Journal of Statistical
Mechanics: Theory and
Experiment

Coarse-graining nonequilibrium diffusions with Markov chains

To cite this article: Ramón Nartallo-Kaluarachchi *et al J. Stat. Mech.* (2026) 033205

OPEN ACCESS

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Classical statistical mechanics, equilibrium and non-equilibrium

Coarse-graining nonequilibrium diffusions with Markov chains

Ramón Nartallo-Kaluarachchi[1,2,*], **Renaud Lambiotte**[1] and **Alain Goriely**[1]

[1] Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, United Kingdom
[2] Centre for Eudaimonia and Human Flourishing, University of Oxford, 7 Stoke Pl, Oxford OX3 9BX, United Kingdom
E-mail: nartallokalu@maths.ox.ac.uk

Received 7 November 2025
Accepted for publication 27 February 2026
Published 31 March 2026

Online at stacks.iop.org/JSTAT/2026/033205
https://doi.org/10.1088/1742-5468/ae4f7d

Abstract. We investigate nonequilibrium steady-state dynamics in both continuous- and discrete-state stochastic processes. Our analysis focuses on planar diffusion dynamics and their coarse-grained approximations by discrete-state Markov chains. Using finite-volume approximations, we derive an approximate master equation directly from the underlying diffusion and show that this discretisation preserves key features of the nonequilibrium steady-state. In particular, we show that the entropy production rate (EPR) of the approximation converges as the number of discrete states goes to the limit. These results are illustrated with analytically solvable diffusions and numerical experiments on nonlinear processes, demonstrating how this approach can be used to explore the dependence of EPR on model parameters. Finally, we address the problem of inferring discrete-state Markov models from continuous stochastic trajectories. We show that discrete-state models significantly underestimate the true EPR.

J. Stat. Mech. (2026) 033205

---

[*] Authors to whom any correspondence should be addressed.

However, we also show that they can provide tests to determine if a stationary planar diffusion is out of equilibrium. This property is illustrated with both simulated data and empirical trajectories from schooling fish.

Keywords: coarse-graining, stationary states, stochastic processes

---

Contents

---

1. Introduction

Many physical processes can be described by the interplay of deterministic forcing and random fluctuations. Such processes are typically described as *diffusive* and have been used to model a plethora of systems including Brownian motion [1], cell motility [2], climate dynamics [3], neural systems [4], and financial markets, amongst others. Mathematically, they can be formulated as Itô *stochastic differential equations* (SDEs), taking the form

$$\mathrm{d}X(t) = f(X(t))\,\mathrm{d}t + \Sigma(X(t))\,\mathrm{d}W(t), \tag{1}$$

where $X(t)$ represents the state of the process at time t, $f(x)$ represents the deterministic forcing, and $\Sigma(x)$ determines the random fluctuations as $W(t)$ is the standard *Wiener process* [5].

It has been argued by Schrödinger that biological systems stave off decay by burning energy to maintain themselves in a *NESS* [6], which has since led to the development of a range of techniques for analysing the properties of a NESS directly from observed, stochastic trajectories of a diffusion [7–10]. Central to many of these techniques, is a *discretisation* of phase-space into discrete *voxels*, where the evolution of a trajectory in a continuum can be approximated as a sequence of discrete states, as shown in figure 1. These methods, and their derivatives, have been used to analyse the nonequilibrium dynamics of many biological systems [7, 8, 11–14]. Nevertheless, such approaches represent a *coarse-graining*, where much of the underlying continuous information is lost. More specifically, it is known that coarse-graining can obscure the properties of a NESS in a complex, and somewhat unpredictable, manner [15]. For example, coarse-graining in both space and time implies that the entropy produced can only be lower bounded from observations of the coarse-grained dynamics [15]. Moreover, coarse-graining introduces memory effects, rendering the dynamics non-Markovian [16]. In turn, Markovian estimates of the entropy produced are also lower-bounds on the true value [17]. Within statistical mechanics, coarse-graining has long been a powerful technique to simplify and reduce complex models through techniques like the *renormalisation group* [18], or *hydrodynamics* [19]. In fact, the stochastic formulation of Brownian motion is itself an abstraction that is obtained by coarse-graining molecular dynamics at a finer scale [20].

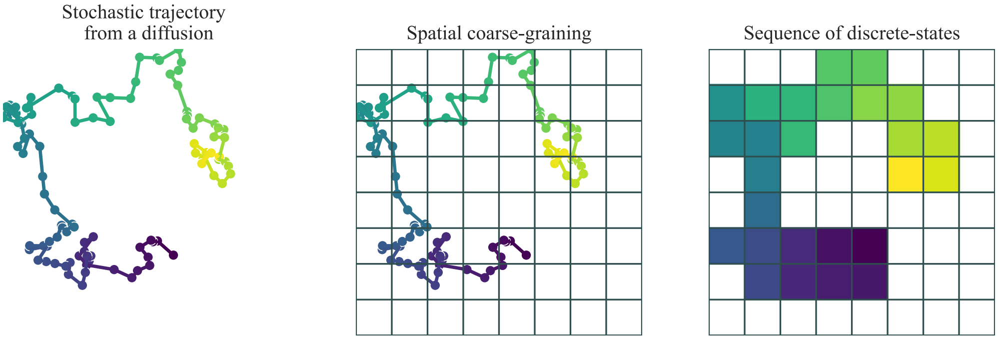


Figure 1. Coarse-graining a stochastic trajectory. A stochastic trajectory from a diffusion can be modelled as a sequence of discrete states by performing a discretisation of state space.

J. Stat. Mech. (2026) 033205

Here, we shed light on the particular problem of coarse-graining a nonequilibrium diffusion with a set of discrete states. Moreover, we focus our work on the properties of the NESS in both the original diffusion and its discrete-approximation. We begin by introducing the mathematical frameworks for both continuous- and discrete-state stochastic processes and their NESS. Following this, we use a *finite-volume approximation* (FVA) to derive an effective discrete-state Markov process from a coarse-grained nonequilibrium diffusion in the plane. We show, analytically, that the entropy production rate (EPR) of a continuous diffusion and its discrete-state approximation converge in the limit, which we confirm with solvable examples. Next, we use our approach to investigate NESS in unsolvable diffusions, including the stochastic van der Pol and frustrated Kuramoto oscillators. We then discuss the statistical inference of a discrete-state process from observed trajectories and show that the underlying distribution of a process can be well-approximated with a simple inference scheme. However, our results show that the EPR of an inferred process is a weak approximation of the true value. Nevertheless, we employ discrete-state models to perform 'testing' of trajectories to indicate if they originate from a process in a NESS. We illustrate this statistical approach on both sampled paths and real-world trajectories from schooling fish, where we determine that the stationary dynamics underlying their movement are not out of equilibrium. Our results provide a general framework for deriving coarse-grained, discrete-state models from an underlying nonequilibrium diffusion. In particular, we show how coarse-grained models can preserve particular features of a NESS, and how such models may be inferred from observed, stochastic trajectories, which has implications for the analysis of biological and physical systems.

2. Stochastic processes and NESSs

We begin by introducing the mathematical frameworks for defining NESSs in both continuous- and discrete-state spaces.

2.1. Diffusions in continuous-space: SDEs and the Fokker–Planck (FP) equation

A continuous-space diffusion, $\{X(t) \in \mathbb{R}^d : t > 0\}$, is defined as the solution to the *SDE*

$$\mathrm{d}X(t) = f(X(t))\,\mathrm{d}t + \Sigma(X(t))\,\mathrm{d}W(t), \tag{2}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ is a vector field that defines the deterministic forcing, known as the *drift*, and $D(x) = \frac{1}{2}\Sigma(x)\Sigma^\top(x) : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is the *diffusion matrix*, which is positive definite [5]. When using spatially-dependent diffusion, we will use the Itô convention for stochastic integration [5][3]. The *density* of the process, $p(x,t)$, is defined as the probability density of the process attaining a value of $x \in \mathbb{R}^d$ at time $t > 0$. Its dynamics are given by the FP equation

$$\partial_t p(x,t) = -\nabla\cdot J(x,t), \tag{3}$$

$$J(x,t) = f(x)\,p(x,t) - \nabla\cdot(D(x)\,p(x,t)), \tag{4}$$

where $J(x,t) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ is known as the *probability flux.*

When the process is *ergodic*, it attains a unique stationary density, $\pi(x)$, which satisfies that the stationary flux is divergence-free, i.e. $\nabla\cdot J_{\mathrm{ss}} = 0$, where

$$J_{\mathrm{ss}}(x) = f(x)\,\pi(x) - \nabla\cdot(D(x)\,\pi(x)). \tag{5}$$

The process is said to be in a *steady-state*. Such a steady-state can either be in or out of thermodynamic equilibrium. If $J_{\mathrm{ss}} \equiv 0$, then the process is in an *equilibrium steady-state* (ESS), otherwise, it is in a *NESS.*

Diffusions in a NESS admit the *Helmholtz–Hodge decomposition* (HHD)

$$f = f_{\mathrm{rev}} + f_{\mathrm{irr}}, \tag{6}$$

$$f_{\mathrm{rev}}(x) = D(x)\,(\nabla\log\pi(x) - \nabla\cdot D(x)), \tag{7}$$

$$f_{\mathrm{irr}}(x) = J_{\mathrm{ss}}(x)/\pi(x), \tag{8}$$

where f_{rev} and f_{irr} represent the *time-reversible* and *time-irreversible* drift respectively [23]. Figure 2 illustrates the HHD and NESS of an Ornstein-Uhlenbeck process in 2D. This implies that they are even and odd under time-reversal respectively.

In keeping with the second law of thermodynamics, NESS thus have time-irreversible dynamics that lead to the production of entropy. The EPR is a key quantity determining the distance of a stationary process from thermodynamic equilibrium. For a diffusion in a NESS, it is given by

$$\Phi = \int_{\mathbb{R}^d} \frac{J_{\mathrm{ss}}^\top(x)\,D^{-1}(x)\,J_{\mathrm{ss}}(x)}{\pi(x)}\,\mathrm{d}x, \tag{9}$$

which is only zero if the process is in a ESS, and is positive otherwise [24]. Using the HHD, the EPR can be alternatively written as

$$\Phi = \int_{\mathbb{R}^d} f_{\mathrm{irr}}^\top(x)\,D^{-1}(x)\,f_{\mathrm{irr}}(x)\,\pi(x)\,\mathrm{d}x, \tag{10}$$

---

[3] It is worth noting that the choice of convention for stochastic integration can lead to important differences in nonequilibrium properties like time-reversibility [21, 22].

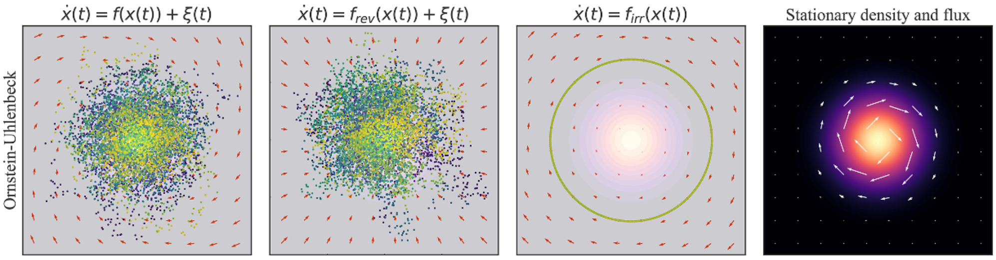


Figure 2. Nonequilibrium steady-state. Processes in a NESS are characterised by the presence of stationary probability flux. A process can be decomposed using the HHD into a reversible component where the drift balances the diffusive fluctuations to maintain the process at stationarity, and the irreversible component drives rotation around the stationary density.

which makes it clear that the EPR only depends on the irreversibility of the process, a central result in stochastic thermodynamics [25].

2.2. Discrete-state processes: continuous-time Markov chains (CTMC)

Discrete-state processes are typically modelled as Markov chains, which can evolve in either continuous or discrete time-steps. We will focus on CTMC. A CTMC, $\{X(t) \in \Omega : t > 0\}$, is a stochastic process taking values in the discrete support Ω which, without loss of generality, can be assumed to be $\Omega \subseteq \mathbb{N}$. Transitions between states occur according to exponentially-distributed waiting times[4]. The distribution of the process, $p_i(t)$, describes the probability that the process is in state i at time t. The vector of such probabilities, $\mathbf{p}(t) = (p_1(t), \ldots)$, evolves according to the *master equation* (ME)

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}(t) = L\mathbf{p}, \tag{11}$$

where L is the *Laplacian* matrix [26]. The off-diagonal entries, $L_{ij} \geqslant 0$, represent the *transition rates* from states j to i. The diagonal entries are then defined to be $L_{ii} = -\sum_{j\neq i} L_{ji}$, which enforces that $\sum_i p_i(t) = 1$ as $\frac{\mathrm{d}}{\mathrm{d}t}(\sum_i p_i(t)) = 0$. Moreover, the waiting distribution in each state is defined to be $\mathrm{Exp}(-L_{ii})$ [27].

If a CTMC is *irreducible* and *aperiodic*, then it is ergodic and will converge to a unique stationary distribution (see reference [28] for definitions), which satisfies $L\pi = 0$. We can define a discrete counterpart to the probability flux vector field

$$J_{ij} = L_{ji}\pi_i - L_{ij}\pi_j, \tag{12}$$

thus the steady-state is in equilibrium if and only if the flux between all pairs of states vanishes [24]. This is more commonly known as the *detailed balance condition*

$$L_{ij}\pi_j = L_{ji}\pi_i. \tag{13}$$

---

[4] One can define a stochastic process with arbitrary, non-exponential waiting time distributions. Later on, we define such processes to be *semi-Markov processes*.

When this condition is violated, the EPR of the NESS can be computed with the Schnakenberg formula[5] [29],

$$\Phi = \frac{1}{2}\sum_{i,j}(L_{ij}\pi_j - L_{ji}\pi_i)\log\frac{L_{ij}\pi_j}{L_{ji}\pi_i}. \tag{14}$$

Similarly, the dynamics of the steady-state are reversible in an ESS and irreversible in a NESS [30].

3. Discrete-state approximations of diffusions

In this next section, we focus on developing and analysing a coarse-grained model of a nonequilibrium diffusion. We will do so by approximating it with a discrete-state Markov chain. Whilst it is important to note that such a coarse-grained model would be non-Markovian, deriving a model with arbitrary waiting-time distributions remains intractable for general multivariate diffusions. Instead, we focus on developing a Markovian approximation of the coarse-grained dynamics using a FVA, thus approximating the FP equation with a ME.

3.1. Non-Markovian approaches

Coarse-graining a Markovian diffusion with discrete-states leads to both memory effects and non-Markovian dynamics, in the same way that coarse-graining a Markov chain via state-aggregation induces memory in discrete-state dynamics [17, 31]. This typically leads to both non-exponential waiting times, as well as non-Markovian transitions[6]. Previous approaches have used first-passage time analysis and asymptotic expansions in an attempt to obtain the waiting-time distributions of a discrete-state approximation of prototypical, univariate diffusions [32–35]. Such an analysis is closely linked to approaches for estimating the EPR from semi-Markov processes[7] using waiting-time distributions [36, 37]. Whilst such approaches more accurately describe the coarse-grained dynamics, they are typically limited in the complexity of the dynamics that they can be applied to. Moreover, the statistical inference of semi-Markov processes from experimental data is more involved, hence data analytic methods typically use a Markovian model [7, 11, 12]. In discrete-time, it is simple to infer higher-order Markov models from sequence data, and thus reveal the effects of memory in real stochastic dynamics [38, 39].

---

[5] At first glance, the discrete and continuous forms of the EPR may not appear to have much in common. However, both are defined as the relative entropy between the forward and backward path-space measures. This can be derived both measure-theoretically, and more heuristically by taking the derivative of the Shannon entropy of either the distribution, in the case of the CTMC, or of the entropy along a stochastic trajectory, in the case of the diffusion. For further details see references [23, 24].

[6] It is important to note that these are separate phenomena. Whilst non-exponential waiting times imply a semi-Markov process, coarse-graining can also induce more disruptive memory effects where transition rates between states depend on the trajectory at large.

[7] A *semi-Markov process* is a discrete-state process with arbitrary waiting-time distributions [28].

Other related works include the coarse-graining of a microscopic Markov process with an SDE in the continuum, or 'diffusion', limit, where the convergence of the EPR has also been studied [40].

3.2. A Markovian approximation

To develop our Markovian approximation, we will consider 2-dimensional diffusions with independent sources of noise,

$$\begin{aligned}\mathrm{d}x(t) &= f^x(x,y)\,\mathrm{d}t + \sigma^x(x,y)\,\mathrm{d}W^x(t),\\ \mathrm{d}y(t) &= f^y(x,y)\,\mathrm{d}t + \sigma^y(x,y)\,\mathrm{d}W^y(t),\end{aligned} \tag{15}$$

which leads to a diagonal diffusion matrix with entries

$$D(x,y) = \begin{pmatrix} D^x(x,y) & 0 \\ 0 & D^y(x,y)\end{pmatrix} = \frac{1}{2}\begin{pmatrix} (\sigma^x(x,y))^2 & 0 \\ 0 & (\sigma^y(x,y))^2\end{pmatrix}. \tag{16}$$

This, in turn, leads to a FP equation of the form

$$\partial_t p(x,y,t) = -\frac{\partial J^x(x,y,t)}{\partial x} - \frac{\partial J^y(x,y,t)}{\partial y}, \tag{17}$$

where the flux field is given by

$$\begin{aligned}J(x,y,t) &= \begin{pmatrix} J^x(x,y,t)\\ J^y(x,y,t)\end{pmatrix},\\ &= \begin{pmatrix} f^x(x,y) - \partial_x D^x(x,y)\\ f^y(x,y) - \partial_y D^y(x,y)\end{pmatrix} p(x,y,t) - \begin{pmatrix} D^x(x,y)\,\partial_x p(x,y,t)\\ D^y(x,y)\,\partial_y p(x,y,t)\end{pmatrix}.\end{aligned} \tag{18}$$

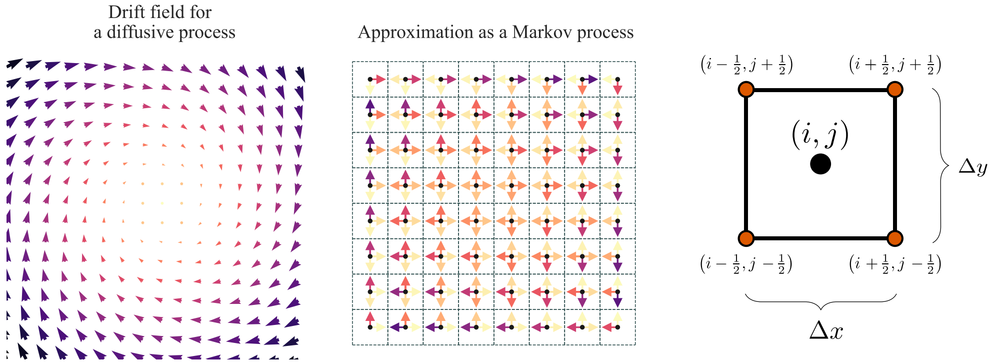
First, we discretise $\mathbb{R}^2$ into a grid of rectangular cells of dimension $\Delta x \times \Delta y$. We denote the midpoint of each cell to be (x_i, y_j), thus its boundary is the rectangle $R_{ij} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$, as shown in figure 3. As described in section 1, the discretisation of a trajectory is performed by assigning a discrete-state to each voxel in the discretised space. In this case, the process is in state $k = (i,j)$ at time t if the underlying diffusion is in the corresponding cell i.e. $(x(t), y(t)) \in B_{ij}$. The corresponding ME describes the evolution of $p_{i,j}(t)$, the probability that the process is in state (i,j).

A FVA of equation (17) yields

$$\frac{\mathrm{d}p_{i,j}}{\mathrm{d}t} = -\frac{J^x_{i+\frac{1}{2},j} - J^x_{i-\frac{1}{2},j}}{\Delta x} - \frac{J^y_{i,j+\frac{1}{2}} - J^y_{i,j-\frac{1}{2}}}{\Delta y}, \tag{19}$$

where the subscript k,l implies the function is evaluated at (x_k, y_l). The expressions on the right of equation (19) represent the discrete flux along the bounding edges of the rectangle. In practice, we assume zero-flux at the boundary, which is chosen sufficiently far from the centre of mass of the stationary density, to simulate an infinite domain.

J. Stat. Mech. (2026) 033205


Drift field for a diffusive process

Approximation as a Markov process

Figure 3. Approximating a diffusion as a Markov process. We aim to derive an approximation of a continuous diffusion as a discrete-state Markov chain using a finite-volume approximation. We use a rectangular grid to coarse-grain $\mathbb{R}^2$ into a set of *volumes*, where we approximate the flux across the boundary, resulting in a CTMC.

It remains to derive expressions for the numerical flux which ultimately leads to a ME of the form

$$\frac{\mathrm{d}p_{i,j}}{\mathrm{d}t} = \sum_{(k,l)} L_{(i,j),(k,l)} p_{k,l}. \tag{20}$$

A numerical scheme yields a valid ME if $L_{(i,j),(k,l)} > 0$ for all off-diagonal entries ($(k,l) \neq (i,j)$), and probability is conserved i.e. the columns of L sum to 0. In reference [41], Strang considered the centred-difference approximation for the FP equation in 1D, which only yields a valid ME under step-size conditions. As a result, the problem of deriving a valid approximation is closely related to that of developing a structure-preserving scheme for the FP equation [42, 43]. We will adopt *Scharfetter–Gummel* (SG) discretisation [44], which alleviates step-size conditions.

3.3. The SG discretisation

The SG discretisation is a structure-preserving scheme for the FP equation that yields the following approximation for the flux across the boundary,

$$J^x_{i+\frac{1}{2},j} = \frac{f^x{}_{i+\frac{1}{2},j}\left(p_{i,j} - \mathrm{e}^{-f^x{}_{i+\frac{1}{2},j}\Delta x/D^x_{i+\frac{1}{2},j}} p_{i+1,j}\right)}{1 - \mathrm{e}^{-f^x{}_{i+\frac{1}{2},j}\Delta x/D^x_{i+\frac{1}{2},j}}}, \tag{21}$$

with analogous expressions for the remaining flux terms [44] (for a derivation, see appendix A). This naturally leads to a ME equation of the form

$$\frac{\mathrm{d}p_{i,j}}{\mathrm{d}t} = L_{(i,j)(i,j)} + \sum_{k=i\pm1} L_{(i,j)(k,j)} + \sum_{k=j\pm1} L_{(i,j)(i,k)}, \tag{22}$$

with off-diagonal transition rates given by

$$L_{(i,j),(i+1,j)} = \frac{{f^x}_{i+\frac{1}{2},j}\,\mathrm{e}^{-{f^x}_{i+\frac{1}{2},j}\Delta x/D^x_{i+\frac{1}{2},j}}}{\Delta x\left(1-\mathrm{e}^{-{f^x}_{i+\frac{1}{2},j}\Delta x/D^x_{i+\frac{1}{2},j}}\right)}, \tag{23}$$

$$L_{(i,j),(i-1,j)} = \frac{{f^x}_{i-\frac{1}{2},j}}{\Delta x\left(1-\mathrm{e}^{-{f^x}_{i-\frac{1}{2},j}\Delta x/D^x_{i-\frac{1}{2},j}}\right)}, \tag{24}$$

$$L_{(i,j),(i,j+1)} = \frac{{f^y}_{i,j+\frac{1}{2}}\,\mathrm{e}^{-{f^y}_{i,j+\frac{1}{2}}\Delta y/D^y_{i,j+\frac{1}{2}}}}{\Delta y\left(1-\mathrm{e}^{-{f^y}_{i,j+\frac{1}{2}}\Delta y/D^y_{i,j+\frac{1}{2}}}\right)}, \tag{25}$$

$$L_{(i,j),(i,j-1)} = \frac{{f^y}_{i,j-\frac{1}{2}}}{\Delta y\left(1-\mathrm{e}^{-{f^y}_{i,j-\frac{1}{2}}\Delta y/D^y_{i,j-\frac{1}{2}}}\right)}, \tag{26}$$

which are necessarily positive[8]. Moreover, the diagonal rates are given by

$$\begin{aligned}
L_{(i,j),(i,j)} = &-\frac{{f^x}_{i+\frac{1}{2},j}}{\Delta x\left(1-\mathrm{e}^{-{f^x}_{i+\frac{1}{2},j}\Delta x/D^x_{i+\frac{1}{2},j}}\right)} - \frac{{f^x}_{i-\frac{1}{2},j}\,\mathrm{e}^{-{f^x}_{i-\frac{1}{2},j}\Delta x/D^x_{i-\frac{1}{2},j}}}{\Delta x\left(1-\mathrm{e}^{-{f^x}_{i-\frac{1}{2},j}\Delta x/D^x_{i-\frac{1}{2},j}}\right)} \\
&-\frac{{f^y}_{i,j+\frac{1}{2}}}{\Delta y\left(1-\mathrm{e}^{-{f^y}_{i,j+\frac{1}{2}}\Delta y/D^y_{i,j+\frac{1}{2}}}\right)} - \frac{{f^y}_{i,j-\frac{1}{2}}\,\mathrm{e}^{-{f^y}_{i,j-\frac{1}{2}}\Delta y/D^y_{i,j-\frac{1}{2}}}}{\Delta y\left(1-\mathrm{e}^{-{f^y}_{i,j-\frac{1}{2}}\Delta y/D^y_{i,j-\frac{1}{2}}}\right)}.
\end{aligned} \tag{27}$$

It is simple to check that the sum

$$L_{(i+1,j),(i,j)} + L_{(i-1,j),(i,j)} + L_{(i,j+1),(i,j)} + L_{(i,j-1),(i,j)} + L_{(i,j),(i,j)} = 0, \tag{28}$$

by replacing i by $i\pm1$ and j by $j\pm1$ into the derived rates. As a result, this scheme yields a valid ME as it has non-negative rates between states and conserves probability.

It is well-known that FVA conserve probability[9] [45]. Moreover, the SG discretisation preserves the stationary distribution and many thermodynamic properties of the continuous solution [46, 47], hence it is well-suited to this application.

3.3.1. Entropy production in the SG discretisation. In this section, we will show that the EPR of the discrete approximation converges to that of the underlying diffusion

---

[8] In cases, where the denominator vanishes, which occurs when the drift vanishes, then we can re-derive the SG discretisation to obtain $L_{(i,j),(i+1,j)} = D^x_{i+\frac{1}{2},j}/\Delta x^2$, and analogous terms for the other transitions [44].

[9] i.e. the numerical solution at every time integrates to 1. This is equivalent to the derived Laplacian having columns that sum to 0.

process. We begin by assuming that the discrete process is stationary with distribution π. As each cell only receives transitions from the four adjacent cells, the EPR of equation (14) reduces to

$$\Phi = \sum_{i,j} \Phi_{(i,j),(i+1,j)} + \Phi_{(i,j),(i-1,j)} + \Phi_{(i,j),(i,j+1)} + \Phi_{(i,j),(i,j-1)}, \tag{29}$$

where each contribution is defined by

$$\Phi_{(i,j)(i+1,j)} = \left(L_{(i,j),(i+1,j)}\pi_{i+1,j} - L_{(i+1,j),(i,j)}\pi_{i,j}\right) \log \frac{L_{(i,j),(i+1,j)}\pi_{i+1,j}}{L_{(i+1,j),(i,j)}\pi_{i,j}}, \tag{30}$$

with analogous expressions for the remaining terms. Substituting in the expressions derived from the SG discretisation, we obtain

$$\begin{aligned}\Phi_{(i,j)(i+1,j)} = {} & \frac{{f^x}_{i+\frac{1}{2},j}}{\Delta x \left(1 - \mathrm{e}^{-{f^x}_{i+\frac{1}{2},j}\Delta x / D^x_{i+\frac{1}{2},j}}\right)} \left(\mathrm{e}^{-{f^x}_{i+\frac{1}{2},j}\Delta x / D^x_{i+\frac{1}{2},j}}\pi_{i+1,j} - \pi_{i,j}\right) \\ & \times \left(-\frac{{f^x}_{i+\frac{1}{2},j}\Delta x}{D^x_{i+\frac{1}{2},j}} + \log \frac{\pi_{i+1,j}}{\pi_{i,j}}\right).\end{aligned} \tag{31}$$

Next, we use the approximation $\pi_{i+1,j} = \pi_{i,j} + \Delta x[\partial_x \pi]_{i,j} + O(\Delta x^2)$ and perform a Taylor expansion, which yields

$$\Phi_{(i,j)(i+1,j)} = \frac{\left(-{f^x}_{i+\frac{1}{2},j}\pi_{i,j} + D^x_{i+\frac{1}{2},j}\left[\partial_x \pi\right]_{i,j}\right)^2}{D^x_{i+\frac{1}{2},j}\pi_{i,j}} + O\left(\Delta x\right). \tag{32}$$

Using that the stationary distribution will converge i.e. $\pi_{i,j} = \Delta x \Delta y \pi(x_i, y_i) + O(\Delta x \Delta y)$, where $\pi(x,y)$ is the stationary density of the diffusion, we have that

$$\Phi_{(i,j)(i+1,j)} + \Phi_{(i,j)(i-1,j)} = \Delta x \Delta y \frac{\left(J^x_{i,j}\right)^2}{D^x_{i,j}\pi_{i,j}} + O\left(\Delta x\right) + O\left(\Delta x \Delta y\right), \tag{33}$$

where J is the probability flux associated with the stationary density π. Thus, the EPR is

$$\Phi = \Delta x \Delta y \sum_{(i,j)} \frac{\left(J^x_{i,j}\right)^2}{D^x_{i,j}\pi_{i,j}} + \frac{\left(J^y_{i,j}\right)^2}{D^y_{i,j}\pi_{i,j}} + O\left(\Delta x\right) + O\left(\Delta y\right). \tag{34}$$

Finally, we can take the limit as $\Delta x, \Delta y$ go to zero,

$$\lim_{\Delta x, \Delta y \to 0} \Phi = \int\!\!\int \frac{\left(J^x\left(x,y\right)\right)^2}{D^x\left(x,y\right)\pi\left(x,y\right)} + \frac{\left(J^y\left(x,y\right)\right)^2}{D^y\left(x,y\right)\pi\left(x,y\right)} \, \mathrm{d}x \, \mathrm{d}y, \tag{35}$$

which is precisely equation (9) in the case of a 2D process with diagonal diffusion matrix i.e. the EPR of the CTMC converges to that of the underlying diffusion process.

3.3.2. Recovering the drift and diffusion coefficients. Given the Laplacian matrix, L, we can recover the drift and diffusion at discrete points in the grid. To show this explicitly, we consider the rates

$$L_{(i,j),(i+1,j)} = \frac{{f^x}_{i+\frac{1}{2},j}\mathrm{e}^{-\xi}}{\Delta x\left(1-\mathrm{e}^{-\xi}\right)}, \qquad L_{(i+1,j),(i,j)} = \frac{{f^x}_{i+\frac{1}{2},j}}{\Delta x\left(1-\mathrm{e}^{-\xi}\right)}, \tag{36}$$

where $\xi = {f^x}_{i+\frac{1}{2},j}\Delta x / D^x_{i+\frac{1}{2},j}$. These are the transition rates from $(i+1,j)$ to (i,j) and its reverse, respectively. We momentarily drop the subscript for convenience, and take the difference in the rates to obtain,

$$L_{(i+1,j),(i,j)} - L_{(i,j),(i+1,j)} = \frac{f^x - f^x\mathrm{e}^{-\xi}}{\Delta x\left(1-\mathrm{e}^{-\xi}\right)} = \frac{f^x}{\Delta x}, \tag{37}$$

thus we recover the drift with

$${f^x}_{i+\frac{1}{2},j} = \Delta x\left(L_{(i+1,j)(i,j)} - L_{(i,j)(i+1,j)}\right). \tag{38}$$

The remaining drift terms can be obtained analogously. To recover the diffusion, we take the log-ratio of the rates to obtain

$$\xi = \log\left(\frac{L_{(i+1,j)(i,j)}}{L_{(i,j)(i+1,j)}}\right) = \frac{f^x\Delta x}{D^x}. \tag{39}$$

We then use our previous expression for the drift to write the diffusion as

$$D^x_{i+\frac{1}{2},j} = \frac{\Delta x^2\left(L_{(i+1,j)(i,j)} - L_{(i,j)(i+1,j)}\right)}{\log\left(L_{(i+1,j)(i,j)}/L_{(i,j)(i+1,j)}\right)}. \tag{40}$$

The remaining diffusion coefficients can be recovered in a similar way.

Remark 3.1 (A variational perspective on convergence from Markov chains to diffusions). It is not immediately obvious that the NESS properties of a CTMC converge to that of the SDE, even when the CTMC is derived from the drift and diffusion of the SDE. In the former, fluctuations have Poissonian statistics stemming from the exponentially-distributed waiting times between transitions, whilst in the latter the fluctuations are driven by Gaussian noise. In order to develop further intuition on this convergence, we consider the *variational approach* to characterising a NESS, which combines macroscopic fluctuation theory [48], and the large deviation principle [49].

It is useful to define the *dissipation potential* and its complex conjugate (Ψ, Ψ^*), which capture the thermodynamic cost of observing a given fluctuation in the flux J. Moreover, at steady-state the EPR is given by $\Phi = \Psi + \Psi^*$ [50–53]. It is well-known that the dissipation potential for a diffusive process is quadratic in the flux, leading to the characteristic quadratic form for the EPR, as in equations (9) and (10). This implies a 'linear response', which, in turn, arises from the Gaussian nature of the fluctuations. On the other hand, the dissipation potential of a CTMC is non-quadratic [50, 53]. However, as discussed and illustrated in appendix C, in the continuum limit the potential, and therefore the EPR, is approximately quadratic in the flux. This argument provides a further intuition for the convergence in section 3.3.1.

4. Example processes and numerical experiments

In this section, we consider our scheme applied to both solvable and unsolvable nonequilibrium diffusions. In the first instance, we confirm with numerical experiments that the stationary distribution and EPR of the discrete process converges for the *Ornstein–Uhlenbeck* (OU) process and stochastic *Hopf oscillator* HO, which have solvable steady-states. Next, we apply our method to investigate the stationary distribution for the stochastic *van der Pol* oscillator and a pair of coupled, frustrated, stochastic *Kuramoto* oscillators, for which there is no analytical form for the steady-state.

J. Stat. Mech. (2026) 033205

4.1. Solvable models

4.1.1. The OU process. The OU process is a multivariate diffusion process that models interacting degrees of freedom under both friction and fluctuations [54]. It is given by the SDE

$$\mathrm{d}X(t) = -BX(t)\,\mathrm{d}t + \Sigma\,\mathrm{d}W(t), \tag{41}$$

where $B \in \mathbb{R}^d \times \mathbb{R}^d$ is the *friction matrix* and $D = \frac{1}{2}\Sigma\Sigma^\top$ is the *diffusion matrix*. If all eigenvalues of B have positive real-parts, the process converges to a zero-mean Gaussian stationary density given by[10]

$$\pi(x) = (2\boldsymbol{\pi})^{-d/2}(\det S)^{1/2}\exp\left(-\frac{1}{2}x^\top S^{-1}x\right), \tag{42}$$

which corresponds to the multivariate normal distribution $\mathcal{N}_d(0,S)$, where S is the covariance matrix satisfying the Lyapunov equation

$$BS + SB^\top = 2D. \tag{43}$$

In general, steady-states of the OU are out of equilibrium. More specifically, the steady-state is in equilibrium if and only if $BD = DB^\top$ which implies $S = B^{-1}D$ [55]. Otherwise, the stationary probability flux is given by

$$J_{\mathrm{ss}}(x) = \mu x\pi(x), \tag{44}$$

where $\mu = DS^{-1} - B$, which then allows for an explicit calculation of the EPR. We first define $Q = BS - D$, which implies that $\mu = -QS^{-1}$ and $Q = -Q^\top$ from the Lyapunov equation, equation (43). With equation (9) and basic results for Gaussian variables, we can obtain $\Phi = \mathrm{Tr}(B^\top D^{-1}Q)$ [23, 54].

In particular, we will consider the 2D system,

$$\mathrm{d}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = -\begin{pmatrix} 2 & -\theta \\ \theta & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\mathrm{d}t + \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}\mathrm{d}W(t), \tag{45}$$

which has covariance $S = \frac{\sigma^2}{4}I$ and $\Phi = \theta^2$.

---

[10] As we denote the stationary density as π, we use $\boldsymbol{\pi}$ to denote the mathematical constant.

4.1.2. The stochastic Hopf oscillator. The *Hopf oscillator*[11] (HO) is a nonlinear process given by the equations

$$\begin{aligned} \mathrm{d}x\left(t\right) &= \left(a - x^2 - y^2\right)x - \omega y\,\mathrm{d}t + \sigma \mathrm{d}W_x\left(t\right), \\ \mathrm{d}y\left(t\right) &= \left(a - x^2 - y^2\right)y + \omega x\,\mathrm{d}t + \sigma \mathrm{d}W_y\left(t\right), \end{aligned} \tag{46}$$

where ω is the natural frequency and σ is the noise intensity. In the absence of noise, for $a < 0$ the system decays to the origin which is an attracting fixed point. At $a = 0$, the system goes through a Hopf bifurcation, resulting in limit-cycle oscillations for $a > 0$ [57]. This system has become a useful model of neural dynamics [58], and hair-cell bundles [59].

To study the NESS of this process in the presence of additive noise, we consider the FP equation in the alternative form

$$\partial_t p = -\left(\nabla \cdot f\right)p - \left(f \cdot \nabla p\right) + \frac{\sigma^2}{2}\nabla^2 p, \tag{47}$$

and replace $\nabla, \nabla\cdot$, and ∇^2 with their expressions in polar coordinates i.e. $x = r\cos\theta$ and $y = r\sin\theta$. We obtain

$$\partial_t p\left(r,\theta,t\right) = \left(4r^2 - 2a\right)p - \left(\left(a - r^2\right)r\frac{\partial p}{\partial r} + \frac{\omega}{r}\frac{\partial p}{\partial \theta}\right) + \frac{\sigma^2}{2}\left(\frac{\partial^2 p}{\partial r^2} + \frac{1}{r}\frac{\partial p}{\partial r} + \frac{1}{r^2}\frac{\partial^2 p}{\partial \theta^2}\right). \tag{48}$$

In order to solve for the steady-state, we use that the limit-cycle behaviour of the deterministic system is radially symmetric, and thus assume that the stationary density is also i.e. $\frac{\partial p}{\partial \theta} = \frac{\partial^2 p}{\partial \theta} = 0$, which implies the distribution takes the form

$$\pi\left(r,\theta\right) = \frac{1}{2\pi}\pi\left(r\right), \tag{49}$$

where $\pi(r)$ satisfies

$$0 = \left(4r^2 - 2a\right)\pi + \left(\frac{\sigma^2}{2r} + r^3 - ar\right)\frac{\mathrm{d}\pi}{\mathrm{d}r} + \frac{\sigma^2}{2}\frac{\mathrm{d}^2\pi}{\mathrm{d}r^2}, \tag{50}$$

which is solved by a *Boltzmann–Gibbs* density[12]

$$\pi\left(r,\theta\right) = \frac{1}{Z}\exp\left(\frac{r^2}{\sigma^2}\left(a - \frac{r^2}{2}\right)\right), \tag{51}$$

where Z is the normalising factor[13]

$$Z = \frac{1}{\sqrt{2}}\exp\left(\frac{a^2}{2\sigma^2}\right)\pi^{3/2}\sigma\left(1 + \operatorname{erf}\left(\frac{a}{\sqrt{2}\sigma}\right)\right). \tag{52}$$

---

[11] Also known as the *Stuart-Landau oscillator* [56].

[12] It is not immediately obvious why a solution of the form of equation (51) solves equation (50). However, it is intuitive that the irreversible, oscillatory circulation is driven by the θ dynamics, which are absent here. This suggests that the dynamics in r may be the negative-gradient of a scalar potential. We can see that $f^r = -\frac{\mathrm{d}U}{\mathrm{d}r}$ for a potential $U = \frac{1}{2}r^2(a - \frac{1}{2}r^2)$. All that remains is to check that $\pi(r) = \exp(-2U/\sigma^2)$ solves equation (50). Whilst the stationary density is a Boltzmann–Gibbs density, the steady-state is in nonequilibrium.

[13] An alternative approach to solving for the stationary density can be found for the case $a = \omega = 1$ in reference [60].

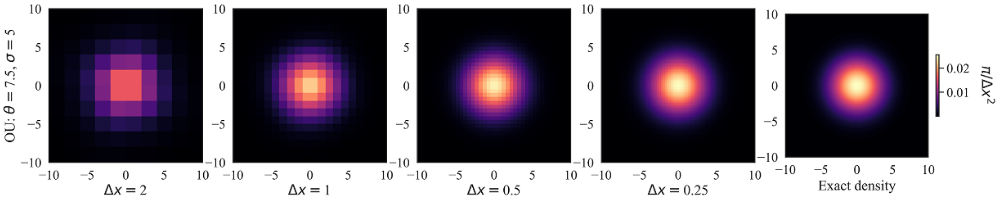


Figure 4. Discrete-approximation of NESS in the OU. For a fixed value of $\theta = 5$, $\sigma = 5$, and a range of discretisation step-sizes, we calculate the stationary distribution of a coarse-grained OU process using the SG discretisation. Even for large step-sizes, the symmetry of the stationary density is preserved, converging to the true density as the step-size decreases.

Given the stationary density, we can perform the HHD to obtain

$$f_{\text{rev}}(x,y) = \begin{pmatrix} x\left(a - x^2 - y^2\right) \\ y\left(a - x^2 - y^2\right) \end{pmatrix}, \tag{53}$$

$$f_{\text{irr}}(x,y) = \begin{pmatrix} -\omega y \\ \omega x \end{pmatrix}, \tag{54}$$

where the irreversible component is driving circular rotation around the stationary density. Moreover, the irreversibility of the process is driven by the frequency of oscillation, and is in equilibrium only when $\omega = 0$. To confirm this, we can use equation (10), to calculate the EPR to be,

$$\Phi = \frac{\pi\omega^2}{\sigma Z}\left(2\sigma + \sqrt{2\pi}\, a e^{\frac{a^2}{4\sigma^2}}\left(1 + \operatorname{erf}\left(\frac{a}{\sqrt{2}\sigma}\right)\right)\right), \tag{55}$$

which scales quadratically in ω. Figure 7 shows examples of the Hopf oscillator in a NESS for $a = -1$ and $a = 2$. Moreover, it shows the EPR as a function of ω and a for a fixed value of σ.

4.1.3. Numerical experiments. We begin our numerical experiments by calculating the stationary distribution for the OU process defined in equation (45) for fixed values of $\theta = 5$ and $\sigma = 5$, whilst varying the discretisation size $\Delta x = \Delta y$ in a finite domain[14]. In figure 4, we see that the stationary distribution of the coarse-grained dynamics preserves the symmetry of the underlying Gaussian distribution, even for larger step-sizes. As the step-size decreases, the distribution converges to the exact density[15].

Next, by varying the parameter θ, we can vary the EPR, $\Phi = \theta^2$, of the underlying OU process. Figure 5 shows the EPR of the coarse-grained process for a range of Δx values as a function of θ. Whilst all the coarse-grained dynamics follow the correct trajectory, increasing the EPR with θ, the smaller step-sizes more closely approximate

---

[14] In this case, we use the domain $[-10, 10]^2$ which we regard as sufficiently far from the centre of mass of the distribution that the no-flux boundaries do not significantly alter the behaviour of the discrete process. For a discussion of boundary conditions and their implementation see appendix A.1.

[15] To compare between a discrete distribution that must *sum* to 1, and a continuous density that must *integrate* to 1, we normalise the discrete distribution π by the area of each volume Δx^2 i.e. the density in cell i can be approximated discretely as $\pi_i/\Delta x^2$.

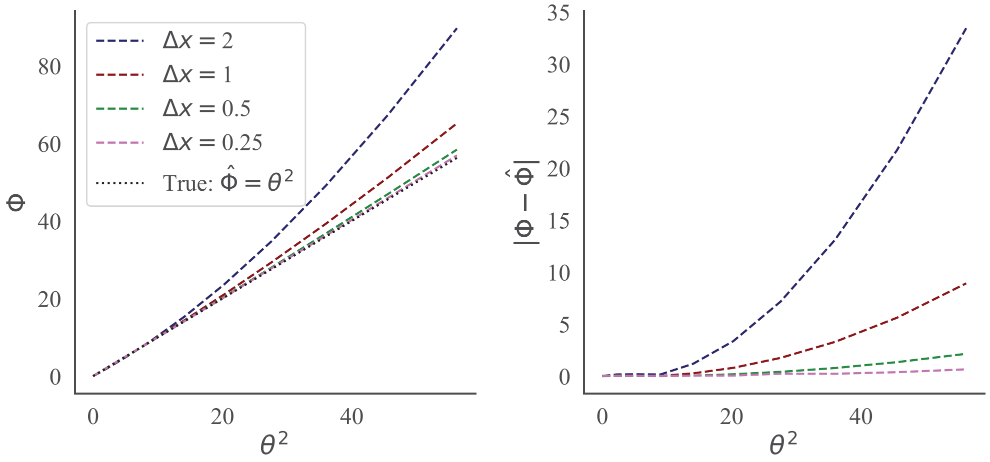


Figure 5. EPR in a coarse-grained OU process. We see that the EPR of the discrete-state approximation also increases approximately quadratically with θ, as in the exact solution. Moreover, we can see that the EPR converges as the step-size goes to 0.

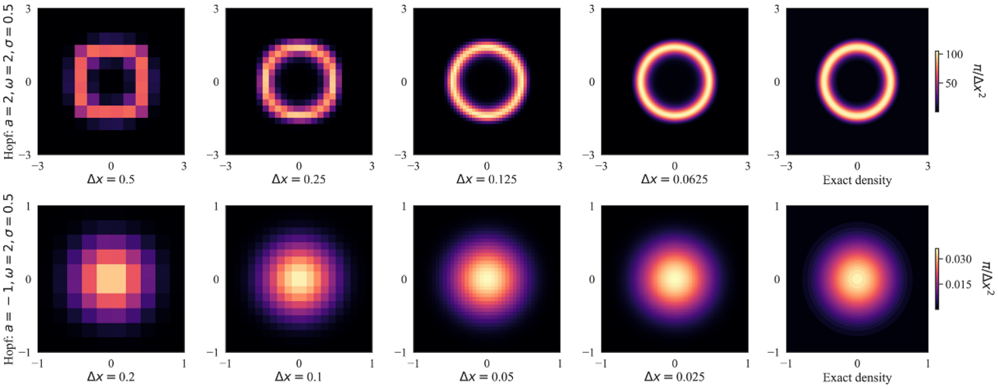


Figure 6. Discrete-approximation of NESS in the Hopf oscillator. For fixed values of $a \in \{-1, 2\}$, $\omega = 2$, $\sigma = 0.5$, and a range of discretisation step-sizes, we calculate the stationary distribution of a coarse-grained Hopf oscillator using the SG discretisation. Even for large step-sizes, the symmetry of the stationary density is preserved, converging to the true density as the step-size decreases.

the true EPR. Moreover, we notice that the EPR of the coarse-grained dynamics may exceed the true EPR. Unlike the EPR of a diffusion observed at a coarse-grained level, the Markov process derived here is not subject to the inequality proved by Esposito in reference [15], where it would necessarily lower-bound the true EPR.

We then turn to the Hopf oscillator, where we perform similar experiments. In figure 6, we consider the stationary distribution of the coarse-grained Hopf oscillator for fixed values of $a = 2$ (top row) and $a = -1$ (bottom row), across a range of step-sizes. As before, the coarse-grained process preserves the symmetry and shape of the

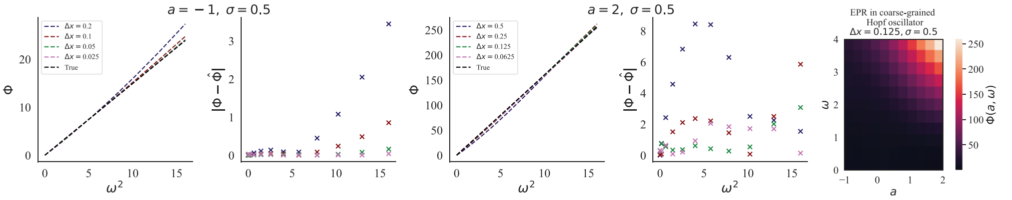


Figure 7. EPR in the coarse-grained Hopf oscillator. For fixed values of a and σ, we can increase ω to increase the EPR of the Hopf oscillator. We show that the EPR of the underlying diffusion is closely approximated by the coarse-grained process. For $a=-1$, we see that the EPR converges as Δx goes to 0. For $a=2$, we use larger step-sizes, and appear to be outside the asymptotic regime. Moreover, we note that for some parameters, calculating the steady-state distribution numerically can induce another source of error not accounted for in the analysis (see appendix A.2). Moreover, for a fixed step-size Δx, we show that the EPR shows the same qualitative behaviour as the true EPR of the underlying diffusion, shown in figure 7.

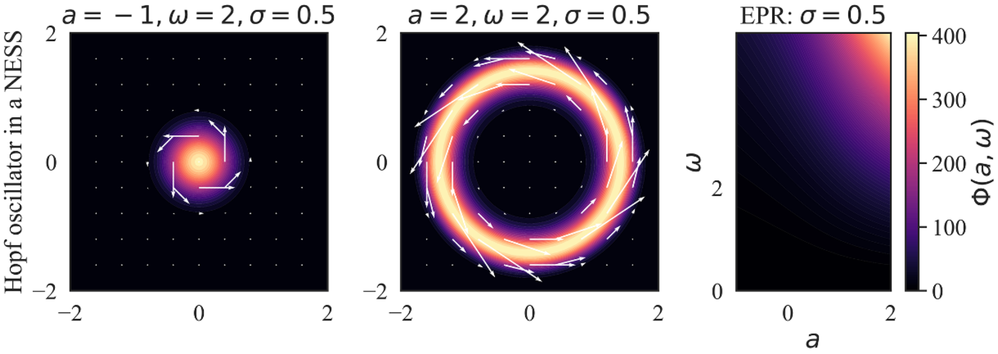


Figure 8. Hopf oscillator in a NESS. The stationary density and flux show that the Hopf oscillator converges to a NESS. For $a<0$, this is a distribution peaked at the origin. For $a>0$, this is a 'Mexican-hat' distribution. Both show rotational probability flux due to the oscillatory dynamics. The EPR varies as a function of ω and a. Both ω and a drive irreversible dynamics, whilst σ drives reversible diffusion.

stationary distribution even for larger step-sizes, for both the peaked and Mexican hat distributions, converging in the limit.

As shown in equation (55), the EPR of the Hopf oscillator increases quadratically in ω. In figure 7, for fixed values of $a=2,-1$ and $\sigma=0.5$, we show that the EPR of the coarse-grained process closely approximates that of the underlying diffusion, converging as the step-size decreases, as expected. Moreover, for a fixed step-size of $\Delta x=0.125$, we explore the EPR of the coarse-grained process as a function of a and ω, showing the same qualitative behaviour as the exact solution in figure 8.

4.2. Unsolvable models

4.2.1. The stochastic van der Pol oscillator (VDP). The VDP is a nonlinear, non-conservative oscillator that converges to an asymmetric limit-cycle in the absence of

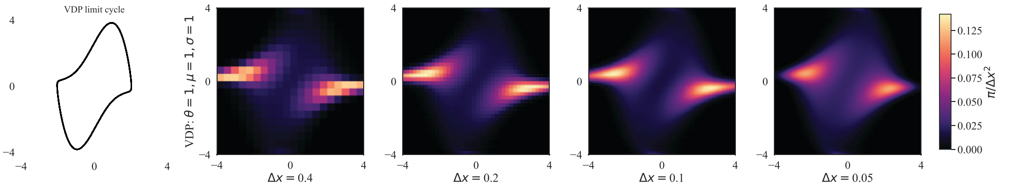


Figure 9. Discrete-approximation of NESS in the VDP. For a fixed value of $\theta = 1$, $\mu = 2$ and $\sigma = 1$, we calculate the stationary distribution of a coarse-grained VDP oscillator using the SG discretisation. Even for small step-sizes, the limit cycle behaviour is present. Moreover, unlike the Hopf oscillator, the rotation is not of uniform speed leading to a build up of probability at the apexes. We consider the domain $[-4,4]^2$.

noise (see figure 9) [57]. We consider a modified, stochastic VDP of the form

$$\begin{aligned} \mathrm{d}x(t) &= \theta y \,\mathrm{d}t + \sigma \,\mathrm{d}W_x(t), \\ \mathrm{d}y(t) &= -\theta x + \mu y \left(1 - x^2\right) \,\mathrm{d}t + \sigma \,\mathrm{d}W_y(t), \end{aligned} \tag{56}$$

where μ is a parameter that controls the nonlinearity and damping. The typical form of the VDP takes $\theta = 1$, which has been added here as a free parameter to control the level of irreversible rotation in the drift field. There is no closed form for the stationary distribution of the VDP, thus we investigate it numerically with our coarse-grained dynamics.

We begin by calculating the stationary distribution for the typical VDP with $\theta = 1$ and $\mu = 2$ for a range of decreasing step-sizes, shown in figure 9. We can see that, even for very coarse dynamics, the limit cycle behaviour is present. Moreover, unlike the Hopf oscillator, rotation around the stationary density does not occur uniformly quickly, thus the distribution is not rotationally invariant around the distribution, with a build up of probability along the apexes of the cycle.

For models without an exact stationary density, it is worth additionally verifying that this is the stationary behaviour seen in the long-time-limit, by numerically integrating the FP equation directly. Using the *Crank–Nicholson* scheme

$$P_{t+1} = B^{-1} A P_t, \tag{57}$$

where $B = I - \frac{\Delta t}{2} L$ and $A = I + \frac{\Delta t}{2} L$ for a time-step Δt [45]. We begin with a initial Gaussian density with zero mean and isotropic deviation $\sigma = 1$ and integrate up to $T = 1.4$ with $\Delta t = 0.2$ and $\Delta x = 0.05$. As shown in figure 10, we see that the FP solution converges to the stationary distribution calculated directly.

The shape and nature of the stationary distribution shifts with variations in parameters. In figure 11, we show three different parameter settings. Increasing μ increases the nonlinearity of the oscillation spreading the distribution over a larger limit cycle, whilst increased θ promotes circular rotation around the origin leading to a concentration of trajectories near the sinusoidal oscillation. Increasing σ increases the diffusivity of trajectories and leads to a wider distribution. Finally, we use our coarse-grained dynamics to investigate the behaviour of the EPR as a function of μ and θ. We sweep

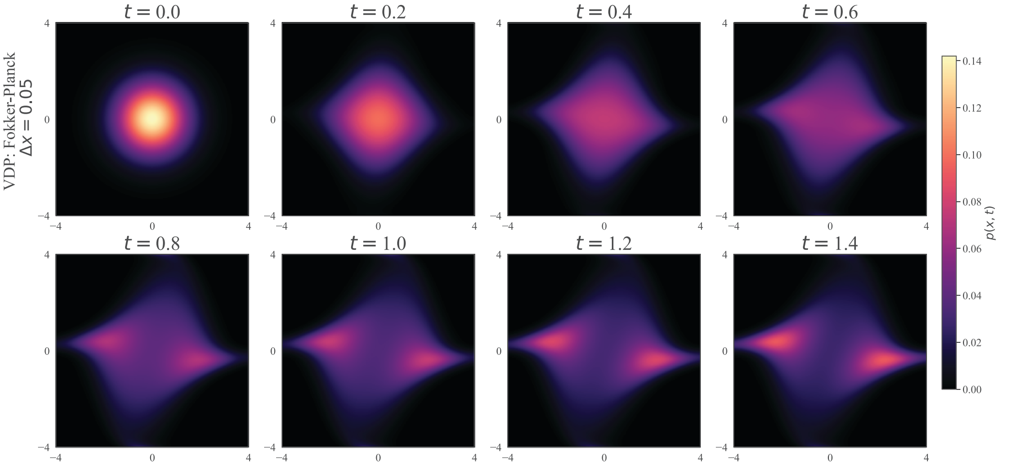


Figure 10. Time-integration of the VDP FP equation. Using the Crank–Nicholson scheme, we integrate the FP equation over time, using our discretised Laplacian. Starting from a Gaussian initial density centred at $(0,0)$ with standard deviation $\sigma = 1$ in each direction, we step through with $\Delta t = 0.2$ up to $T = 1.4$, we can see that the distribution equilibrates to the same distribution calculated directly in figure 9.

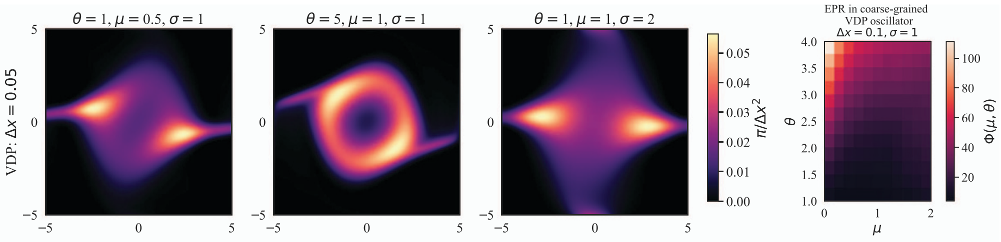


Figure 11. NESS of the VDP oscillator. The shape and nature of the stationary distribution shifts in response to changes in parameters. We illustrate this by calculating the stationary distribution of the VDP for an increased value of $\mu = 3$, $\theta = 5$ and $\sigma = 2$ respectively. An increased value of μ promotes nonlinearity in the oscillation, whilst θ promotes circular rotation around the origin. Finally, σ increases the diffusivity of the process, leading to a less peaked distribution. Additionally, we can investigate how the EPR varies with μ and θ in the coarse-grained dynamics. We find that θ increases the EPR whilst μ decreases it.

between $\mu \in [0,2]$ and $\theta \in [1,4]$ as shown in figure 11. The EPR decreases as a function of μ, which promotes nonlinearity in the oscillations. On the other hand, the EPR increases as a function of θ, which promotes irreversible, circular rotation in the drift field.

4.2.2. A pair of frustrated Kuramoto oscillators. The *Kuramoto* oscillator is a model that described the phase of interacting limit-cycle oscillators [61], and is typically used to model synchronisation. It has found a range of applications describing neural dynamics [58], the flashing of fireflies, or the voltage oscillations of Josephson junctions [57].

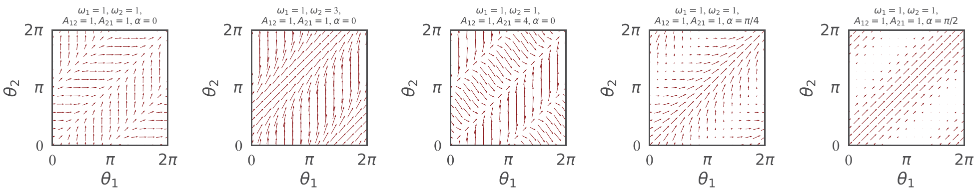


Figure 12. Drift field of coupled Kuramoto oscillators. We consider the drift field of a pair of Kuramoto oscillators. When the model is symmetric, the synchronised state is the only stable equilibrium, but this can be disrupted with asymmetry in the frequencies, interaction strengths or through frustration.

We first consider the case of a single, isolated Kuramoto oscillator

$$\mathrm{d}\theta\left(t\right) = \omega\ \mathrm{d}t + \sigma\ \mathrm{d}W\left(t\right), \tag{58}$$

oscillating at a natural frequency of $\omega > 0$ where $\theta(t) \in [0, 2\boldsymbol{\pi})$. More explicitly, we arrive at the FP equation

$$\partial_t p = -\omega \partial_\theta p + \frac{\sigma^2}{2} \partial_\theta^2 p, \tag{59}$$

which yields the constant stationary density $\pi(\theta) = 1/2\boldsymbol{\pi}$, and stationary flux $J_{\mathrm{ss}}(\theta) = \omega/2\boldsymbol{\pi}$ (Recall that we defined $\boldsymbol{\pi}$ to be ratio of a circle's perimeter to its diameter, whilst π is the stationary density). This leads to an EPR of $\Phi = 2\omega^2/\sigma^2$, from equation (9), but constraining the integral to the domain $[0, 2\boldsymbol{\pi})$. This indicates that a single Kuramoto oscillator is in a NESS, with irreversibility driven by oscillation and mitigated by diffusion.

Next we focus on a system with a pair of Kuramoto oscillators with additive noise. In addition, their interaction is *frustrated*, a feature that can lead to nonequilibrium behaviour [62]. We consider

$$\begin{aligned} \mathrm{d}\theta_1\left(t\right) &= \omega_1 + A_{12}\sin\left(\theta_2 - \theta_1 + \alpha\right)\ \mathrm{d}t + \sigma\ \mathrm{d}W_1\left(t\right), \\ \mathrm{d}\theta_2\left(t\right) &= \omega_2 + A_{21}\sin\left(\theta_1 - \theta_2 + \alpha\right)\ \mathrm{d}t + \sigma\ \mathrm{d}W_2\left(t\right), \end{aligned} \tag{60}$$

where $\theta_i \in [0, 2\boldsymbol{\pi})$ are the phases, $\omega_i > 0$ are the natural frequencies, α is the frustration parameter, and (A_{12}, A_{21}) are the coupling strengths. Figure 12 shows the drift field for a range of different values. In the symmetric case, we see that the synchronised state is a stable equilibrium, which can be disrupted with asymmetry in the natural frequencies, interaction strengths or through frustration[16].

When coarse-graining this Kuramoto model with discrete states, we must enforce the periodic boundaries into our dynamics. Unlike the 'no-flux' boundaries that we considered for the other models, we now enforce a periodic boundary on the discrete lattice

---

[16] In fact, for $A_{12} = A_{21}$ and $\alpha = \pi/2$, the NESS is explicitly solvable. The stationary distribution is uniform, $\pi = 1/4\boldsymbol{\pi}^2$ and the EPR is given by $\Phi = 2(\omega_1^2 + \omega_2^2 + A_{12}^2)/\sigma^2$. This shows that frustration can lead to a NESS, even in the case of symmetric coupling and no internal oscillations.

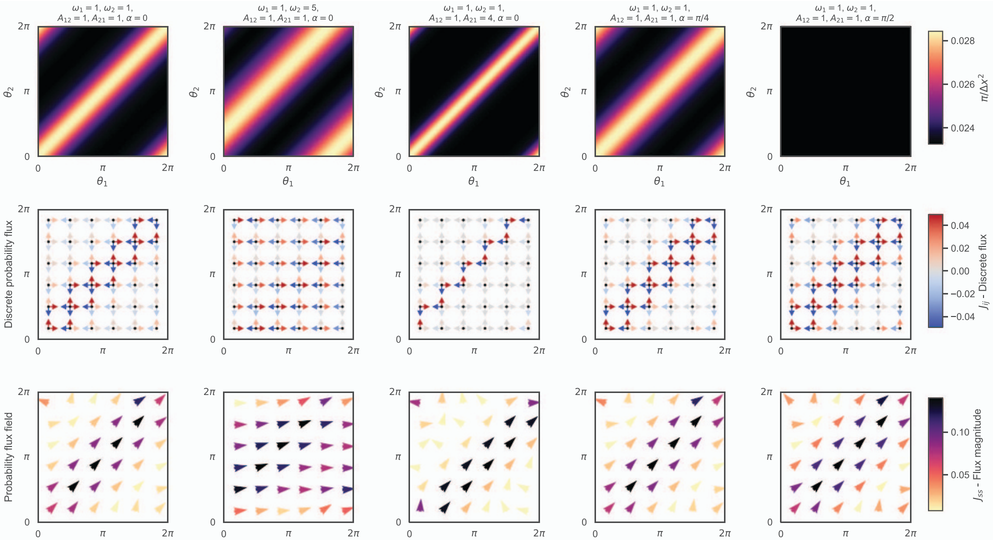


Figure 13. NESS in the Kuramoto model. For a range of different parameter values, we can compute the stationary density, discrete flux and continuous flux field in the NESS. We can see that the coupled Kuramoto system is in a NESS with most of the flux aligned along the synchronised oscillation. Nevertheless, the shape of the stationary distribution depends on the parameters.

such that it approximates $\mathbb{T}^2 = [0, 2\pi)^2$. Our coarse-grained dynamics still preserve probability and therefore still yield a valid ME. Figure 13 shows the stationary distribution, discrete probability flux, and continuous probability flux field for the Kuramoto model in a range of NESS. The parameters can shift the shape and nature of the stationary distribution, and both the density and flux are predominantly aligned with the synchronised oscillation.

Finally, we investigate the EPR as a function of the various parameters. In the first panel of figure 14, we calculate the EPR as a function of various values of the coupling strengths. Whilst the system remains out of equilibrium for symmetric coupling, due to the internal oscillations, asymmetry in coupling strength drives the system further from equilibrium, in keeping with previous results on non-reciprocity in collective dynamics [63, 64]. On the other hand, from the second panel, we see that only the absolute value, not the asymmetry, of natural frequencies drives the system from equilibrium, by increasing the speed of the collective oscillation. Finally, in the third panel, we investigate the effect of frustration and noise. Unsurprisingly, noise reduces the EPR by increasing reversible motion. Interestingly, the effect of frustration does not appear to be constant with respect to noise, with the EPR-maximising value of α varying slightly with the noise intensity, σ.

5. Statistical inference for discrete-state processes

When analysing NESS in physical and biological systems, we typically do not have access to the function forms of the drift or diffusion. Instead, we must analyse a diffusive

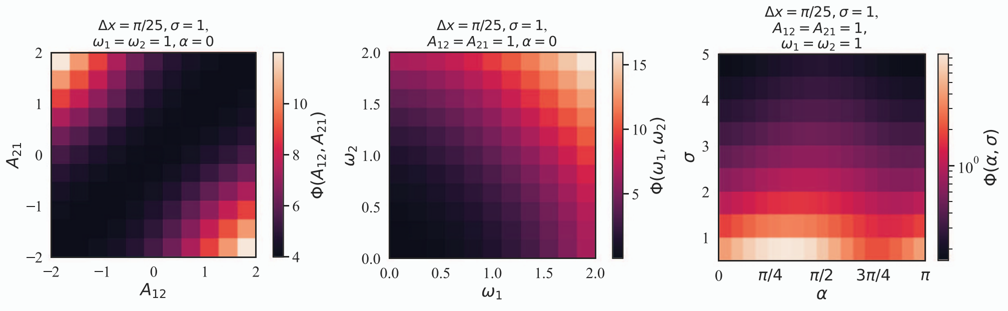


Figure 14. EPR in coarse-grained Kuramoto oscillators. We investigate how the EPR varies as a function of the parameters. We find that coupling asymmetry increases the EPR, whilst the absolute value of the natural frequencies drives the EPR, regardless of asymmetry. Finally, we find that the EPR-maximising value of the frustration depends on the noise-intensity.

process from observations of stochastic trajectories. In order to analyse the observed trajectories of a diffusion, it is common to infer a stochastic model. Whilst some methods attempt to directly infer a SDE [9, 65–67], we will instead focus on the inference of a discrete-state Markov process from coarse-grained observations, which is relevant to a number of applied methods [7, 11–14]. In this case, data takes the form of a sequence of coarse-grained states, but bifurcates into two possible situations.

In the first, we have information about K 'jumps' and data takes the form, $\{(X_k, t_k) : k \in \{1, \ldots, K\}\}$, where X_k is the state the process jumped into at 'jump time' t_k. Given this knowledge about the exact jump times, the *maximum likelihood estimator* (MLE) of the transition rate from state j to state i is given by

$$\hat{L}_{ij} = \frac{N_{j\to i}}{T_j}, \tag{61}$$

where $N_{j\to i}$ is the observed number of jumps from state j to state i that occurred during the K steps, and T_j is the *holding time* i.e. the total time the process spent in state j during the trajectory [68].

In the second, more typical, case, data comes in the form $\{X_t : t = k\Delta t, k \in \{1, \ldots, K\}\}$, where we observe the process at a series of discrete snapshots, typically equispaced. Given such observations in discrete time, equation (61) is no longer the MLE of the transition rate of the CTMC, but instead an approximation whose accuracy depends on the time-step. Instead, we must infer a *discrete-time Markov chain* (DTMC). In discrete time, the MLE of the *transition probability matrix*[17] (TPM), P, is given by

$$\hat{P}_{ij} = \frac{N_{j\to i}}{\sum_k N_{j\to k}}, \tag{62}$$

---

[17] A TPM, P, defines a DTMC, where the entry P_{ij} is the probability that the process will transition from state j to state i in a single time-step [28].

where $N_{j\to k}$ is the number of observed transitions from state j to state k[18][68].

The inference of a CTMC from P is known as the *embedding problem* for Markov matrices [69]. When monitoring a CTMC, with Laplacian L, at discrete time-steps, with interval Δt, the observation is a DTMC with TPM

$$P = \exp\left(L\Delta t\right). \tag{63}$$

Given P, the embedding problem is to find L such that equation (63) is satisfied [69][19]. Importantly, if L solves the embedding problem for $\hat{P}$, then it is the MLE for the CTMC. However, it need not be unique as the matrix exponential is not injective. One simple condition that guarantees the existence and uniqueness of L is

$$\inf_i \left(P_{ii}\right) \geqslant \frac{1}{2}. \tag{64}$$

A looser condition is that

$$\inf_i \left(P_{ii}\right) \cdot \det\left(P\right) > \exp\left(-\boldsymbol{\pi}\right) \prod_i P_{ii}, \tag{65}$$

where L can be calculated using the unique logarithm of P (see reference [68] and references therein). The embedding problem has significant implications for applications with coarse temporal observations including in sociology [70], economics [71], and evolutionary biology [72], where direct *expectation maximisation* algorithms are a sensible approach [68]. In the case of a diffusion and its coarse-graining, temporal granularity is typically high, thus we opt to use the approximate MLE of equation (61), which proved more numerically robust than computing the matrix logarithm.

When inferring the transition rates with equation (61), we only count transitions between adjacent states, discarding transitions between non-adjacent states. In practice, we choose time and spatial resolution parameters such that at least 70% of observed transitions are valid. In appendix E.1, we consider an alternative optimisation-based approach to the embedding problem, which uses all transitions, but show that the results are comparable, whilst the method is significantly less efficient.

5.1. Numerical estimation of the EPR

We sample trajectories from the OU and HO processes (see appendix D.1 and D.2 for sampling methods). Figure 15 shows the stationary distributions of the inferred CTMCs, which are a close approximation of the true density (see section 3). Moreover, we investigate the EPR estimates using trajectories of increasing length and decreasing grid-size. We find that longer trajectories and finer grid-sizes give more accurate estimates of the EPR, which are very stable over 100 iterations. However, in both cases, the inferred EPR is a substantial underestimate of the true EPR, highlighting the challenge of estimating the EPR from finite data, even for fine-grids.

---

[18] In practice, we must also avoid singularities, thus we employ a Bayesian prior. This means we add a 'pseudo-count' for all valid transitions [39].

[19] In our case, $P = \hat{P}$ is calculated from the observations using equation (62).

J. Stat. Mech. (2026) 033205

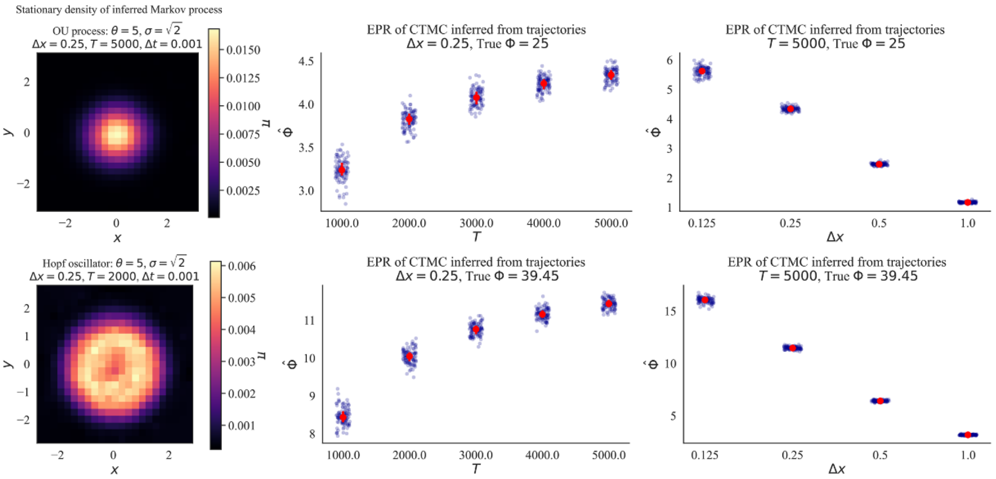


Figure 15. Inferring NESS with Markov chain approximations. Using samples from the OU process in equation (45), we infer a CTMC and calculate the stationary distribution, which is a good approximation of the true density. We also compute the EPR for trajectories of different lengths and using different grid-sizes. We find that using longer trajectories and smaller grids leads to higher accuracy, but that the inferred EPR is substantially lower than the true value.

5.2. Determining if trajectories are from a NESS

A key drawback of inferring the EPR from sampled trajectories is that, by only considering finite data, we may obtain an unreliable estimate of the true EPR. In particular, we may naturally overestimate the irreversibility and flux in the stochastic process, and incorrectly determine that a reversible stationary process is out of equilibrium. For example, if we witness an odd number of transitions between two states, we will naturally observe an asymmetry, even if the true flux between the states is zero. As a result it is necessary to perform ‘surrogate testing’ to determine if the measured irreversibility and EPR is a genuine statistical feature, or an artefact arising from the ‘noise-floor’, the level of irreversibility that exists due to finite data. A typical approach to surrogate testing for time-irreversibility and entropy production is to shuffle trajectories in time, breaking the temporal order, and then re-inferring the EPR to obtain an estimate of the noise-floor [7, 11, 58]. Whilst this approach works for some *direct* measures of time-irreversibility, it is unsuitable for this Markov inference approach. This is because during the inference, we only consider transitions between adjacent cells—choosing the resolution of our grid such that this accounts for a sufficient number of time-points. When the trajectory is randomly shuffled, transitions between neighbouring cells become extremely rare, thus there are not enough valid transitions to infer the surrogate model.

Instead, in order to test if the dynamics of the stochastic data are in fact from a NESS, we analyse the trajectory as before, with one crucial difference. For each valid transition in the trajectory, we flip the direction of the original transition with probability 1/2, thus erasing the net flux in the system that results in entropy production and irreversibility. Then, the EPR of the genuine trajectory can be compared to a

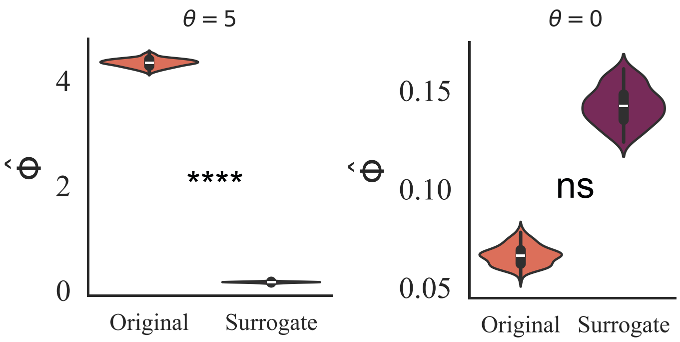


Figure 16. Surrogate testing for NESS. Using the shuffling procedure we can create surrogate Markov models from observed sequence data, and use this to perform testing to identify if real-world trajectories are from an ESS or a NESS. We illustrate this with trajectories from the OU process at $\theta = 5$, a NESS with $\Phi = 25$, and $\theta = 0$, an ESS. Using a one-sided t-test, we find that at $\theta = 5$ the process has significant (****, $p < 0.0001$) entropy production, thus is in a NESS, whilst at $\theta = 0$, the process does not have significant entropy production, thus is not, as expected.

surrogate distribution obtained from different realisations of the surrogate model. Under the null hypothesis that the genuine trajectory is from an ESS, we can test if the trajectory is from a NESS using a one-sided $t-$test. We illustrate this approach on trajectories from the OU process with $\theta = 5$, an irreversible process with $\Phi = 25$, and $\theta = 0$, a reversible process. Figure 16 shows a comparison between the EPR of 100 trajectories sampled from the OU process compared with 100 surrogate models. At $\theta = 5$, we find that the EPR of the original trajectories is significantly higher (one-sided independent t-test; ****, $p < 0.0001$), than the surrogate trajectories, thus we can conclude that the original trajectories are sampled from a NESS. On the other hand, at $\theta = 0$, we find that the original trajectories are not significantly higher than the surrogate models (one-sided independent $t-$test; ns, $p > 0.05$), thus we must conclude, correctly, that the original process is reversible[20].

5.2.1. An application to real-world trajectories: group-polarisation in schooling fish. Finally, we apply this approach to a real-world case of a stationary diffusion, the polarisation trajectories of schooling fish, and determine that such a process is *not* in a NESS. We follow the analysis of Jhawar *et al* [73], who analyse populations of $N = 15, 30$ and 60 fish moving in a tank. In particular, the positions, $\mathbf{x}_i(t)$, and velocities, $\mathbf{v}_i(t)$, of the fish can be extracted from video recordings. The *group polarisation* is an order parameter describing their collective alignment given by

$$\mathbf{M}(t) = \frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{v}}_i(t), \tag{66}$$

---

[20] We note that in this case, the surrogate trajectories appear to have a significantly higher EPR than the original time-series. In theory, this shuffling approach should bring a NESS closer to detailed balance. On the other hand, if a trajectory is genuinely reversible, flipping a number of transitions is likely to *increase* the EPR. Although this effect should dissipate as the number of transitions increases.

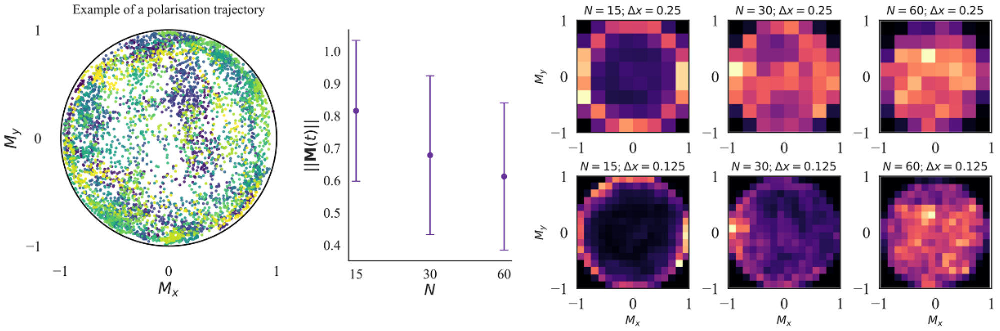


Figure 17. Stationary dynamics in schooling fish. We consider the group polarisation vector $\mathbf{M} = (M_x, M_y)$ for populations of schooling fish of size $N = 15, 30$ and 60. We find that the degree of collective alignment increases as group-size decreases, which can be seen by computing the mean and standard deviation of $||\mathbf{M}(t)||$ over time. Additionally, we can infer a discrete-state Markov process from the trajectories and calculate its stationary density, which again illustrates the higher level of collective alignment for smaller populations of fish.

where $\hat{\mathbf{v}}_i = \mathbf{v}_i/||\mathbf{v}_i||$, is the normalised velocity of fish i. A value of $||\mathbf{M}||$ close to 1 represents a coherent collective direction, whilst $||\mathbf{M}|| \approx 0$ implies no collective direction, and isotropic individual moment. As $\mathbf{M} \in D_1(0) \subset \mathbb{R}^2$, by inferring an SDE, Jhawar *et al* show that the group polarisation vector can be modelled as a stationary planar diffusion [73]. Figure 17 shows an example trajectory of the group polarisation vector. Moreover, using our discrete-state approach, we can infer the stationary densities for different grids and population sizes. In particular, the results shown in figure 17 align with the findings of Jhawar *et al* indicating that 'schooling', collective alignment, increases as population size decreases. This can be illustrated by both the mean value of $||\mathbf{M}(t)||$ over time, as well as the higher densities at the boundary of the unit disk (or its approximation with a square grid) in the steady-state solution of the inferred Markov chain.

Next, we investigate if the stationary dynamics of the schooling fish are best described as a NESS or an ESS. For each of the population sizes, we infer the EPR of the original trajectory as well as the EPR of 100 surrogate models, obtained using the shuffling procedure. Figure 18 shows that, across population size, the EPR is not significantly higher for the original trajectories when compared to the surrogate model, from which we conclude that the process is not in a NESS. This result supports a well-known phenomenon in collective dynamics and active matter. Whilst collective dynamics are inherently *active* at a local level [74], as they dissipate energy in order for agents to adjust their states dynamically, their collective behaviour can be described by ESS dynamics respecting time-reversal symmetry [75] and maximum-entropy distributions [76, 77].

6. Discussion

NESSs are ubiquitous phenomena in physical and biological systems [8, 58, 78]. Approximating the continuous-state dynamics of these diffusions with discrete-state

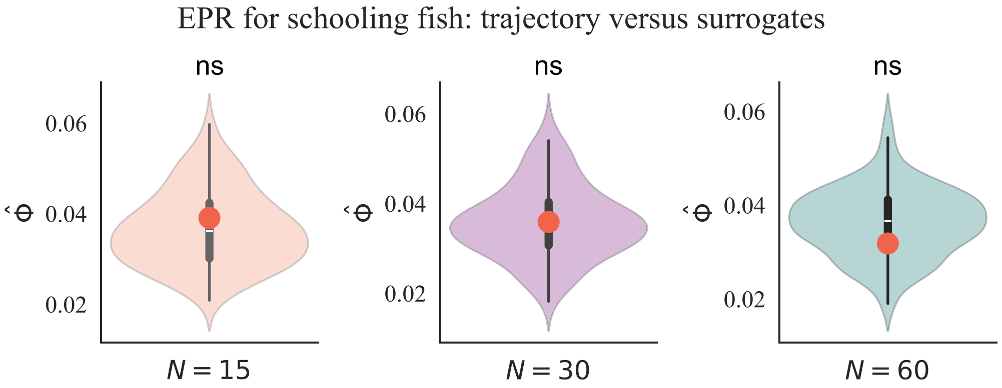


Figure 18. EPR for populations of schooling fish compared to surrogate models. Using the discrete-state model, we can measure the EPR of the trajectories of schooling fish. This is represented by the single orange dot in each panel. Using the surrogate model procedure, we obtain a 'null' distribution (the violin plot) of the EPR in the surrogate model. We perform hypothesis testing to compare the single empirical trajectory with 100 surrogate models and find that the result is not significant across all population sizes.

models is a simple and effective way to analyse their dynamics, and has found applications to physical and biological data at multiple scales [7, 10–12]. Nevertheless, the consistency of the properties of the NESS between the continuous and discrete spaces was not well understood. Moreover, these approaches suffer the complex and obscurant effects of coarse-graining which can hide probability flux leading to underestimates of the EPR [15] as well as induce non-Markovian effects that lead to artefactual scaling relationships [17].

In this paper, we presented a new approach to analytically derive a discrete-state Markov process from stationary planar diffusion, using finite-volume methods. Moreover, we showed that the EPR of the approximation converges in the infinite-state limit, which we illustrated with both solvable and unsolvable examples of stationary diffusions. In the case of unsolvable processes, this approximation gives an accurate way to investigate the behaviour of the EPR under changes to the parameters, which we illustrated with the stochastic van der Pol and Kuramoto models. Finally, we considered the problem of statistical inference from trajectory data. Using sampled paths, our results show that discrete-state models significantly underestimate the EPR of a NESS, even for long trajectories and fine grids. Nevertheless, discrete-state models can be used to define a hypothesis test that can be used to determine if a trajectory is from a NESS or not, which we, again, illustrated with sampled trajectories. Finally, we considered an illustrative application to real-world trajectories by analysing the group-polarisation trajectories from schooling fish [73].

Our theoretical analysis and data-analytic methods offer a much needed framework to investigate the intersection and the consistency of discrete and continuous nonequilibrium stochastic processes. This approach has applications to both theoretical investigations of stochastic processes, as well as the data analysis of nonequilibrium stochastic trajectories.

Acknowledgements

The authors would like to thank Alexander Strang, Tassilo Schwarz and Jean-Charles Delvenne for important conversations about coarse-graining and stochastic processes. Additionally, they would like to thank Endre Süli, José Antonio Carrillo, Maria Bruna, and Jaime Agudo-Canalejo for their advice on finite-volume discretisations. Finally, they would like to thank Jhawar *et al* for making their data on schooling fish publicly available. R.N.K acknowledge support in the form of an EPSRC Doctoral Scholarship from Grants Nos. EP/T517811/1 and EP/R513295/1 and an Enrichment Community Award from The Alan Turing Institute. R L acknowledges support from the EPSRC Grants EP/V013068/1, EP/V03474X/1 and EP/Y028872/1.

Code and data availability

Python code, Jupyter notebooks, and trajectory data are available at https://github.com/rnartallo/coarse_grained_diffusion. The full dataset for the schooling fish has been made available by Jhawar *et al* at https://zenodo.org/records/3632470.

Author contributions

R.N.K completed research and wrote the manuscript. R.L. and A.G. supervised research and edited the manuscript.

Appendix A. The Scharfetter–Gummel (SG) discretisation

To derive the SG discretisation in equation (21), we follow the analysis of reference [44], extending to the 2D case with non-constant drift. We begin with the FVA of the FP equation

$$\frac{\mathrm{d}p_{i,j}}{\mathrm{d}t} = -\frac{J^x_{i+\frac{1}{2},j} - J^x_{i-\frac{1}{2},j}}{\Delta x} - \frac{J^y_{i,j+\frac{1}{2}} - J^y_{i,j-\frac{1}{2}}}{\Delta y}, \tag{67}$$

which is equation (19) of the main text. Next, we assume that the flux $J^x_{i+\frac{1}{2}}$, drift $f^x{}_{i+\frac{1}{2}}$ and diffusion $D^x_{i+\frac{1}{2}}$ are constant over the interval $[x_i, x_{i+1}]$. As a result, $p^j(x) := p(x, y_j)$ satisfies the first-order ODE with constant coefficients

$$J^x_{i+\frac{1}{2},j} = f^x{}_{i+\frac{1}{2},j}\, p^j(x) - D^x_{i+\frac{1}{2},j} \frac{\mathrm{d}p^j}{\mathrm{d}x}(x), \tag{68}$$

with boundary values $p^j(x_i) = p_{i,j}$ and $p^j(x_{i+1}) = p_{i+1,j}$. For readability, we now drop the sub- and superscripts. We can solve this boundary value problem using an integrating factor,

$$J\mathrm{e}^{-f(x-x_i)/D} = \left(fp(x) - D\frac{\mathrm{d}p}{\mathrm{d}x}\right)\mathrm{e}^{-f(x-x_i)/D}, \tag{69}$$

$$= -D\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-f(x-x_i)/D}p(x)\right). \tag{70}$$

We then integrate this over the range $[x_i, x_{i+1}]$ where $\Delta x = x_{i+1} - x_i$. This yields

$$\int_{x_i}^{x_i+\Delta x} J\mathrm{e}^{-f(x-x_i)/D} = -D\left(p(x)\,\mathrm{e}^{-f(x-x_i)/D}\right)\Big|_{x_i}^{x_{i+1}}, \tag{71}$$

$$J\frac{D}{f}\left(1 - \mathrm{e}^{-f\Delta x/D}\right) = D\left(p_{i,j} - p_{i+1,j}\mathrm{e}^{-f\Delta x/D}\right). \tag{72}$$

Reintroducing the subscripts, we obtain

$$J^x_{i+\frac{1}{2},j} = \frac{f^{\,x}{}_{i+\frac{1}{2},j}\left(p_{i,j} - \mathrm{e}^{-f^{\,x}{}_{i+\frac{1}{2},j}\Delta x/D^x_{i+\frac{1}{2},j}}p_{i+1,j}\right)}{1 - \mathrm{e}^{-f^{\,x}{}_{i+\frac{1}{2},j}\Delta x/D^x_{i+\frac{1}{2},j}}}, \tag{73}$$

as in equation (21). A similar derivation can be made for the flux terms $J^x_{i-\frac{1}{2},j}, J^y_{i,j+\frac{1}{2}}, J^y_{i,j-\frac{1}{2}}$. These can then be substituted into equation (17) where the coefficients of $p_{i+1,j}, p_{i-1,j}, p_{i,j+1}, p_{i,j-1}$ and $p_{i,j}$ can be derived.

Whilst not reported here, we found that other structure-preserving schemes such as the *Chang-Cooper* discretisation [79] yielded valid MEs for these diffusions, whilst the centred-difference scheme only did so under step-size conditions. A number of structure-preserving schemes for the FP equation can be parametrised as a single family as described in reference [43].

A.1. Boundary conditions

The derivations of the transition rates in section 3.2 assume an infinite grid that approximates $\mathbb{R}^2$. In practice, this is not possible and we must instead enforce a finite boundary condition. There are three kinds of boundaries we can enforce with this grid-based set up: 1) absorbing 2) reflecting 3) periodic. In the first case, there is a net flux out of the grid, thus the process does not conserve probability. This creates significant problems when attempting to solve for the stationary distribution. In the second case, a reflective boundary implies ‘no-flux’ at the border. This does preserve probability and is the most sensible choice for unbounded processes on $\mathbb{R}^2$. For a cell at the boundary, a transition (or two for the corner cells) is not possible, thus we must correct the diagonal entry of the Laplacian correspondingly in order to conserve probability. In all cases except the Kuramoto oscillator, we opt for this boundary condition. In the case of the Kuramoto oscillator, we have a periodic domain, thus we mirror this with a periodic boundary condition. This preserves probability and transitions are possible from one side of the grid to the other side. The transition rates must be amended accordingly.

A.2. Solving for the stationary distribution

Solving for the stationary distribution involves solving the equation $L\pi = 0$. Whilst for small number of discrete-states this can be calculated numerically in an efficient and exact way, for large state spaces (fine grids) this calculation becomes prohibitively slow. Instead, we solve

$$\min_{\pi:\sum_j \pi_j = 1} ||L\pi||^2, \tag{74}$$

using LSQR [80], which capitalises on the sparsity of L. Whilst our theoretical derivations confirm that entries of π should have the same sign, numerical instabilities may occur. For extremely fine grids, the discretisation error can be comparable to the error incurred through this approximation, thus convergence analysis becomes challenging.

Appendix B. Non-diagonal diffusion

In this paper we focus on the case of diagonal diffusion, and use the SG discretisation with a five-point stencil. This implies that the $\mathrm{d}p_{i,j}/\mathrm{d}t$ depends only on the four adjacent neighbours $\{p_{i+1,j}, p_{i-1,j}, p_{i,j+1}, p_{i,j-1}\}$. Here we will show that using a non-diagonal diffusion matrix requires a nine-point stencil, including the additional terms $\{p_{i+1,j+1}, p_{i-1,j+1}, p_{i+1,j-1}, p_{i-1,j-1}\}$. However, we show that this discretisation does not yield a valid ME. Finally, we consider the use of shifted grids for coarse-graining processes which are diagonal under a coordinate transform, which includes the case of constant diffusion.

B.1. Discretising non-diagonal diffusion with a nine-point stencil

For simplicity, we will consider an SDE with no drift and non-diagonal, spatially-dependent diffusion, where $D(x,y)$ is given by

$$D(x,y) = \begin{pmatrix} D^x(x,y) & D^{xy}(x,y) \\ D^{xy}(x,y) & D^y(x,y) \end{pmatrix}, \tag{75}$$

where D is positive definite. The FP flux is then given by $J = -\nabla\cdot(\mathrm{d}p)$. This process cannot be discretised with a generalisation of the SG discretisation due to these off-diagonal entries producing cross derivatives. We first note that

$$\nabla\cdot(\mathrm{d}p) = \begin{pmatrix} \partial_x(D^x p) + \partial_y(D^{xy} p) \\ \partial_x(D^{xy} p) + \partial_y(D^y p) \end{pmatrix}. \tag{76}$$

To obtain an expression for the discrete flux, we will use traditional centred differences [45], for example we have

$$\partial_x(D^x p)|_{i+\frac{1}{2},j} \approx \frac{D^x_{i+1,j} p_{i+1,j} - D^x_{i,j} p_{i,j}}{\Delta x}, \tag{77}$$

$$\partial_y \left(D^{xy} p\right)\big|_{i+\frac{1}{2},j} \approx \frac{D^{x}_{i+\frac{1}{2},j+1} p_{i+\frac{1}{2},j+1} - D^{x}_{i+\frac{1}{2},j-1} p_{i+\frac{1}{2},j-1}}{2\Delta y}, \tag{78}$$

$$\approx \frac{D^{x}_{i+\frac{1}{2},j+1}\left(p_{i+1,j+1}+p_{i,j+1}\right) - D^{x}_{i+\frac{1}{2},j-1}\left(p_{i+1,j-1}+p_{i,j-1}\right)}{4\Delta y}. \tag{79}$$

Repeating this calculation at each boundary, we can obtain the transition rates between states,

$$L_{(i,j)(i+1,j)} = \frac{D^{xx}_{i+1,j}}{\Delta x^2} + \frac{D^{xy}_{i+1,j+\frac{1}{2}} - D^{xy}_{i+1,j-\frac{1}{2}}}{4\Delta x \Delta y}, \tag{80}$$

$$L_{(i,j)(i-1,j)} = \frac{D^{xx}_{i-1,j}}{\Delta x^2} - \frac{D^{xy}_{i-1,j+\frac{1}{2}} - D^{xy}_{i-1,j-\frac{1}{2}}}{4\Delta x \Delta y}, \tag{81}$$

$$L_{(i,j)(i,j+1)} = \frac{D^{yy}_{i,j+1}}{\Delta y^2} + \frac{D^{xy}_{i+\frac{1}{2},j+1} - D^{xy}_{i-\frac{1}{2},j+1}}{4\Delta x \Delta y}, \tag{82}$$

$$L_{(i,j)(i,j-1)} = \frac{D^{yy}_{i,j-1}}{\Delta y^2} - \frac{D^{xy}_{i+\frac{1}{2},j-1} - D^{xy}_{i-\frac{1}{2},j-1}}{4\Delta x \Delta y}, \tag{83}$$

$$L_{(i,j)(i+1,j+1)} = \frac{D^{xy}_{i+\frac{1}{2},j+1} + D^{xy}_{i+1,j+\frac{1}{2}}}{4\Delta x \Delta y}, \tag{84}$$

$$L_{(i,j)(i-1,j+1)} = -\frac{D^{xy}_{i-\frac{1}{2},j+1} + D^{xy}_{i-1,j+\frac{1}{2}}}{4\Delta x \Delta y}, \tag{85}$$

$$L_{(i,j)(i+1,j-1)} = -\frac{D^{xy}_{i+\frac{1}{2},j-1} + D^{xy}_{i+1,j-\frac{1}{2}}}{4\Delta x \Delta y}, \tag{86}$$

$$L_{(i,j)(i-1,j-1)} = \frac{D^{xy}_{i-\frac{1}{2},j-1} + D^{xy}_{i-1,j-\frac{1}{2}}}{4\Delta x \Delta y}. \tag{87}$$

Examining these rates we see that the positivity constraint is violated e.g. if $L_{(i,j)(i+1,j+1)} > 0$ then $L_{(i+1,j)(i,j+1)} < 0$. As a result, we cannot define a valid Markov chain with this rectangular grid unless $D^{xy} = 0$.

B.2. Diffusion which is diagonal under a coordinate transform

Whilst we have showed that any process with non-diagonal diffusion does not yield a valid ME, we now consider the special case of diffusion which is diagonal under a coordinate transform. Consider a process where the diffusion matrix can be written as,

$$D(x,y) = V\Lambda(x,y)V^{-1}, \tag{88}$$

where $V \in \mathbb{R}^{d\times d}$ is a matrix whose columns are the eigenvectors of D and $\Lambda(x,y) = \mathrm{diag}(\lambda_1(x,y),\ldots,\lambda_d(x,y))$ is the diagonal matrix of eigenvalues i.e. the eigenvalues may change with space, but the eigenvectors remain constant. Clearly such a matrix-valued function can be diagonalised, $\Lambda(x,y) = V^{-1}D(x,y)V$, where the original approach can

J. Stat. Mech. (2026) 033205

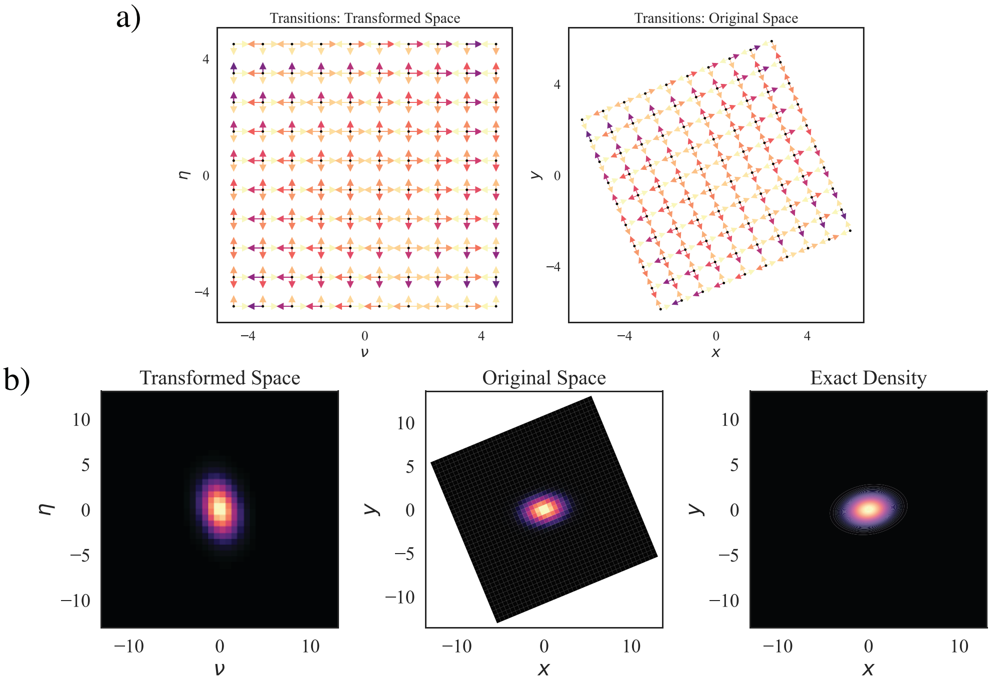


Figure 19. Approximating processes with non-diagonal diffusion. (a) We can apply the original approach, using a square grid, in the transformed space where the diffusion is diagonal. We can then return to the original coordinates where the grid is then rotated. (b) We illustrate this with the stationary distribution of the approximation and the exact density in both the transformed and original space.

be applied in the transformed coordinates, which we call (ν, η) in $\mathbb{R}^2$. This is equivalent to constructing a grid of *rotated rectangles* where the nodes are placed at

$$(x_i, y_j) = \mathbf{v}_1 i\Delta\nu + \mathbf{v}_2 j\Delta\eta, \tag{89}$$

where $\mathbf{v}_1, \mathbf{v}_2$ are the columns of V.

Example: OU process. We illustrate this coordinate transform approach on an OU process with

$$B = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} 4 & 2 \\ 1 & 2 \end{pmatrix}, \tag{90}$$

i.e. where the diffusion is not diagonal, but diagonalisable and constant. We transform into the coordinate system where the diffusion is diagonal, and apply the SG discretisation with a square grid. The transition rates can then be considered as acting on a rotated grid in the original coordinates. Figure 19 illustrates the transition rates (Panel (a)), and stationary distribution (Panel (b)) compared to the exact density in both the original and transformed coordinates.

For a diffusion matrix which does not have constant eigenvectors, it would be best approximated with an adaptive, irregular grid. However, this severely complicated both theoretical and numerical analysis.

Appendix C. A variational approach to nonequilibrium steady-states

Here, we will briefly consider the *energy dissipation principle*, which is central to variational approaches to NESS [48–53].

We begin with a diffusive process of the form in equation (1), and define the *thermodynamic force*,

$$F(x) = \chi^{-1}(x)\, J(x) = D^{-1}(x)\, f_{\mathrm{irr}}(x), \tag{91}$$

where $\chi(x) = D(x)p(x)$ is the *mobility*. This force is the conjugate variable to the flux, allowing us to write the EPR at stationarity in the form $\Phi = \int J_{\mathrm{ss}} \cdot F_{\mathrm{ss}}\, \mathrm{d}x$. Crucially, the variational structure is defined by two convex conjugate functionals. First, we define the *dissipation potential* $\Psi(J)$, which captures the thermodynamic cost of observing a given fluctuation in the flux J [48],

$$\Psi(J) = \frac{1}{2}\int J(x)^{\top} \chi^{-1}(x)\, J(x)\ \mathrm{d}x. \tag{92}$$

Second, we define the *dual dissipation potential*, $\Psi^{*}(F)$, which acts on the forces,

$$\Psi^{*}(F) = \frac{1}{2}\int F(x)^{\top} \chi(x)\, F(x)\ \mathrm{d}x = \frac{1}{2}\int p(x)\, F(x)^{\top} D(x)\, F(x)\ \mathrm{d}x. \tag{93}$$

The central result of the energy dissipation principle is that the physical steady state minimizes the functional $\Psi(J) + \Psi^{*}(F) - \int J \cdot F \mathrm{d}x$. Consequently, at the physical steady state (where $J_{\mathrm{ss}} = \chi F_{\mathrm{ss}}$), the total entropy production decomposes symmetrically into the sum of the potentials i.e. $\Phi = \Psi(J_{\mathrm{ss}}) + \Psi^{*}(F_{\mathrm{ss}})$. It is clear that the EPR has a *quadratic* relationship with the flux, stemming from the quadratic nature of the dissipation potential.

On the other hand, the dissipation potential of a CTMC is non-quadratic [50, 53]. To show this, we consider a CTMC and write the discrete flux, equation (12), as

$$J_{ij} = \Lambda_{ij} \sinh\left(\frac{F_{ij}}{2}\right), \tag{94}$$

where $\Lambda_{ij} = 2\sqrt{L_{ij}\pi_j L_{ji}\pi_i}$ is the discrete analogue of the *mobility*, and $F_{ij} = \log\left(L_{ji}\pi_i / L_{ij}\pi_j\right)$ is the discrete *force* [53]. At stationarity, we have that $\Phi = \frac{1}{2}\sum_{i,j} J_{ij}F_{ij}$[21]. To derive the dissipation potential and its conjugate, we note that $J_{ij} = \frac{\partial \Psi^{*}_{ij}}{\partial F_{ij}}$, thus

$$\Psi^{*}_{ij} = 2\Lambda_{ij}\left(\cosh\left(\frac{F_{ij}}{2}\right) - 1\right). \tag{95}$$

Via the Legendre transform, we have that

$$\Psi_{ij}(J_{ij}) = \sup_{F_{ij}}\left(J_{ij}F_{ij} - \Psi^{*}_{ij}(F_{ij})\right), \tag{96}$$

---

[21] The factor of $\frac{1}{2}$ stems from double counting each directed edge (i, j) which has the same contribution to the EPR.

which is achieved at $F_{ij} = 2\mathrm{arcsinh}\left(\frac{J_{ij}}{\Lambda_{ij}}\right)$, thus we have that[22]

$$\Psi_{ij}\left(J_{ij}\right) = 2J_{ij}\mathrm{arcsinh}\left(\frac{J_{ij}}{\Lambda_{ij}}\right) - 2\Lambda_{ij}\left(\sqrt{1+\frac{J_{ij}^2}{\Lambda_{ij}^2}} - 1\right). \tag{97}$$

The EPR along a directed edge (i,j), $\mathit{\Phi}_{ij}$, can be written as the sum

$$\Phi_{ij} = \Psi_{ij}^*\left(F_{ij}\right) + \Psi_{ij}\left(J_{ij}\right), \tag{98}$$

$$= 2J_{ij}\mathrm{arcsinh}\left(\frac{J_{ij}}{\Lambda_{ij}}\right). \tag{99}$$

Clearly, this is function is not quadratic in the flux. However, in the continuum limit that we consider in section 3.3.1, the quantity $J_{ij}/\Lambda_{ij} \to 0$, as the rates increase. As a result, we can use the expansion $\mathrm{arcsinh}(x) \approx x + O(x^3)$, to obtain

$$\Phi_{ij} \approx \frac{2J_{ij}^2}{\Lambda_{ij}}, \tag{100}$$

where the EPR is a quadratic function of the flux, normalised by the mobility, in direct analogy with the expression for a diffusive process.

Appendix D. Sampling methods for diffusions

D.1. Exact sampling for the Ornstein–Uhlenbeck (OU) process

In section 5, we sample from the OU process. Unlike nonlinear processes, the OU process admits an exact simulation from its transition kernel [23]. Given an OU process

$$\mathrm{d}x\left(t\right) = -Bx\left(t\right) + \Sigma\,\mathrm{d}W\left(t\right), \tag{101}$$

with $D = \frac{1}{2}\Sigma\Sigma^\top$, the solution is given by,

$$x\left(t\right) = \mathrm{e}^{-Bt}x\left(0\right) + \int_0^t \mathrm{e}^{-B(t-s)}\Sigma\,\mathrm{d}W\left(s\right), \tag{102}$$

Therefore, the transition kernel is given by

$$x\left(t\right) \sim \mathcal{N}\left(\mathrm{e}^{-Bt}x\left(0\right), S\left(t\right)\right), \tag{103}$$

where,

$$S\left(t\right) = \int_0^t 2\mathrm{e}^{-B(t-s)}D\left(\mathrm{e}^{-B(t-s)}\right)^\top \mathrm{d}s, \tag{104}$$

thus the process can be simulated exactly.

---

[22] We use the identity $\cosh(\mathrm{arcsinh}(x)) = \sqrt{1+x^2}$.

D.2. Splitting method for the stochastic Hopf oscillator

In section 5, we sample from the Hopf oscillator. This process does not admit exact simulation and we must employ numerical methods to simulate trajectories. We employ a geometric splitting integrator for the Hopf oscillator which preserves the underlying dynamics [81]. Given an SDE,

$$\mathrm{d}x = f(x)\ \mathrm{d}t + \Sigma\,\mathrm{d}W(t), \tag{105}$$

we split the drift term into a nonlinear and linear part where we consider the additive noise alongside the linear drift,

$$\mathrm{d}x^{[1]}(t) = Ax^{[1]}(t)\ \mathrm{d}t + \Sigma\,\mathrm{d}W(t), \tag{106}$$

$$\mathrm{d}x^{[2]}(t) = G\left(x^{[2]}(t)\right)\ \mathrm{d}t, \tag{107}$$

where $f(x) = Ax + G(x)$. The linear part, $x^{[1]}$, defines an OU process which can be simulated exactly. If the splitting is chosen such that the ordinary differential equation for $x^{[2]}$ can be solved explicitly, then the solution of the full process can be approximated with the Strang approximation,

$$x(t+h) \approx \varphi_1^{[h/2]} \circ \varphi_2^{[h]} \circ \varphi_1^{[h/2]}(x(t)), \tag{108}$$

where $\varphi_i^{[h]}$ is a simulation from the exact h-time transition kernel of $x^{[i]}$.

For the Hopf oscillator, we use the decomposition,

$$\mathrm{d}x^{[1]}(t) = \begin{pmatrix} a & -\omega \\ \omega & a \end{pmatrix} x^{[1]}(t)\ \mathrm{d}t + \Sigma\,\mathrm{d}W(t), \tag{109}$$

$$\mathrm{d}x^{[2]}(t) = \begin{pmatrix} -\left(x^2+y^2\right)x \\ -\left(x^2+y^2\right)y \end{pmatrix}\ \mathrm{d}t, \tag{110}$$

where $x^{[2]} = (x, y)$. The nonlinear part can be solved easily in polar coordinates.

Appendix E. Statistical inference

E.1. Solving the embedding problem without discarding transitions

As mentioned, the embedding problem aims to find a Laplacian, L, whose observation in discrete-time has TPM,

$$P = \exp(L\Delta t). \tag{111}$$

Whilst L is assumed to only admit certain transitions, P can have non-zero entries in any position. We can formulate a solution to the embedding problem as a non-convex optimisation problem,

$$L^* = \mathrm{argmin}_{L\in\mathcal{L}} ||P - \exp(L\Delta t)||_F, \tag{112}$$

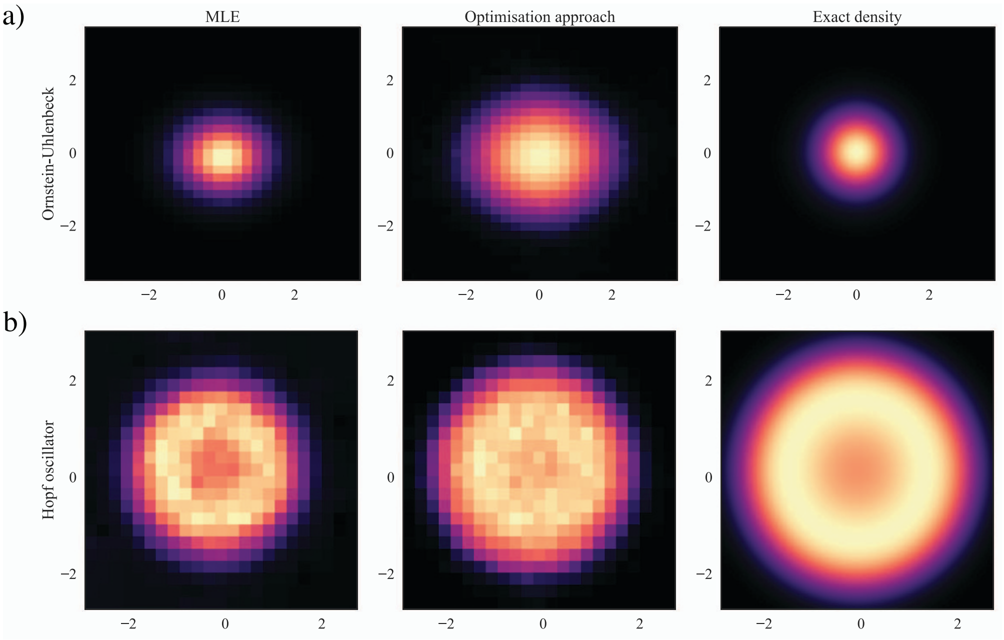


Figure 20. Maximum-likelihood vs. optimisation for the embedding problem. We compare the MLE estimator of equation (61) with the non-convex optimisation approach in equation (112). We reconstruct the Laplacian matrix, and then solve for the stationary distribution, which we compare with the exact density for the (a) Ornstein–Uhlenbeck process, and (b) Hopf oscillator.

where $\mathcal{L}$ is the set of valid Laplacian matrices which conserve probability, and only have non-zero transition rates for valid transitions. Due to presence of the matrix-exponential, this problem is non-convex.

We can compare this approach to the MLE in equation (61), infer the Laplacian, and solve for the stationary distribution. In practice, we use the standard `L-BFGS-B` method from `scipy.optimize`, but we find that this approach is far more inefficient than the direct estimation using equation (61).

Figure 20 shows the result of a numerical experiment comparing the direct MLE with the optimisation-based approach for a trajectory from *a*) the OU process, and *b*) the Hopf oscillator, compared to the exact stationary density. We find that the methods yield comparably accurate stationary distributions.

E.2. Taking the time-step to zero

In section 5, we show that as we increase the length of the stochastic trajectory, or increase the spatial resolution of the grid, our estimate of the EPR improves (increases), as shown in figure 15 . It is tempting to assume that a similar behaviour will occur when we take the limit of the time-step in the stochastic trajectory to zero. However, as shown in figure 21, this is not the case. Whilst we would expect convergence when taking both the time and spatial resolution to the limit, taking the time resolution to the limit whilst keeping the spatial resolution is a more complex scenario. It is true that as the

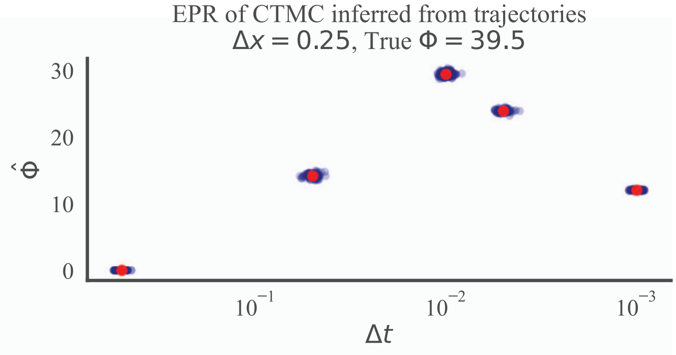


Figure 21. EPR for decreasing time-steps. We infer a CTMC from a trajectory sampled from the Hopf oscillator and measure the EPR, for a series of decreasing time-steps and a fixed spatial resolution. We find that the EPR does not converge as the drift to diffusion ratio goes to zero.

time-step decreases initially, the transition rates are better estimated leading to an improved estimate of the EPR. However, beyond a certain point, this process begins to introduce increasingly prevalent non-Markovian effects. As a result our Markovian estimator becomes increasingly inaccurate.

It is also worth noting that drift scales as $O(\Delta t)$, whilst diffusion scales as $O(\sqrt{\Delta t})$, thus, as the time-step decreases, reversible diffusion dominates irreversible drift leading to increased number of reversible transitions i.e. the per-step drift to diffusion ratio goes to zero.

ORCID iDs

Ramón Nartallo-Kaluarachchi 0000-0003-0396-5783
Renaud Lambiotte 0000-0002-0583-4595
Alain Goriely 0000-0002-6436-8483

References

[1] Brown R 1828 A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies *Phil. Mag.* 4 161–73

[2] Li L, Cox E C and Flyvbjerg H 2011 Dicty dynamics': dictyostelium motility as persistent random motion *Phys. Biol.* **8** 046006

[3] Franzke C L E and O'Kane T J 2017 *Nonlinear and Stochastic Climate Dynamics* (Cambridge University Press)

[4] Harrison L, David O and Friston K 2005 Stochastic models of neuronal dynamics *Phil. Trans. R. Soc. B* **360** 1075–91

[5] Pavliotis G 2014 *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations* (Springer)

[6] Schrödinger E 1944 *What is Life? The Physical Aspect of the Living Cell and Mind* (Cambridge University Press)

[7] Battle C, Broedersz C P, Fakhri N, Geyer V F, Howard J, Schmidt C F and Mackintosh F C 2016 Broken detailed balance at mesoscopic scales in active biological systems *Science* **352** 604–7

[8] Gnesotto F S, Mura F, Gladrow J and Broedersz C P 2017 Broken detailed balance and non-equilibrium dynamics in living systems: a review *Rep. Prog. Phys.* **81** 066601

[9] Frishman A and Ronceray P 2020 Learning force fields from stochastic trajectories *Phys. Rev. X* **10** 021009

[10] Li J, Horowitz J M, Gingrich T R and Fakhri N 2019 Quantifying dissipation using fluctuating currents *Nat. Commun.* **10** 1666

[11] Lynn C W, Cornblath E J, Papadopoulos L and Bassett D S 2021 Broken detailed balance and entropy production in the human brain *Proc. Natl Acad. Sci.* **118** e2109889118

[12] Nartallo-Kaluarachchi R, Expert P, Beers D, Strang A, Kringelbach M L, Lambiotte R and Goriely A 2024 Decomposing force fields as flows on graphs reconstructed from stochastic trajectories *Proc. 3rd Learning on Graphs Conf. (LoG 2024)* vol 269

[13] Paijmans J, Lubensky D K and ten Wolde P R 2017 A thermodynamically consistent model of the post-translational Kai circadian clock *PLOS Comput. Biol.* **13** e1005415

[14] Kimmel J C, Chang A Y, Brack A S and Marshall W F 2018 Inferring cell state by quantitative motility analysis reveals a dynamic state system and broken detailed balance *PLOS Comput. Biol.* **14** e1005927

[15] Esposito M 2012 Stochastic thermodynamics under coarse graining *Phys. Rev. E* **85** 041125

[16] Seifert U 2019 From stochastic thermodynamics to thermodynamic inference *Annu. Rev. Condens. Matter Phys.* **10** 171–92

[17] Schwarz T, Kolomeisky A B and Godec A 2024 Mind the memory: consistent time reversal removes artefactual scaling of energy dissipation rate and provides more accurate and reliable thermodynamic inference, (arXiv:2410.11819v4)

[18] Wilson K G 1983 The renormalization group and critical phenomena *Rev. Mod. Phys.* **55** 583

[19] Marchetti M C, Joanny J F, Ramaswamy S, Liverpool T B, Prost J, Rao M and Simha R A 2013 Hydrodynamics of soft active matter *Rev. Mod. Phys.* **85** 1143

[20] Schilling T 2022 Coarse-grained modelling out of equilibrium *Phys. Rep.* **972** 1–45

[21] van Kampen N G 1981 Itô versus Stratonovich *J. Stat. Phys.* **24** 175–87

[22] Ayala M, Dirr N, Pavliotis G A and Zimmer J 2025 Reversibility, covariance and coarse-graining for Langevin dynamics: On the choice of multiplicative noise (arXiv:2511.03347)

[23] Costa L D and Pavliotis G A 2023 The entropy production of stationary diffusions *J. Phys. A: Math. Theor.* **56** 365001

[24] Jiang D-Q, Qian M and Qian M-P 2004 *Mathematical Theory of Nonequilibrium Steady States: On the Frontier of Probability and Dynamical Systems* (Springer)

[25] Seifert U 2012 Stochastic thermodynamics, fluctuation theorems and molecular machines *Rep. Prog. Phys.* **75** 126001

[26] Gardiner C 2010 *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer)

[27] Masuda N, Porter M and Lambiotte R 2017 Random walks and diffusion on networks *Phys. Rep.* **716-717** 1–58

[28] Ross S M 2019 *Introduction to Probability Models* (Academic)

[29] Schnakenberg J 1976 Network theory of microscopic and macroscopic behavior of master equation systems *Rev. Mod. Phys.* **48** 571

[30] Roldán E 2014 *Irreversibility and Dissipation in Microscopic Systems* (Springer)

[31] Faccin M, Schaub M T and Delvenne J-C 2021 State aggregations in Markov chains and block models of networks *Phys. Rev. Lett.* **127** 078301

[32] Gernert R, Emary C and Klapp S H L 2014 Waiting time distribution for continuous stochastic systems *Phys. Rev. E* **90** 062115

[33] Falasco G and Esposito M 2021 Local detailed balance across scales: From diffusions to JUMP processes and beyond *Phys. Rev. E* **103** 042114

[34] Ghosala A and Bisker G 2022 Inferring entropy production rate from partially observed Langevin dynamics under coarse-graining *Phys. Chem. Chem. Phys.* **24** 24021–31

[35] Meyberg E, Degünther J and Seifert U 2024 Entropy production from waiting-time distributions for overdamped Langevin dynamics *J. Phys. A: Math. Theor.* **57** 25LT01

[36] Martínez I A, Bisker G, Horowitz J M and Parrondo J M R 2019 Inferring broken detailed balance in the absence of observable currents *Nat. Commun.* **10** 3542

[37] Skinner D J and Dunkel J 2021 Estimating entropy production from waiting time distributions *Phys. Rev. Lett.* **127** 198101

[38] Rosvall M, Esquivel A V, Lancichinetti A, West J D and Lambiotte R 2014 Memory in network flows and its effects on spreading dynamics and community detection *Nat. Commun.* **5** 4630

[39] Sahasrabuddhe R, Lambiotte R and Rosvall M 2025 Concise network models of memory dynamics reveal explainable patterns in path data *Sci. Adv.* **11** eadw4544

[40] Busiello D M, Hidalgo J and Maritan A 2019 Entropy production for coarse-grained dynamics *New J. Phys.* **21** 073004

[41] Strang A 2020 Applications of the Helmholtz-Hodge decomposition to networks and random processes *PhD Thesis* Case Western Reserve University

[42] Holubec V, Kroy K and Steffenoni S 2019 Physically consistent numerical solver for time-dependent Fokker-Planck equations *Phys. Rev. E* **99** 032117

[43] Heida M, Kantner M and Stephan A 2021 Consistency and convergence for a family of finite volume discretizations of the Fokker-Planck operator *ESAIM: Math. Modelling Numer. Anal.* **55** 3017–42

[44] Frensley W R 2004 Scharfetter-Gummel discretization scheme for drift-diffusion equations *Technical Report* (The University of Texas at Dallas)

[45] LeVeque R J 2007 *Finite Difference Methods for Ordinary and Partial Differential Equations* (SIAM)

[46] Bessemoulin-Chatard B 2012 A finite volume scheme for convection–diffusion equations with nonlinear diffusion derived from the Scharfetter–Gummel scheme *Numerische Mathematik* **121** 637–70

[47] Schlichting A and Seis C 2022 The Scharfetter-Gummel scheme for aggregation-diffusion equations *IMA J. Numer. Anal.* **42** 2361–402

[48] Bertini L, Sole A D, Gabrielli D, Jona-Lasinio G and Landim C 2015 Macroscopic fluctuation theory *Rev. Mod. Phys.* **87** 593

[49] Touchette H and Harris R J 2013 Large deviation approach to nonequilibrium systems *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond* (Wiley)

[50] Mielke A, Peletier M A and Renger D R M 2014 On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion *Potential Anal.* **41** 1293–327

[51] Mielke A, Renger D R M and Peletier M A 2016 A generalization of Onsager's reciprocity relations to gradient flows with nonlinear mobility *J. Non-Equilib. Thermodyn.* **41** 141–9

[52] Mielke A 2023 Non-equilibrium steady states as saddle points and EDP convergence for slow-fast gradient systems *J. Math. Phys.* **64** 123502

[53] Kaiser M, Jack R L and Zimmer J 2018 Canonical structure and orthogonality of forces and currents in irreversible Markov chains *J. Stat. Phys.* **170** 1019–50

[54] Godrèche C and Luck J 2018 Characterising the nonequilibrium stationary states of Ornstein-Uhlenbeck processes *J. Phys. A: Math. Theor.* **52** 035002

[55] Lax M 1960 Fluctuations from the nonequilibrium steady state *Rev. Mod. Phys.* **32** 25

[56] Matthews P C and Strogatz S H 1990 Phase diagram for the collective behavior of limit-cycle oscillators *Phys. Rev. Lett.* **65** 1701

[57] Strogatz S H 2008 *Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry and Engineering* (Westview Press)

[58] Nartallo-Kaluarachchi R, Kringelbach M, Deco G, Lambiotte R and Goriely A 2026 Nonequilibrium physics of brain dynamics *Phys. Rep.* **1152** 1–43

[59] Sheth J, Meenderink S W F, Quiñones P M, Bozovic D and Levine A J 2018 Nonequilibrium limit-cycle oscillators: fluctuations in hair bundle dynamics *Phys. Rev. E* **97** 062411

[60] Yuan R, Tang Y and Ao P 2017 SDE decomposition and A-type stochastic interpretation in nonequilibrium processes *Front. Phys.* **12** 0201

[61] Kuramoto Y 1984 *Chemical Oscillations, Waves and Turbulence* (Springer)

[62] Gupta S, Campa A and Ruffo S 2014 Kuramoto model of synchronization: equilibrium and nonequilibrium aspects *J. Stat. Mech.* **2014** R08001

[63] Fruchart M, Hanai R, Littlewood P B and Vitelli V 2021 Non-reciprocal phase transitions *Nature* **592** 363–9

[64] Nartallo-Kaluarachchi R, Asllani M, Deco G, Kringelbach M L, Goriely A and Lambiotte R 2024 Broken detailed balance and entropy production in directed networks *Phys. Rev. E* **110** 034313

[65] Friedrich R, Peinke J, Sahimi M and Tabar M R R 2011 Approaching complexity by stochastic methods: From biological systems to turbulence *Phys. Rep.* **506** 87–162

[66] Brückner D B, Ronceray P and Broedersz C P 2020 Inferring the dynamics of underdamped stochastic systems *Phys. Rev. Lett.* **125** 058103

[67] Beheiry M E, Dahan M and Masson J-B 2015 InferenceMAP: mapping of single-molecule dynamics with Bayesian inference *Nat. Methods* **12** 594–5

[68] Bladt M and Sørensen M 2005 Statistical inference for discretely observed Markov jump processes *J. R. Stat. Soc. Ser. B* **67** 395–410

[69] Casanellas M, Fernández-Sánchez J and Roca-Lacostena J 2023 The embedding problem for Markov matrices *Publ. Mát.* **67** 411–45

[70] Singer B and Spilerman S 1976 The representation of social processes by Markov models *Am. J. Soc.* **82** 1–54

[71] Geweke J, Marshall R C and Zarkin G A 1986 Mobility indices in continuous time Markov chains *Econometrica* **54** 1407–23

[72] Verbyla K L, Yap V B, Pahwa A, Shao Y and Huttley G A 2013 The embedding problem for Markov models of nucleotide substitution *PLoS One* **8** e69187

[73] Jhawar J, Morris R G, Amith-Kumar U R, Raj M D, Rogers T, Rajendran H and Guttal V 2020 Noise-induced schooling of fish *Nat. Phys.* **16** 488–93

[74] Toner J and Tu Y 1998 Flocks, herds and schools: a quantitative theory of flocking *Phys. Rev. E* **58** 4828

[75] Fodor E, Nardini C, Cates M E, Tailleur J, Visco P and van Wijland F 2016 How far from equilibrium is active matter? *Phys. Rev. Lett.* **117** 038103

[76] Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M and Walczak A M 2012 Statistical mechanics for natural flocks of birds *Proc. Natl Acad. Sci.* **109** 4786–91

[77] Mora T, Walczak A M, Castello L D, Ginelli F, Melillo S, Parisi L, Viale M, Cavagna A and Giardina I 2016 Local equilibrium in bird flocks *Nat. Phys.* **12** 1153–7

[78] Fang X, Kruse K, Lu T and Wang J 2019 Nonequilibrium physics in biology *Rev. Mod. Phys.* **91** 045004

[79] Pareschi L and Zanella M 2018 Structure preserving schemes for nonlinear Fokker-Planck equations and applications *J. Sci. Comput.* **74** 1575–600

[80] Paige C C and Saunders M A 1982 LSQR: sparse linear equations and least squares problems *ACM Trans. Math. Softw.* **8** 195–209

[81] Buckwar E, Samson A, Tamborrino M and Tubikanec I 2022 A splitting method for SDEs with locally Lipschitz drift: Illustration on the FitzHugh-Nagumo model *Appl. Numer. Math.* **179** 191–220

https://doi.org/10.1088/1742-5468/ae4f7d