We denote the rate for transitioning from *y* to *z* by $r(y, z)$ and require a non-vanishing rate $r(z, y)$ for the reverse transition.

The probability of occupying state *y* at time *t*, $p_t(y)$, evolves according to the master equation

$$\frac{\partial p_t(y)}{\partial t} = -\sum_{z \neq y} j^{p_t}(y, z), \tag{1}$$

where $j^p(y, z) = p(y) r(y, z) - p(z)r(z, y)$ is the current passing from *y* to *z*. Following Bertini *et al* [22, 23], the superscript *p* highlights that this current is a function of the density. At long times, the probability of occupying mesostate *y* approaches the steady-state value,

$$\pi(y) \equiv \lim_{t \to \infty} p_t(y). \tag{2}$$

Similarly, the current passing from state *y* to *z* approaches $j^\pi(y, z) = \pi(y)r(y, z) - \pi(z)r(z, y)$ in the steady state. We will use bold symbols to represent vectors and let context distinguish whether the vector is indexed over vertices (as in π) or edges (as in j^π and r).

An equilibrium system obeys detailed balance, meaning that all steady-state currents vanish, $j^\pi = 0$. Non-vanishing currents may be generated if the system couples to multiple external reservoirs with differing intensive parameters, e.g. temperatures, pressures, or chemical potentials. The environment, which we model as a set of infinitely-large reservoirs, can

therefore lead the dynamics to break detailed balance, i.e. it need not be the case that mesostates y and z satisfy $\pi(y)r(y,z) = \pi(z)r(z,y)$.

Because the mesoscopic Markov dynamics does not explicitly model the state of the thermodynamic reservoirs (how much energy and how many particles are in the baths), the impact of the reservoirs appears implicitly in the rates $\boldsymbol{r}$. As an example, suppose mesostate y has a (free) energy lower than that of mesostate z by an amount ΔE. The transfer of energy from a thermal reservoir at inverse temperature β can induce an uphill transition from y to z. If we assume the mesoscopic configurations are locally equilibrated with the thermal reservoir, then the transitions must obey detailed balance with respect to the reservoir's equilibrium distribution p_{eq} [35]. This local detailed balance condition, $p_{\mathrm{eq}}(y)r(y,z) = p_{\mathrm{eq}}(z)r(z,y)$, yields

$$\frac{r(y,z)}{r(z,y)} = \frac{p_{\mathrm{eq}}(z)}{p_{\mathrm{eq}}(y)} = \exp(-\beta\Delta E), \tag{3}$$

which relates the ratio of transition rates to the energy flux from the reservoir. These same arguments apply to other types of reservoirs, e.g. a particle reservoir at constant chemical potential. The ratio of transition rates, more generally, is expressed in terms of the entropy change in the reservoir, ΔS:

$$\frac{r(y,z)}{r(z,y)} = \exp(\Delta S). \tag{4}$$

Satisfying local detailed balance with the various reservoirs does not imply that the Markov dynamics obeys detailed balance. When coupled to multiple reservoirs with incompatible equilibrium states, the Markov dynamics breaks detailed balance, but equation (4) still relates the mesoscopic rates to the flow of entropy from the reservoirs. For density $\boldsymbol{p}$, we quantify the broken detailed balance by

$$F^p(y,z) = \ln\frac{p(y)r(y,z)}{p(z)r(z,y)}, \tag{5}$$

which we call the thermodynamic force because it measures how much free energy the baths must provide in order for the system to transition from y to z[4]. In the steady state, the rate of that free energy transfer is given by

$$\sigma^\pi(y,z) = j^\pi(y,z)F^\pi(y,z). \tag{6}$$

We call this quantity the dissipation rate for the edge connecting y and z. The total dissipation rate for the system is computed by summing the dissipation rates for all edges:

$$\Sigma^\pi = \boldsymbol{j}^\pi \cdot \boldsymbol{F}^\pi \equiv \sum_{y<z} j^\pi(y,z)F^\pi(y,z). \tag{7}$$

The sum runs over pairs of states with $y < z$ to avoid double counting.

Computing the dissipation rate via equation (7) is all-but-impossible for complicated problems. The computation requires that all possible mesostates be identified, that the rates for transitions between these mesostates be measured, and that the steady-state density be computed. Even the simplest of these tasks, identifying the set of mesostates, is frequently impractical. More commonly, it is only possible to monitor transitions between some coarse-grained macrostates. With this limitation in mind, we set out to infer the dissipation rate on the basis of fluctuations in finite-time stochastic trajectories. Our strategy merely provides a lower bound

---

[4] $F^p(y,z)$ is the difference between the entropy gain of the bath $\Delta S = \ln r(y,z)/r(z,y)$ and the gain in Shannon entropy of the system $\ln p(z)/p(y)$.

on the dissipation rate, but may be applied even when we only observe macrostate transitions, a point we return to in section 7.

2.2. Fluctuations

An infinitely long trajectory samples all configurations in proportion to the steady-state distribution, but a single finite-time trajectory has fluctuations. Consider one realization of the jump process, initialized in the steady state and observed for a long but finite time T. We let $x(t)$ denote the identity of the occupied mesostate at time t. Given this trajectory, an unbiased estimate of the steady-state density at mesostate y is found by measuring the fraction of time spent in y:

$$p(y) = \frac{1}{T} \int_0^T dt \ \delta_{x(t),y},$$

(8)

with Kronecker delta $\delta_{\alpha,\beta}$. This time-averaged density $\boldsymbol{p}$ is the *empirical density*. Despite similar notation, the empirical density should not be confused with the instantaneous probability of occupying state y at time t, which we have denoted $p_t(y)$. The instantaneous density $p_t(y)$ depends explicitly on time t, whereas the empirical density $p(y)$ depends on the timescale T over which the density was averaged.

Analogous to the empirical density, the *empirical current* from y to z counts the rate of transitions from y to z, less those from z to y:

$$j(y,z) \equiv \frac{1}{T} \int_0^T dt \ \delta_{x(t^-),y}\delta_{x(t^+),z} - \delta_{x(t^-),z}\delta_{x(t^+),y}.$$

(9)

The notation $x(t^\pm)$ is shorthand for the configuration immediately before or immediately after time t. Note that $j(y,z)$ differs from the density-dependent current $j^p(y,z)$. The former reflects the number of transitions observed in a stochastic trajectory while the latter is the current that one would expect given the empirical density $\boldsymbol{p}$ and the transition rates $\boldsymbol{r}$.

The probability distribution for the empirical density and current reflects the fluctuations anticipated in finite-time experiments or simulations. For large T, this distribution adopts the large deviation form

$$P(\boldsymbol{p},\boldsymbol{j}) \asymp e^{-TI(\boldsymbol{p},\boldsymbol{j})}$$

(10)

with the joint rate function $I(\boldsymbol{p},\boldsymbol{j})$ measuring, on an exponential scale, the chance of observing fluctuations away from steady-state density $\boldsymbol{\pi}$ and steady-state current $\boldsymbol{j^\pi}$ [36]. We use $\asymp$ to indicate asymptotic equivalence, meaning

$$-\frac{1}{T} \ln P(\boldsymbol{p},\boldsymbol{j}) = I(\boldsymbol{p},\boldsymbol{j}) + \mathcal{O}\left(\frac{1}{T}\right).$$

(11)

The rate function attains a minimum at the steady-state values ($I(\boldsymbol{\pi},\boldsymbol{j^\pi}) \equiv 0$) since the steady-state behavior dominates in the long-time limit. The local curvature of I around its minimum, the Hessian, reveals the variance of the empirical fluctuations away from the steady-state values. Remarkably, the local curvature can be computed analytically because an explicit expression for $I(\boldsymbol{p},\boldsymbol{j})$ is known [20–23].

3. Coarse-graining in time

The Markov jump process has rich temporal correlations—after a short time, a system initialized in mesostate y is more likely to be in a neighboring state z than the steady-state probability $\pi(z)$ would predict. Though these correlations add complexity to transient dynamics,

they die out after some finite correlation time. The empirical density and empirical current can be averaged over sufficiently long times T such that the transient dynamics becomes unimportant. In this way, $\boldsymbol{p}$ and $\boldsymbol{j}$ can be thought of as the result of temporal coarse-graining with a coarse-graining timescale exceeding natural correlation times. In the limit of large T, the empirical fluctuations can therefore be mimicked by Poisson point processes, which lack all temporal correlations. More specifically, the coarse-grained Markov jump process resembles the behavior of a collection of independent Poisson point processes, one per directed edge of the graph, which count the number of transition events for each edge in time T. The large deviation structure of Poisson processes are particularly simple, providing a heuristic route to $I(\boldsymbol{p},\boldsymbol{j})$ for a general nonequilibrium jump process. Rigorous derivations of this result utilize a technique called Cramér tilting [21]. Rather than recapitulate that argument, we illustrate how the form of the jump process rate function originates from a Poisson point process.

3.1. Flows

To explicitly demonstrate that the Markov jump process' long-time behavior resembles that of a collection of Poisson point processes, we first focus on a single type of dynamical event—hops from mesostate y to z in figure 1. We record the time of each such hop by a blue tick on a timeline in figure 2. The average density of ticks is

$$q(y,z) = \frac{1}{T} \int_0^T dt \, \delta_{x(t^-),y} \delta_{x(t^+),z}, \tag{12}$$

the empirical flow rate from y to z [22, 23]. Due to the stochastic nature of the trajectories, the spacing between tick marks is variable. Immediately following each $y \to z$ event, the system is localized in mesostate z, and the next $y \to z$ event cannot occur until the system finds its way back to mesostate y. This need to reset introduces temporal correlations, quantified by the pair-correlation function

$$g(t) = \left\langle \frac{\sum_{i<j} \delta \left(t - (t_j - t_i) \right)}{q(y,z)} \right\rangle, \tag{13}$$

with t_i denoting the time of the ith $y \to z$ hop. The average in the definition of $g(t)$ is taken over the ensemble of long trajectories of length T.

Akin to the pair correlation function of liquid state theory [37], $g(t)$ captures the probability that, given an event at time 0, another event occurs after a delay time t. The blue curve in figure 2 shows $g(t)$ for the flows from state y to z in figure 1. At short times, $g(t)$ exceeds one, indicating a propensity for bursty repeated events since the return rate $r(z,y)$ can allow multiple $y \to z$ events in rapid succession. After slightly longer times, $g(t)$ is less than one; trajectories which do not backtrack from z to y only enable the next event upon returning to y via a slower, circuitous route. Finally, $g(t)$ approaches 1 for $t > \tau_{\text{corr}}$, the timescale on which correlations are lost.

These temporal correlations are an important feature of a Markov jump process at short times. However, the two-time correlations encoded in $g(t)$ become insignificant at long times, for which $g(t) \to 1$, resembling a Poisson point process[5]. This loss of correlations suggests that the probability of observing Q transitions from y to z over a long time T should be asymptotically given by a Poisson distribution [30],

---

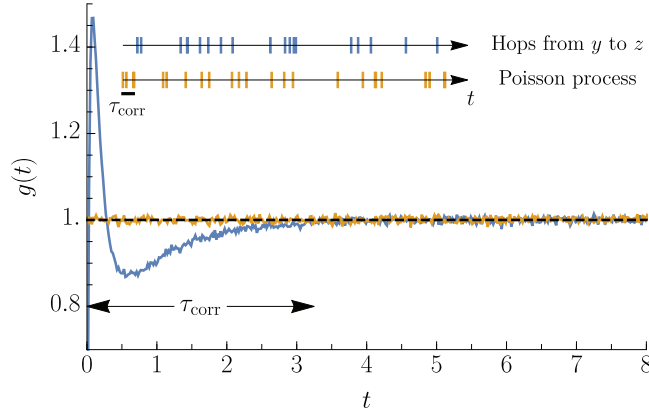[5] The Poisson point process lacks correlations, so $g(t) = 1$.

Figure 2. Pair correlation functions $g(t)$ for Markov dynamics on the graph in figure 1 (blue) and for an Poisson point process (orange). Time is reported in the same units as the inverse of the rate constants in figure 1. The colored lines are collected by sampling from 5000 trajectories, each of length 10 000; the dashed black line shows the exact result for a Poisson point process. The inset shows a timeline for the events observed in a representative trajectory, with each tick corresponding to the time of a hopping event. The Poisson point process is constructed so that it has the same average density of tick marks as the on the graph. For times much larger than a correlation time τ_{corr}, the hops become uncorrelated, so the fluctuations in the number of events asymptotically approaches that of the Poisson point process, equation (15).

$$P(Q) = \frac{\lambda^Q \mathrm{e}^{-\lambda}}{Q!}. \tag{14}$$

The Poisson parameter λ must be chosen to match the steady-state flows in the long time limit, $\langle Q \rangle = \pi(y)r(y,z)T$. We note that this condition can be met by choosing $\lambda = Tp(y)r(y,z)$ if we average over both empirical densities and flows. This observation is suggestive of a long-time Poisson form with an effective rate that depends on the empirical density $p(y)$:

$$P_{\mathrm{ind}}(p(y), q(y,z)) \asymp \frac{(Tp(y)r(y,z))^{Tq(y,z)} \mathrm{e}^{-Tp(y)r(y,z)}}{(Tq(y,z))!}, \tag{15}$$

with the subscript 'ind' denoting that we have considered the single edge as independent of the other edges. In reality, the statistics of neighboring edges are coupled by a conservation law: a trajectory must leave one state to enter another, so the empirical flow conserves probability at every mesostate[6], which requires

$$\sum_z (q(y,z) - q(z,y)) = 0 \ \forall y. \tag{16}$$

P_{ind} can be thought of as the effective single-edge distribution in the absence of the conservation law constraint. We stress that P_{ind} should not be confused with the marginal distribution for the single-edge statistics, which we will see is much more complicated.

The joint statistics of density and flow across all edges, however, can be simply expressed. This simplicity arises because the long-time flow fluctuations on the various edges are coupled

---

[6] The initial and final mesostates do not conserve probability—there is a source where the system starts at time zero and a sink where it ends at time T, but in the long time limit this effect is insignificant.

only by the constraint equation (16), which requires that the empirical flow q must conserve probability. When the empirical flow is conservative, the long-time joint probability of p and q can be written as a product over the independent edge probabilities P_{ind}. Any vector q that does not satisfy the conservation law, of course, has vanishing probability in the long-time limit. To write this claim in large deviation form we first rewrite P_{ind} in large deviation form, $P_{ind}(p(y), q(y,z)) \asymp e^{-T I_{ind}(p(y), q(y,z))}$, with the rate function

$$I_{ind}(p(y), q(y,z)) = p(y)r(y,z) - q(y,z) + q(y,z)\ln\frac{q(y,z)}{p(y)r(y,z)}. \tag{17}$$

The joint density and flow fluctuations are then given by

$$I(\boldsymbol{p}, \boldsymbol{q}) = \begin{cases} \sum_{y<z} I_{ind}(p(y), q(y,z)), & \boldsymbol{q} \text{ satisfies equation (16)} \\ \infty, & \text{otherwise.} \end{cases} \tag{18}$$

This form of $I(\boldsymbol{p}, \boldsymbol{q})$ reveals why P_{ind} is not the marginal distribution for single-edge flow statistics. The conservation law constraint complicates any attempts to simply express the marginal $P(p(y), q(y,z))$ by integrating out the other densities and flows.

Our heuristic arguments for the form of $I(\boldsymbol{p}, \boldsymbol{q})$ are suggestive, but they do not constitute a proof so much as a motivation. Equation (18), known in mathematics literature as the Level 2.5 large deviation function, can be proven using more sophisticated arguments that construct an effective process to generate rare densities and flows [21, 23]. While that construction offers rigor and alternative insight, we find our simple, heuristic derivation to be instructive since it clearly identifies the essential physics: we may discard temporal correlations in a complicated Markov jump process to obtain an asymptotically equivalent collection of Poisson point processes.

3.2. Currents

Suppose now that we are interested in the empirical currents rather than the flows. Unlike $q(y, z)$, $j(y, z)$ deducts the rate of reversed hops from z to y. The probability of measuring a current $j(y, z)$ is thus given by marginalizing over $q(y, z)$ and $q(z, y)$ with the constraint $j(y, z) = q(y, z) - q(z, y)$. As with the flows, currents across different edges are coupled by probability conservation:

$$\sum_z j(y,z) = 0 \ \forall y. \tag{19}$$

For large T, this marginalization of flows proceeds via a saddle point approximation [36] to yield $P(\boldsymbol{p}, \boldsymbol{j}) \asymp e^{-T I(\boldsymbol{p}, \boldsymbol{j})}$ with

$$I(\boldsymbol{p}, \boldsymbol{j}) = \begin{cases} \sum_{y<z} \Psi(p(y), p(z), j(y,z)), & \boldsymbol{j} \text{ satisfies equation (19)} \\ \infty, & \text{otherwise.} \end{cases} \tag{20}$$

The function Ψ follows from treating each edge independently. For example, integrating out $q(y, z)$ and $q(z, y)$ yields

$$\begin{aligned} \Psi(p(y), &p(z), j(y,z)) \\ &= \inf_{q(y,z)} I\big(p(y), q(y,z)\big) + I\big(p(z), q(y,z) - j(y,z)\big). \end{aligned} \tag{21}$$

The minimizer $q_\star(y,z)$ is the root of a quadratic,

$$q_\star(y,z) = \frac{1}{2}\left(j(y,z) + \sqrt{j(y,z)^2 + a^p(y,z)^2}\right), \qquad (22)$$

where $a^p(y,z) \equiv 2\sqrt{p(y)\,p(z)r(y,z)r(z,y)}$. Thus Ψ may be expressed in terms of $q_\star$ as

$$\begin{aligned}
&\Psi(p(y),p(z),j(y,z)) \\
&\quad = I\big(p(y),q_\star(y,z)\big) + I\big(p(z),q_\star(z,y)\big) \\
&\quad = \Bigg[p(y)r(y,z) + p(z)r(z,y) - q_\star(y,z) - q_\star(z,y) \\
&\qquad + q_\star(y,z)\ln\frac{q_\star(y,z)}{p(y)r(y,z)} + q_\star(z,y)\ln\frac{q_\star(z,y)}{p(z)r(z,y)} \Bigg].
\end{aligned} \qquad (23)$$

In appendix A we carry out straightforward algebraic manipulations to bring Ψ into the form of Bertini *et al* [23],

$$\Psi = \sqrt{j^p(y,z)^2 + a^p(y,z)^2} - \sqrt{j(y,z)^2 + a^p(y,z)^2} + j(y,z)\left(\mathrm{arcsinh}\frac{j(y,z)}{a^p(y,z)} - \mathrm{arcsinh}\frac{j^p(y,z)}{a^p(y,z)}\right). \qquad (24)$$

4. Rate function bound

To relate the fluctuations described by the rate function $I(\boldsymbol{p},\boldsymbol{j})$ to the thermodynamic dissipation, it is useful to write I in terms of the currents and thermodynamic forces F^p. For notational compactness we suppress the labels of the vertices. We note that $j^p/\sinh(F^p/2) = a^p \geqslant 0$, implying that j^p has the same sign as F^p. Furthermore, it is natural to measure the empirical current j relative to the current j^p, so we introduce $\bar{j} \equiv j/j^p$. With this new notation,

$$\Psi = j^p\left[\coth\frac{F^p}{2} - \frac{\bar{j}F^p}{2} + \bar{j}\,\mathrm{arcsinh}\left(\bar{j}\sinh\frac{F^p}{2}\right) - \sqrt{\bar{j}^2 + \mathrm{csch}^2\frac{F^p}{2}}\right]. \qquad (25)$$

Taylor expanding in powers of F^p, we obtain

$$\Psi = j^p\left[\frac{(\bar{j}-1)^2 F^p}{4} - \frac{(\bar{j}^2-1)^2(F^p)^3}{192} + \frac{(\bar{j}^2-1)^2(3\bar{j}^2+1)(F^p)^5}{7680} + \mathcal{O}((F^p)^7)\right]. \qquad (26)$$

The low-order partial sums of this series alternate, first overestimating, then underestimating Ψ (see figure 3). In particular, the first partial sum bounds the rate function by a quadratic [26],

$$\Psi \leqslant \Psi_{\mathrm{quad}} \equiv \frac{(\bar{j}-1)^2 \sigma^p}{4}, \qquad (27)$$

where $\sigma^p = j^p F^p$ is the local dissipation for the edge.

Truncation of a Taylor series, of course, does not necessarily yield a bound. To prove inequality (27) we must confirm the positivity of the residual $\Delta \equiv \Psi_{\mathrm{quad}} - \Psi$. Since Δ is symmetric about $\bar{j} = 0$, it suffices to consider positive $\bar{j}$. We note that Δ vanishes when $\bar{j} = 1$, and
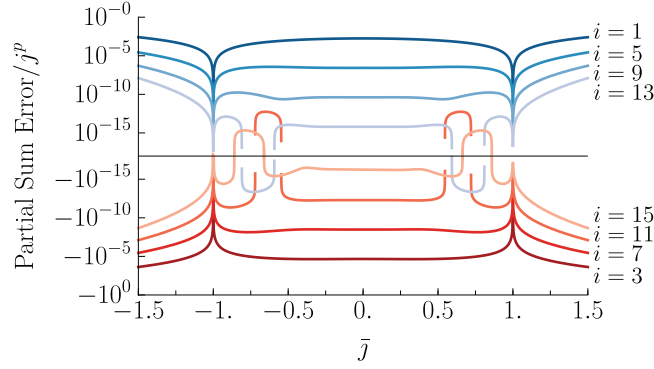
Figure 3. Residuals from approximating Ψ by the partial sums of the Taylor series expansion in F^p, equation (26), plotted for $F^p = 0.75$. The label i indicates that the partial sum is computed up through (and including) the $(F^p)^i$ term in the series. The residual is this i^{th} partial sum minus Ψ. When $\bar{j}$ is in the neighborhood of ± 1, the Taylor expansion for Ψ is an alternating series, so the partial sums provide upper and lower bounds on Ψ. For low-order partial sums and for small F^p the partial sums provide bounds for all values of $\bar{j}$. As an example, the $i = 1$ partial sum is the quadratic bound given by equation (27), which holds for all values of $\bar{J}$ and F^p.

$$\frac{\partial \Delta}{\partial \bar{j}} = j^p \left[\text{arcsinh}\left(\bar{j} \sinh \frac{F^p}{2} \right) - \frac{\bar{j} F^p}{2} \right], \tag{28}$$

implying that Δ strictly increases as $|\bar{j} - 1|$ grows[7]. Since $\Psi \leqslant \Psi_{\text{quad}}$ is bounded for every edge of the graph, $I(\boldsymbol{p},\boldsymbol{j})$ is bounded by a quadratic form whose curvature is determined by the local dissipation rates:

$$I(\boldsymbol{p},\boldsymbol{j}) \leqslant I_{\text{quad}}(\boldsymbol{p},\boldsymbol{j}) \equiv \sum_{y<z} \frac{\sigma^p(y,z)}{4 j^p(y,z)^2} \left(j(y,z) - j^p(y,z) \right)^2, \tag{29}$$

Equation (29) is restricted to conservative currents, or alternatively we take $I_{\text{quad}} \equiv \infty$ for nonconservative $\boldsymbol{j}$.

Recall that by construction $I(\boldsymbol{\pi},\boldsymbol{j}^{\pi}) = 0$. Our quadratic bound shares this minimum. Furthermore, $I(\boldsymbol{\pi},-\boldsymbol{j}^{\pi}) = I_{\text{quad}}(\boldsymbol{\pi},-\boldsymbol{j}^{\pi})$, so I_{quad} is the tightest quadratic upper bound which has a minimum at $(\boldsymbol{\pi},\boldsymbol{j}^{\pi})$ and can be expressed as a single sum over edges (i.e. is diagonal in the edge basis) [27]. The curvature of I, evaluated at $\boldsymbol{\pi}$ and $\boldsymbol{j}^{\pi}$, reflects the variance of long-time current fluctuations. The tighter inequality (29) is, the more precisely we may relate this variance to the local dissipation rates. Truncation after the first-order term of the series expansion equation (26) becomes exact for small thermodynamic forces, meaning that the quadratic bound is accurate when F^p is small.

5. Scalar current fluctuations

The large deviation function $I(\boldsymbol{p},\boldsymbol{j})$ describes the joint distribution for the density in all meso-states and currents between these states. Experiments, however, cannot hope to resolve the statistical details of a large number of degrees of freedom. To make practical use of inequality

---

[7] Using the monotonicity of arcsinh and the concavity of sinh, equation (28) implies that $\partial \Delta / \partial \bar{j} > 0$ for $\bar{j} > 1$ and $\partial \Delta / \partial \bar{j} < 0$ for $0 \leqslant \bar{j} \leqslant 1$.

(29), we must project the result onto a smaller probability space. Rather than monitoring $\boldsymbol{p}$ and $\boldsymbol{j}$, suppose we only measure a scalar current of the form

$$j_d = \boldsymbol{j} \cdot \boldsymbol{d} \equiv \sum_{y<z} j(y,z) d(y,z). \tag{30}$$

The generalized current j_d is a linear combination of the currents between mesostates with expansion coefficients $\boldsymbol{d}$ [24]. This construction affords significant flexibility. By choosing $\boldsymbol{d} = \boldsymbol{F}^\pi$, the generalized current is the same as the dissipation rate. Other choices of $\boldsymbol{d}$ can highlight the current across a single edge or the current associated with transitions between macrostates, as illustrated in section 7.

In the long-time limit, the distribution for this generalized current adopts the large deviation form $P(j_d) \asymp \mathrm{e}^{-TI(j_d)}$. Because $P(j_d)$ can be constructed from $P(\boldsymbol{p}, \boldsymbol{j})$, the rate function for the generalized current is related to $I(\boldsymbol{p}, \boldsymbol{j})$[8]:

$$I(j_d) = \inf_{\boldsymbol{p}, \boldsymbol{j} | \boldsymbol{j} \cdot \boldsymbol{d} = j_d} I(\boldsymbol{p}, \boldsymbol{j}). \tag{31}$$

This infimum is bounded from above by $I(\boldsymbol{p}^*, \boldsymbol{j}^*)$ for any choice of $\boldsymbol{p}^*$ and conservative $\boldsymbol{j}^*$ such that $\boldsymbol{j}^* \cdot \boldsymbol{d} = j_d$. We choose $\boldsymbol{p}^* = \boldsymbol{\pi}$ and $\boldsymbol{j}^* = (j_d/j_d^\pi) \boldsymbol{j}^\pi$, where $j_d^\pi \equiv \boldsymbol{j}^\pi \cdot \boldsymbol{d}$. As a multiple of the conservative steady-state current $\boldsymbol{j}^\pi$, $\boldsymbol{j}^*$ is guaranteed to be conservative. Hence

$$\begin{aligned}
I(j_d) &\leqslant I(\boldsymbol{p}^*, \boldsymbol{j}^*) \\
&= \frac{1}{4} \left(\frac{j_d}{j_d^\pi} - 1 \right)^2 \sum_{y<z} \sigma^\pi(y,z) \\
&= \frac{(j_d - j_d^\pi)^2 \Sigma^\pi}{4(j_d^\pi)^2},
\end{aligned} \tag{32}$$

where we have used inequality (29). The bound on the large deviation function translates to a bound on the variance of j_d since $\mathrm{var}(j_d) = 1/I''(j_d^\pi)$, so measuring the mean and variance of any scalar current provides a lower bound on the dissipation rate:

$$\frac{2(j_d^\pi)^2}{\mathrm{var}(j_d)} \leqslant \Sigma^\pi. \tag{33}$$

The bound on Σ^π is most useful if it is tight, but there are two distinct reasons it might be loose: (1) Ψ for the edges could deviate significantly from the quadratic bound Ψ_{quad} or (2) our choice of $\boldsymbol{p}^*$ and $\boldsymbol{j}^*$ could be suboptimal. In the remainder of the paper we study diffusion processes as a limit of Markov jump processes. That limiting procedure yields $\Psi = \Psi_{\mathrm{quad}}$, implying that the diffusion process bound is weakened only by our suboptimal $\boldsymbol{p}^*$ and $\boldsymbol{j}^*$. For diffusions, the high-dimensional rate function $I(\boldsymbol{p}, \boldsymbol{j})$ is exactly given by the quadratic $I_{\mathrm{quad}}(\boldsymbol{p}, \boldsymbol{j})$, but the low-dimensional rate function $I(j_d)$ can still differ significantly from the upper bound in inequality (32). Section 7 considers this scenario in greater detail.

6. Diffusion processes

Though our results have been derived for Markov jump processes, we may translate them into appropriate forms to describe fluctuations in diffusion processes. For simplicity, we consider this diffusive limit only for a single particle moving in two dimensions with mobility μ. Let $\boldsymbol{x} \equiv (x_1, x_2)$ denote the particle's position, which evolves according to an overdamped

---

[8] The restriction to conservative currents is built into I because $I(\boldsymbol{p}, \boldsymbol{j}) = \infty$ for non-conservative $\boldsymbol{j}$
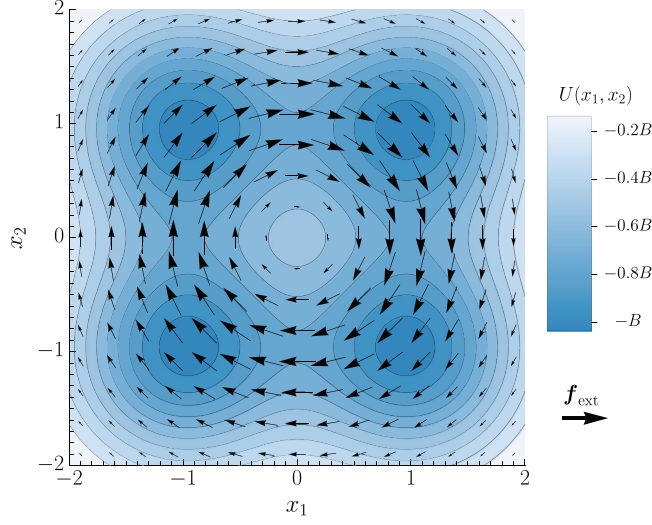
Figure 4. Contour plot of a free energy $U(\boldsymbol{x}) = -B \sum_{i=1}^{4} \mathrm{e}^{-(\boldsymbol{x}-\boldsymbol{c}_i)^2}$, with the $\boldsymbol{c}_i$'s setting the center of the Gaussian wells at $(\pm 1, \pm 1)$. The parameter B controls the barrier heights (or equivalently the well depths). The black arrows represent a vector field of the non-gradient external field, $\boldsymbol{f}_{\mathrm{ext}}(\boldsymbol{x}) = A\boldsymbol{x}^2 \mathrm{e}^{-3|\boldsymbol{x}|}(x_2, -x_1)$, which drives cycles around the origin. The external field's amplitude is regulated by the parameter A.

Langevin equation with deterministic force $\boldsymbol{f}(\boldsymbol{x}) \equiv (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}))$ and random force $\boldsymbol{\eta}$. We further decompose the deterministic force into a contribution from a free energy gradient $\nabla U(\boldsymbol{x})$ and one from a non-gradient external field $\boldsymbol{f}_{\mathrm{ext}}(\boldsymbol{x})$,

$$\frac{\partial \boldsymbol{x}}{\partial t} = -\mu \nabla U(\boldsymbol{x}) + \mu \boldsymbol{f}_{\mathrm{ext}}(\boldsymbol{x}) + \boldsymbol{\eta}. \tag{34}$$

The random force at time t, $\boldsymbol{\eta}(t) \equiv (\eta_1(t), \eta_2(t))$ is a vector of Gaussian random variables satisfying $\langle \eta_k(t) \rangle = 0$ and $\langle \eta_k(t) \eta_{k'}(t') \rangle = 2D\delta_{k,k'}\delta(t-t')$, where D is the diffusion constant. In section 7 we will consider the particular driven diffusive process shown in figure 4 as an illustrative example.

Associated to the Langevin equation is a Fokker–Planck equation describing the evolution of probability density at position $\boldsymbol{x}$, $\rho(\boldsymbol{x})$:

$$\frac{\partial \rho(\boldsymbol{x})}{\partial t} = -\nabla \cdot \boldsymbol{\mathcal{J}}^\rho(\boldsymbol{x}), \tag{35}$$

where

$$\boldsymbol{\mathcal{J}}^\rho(\boldsymbol{x}) \equiv (\mathcal{J}_1^\rho(\boldsymbol{x}), \mathcal{J}_2^\rho(\boldsymbol{x})) = \boldsymbol{f}(\boldsymbol{x})\rho(\boldsymbol{x}) - D\nabla\rho(\boldsymbol{x}) \tag{36}$$

is the ρ-dependent current at $\boldsymbol{x}$. We denote the steady-state density ρ^π and the steady-state current $\boldsymbol{\mathcal{J}}^\pi$. As in the jump process, we define an empirical density field and empirical current field, ρ and $\boldsymbol{\mathcal{J}}$ respectively, and measure the probability of fluctuations away from ρ^π and $\boldsymbol{\mathcal{J}}^\pi$ using the rate function $\mathcal{I}[\rho, \boldsymbol{\mathcal{J}}]$. The square brackets highlight that $\mathcal{I}$ is now a functional of the density and current fields. The exact form of $\mathcal{I}$ is known [27, 38]. In this section, we give a complementary derivation of the result using a limit of the Markov jump process results. The limiting procedure serves to clarify the origin of quadratic current fluctuations in diffusion processes.
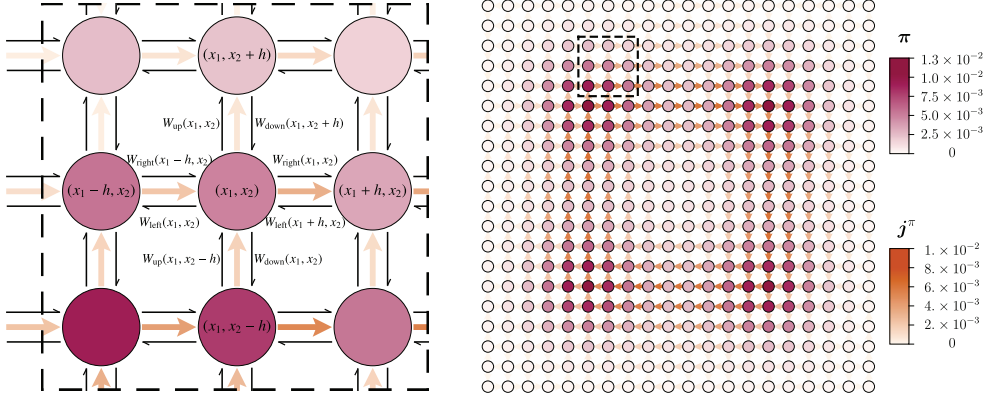
Figure 5. A jump process approximates the diffusion in figure 4. On the left, a close-up view of a sub-system shows the transition rates on the lattice. On the right, the steady-state behavior of the full state space ($x \in [-2, 2] \times [-2, 2]$) is shown for $A = 8, B = 5, D = 1, \mu = 1$. The steady-state density primarily resides in the four wells, but the external field drives some clockwise current around the origin.

To leverage our previous results, we approximate the diffusion process by a jump process on a square lattice, as depicted in figure 5. There is no unique way to coarse-grain a diffusion to a jump process [39]. Thus we have freedom in how we construct our model so long as it yields a diffusive limit that matches the Fokker–Planck equation, equation (35), when the lattice spacing becomes infinitesimal. We construct a simple nearest-neighbor jump process which is entirely characterized by four space-dependent transition rates for hopping from a grid point at x to a nearest-neighbor site: $W_{\text{up}}(x)$, $W_{\text{right}}(x)$, $W_{\text{down}}(x)$, and $W_{\text{left}}(x)$. These rates must scale with the lattice spacing h in such a way that the first two jump moments give the correct drift and diffusion [39], requiring

$$
\begin{aligned}
h(W_{\text{right}}(x) - W_{\text{left}}(x)) &= \mu f_1(x) \\
h(W_{\text{up}}(x) - W_{\text{down}}(x)) &= \mu f_2(x) \\
h^2(W_{\text{right}}(x) + W_{\text{left}}(x)) &= 2D \\
h^2(W_{\text{up}}(x) + W_{\text{down}}(x)) &= 2D,
\end{aligned}
\tag{37}
$$

as $h \to 0$. From these constraints the hopping rates are

$$
\begin{aligned}
W_{\text{right}}(x) &= \frac{\mu f_1(x)}{2h} + \frac{D}{h^2} \\
W_{\text{left}}(x) &= \frac{-\mu f_1(x)}{2h} + \frac{D}{h^2} \\
W_{\text{up}}(x) &= \frac{\mu f_2(x)}{2h} + \frac{D}{h^2} \\
W_{\text{down}}(x) &= \frac{-\mu f_2(x)}{2h} + \frac{D}{h^2}.
\end{aligned}
\tag{38}
$$

We now identify each vertex of the graph by its location x, writing the density at that vertex as $p(x)$. The edge connecting grid point x to its neighbor on the right has current $j^p_{\text{right}}(x) \equiv j^p(x, x + (h, 0))$, and the thermodynamic force associated to that edge is

$F_{\text{right}}^p(\boldsymbol{x}) \equiv F^p(\boldsymbol{x}, \boldsymbol{x} + (h, 0))$. The 'up', 'down', and 'left' directions are treated analogously. Using this notation, we may rewrite I_{quad} of equation (29) as

$$I_{\text{quad}}(\boldsymbol{p},\boldsymbol{j}) = \frac{1}{8} \sum_{\boldsymbol{x}} \left[\frac{F_{\text{left}}^p(\boldsymbol{x})}{j_{\text{left}}^p(\boldsymbol{x})} (j_{\text{left}}(\boldsymbol{x}) - j_{\text{left}}^p(\boldsymbol{x}))^2 + \frac{F_{\text{right}}^p(\boldsymbol{x})}{j_{\text{right}}^p(\boldsymbol{x})} (j_{\text{right}}(\boldsymbol{x}) - j_{\text{right}}^p(\boldsymbol{x}))^2 \right.$$
$$\left. + \frac{F_{\text{down}}^p(\boldsymbol{x})}{j_{\text{down}}^p(\boldsymbol{x})} (j_{\text{down}}(\boldsymbol{x}) - j_{\text{down}}^p(\boldsymbol{x}))^2 + \frac{F_{\text{up}}^p(\boldsymbol{x})}{j_{\text{up}}^p(\boldsymbol{x})} (j_{\text{up}}(\boldsymbol{x}) - j_{\text{up}}^p(\boldsymbol{x}))^2 \right]. \tag{39}$$

To simplify the expression further, we must convert from the discrete-space density $\boldsymbol{p}$ and current $\boldsymbol{j}$ to the fields ρ and $\boldsymbol{\mathcal{J}}$. In the continuum limit, $\boldsymbol{p}(\boldsymbol{x})$ vanishes while $\rho(\boldsymbol{x})$ remains finite such that $\boldsymbol{p}(\boldsymbol{x}) \to h^2 \rho(\boldsymbol{x})$. The current field $\boldsymbol{\mathcal{J}}^\rho$ is also finite, but the current on any edge of the lattice vanishes with order h:

$$j_{\text{right}}^p(\boldsymbol{x}) = p(\boldsymbol{x}) W_{\text{right}}(\boldsymbol{x}) - p\Big(\boldsymbol{x} + (h, 0)\Big) W_{\text{left}}\Big(\boldsymbol{x} + (h, 0)\Big)$$
$$= \left(\mu f_1(\boldsymbol{x}) \rho(\boldsymbol{x}) - D \frac{\partial \rho(\boldsymbol{x})}{\partial x_1} \right) h + \mathcal{O}(h^2)$$
$$= \mathcal{J}_1^\rho(\boldsymbol{x}) h + \mathcal{O}(h^2). \tag{40}$$

Likewise, the thermodynamic force on each edge vanishes in proportion to h:

$$F_{\text{right}}^p(\boldsymbol{x}) = \ln \frac{p(\boldsymbol{x}) W_{\text{right}}(\boldsymbol{x})}{p\Big(\boldsymbol{x} + (h, 0)\Big) W_{\text{left}}\Big(\boldsymbol{x} + (h, 0)\Big)}$$
$$= \left(\frac{\mu f_1(\boldsymbol{x})}{D} - \frac{1}{p(\boldsymbol{x})} \frac{\partial p(\boldsymbol{x})}{\partial x_1} \right) h + \mathcal{O}(h^2)$$
$$= \mathcal{F}_1^\rho(\boldsymbol{x}) h + \mathcal{O}(h^2), \tag{41}$$

with a finite thermodynamic force field given by

$$\boldsymbol{\mathcal{F}}^\rho(\boldsymbol{x}) \equiv (\mathcal{F}_1^\rho(\boldsymbol{x}), \mathcal{F}_2^\rho(\boldsymbol{x})) = \frac{\mu \boldsymbol{f}(\boldsymbol{x})}{D} - \nabla \ln \rho(\boldsymbol{x}). \tag{42}$$

Hence the ratio of force to current on an edge, which appears in each term of equation (39), remains finite and independent of the jump direction,

$$\lim_{h \to 0} \frac{F_{\text{right}}^p(\boldsymbol{x})}{j_{\text{right}}^p(\boldsymbol{x})} = \frac{\mathcal{F}_1^\rho(\boldsymbol{x})}{\mathcal{J}_1^\rho(\boldsymbol{x})} = \frac{\frac{\mu f_1}{D} - \frac{\partial}{\partial x_1} \ln \rho(\boldsymbol{x})}{\mu f_1 \rho - D \frac{\partial}{\partial x_1} \rho(\boldsymbol{x})} = \frac{1}{D \rho(\boldsymbol{x})}. \tag{43}$$

Recall from equation (26) that the rate function bound (29) becomes tighter for small thermodynamic forces. Since the forces vanish in the $h \to 0$ limit, the jump process on an infinitesimal grid thus saturates inequality (29), yielding the equality

$$\mathcal{I}[\rho, \boldsymbol{\mathcal{J}}] = \lim_{h \to 0} I(\boldsymbol{p},\boldsymbol{j}) = \lim_{h \to 0} I_{\text{quad}}(\boldsymbol{p},\boldsymbol{j}). \tag{44}$$

In other words, approximating I by I_{quad} becomes exact in the continuum limit. We convert the sum to an integral using

$$\sum_{\boldsymbol{x}} h^2 \to \int d\boldsymbol{x} \tag{45}$$

and insert equations (40) and (43) into (39) to obtain the rate function for diffusions

$$\mathcal{I}[\rho, \mathcal{J}] = \int d\mathbf{x} \frac{(\mathcal{J}(\mathbf{x}) - \mathcal{J}^\rho(\mathbf{x}))^2}{4D\rho(\mathbf{x})}. \tag{46}$$

Note that the current fluctuations are strictly quadratic, a fact that can be traced back to the vanishingly small thermodynamic force on each infinitesimal edge of the graph.

The rate function may alternatively be expressed with respect to the local dissipation rates,

$$\boldsymbol{\sigma}^\rho(\mathbf{x}) = \mathcal{J}^\rho(\mathbf{x}) \cdot \mathcal{F}^\rho(\mathbf{x}) = \frac{\mathcal{J}^\rho(\mathbf{x})^2}{D\rho(\mathbf{x})}. \tag{47}$$

Interpreting $\boldsymbol{\sigma}^\rho$ as a physical dissipation requires that the particle exchanges energy with a thermal reservoir at inverse temperature β and satisfies the Einstein relation $\beta D = \mu$. The second equality follows from a rearrangement of equation (43),

$$\mathcal{J}^\rho(\mathbf{x}) = D\mathcal{F}^\rho(\mathbf{x})\rho(\mathbf{x}), \tag{48}$$

which is an expression of linear response; the current at any point in space is linearly proportional to the thermodynamic force at that point. This linear-response feature of diffusion processes implies that the current fluctuations are specified by local dissipation rates:

$$\mathcal{I}[\rho, \mathcal{J}] = \int d\mathbf{x} \; \frac{\boldsymbol{\sigma}^\rho(\mathbf{x})}{4\mathcal{J}^\rho(\mathbf{x})^2} \left(\mathcal{J}(\mathbf{x}) - \mathcal{J}^\rho(\mathbf{x}) \right)^2. \tag{49}$$

Our form for $\mathcal{I}$ is analogous to the right-hand side of inequality (29), but for diffusion processes the inequality has become an equality.

7. Coarse-graining in space

Up to this point we have assumed a mesoscopic Markov process that is monitored with complete precision. More commonly, it is only possible to observe coarse-grained macrostates, and transitions between them might not be Markovian. Rather, the macroscopic dynamics is described by a hidden Markov model [40]. Remarkably, current fluctuations in the non-Markovian macroscopic dynamics carry information about the dissipation rate of the underlying mesoscopic degrees of freedom.

To illustrate the connection between dissipation and a macroscopic current, we study the generalized scalar current j_d, defined in equation (30). The vector $\mathbf{d}$ determines which microscopic transitions are observable, so for a suitable choice of $\mathbf{d}$, j_d can be any macroscopic current. Consider, for example, the two-dimensional driven diffusion process in figure 4. A particle evolves on a free energy landscape

$$U(\mathbf{x}) = -B \sum_{i=1}^{4} e^{-(\mathbf{x}-\mathbf{c}_i)^2} \tag{50}$$

with B controlling the depth of four Gaussian wells centered at $\mathbf{c} = (\pm 1, \pm 1)$. At a coarse-grained level, these wells define four possible macrostates, the quadrants of the coordinate plane. In addition to the free energy landscape, we introduce a non-gradient external field with amplitude A,

$$\mathbf{f}_{\text{ext}}(\mathbf{x}) = A\mathbf{x}^2 e^{-3|\mathbf{x}|}(x_2, -x_1), \tag{51}$$
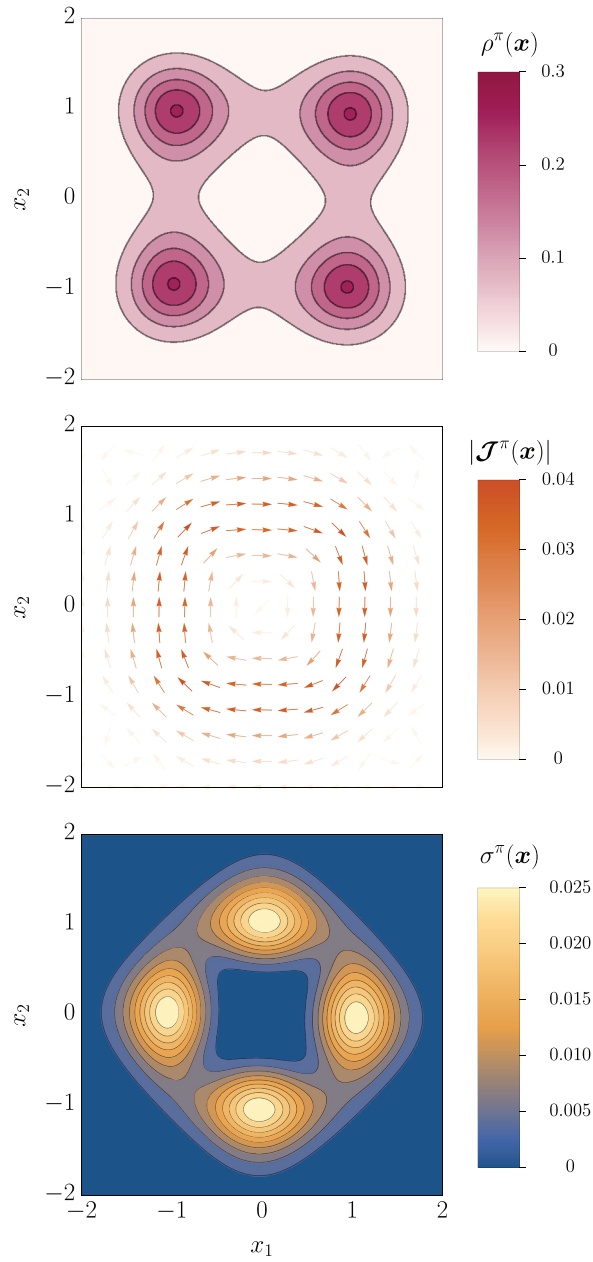
Figure 6. Steady-state density field, current field, and local dissipation rate (top to bottom) for the example from figure 4 with $A = 8, B = 5, D = 1, \mu = 1$. The diffusion process was approximated by a jump process on a 400×400 grid, and the master equation was numerically solved in discrete space. Note that the local dissipation rate is largest at the barriers, where the density is low and the current high.

which drives transitions between the macrostates. By using equation (38) to discretize the state space, we can numerically solve for the steady state, as plotted in figure 6. We focus on the macroscopic current given by the rate of clockwise (CW) transitions between these
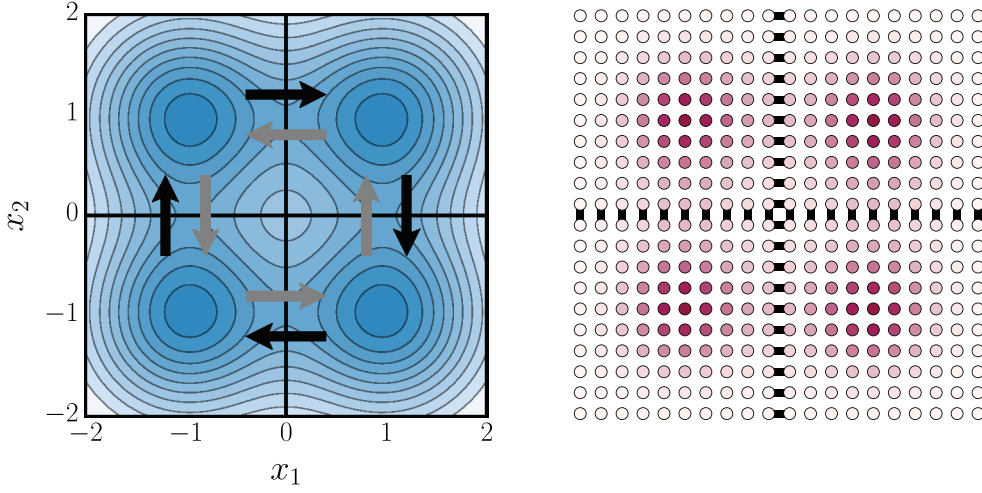
Figure 7. Left: The driven diffusion process in figure 4 is naturally coarse-grained into four states, and the macroscopic current is constructed by monitoring transitions between them. Clockwise transitions (black arrows) are given the weight $+1$ while counterclockwise transitions (gray arrows) carry weight -1. All motion within a coarse-grained macrostate is undetected. Right: In practice, we compute large deviation properties of the coarse-grained current by solving the diffusion process on a grid, as in section 6. On the grid, j_d is constructed by weighting the black edges by ± 1 and all other edges by 0 as specified in equation (52).
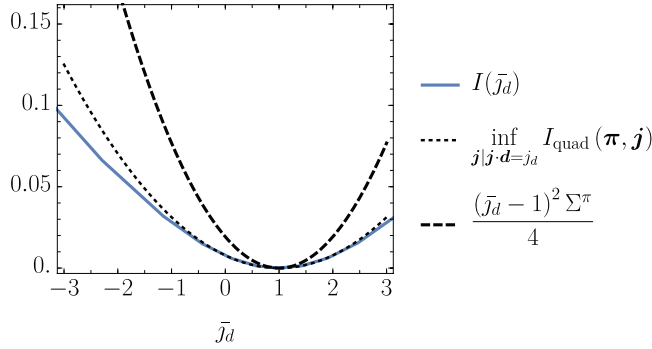

Figure 8. Large deviation function for the macroscopic current with the four coarse-grained states in figure 7. U and $\boldsymbol{f}_{\text{ext}}$ are given by equations (50) and (51), respectively, with $A = 8$, $B = 5$. Macroscopic currents are expressed in relation to the steady-state value: $\bar{j}_d \equiv j_d / j_d^{\pi}$. The blue line shows the current fluctuations for the diffusion process solved on a grid of 400×400 states. $I(j_d)$ is the numerical Legendre transform of a scaled cumulant generating function, computed as the maximum eigenvalue of a tilted rate matrix [36, 41, 42]. The dotted black line is the result of the constrained minimization of the quadratic form, computed as described in appendix B. This constrained minimization very closely approximates $I(\bar{j}_d)$ for small deviations. The dashed black line is the quadratic rate function bound equation (32) with curvature regulated by the total dissipation rate.

coarse-grained states less the rate of counter-clockwise (CCW) transitions. As illustrated in figure 7, we measure this current, $j_d = \boldsymbol{j} \cdot \boldsymbol{d}$, by choosing
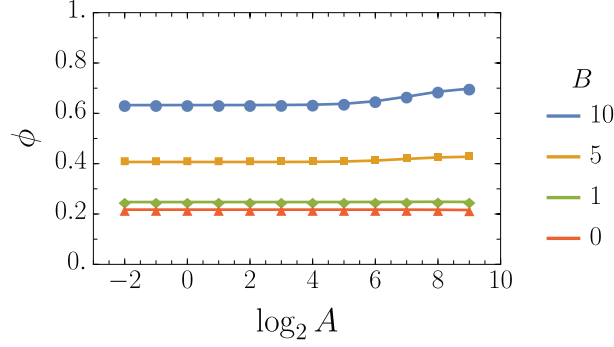
Figure 9. Fraction of the total dissipation rate which can be deduced from the macroscopic current fluctuations in the four-well diffusion model defined by equations (50) and (51). *A* controls the amplitude of the external driving field and *B* controls the well depths. Equation (54) defines ϕ, which takes the value 1 when the current fluctuations account for all of the dissipation. Values of ϕ were computed by representing the diffusion process as a jump process on a 400×400 grid. For large *B*, the wells become deep metastable states which are amenable to coarse-graining, so the macroscopic current accounts for a significant fraction of the dissipation.

$$
d(y,z) = \begin{cases} 1, & y \to z \text{ is a CW macrostate transition} \\ -1, & y \to z \text{ is a CCW macrostate transition} \\ 0, & \text{otherwise.} \end{cases} \tag{52}
$$

The results from section 5 directly apply to this construction. In particular, inequality (33) implies $\Sigma^\pi \geqslant 2(j_d^\pi)^2/\text{var}(j_d)$. The right-hand side of this inequality depends on the first two moments of the *macroscopic* empirical current distribution. By measuring these moments with coarse-grained observations, we therefore bound the total entropy production of the *mesoscopic* system, Σ^π.

To assess the inequality's tightness, we used a 400×400 grid to numerically compute $I(j_d)$ for the four-well model with $D = \mu = 1$, and with various choices of *A* and *B*. We saw in section 6 that I_{quad} becomes exact in the diffusive limit, so

$$
I(j_d) = \inf_{\boldsymbol{p}, \boldsymbol{j} | \boldsymbol{j} \cdot \boldsymbol{d} = j_d} I_{\text{quad}}(\boldsymbol{p}, \boldsymbol{j}) \leqslant \inf_{\boldsymbol{j} | \boldsymbol{j} \cdot \boldsymbol{d} = j_d} I_{\text{quad}}(\boldsymbol{\pi}, \boldsymbol{j}). \tag{53}
$$

The numerical results, plotted in figure 8, reflect that inequality (53) is a tight bound. In contrast, the dissipation rate bound, inequality (32), is noticeably weaker. A measure of this weakness is the ratio

$$
\phi = \left(\frac{2(j_d^\pi)^2}{\text{var}(j_d)} \right) / \Sigma^\pi, \tag{54}
$$

which ranges from 0 to 1. This ratio, plotted in figure 9 for a range of driving amplitudes and well depths, assesses the fraction of the dissipation rate that can be inferred from the macroscopic current fluctuations.

It has been shown that our inference strategy detects the greatest fraction of dissipation if the macroscopic current is constructed to be proportional to the thermodynamic force [26]. More precisely, $\phi = 1$ when $\boldsymbol{d} \propto \boldsymbol{F}^\pi$, and ϕ decreases when the vector $\boldsymbol{d}$ is less aligned with $\boldsymbol{F}^\pi$. In our coarse-graining example, the elements of $\boldsymbol{d}$ are set to 0 or ± 1 based on
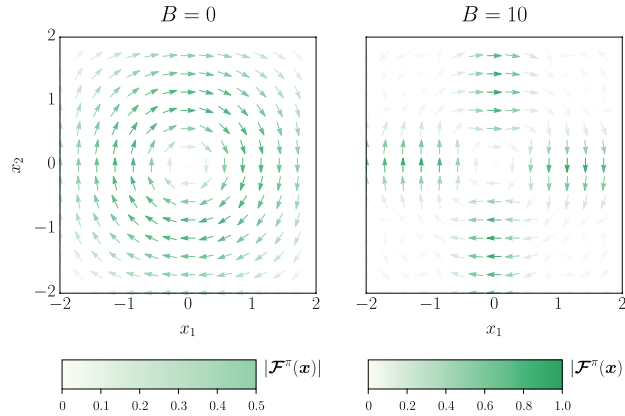
Figure 10. Thermodynamic force for the four-well model with $A = 8$, solved on a 400×400 grid. For large B, the free energy landscape has deep wells, and the thermodynamic forces are large along the barriers separating the wells. In this large-B limit, the thermodynamic force resembles $\boldsymbol{d}$, which vanishes everywhere except for along the axes.

whether each mesoscopic transition is observable. With $\boldsymbol{d}$ constrained by which transitions are observable, ϕ can only be varied if $\boldsymbol{F}^\pi$ is altered, something that can be achieved by tuning A and B.

We can understand ϕ's dependence on these model parameters by analyzing the degree to which $\boldsymbol{F}^\pi$ aligns with $\boldsymbol{d}$. Figure 10 illustrates how the continuous thermodynamic force field $\boldsymbol{\mathcal{F}}^\pi(\boldsymbol{x})$ varies with the well depth. In the absence of wells ($B = 0$), the thermodynamic force is radially symmetric about the origin. As the wells are made deeper, $\boldsymbol{\mathcal{F}}^\pi$ is amplified along the axes, more closely resembling $\boldsymbol{d}$ of figure 7. This resemblance explains the trend that ϕ increases with increasing B. Interestingly, ϕ depends only weakly on the external field amplitude, a fact we rationalize with a linear-response argument at the end of appendix B.

8. Conclusions

It is typically impossible to resolve detailed mesoscopic dynamics, even with the most sophisticated experimental tools. Because the dissipation rate is defined in terms of such immeasurable mesoscopic transitions, it may seem that there is no hope of experimentally measuring the dissipation, short of explicitly detecting the free energetic flows from the thermodynamic reservoirs. Our models and analysis demonstrate that, in fact, a bound on the total dissipation can be inferred by monitoring only the fluctuations in macroscopic currents. The procedure we use to infer this bound is extremely adaptable; we make essentially no assumptions about the mesoscopic details of the system nor about the macroscopic currents that will be observed.

Crucially, it is the macroscopic current fluctuations, not their averages, that reveal the dissipation of the unobserved mesoscopic dynamics. Motivated by the inability to fully resolve mesoscopic dynamics in an experiment, we have constructed a coarse-grained macroscopic current by recording a small subset of the mesoscopic transitions. The fraction of total dissipation due to the irreversibiilty of one of these observed transitions is small and in fact vanishes in the diffusive limit. Were we to measure only the average current through the observed links,

we would at best deduce this infinitesimally small portion of the total dissipation. However, if we use the fluctuation-dissipation relation for the observed currents, we tacitly constrain the average behavior of all the unobserved transitions. Hence the fluctuations in a tiny subset of the mesoscopic transitions carry information about an appreciable fraction of the total dissipation. This useful fact stems from a fundamental bound on the extent of fluctuations: the total dissipation rate sets a bound for the variance in *any* generalized current.

We anticipate that our inference scheme will be a robust strategy for analyzing dissipation in stochastic, biophysical systems. The extent to which coarse-grained fluctuations reveal the total dissipation, measured by ϕ, depends on the fidelity of the coarse-graining. For the model we studied in section 7, ϕ nears 1 when the wells are sufficiently deep that the coarse-grained states are long-lived metastable macrostates. The step-like dynamics of dissipative biological machines [43, 44] demonstrate precisely this type of metastability and time-scale separation [45]. Indeed, it is often the goal of biological studies to assign the experimentally observed metastable macrostates to coarse-grained descriptions, e.g. ligation states. Together, these observations bolster the prospect of accurately inferring dissipation rates using data from single molecule experiments.

Acknowledgments

Appendix A. Derivation of equivalent forms of Ψ

In the main text we have extensively discussed the time-asymmetric empirical current from state y to z, $j(y,z)$. This quantity is time-asymmetric because it acquires a negative sign if the trajectory is run backwards in time. Maes and coworkers have highlighted the importance of the time-symmetric empirical traffic, which counts the total number of hops across the yz edge in either direction [12, 21, 46–48]. As in our study of the currents, we monitor two different forms of the traffic: the expected traffic given $\boldsymbol{p}$, $t^p(y, z) = p(y)r(y, z) + p(z)r(z, y)$, and the traffic $t_\star(y,z) = q_\star(y,z) + q_\star(z,y)$. Ψ naturally decomposes into contributions from the traffic and contributions from the currents, $\Psi = \Psi_{\text{traffic}} + \Psi_{\text{current}}$ [21], with

$$\Psi_{\text{traffic}} \equiv p(y)r(y, z) + p(z)r(z, y) - q_\star(y, z) - q_\star(z, y)$$

and

$$\Psi_{\text{current}} \equiv q_\star(y, z) \ln \frac{q_\star(y, z)}{p(y)r(y, z)} + q_\star(z, y) \ln \frac{q_\star(z, y)}{p(z)r(z, y)}. \tag{A.1}$$

Ψ_{traffic} is simply the difference between the expected and the observed number of hops, $\Psi_{\text{traffic}} = t^p(y, z) - t_\star(y, z)$. To bring Ψ_{traffic} into the form that appears in equation (24), we recognize that $t_\star(y, z) = \sqrt{j(y, z)^2 + a^p(y, z)^2}$ and that $t^p(y, z) = \sqrt{j^p(y, z)^2 + a^p(y, z)^2}$. Simplifying Ψ_{current} is slightly more involved. Using the solution for $q_\star$, equation (22), Ψ_{current} can be rewritten in the form of equation (24),

$$\Psi_{\text{current}} = \frac{j(y,z)}{2} \ln \frac{q_\star(y,z)p(z)r(z,y)}{q_\star(z,y)p(y)r(y,z)} + \frac{\sqrt{j(y,z)^2 + a^p(y,z)^2}}{2} \ln \underbrace{\frac{q_\star(y,z)q_\star(z,y)}{p(y)r(y,z)p(z)r(z,y)}}_{0}$$

$$= \frac{j(y,z)}{2} \left(\ln \frac{q_\star(y,z)}{q_\star(z,y)} + \ln \frac{p(z)r(z,y)}{p(y)r(y,z)} \right) = \frac{j(y,z)}{2} \ln \frac{j(y,z) + \sqrt{j(y,z)^2 + a^p(y,z)^2}}{-j(y,z) + \sqrt{j(y,z)^2 + a^p(y,z)^2}}$$

$$+ \frac{j(y,z)}{2} \ln \frac{j^p(y,z) + \sqrt{j^p(y,z)^2 + a^p(y,z)^2}}{-j^p(y,z) + \sqrt{j^p(y,z)^2 + a^p(y,z)^2}} = j(y,z) \left(\operatorname{arcsinh} \frac{j(y,z)}{a^p(y,z)} - \operatorname{arcsinh} \frac{j^p(y,z)}{a^p(y,z)} \right).$$

$$(A.2)$$

Appendix B. Tight quadratic current fluctuation bound

In the main text, an upper bound for $I(j_d)$ was obtained by evaluating I at $\boldsymbol{p}^* = \boldsymbol{\pi}$ and $\boldsymbol{j}^* = (j_d/j_d^\pi) \boldsymbol{j}^\pi$. This choice implies that the chance of measuring a value j_d is at least as likely as it would be if we simultaneously scaled all of the steady-state mesoscopic currents so as to make $\boldsymbol{j}^* \cdot \boldsymbol{d} = j_d$. There can, however, be more likely ways to obtain j_d which do not equally scale the different mesoscopic currents. Identifying the optimum is a straightforward exercise in linear algebra: the quadratic form I_{quad} must be extremized subject to the linear constraint $\boldsymbol{j} \cdot \boldsymbol{d} = j_d$ and to current conservation. The restriction to conservative currents may be imposed by a collection of Lagrange multipliers, one per node, as in [26] or by expressing the currents in a cycle basis as in [27, 49]. We use the second strategy here. Since both strategies exactly solve the same convex optimization problem, they must yield the same solution.

Let $a(i)$ be the current around the ith cycle in a complete cycle basis. This basis may be constructed from spanning trees [50], but for now the particular cycle basis is immaterial. We introduce

$$\chi_{yz}^i = \begin{cases} 1, & \text{cycle } i \text{ contains edge } yz \text{ with } y < z \\ -1, & \text{cycle } i \text{ contains edge } yz \text{ with } z < y \\ 0, & \text{otherwise} \end{cases} \qquad (B.1)$$

to convert between the edge basis and the cycle basis so that

$$j(y,z) = \sum_i \chi_{yz}^i a(i). \qquad (B.2)$$

Recall that

$$I_{\text{quad}}(\boldsymbol{p},\boldsymbol{j}) = \begin{cases} \sum_{y<z} \frac{\sigma^p(y,z)}{4} \left(\frac{j(y,z)}{j^p(y,z)} - 1 \right)^2, & \text{conservative } \boldsymbol{j} \\ \infty, & \text{otherwise.} \end{cases} \qquad (B.3)$$

In terms of the cycle basis the restriction to conservative currents is more natural:

$$I_{\text{quad}}(\boldsymbol{p},\boldsymbol{a}) = \sum_{y<z} \left(\sum_i \chi_{yz}^i (a(i) - a^p(i)) \right)^2 \frac{F^p(y,z)}{4j^p(y,z)}$$

$$= \frac{1}{2} \sum_{i,k} (a(i) - a^p(i)) G_{ik} (a(k) - a^p(k)), \qquad (B.4)$$

with $G_{ik} = \sum_{y<z} \chi_{yz}^i \chi_{yz}^k \frac{F^p(y,z)}{2j^p(y,z)}$ [27]. We also translate the linear constraint into the cycle basis:

$$j_d = \sum_{y<z} \sum_i \chi_{yz}^i a(i) d(y,z) = \sum_i a(i) d(i) \equiv \boldsymbol{a} \cdot \boldsymbol{d}, \tag{B.5}$$

where we have defined $d(i) \equiv \sum_{y<z} \chi_{yz}^i d(y,z)$. In words, $d(i)$ is now the expansion coefficient that indicates how much cycle i contributes to the generalized current j_d.

In analogy with the main text, $I(j_d) \leqslant I_{\text{quad}}(\boldsymbol{p}^*, \boldsymbol{a}^*)$ for any choice of $\boldsymbol{p}^*$ and $\boldsymbol{a}^*$ which satisfy the constraint $\boldsymbol{a} \cdot \boldsymbol{d} = j_d$. We again choose $\boldsymbol{p}^* = \boldsymbol{\pi}$, but now we compute the exact minimizer of the quadratic form I_{quad}:

$$\boldsymbol{a}^* = \boldsymbol{a}^\pi + \left(\frac{j_d - j_d^\pi}{\boldsymbol{d}^T G^{-1} \boldsymbol{d}} \right) G^{-1} \boldsymbol{d}. \tag{B.6}$$

Hence,

$$I(j_d) \leqslant \inf_{\boldsymbol{j} \mid \boldsymbol{j} \cdot \boldsymbol{d} = j_d} I_{\text{quad}}(\boldsymbol{\pi}, \boldsymbol{j}) = \frac{(j_d - j_d^\pi)^2}{2 \, \boldsymbol{d}^T G^{-1} \boldsymbol{d}}. \tag{B.7}$$

For the coarse-graining problem of section 7, this denominator is simple if we use the cycle basis consisting of the square lattice's plaquettes. Each plaquettes is one of the squares of side length h. Notice that for our coarse-graining procedure, $d(i) = 0$ for all but one of the plaquettes. Only the central plaquette that encircles the origin has a nonzero d. Since we can order the plaquettes arbitrarily, we designate the central plaquette by the label 'o' to highlight that it is the plaquette at the origin. Because $d(o) = 4$, $\boldsymbol{d}^T G^{-1} \boldsymbol{d} = 16 G_{oo}^{-1}$ with G_{oo}^{-1} denoting the diagonal matrix element of G^{-1} for the plaquette at the origin.

Finally, we can see why ϕ is nearly independent of the driving field amplitude A in figure 9. In terms of the G matrix,

$$\phi = \left(\frac{2(j_d^\pi)^2}{\text{var}(j_d)} \right) / \Sigma^\pi = \frac{(j_d^\pi)^2}{8 G_{oo}^{-1} \Sigma^\pi}. \tag{B.8}$$

For sufficiently small driving amplitude, we anticipate that the mesoscopic currents will respond linearly to the external force: $\boldsymbol{j}^\pi \propto A$. In this linear regime we also have $\boldsymbol{F}^\pi \propto A$, so $\Sigma^\pi \propto A^2$. Furthermore, since matrix elements of G involve ratios of F^p and j, the linear-response value of G_{oo}^{-1} is independent of A. Combining these effects, we see that ϕ does not vary with A in the linear-response regime. Empirically, we observe in figure 9 that insensitivity to A in fact extends far beyond the linear-response regime.

References

[1] Smith S B, Finzi L and Bustamante C 1992 *Science* **258** 1122
[2] Blickle V and Bechinger C 2012 *Nat. Phys.* **8** 143
[3] Martínez I A, Roldán E, Dinis L, Petrov D, Parrondo J M R and Rica R A 2016 *Nat. Phys.* **12** 67
[4] Marconi U M B, Puglisi A, Rondoni L and Vulpiani A 2008 *Phys. Rep.* **461** 111
[5] Einstein A 1905 *Ann. Phys.* **322** 549
[6] Onsager L 1931 *Phys. Rev.* **37** 405
[7] Onsager L 1931 *Phys. Rev.* **38** 2265
[8] Callen H B and Welton T A 1951 *Phys. Rev.* **83** 34
[9] Kubo R 1966 *Rep. Prog. Phys.* **29** 255
[10] Harada T and Sasa S I 2005 *Phys. Rev. Lett.* **95** 130602

[11] Speck T and Seifert U 2006 *Europhys. Lett.* **74** 391
[12] Baiesi M, Maes C and Wynants B 2009 *Phys. Rev. Lett.* **103** 010602
[13] Prost J, Joanny J F and Parrondo J M R 2009 *Phys. Rev. Lett.* **103** 090601
[14] Seifert U and Speck T 2010 *Europhys. Lett.* **89** 10007
[15] Bohec P, Gallet F, Maes C, Safaverdi S, Visco P and Van Wijland F 2013 *Europhys. Lett.* **102** 50005
[16] Fodor É, Guo M, Gov N, Visco P, Weitz D and van Wijland F 2015 *Europhys. Lett.* **110** 48005
[17] Fodor É, Ahmed W W, Almonacid M, Bussonnier M, Gov N S, Verlhac M H, Betz T, Visco P and van Wijland F 2016 *EPL* **116** 30008
[18] Moffitt J R and Bustamante C 2014 *FEBS J.* **281** 498
[19] Barato A C and Seifert U 2015 *J. Phys. Chem.* B **119** 6555
[20] Kesidis G and Walrand J 1993 *IEEE Trans. Inf. Theory* **39** 1056
[21] Maes C and Netočný K 2008 *Europhys. Lett.* **82** 30003
[22] Bertini L, Faggionato A and Gabrielli D 2015 *Ann. Inst. Henri Poincare* B **51** 867
[23] Bertini L, Faggionato A and Gabrielli D 2015 *Stoch. Process. Appl.* **125** 2786
[24] Barato A C and Seifert U 2015 *Phys. Rev. Lett.* **114** 158101
[25] Pietzonka P, Barato A C and Seifert U 2016 *Phys. Rev.* E **93** 052145
[26] Gingrich T R, Horowitz J M, Perunov N and England J L 2016 *Phys. Rev. Lett.* **116** 120601
[27] Polettini M, Lazarescu A and Esposito M 2016 *Phys. Rev.* E **94** 052104
[28] Kubo R, Toda M and Hashitsume N 2012 *Statistical Physics II: Nonequilibrium Statistical Mechanics* vol 31 (New York: Springer)
[29] Roldán É and Parrondo J M 2010 *Phys. Rev. Lett.* **105** 150607
[30] Van Kampen N 1992 *Stochastic Processes in Physics and Chemistry* vol 1 (New York: Elsevier)
[31] Hoffmann K H and Salamon P 2011 *Phys. A: Stat. Mech. Appl.* **390** 3086
[32] Esposito M 2012 *Phys. Rev.* E **85** 041125
[33] Chodera J D, Singhal N, Pande V S, Dill K A and Swope W C 2007 *J. Chem. Phys.* **126** 155101
[34] Bowman G R, Huang X and Pande V S 2009 *Methods* **49** 197
[35] Esposito M and Van den Broeck C 2010 *Phys. Rev.* E **82** 011143
[36] Touchette H 2009 *Phys. Rep.* **478** 1
[37] Chandler D 1987 *Introduction to Modern Statistical Mechanics* (Oxford: Oxford University Press)
[38] Barato A C and Chetrite R 2015 *J. Stat. Phys.* **160** 1154
[39] Gardiner C 2009 *Stochastic Methods* (Berlin: Springer)
[40] Noé F, Wu H, Prinz J H and Plattner N 2013 *J. Chem. Phys.* **139** 184114
[41] Lebowitz J L and Spohn H 1999 *J. Stat. Phys.* **95** 333
[42] Lecomte V, Appert-Rolland C and van Wijland F 2007 *J. Stat. Phys.* **127** 51
[43] Smith D E, Tans S J, Smith S B, Grimes S, Anderson D L and Bustamante C 2001 *Nature* **413** 748
[44] Hayashi K, Ueno H, Iino R and Noji H 2010 *Phys. Rev. Lett.* **104** 218103
[45] Wang S W, Kawaguchi K, Sasa S I and Tang L H 2016 *Phys. Rev. Lett.* **117** 070601
[46] Maes C and Netočný K 2007 *J. Math. Phys.* **48** 053306
[47] Maes C, Netocny K and Wynants B 2008 *Markov Process. Relat. Fields* **14** 445
[48] Maes C 2016 arXiv:1603.05147
[49] Pietzonka P, Barato A C and Seifert U 2016 *J. Phys. A: Math. Theor.* **49** 34LT01
[50] Schnakenberg J 1976 *Rev. Mod. Phys.* **48** 571