PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Jhawar J, Guttal V. 2020 Noise-induced effects in collective dynamics and inferring local interactions from data. *Phil. Trans. R. Soc. B* **375**: 20190381.
http://dx.doi.org/10.1098/rstb.2019.0381

Accepted: 9 February 2020

One contribution of 14 to a theme issue 'Multi-scale analysis and modelling of collective migration in biological systems'.

Subject Areas:
behaviour, ecology, theoretical biology

Keywords:
stochastic differential equations, mesoscopic dynamics, collective behaviour, finite-size effects, noise-induced transitions, fish

Author for correspondence:
Jitesh Jhawar
e-mail: jiteshjhawar@gmail.com

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5025683.

THE ROYAL SOCIETY PUBLISHING

Noise-induced effects in collective dynamics and inferring local interactions from data

Jitesh Jhawar and Vishwesha Guttal

Centre for Ecological Sciences, Indian Institute of Science, Bengaluru 560012, India

JJ, 0000-0002-8774-2351

In animal groups, individual decisions are best characterized by probabilistic rules. Furthermore, animals of many species live in small groups. Probabilistic interactions among small numbers of individuals lead to a so-called *intrinsic noise* at the group level. Theory predicts that the strength of intrinsic noise is not a constant but often depends on the collective state of the group; hence, it is also called a *state-dependent noise* or a *multiplicative noise*. Surprisingly, such noise may produce collective order. However, only a few empirical studies on collective behaviour have paid attention to such effects owing to the lack of methods that enable us to connect data with theory. Here, we demonstrate a method to characterize the role of stochasticity directly from high-resolution time-series data of collective dynamics. We do this by employing two well-studied individual-based toy models of collective behaviour. We argue that the group-level noise may encode important information about the underlying processes at the individual scale. In summary, we describe a method that enables us to establish connections between empirical data of animal (or cellular) collectives and the phenomenon of noise-induced states, a field that is otherwise largely limited to the theoretical literature.

This article is part of the theme issue 'Multi-scale analysis and modelling of collective migration in biological systems'.

1. Introduction

Collective behaviour is an emergent property arising from repeated local interactions among organisms [1–4]. A number of empirical studies over the last decade have offered us novel insights on the challenging problems of characterizing collective motion and on inferring underlying local interactions [5–7]. Much of this success has been possible owing to the availability of high-resolution spatio-temporal data of animal groups in motion, and thus in being able to reconstruct fine-scale movement of organisms. However, many of the studies only consider the *average* or *mean* properties of the group, for example average group polarization or average degree of consensus among group members. Consequently, these studies inadvertently ignore variability of group properties, or more broadly the role of stochasticity. The conventional wisdom dictates that stochasticity often destroys order. However, this is not always the case; stochasticity may sometimes create counterintuitive phenomena in complex systems [8–10] and thus deserves careful attention both in theoretical and in empirical studies.

Stochasticity in collective behaviour arises from a number of factors. Here, focusing only on factors internal to the system, we note that organisms' decisions are likely to be inherently probabilistic, either when acting on their own or when interacting with other organisms. Additionally, animal groups are finite in size, and in many taxa, groups are often relatively small. In such systems, the resulting group-level stochasticity, also called the *intrinsic noise*, can produce nontrivial collective dynamics [8,11–13].

We illustrate this concept with a simple example. Consider a colony of ants choosing between two equally good nests [12]. Assume a simple scenario in

which each ant may either pick one of the two nests randomly or copy the nest choice of a randomly chosen ant. Clearly, there is no preference for ants to pick one nest over the other. We may, therefore, expect that ant colony members will be divided equally between the two choices and hence fail to arrive at a consensus. However, such an expectation is true only when the colony size is very large, formally called the *deterministic limit*. Theory predicts that if we account for stochasticity in the system, the colony does reach a consensus, but only when the colony size is smaller than a threshold value [12]. This consensus is possible, intriguingly, because smaller groups exhibit more fluctuations. Therefore, in the physics literature, the collective order or consensus in this simple system is also called *(intrinsic-) noise-induced order* [11,12].

The literature on noise-induced collective behaviour is relatively small and remains largely theoretical. Apart from a recent work which demonstrates that schooling in fish is a noise-induced state [14], empirical work analysing stochasticity and its role in shaping collective behaviour remains at the margins of collective behaviour research [15–19]. Given that many animals live in small groups and that behavioural interactions are inherently stochastic, we assert that there is a vast scope for applying these intriguing theoretical ideas to empirical research on collective behaviour.

In this article, we describe a method to characterize intrinsic noise in collective dynamics of animal groups [15,18,19]. We argue how such an analysis may also help us reveal local interactions that underlie the emergent patterns of collectives. The method can be applied to a highly resolved time series of the collective state of interest; for example, the collective state could be group polarization (or group consensus), which quantifies the degree of directional alignment (or agreement among many choices) among group members. The method we describe can be traced to van Kampen [20,21] in the general context of stochastic processes in physics and chemistry, but was later developed further [13,22,23] and applied even in some biological studies [15,24–27]. However, many important issues about the method—especially, the appropriate time scale needed to characterize such dynamics—although crucial, remain unresolved. Here, we aim not only to address such methodological issues but also to open up the potential role of stochasticity in collective dynamics of biological systems.

2. Noise-induced collective behaviour: a brief introduction

In the field of collective behaviour, we are interested in how individual-level interactions (which are often stochastic) scale to emergent collective properties. To understand the role of noise in collectives, we employ the so-called *mesoscopic* models; this refers to a description of collective dynamics at an intermediate scale while explicitly accounting for the finite size (N) of the groups. At this (group-level) scale, probabilistic interactions produce a mean effect on the dynamics of a collective state. In addition, owing to finite size of groups, there could be substantial variations (or 'noise') around the mean effect. Typically, noise is expected to merely create fluctuations around the mean (e.g. a Gaussian distribution). However, when such *group-level noise creates new states* in the system, they are called noise-induced states [8].

In formal terms, mesoscopic dynamics of a collective state of the group, denoted by m, may be written in terms of a stochastic differential equation (SDE) [11,12,18]

$$\dot{m} = f(m) + g(m)\eta(t), \tag{2.1}$$

(understood according to the Itô convention [21]) where $\eta(t)$ represents the noise term. Some key assumptions of this framework are:

(i) The underlying process is Markovian, i.e. the current state depends only on the previous state but not on how the previous state was reached.
(ii) The noise $\eta(t)$ is uncorrelated and follows a Gaussian distribution with mean zero and unit variance ($\langle\eta(t)\rangle = 0$; $\langle\eta(t)\eta(t')\rangle = \delta(t - t')$).
(iii) The mesoscopic state of the collective can be described via a single coarse-grained dynamical variable (m).
(iv) The spatial extent of the group is sufficiently small or the system is fully connected.

In the Discussion, we will revisit the above assumptions of our framework in the context of applications to real systems.

In equation (2.1), the first term $f(m)$, or more generally the *deterministic term*, arises from the mean effect of individual-level probabilistic interactions among group members. On the other hand, the *stochastic term* $g(m)$ is a consequence of variations, typically due to finite size of the system, around this mean.

In very large groups ($N \to \infty$), the stochastic term can be ignored and only the deterministic term $f(m)$ drives the collective dynamics. In this limit, also called the thermodynamic limit in the physics literature, collective states are given by the stable fixed points of the ordinary differential equation $\dot{m} = f(m)$.

For finite-sized groups, however, the stochastic term $g(m)$ is proportional to $1/\sqrt{N}$; thus, the strength of the stochastic term is not negligible for smaller groups. We say that a system exhibits a *noise-induced state* when the dynamics of the finite-sized collective are qualitatively different from its deterministic limit (box 1).

We emphasize that the noise in the SDE (equation (2.1)) is at the mesoscopic or group level. Therefore, a noise-induced state refers to a nontrivial state arising from group-level noise. Furthermore, the noise-induced state is not merely a spread/fluctuations observed around the deterministic stable state but is a new state that is absent in the deterministic limit (see box 1 for an example). Furthermore, a mere presence of noise at the level of individuals need not create a noise-induced state. It is often an interplay of deterministic and stochastic terms at the group level that creates a noise-induced state. In the context of collective behaviour, if a group-level noise ($g(m)$) creates order (e.g. collective motion or consensus) that was absent in the deterministic limit, we say the system exhibits *noise-induced order*.

We demonstrate these principles using two simple individual-based non-spatial stochastic models of collective behaviour from the literature [12,18]. Here, individual rules are described via stochastic interaction rates (or probabilities). The models we have chosen have contrasting collective properties, with the collective order being driven *stochastically* (i.e. noise-induced order) in one model whereas it is being driven *deterministically* in the other model.

Box 1. Noise-induced order.

Consider m to be a quantitative descriptor of collective order, such as degree of consensus or polarization among group members, as described in §2a. Consider a hypothetical dynamic of m given by the stochastic differential equation

$$\dot{m} = -\alpha m + \sigma\,\eta(t), \tag{2.2}$$

where α is a constant and $\eta(t)$ represents uncorrelated Gaussian noise with mean zero (i.e. $\langle\eta(t)\rangle = 0$ and $\langle\eta(t)\eta(t')\rangle = \delta(t - t')$, where $\delta(t)$ is a Dirac-delta function) and σ is the strength of the noise. This equation is also known as the Langevin equation or Ornstein–Uhlenbeck process.

In the deterministic limit, i.e. $\sigma = 0$, the only fixed point of the system is $m^* = 0$ (i.e. disorder) and is stable (figure 1*a*). When σ is a nonzero constant, i.e. independent of m (figure 1*b*), it is referred to as the *additive noise*. Perturbations to the fixed point arising from the additive noise term are damped because the deterministic term pulls the system back to the fixed point (i.e. $m^* = 0$). Therefore, for all finite σ, the steady-state probability density function $\mathcal{P}(m)$ shows a mode at zero with a width proportional to σ (figure 1*c*). In other words, the additive noise plays the expected role of merely 'adding noise' to the deterministic stable state.

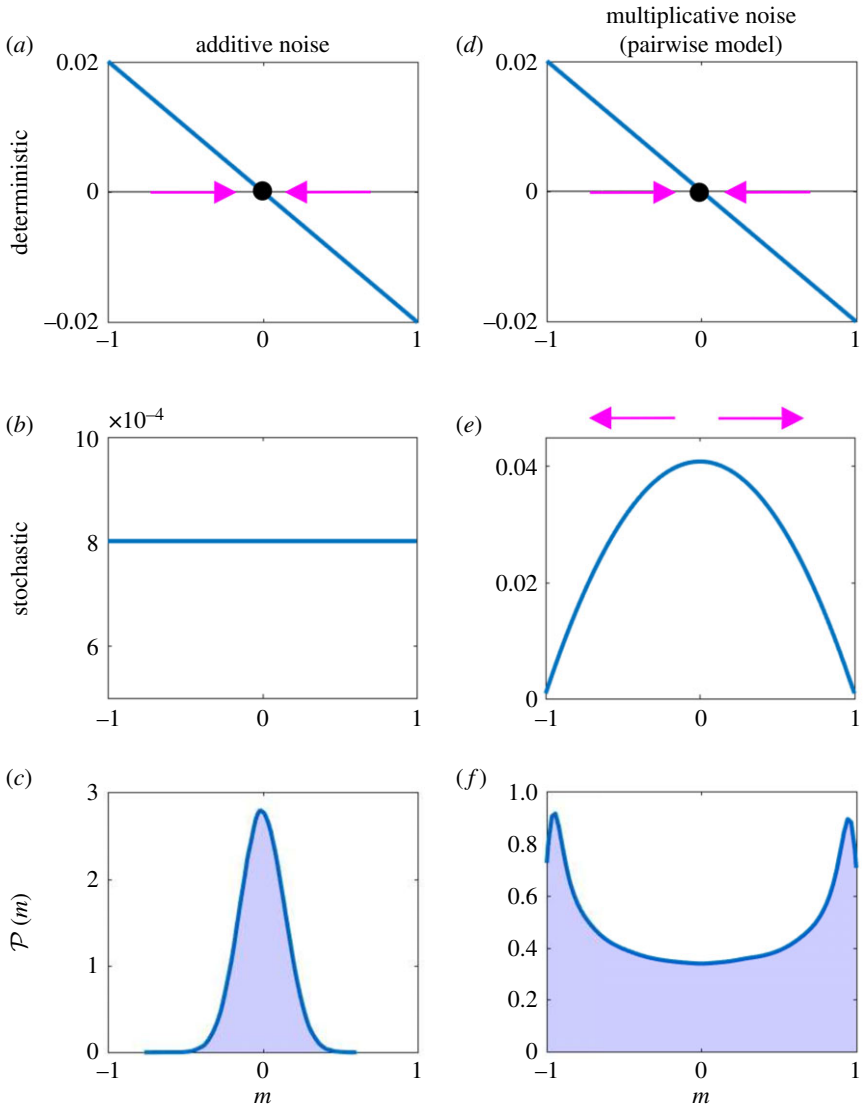

Figure 1. An example of how noise can create new ordered states. In the first column, corresponding to equation (2.2), stochasticity merely creates a distribution around the deterministic stable state of disorder ($m^* = 0$). In the right column, corresponding to equation (2.3), stochasticity creates new states, i.e. modes or most likely states, around $m = \pm 1$. (Online version in colour.)

Let us now consider the dynamics of the collective state given by the stochastic differential equation [11,12,14],

$$\dot{m} = -\alpha m + \sigma\sqrt{\alpha + \beta(1 - m^2)}\,\eta(t), \tag{2.3}$$

where α and β are constants. This equation is inspired by the mesoscopic dynamics of the pairwise copying model introduced in §2a. Here, the deterministic part is identical to that of equation (2.2) (figure 1*d*) and hence pulls the system towards

disorder ($m = 0$). However, the strength of the noise depends on the current value of the state ($m(t)$) and is also referred to as *state-dependent* or *multiplicative noise* (figure 1*e*).

Here, when the system approaches the deterministic stable state of $m = 0$, the noise strength is highest and thus, pushes the system away from disorder, $m = 0$. Consequently, when σ is above a threshold value, the most likely states of the system are in the proximity of $m = \pm 1$ (figure 1*f*). These new most likely states in the probability density function $\mathcal{P}(m)$, which were absent in the deterministic limit, are called noise-induced states. In this case, where m refers to collective dynamics, we refer to the most likely states ($m = \pm 1$, corresponding to a consensus or group order) as the *noise-induced order*.

(a) Individual-based models of binary choice and their mesoscopic descriptions

We consider a simple scenario of decision making in a binary choice set-up. Binary choices, for example, can be used to represent the nest (or food) choice of ants, as described in the Introduction. We emphasize that we have deliberately chosen a simple framework—a non-spatial system with only two states—for our study since our intention is to highlight the key principles of noise-induced states and to demonstrate a method of how to infer noise-induced states from data. Despite the simplicity, the model can be applied to contexts of decision making and even collective motion—for example, a group moving in an annulus. Indeed, this model and its extensions have been applied in a wide range of contexts, such as marching locusts [18], fish schooling [14], decision making in animals [28,29] recruitment of cell signalling molecules [30] and even financial markets [31,32].

Here, each individual of a group of finite size N can be in one of the two states X_1 or X_2, representing their choice of the nest 1 or 2, respectively. We denote the proportion of group members choosing nest i ($i = 1, 2$) as $x_i = N_i/N$, where N_i is simply the number of individuals choosing the nest i.

The collective state of interest (also termed the *order parameter*) is the degree of consensus among group members, defined as $m = x_1 - x_2$. A high degree of consensus (collective order) corresponds to either $m = \pm 1$. The disordered state, in which group members are split between two nests and hence do not arrive at a consensus, corresponds to $m = 0$.

We now define two models in which group members attempt to arrive at the consensus via different sets of microscopic rules.

Pairwise copying model: In this model, individuals update their states via two mechanisms. First is a *spontaneous switching* where individuals change their state randomly, i.e. with no interactions with other group members, at a rate r_1. Using the notation of chemical reactions, this may be written as

$$X_1 \xrightarrow{r_1} X_2 \tag{2.4a}$$

and

$$X_2 \xrightarrow{r_1} X_1, \tag{2.4b}$$

showing that spontaneous switching is unbiased. Second is the *pairwise copying interaction*, where a focal individual, at a rate r_2, copies the state of a randomly chosen individual from the rest of the group. In terms of chemical reactions, this may be written as

$$X_1 + X_2 \xrightarrow{r_2} 2X_1 \tag{2.5a}$$

and

$$X_2 + X_1 \xrightarrow{r_2} 2X_2, \tag{2.5b}$$

which appears to create consensus among individuals [33–37], but nevertheless remains unbiased between two choices.

With these individual-level probabilistic rules, we now turn our attention to the dynamics of the collective, which are the degree of consensus (m) among individuals for this model. One approach to investigating collective behaviour in these models is via computer simulations of the above probabilistic rules. However, as discussed earlier, we need the analytical framework of SDEs to decipher the role of stochasticity (box 1). Recall that this in turn requires a mesoscopic description of collective dynamics, which accounts for both probabilistic interactions and finite group sizes, via stochastic differential equations. We refer the mathematically inclined readers to [12,16–18] (also see [11] for a pedagogical review) for further details on deriving mesoscopic models of collective behaviour.

For the pairwise copying model, the mesoscopic dynamics of m follow the stochastic differential equation [12]

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -2r_1 m + \frac{1}{\sqrt{N}}\sqrt{2r_1 + (1-m^2)r_2}\ \eta(t), \tag{2.6}$$

where $\eta(t)$ represents uncorrelated Gaussian noise.

In this SDE, the first term captures how the dynamics of consensus are shaped deterministically (i.e. mean effect), in a putative $N \to \infty$ limit. The second term captures the residual stochasticity associated with behaviour of the finite group size. The above equation can be solved analytically to obtain the steady-state probability density function of m [12]. Here, we focus on the intuition of dynamics driven by the above two terms.

In the limit of large group sizes ($N \to \infty$), where the stochastic term becomes negligible, the dynamics of order are given by $\dot{m} = -\alpha m$. This is a simple and well-known differential equation whose stable solution is $m^* = 0$. Any perturbation $|m| > 0$ decays exponentially to $m = 0$. In other words, any degree of consensus ($|m| > 0$) will quickly decay ($|m| \to 0$) and the system becomes disordered. Hence, the deterministic (or the large group size) limit of the system does not admit consensus within groups.

By contrast, for small group sizes the magnitude of the stochastic term—given by $1/\sqrt{N}\sqrt{(2r_1 + r_2(1-m^2))}$—is not negligible. Moreover, stochasticity is maximum when the group is disordered ($m = 0$) while it is least when there is consensus ($|m| = 1$). Consequently, when N is sufficiently small, stochasticity pushes the system away from the disordered state at a rate that is larger than the rate of deterministic pull towards disorder. Thus, the system achieves consensus ($|m| = 1$).

In other words, in the pairwise copying model, a curious interplay of deterministic and stochastic terms maintains order or consensus in small groups. Such a group consensus or collective order, which arises from stochasticity and is away from deterministic stable state, is termed *noise-induced order*.

Box 2. Linking individual-level probabilistic rules to group-level dynamics.

Individual animal interactions and decisions are best modelled as probabilistic. It is not always obvious how these individual-level probabilistic interactions scale to the group level or the mesoscopic dynamics. To understand this, recall that while the deterministic term in the mesoscopic SDE is a mean effect of interactions, the stochastic term captures the residual variations around the mean. We now discuss this scaling of individual-level rules to group dynamics in the contexts of pairwise and ternary copying models.

Spontaneous switching: The spontaneous switches of states (r_1) are random changes in individuals' state, without interaction with any other individuals. At the group level, the mean effect of such random state-changes is to reduce the order or consensus within groups (captured by the term $-2r_1m$ in the deterministic term of equation (2.6)). As expected, individual-level randomness also leads to stochasticity at the group level ($2r_1$ in the stochastic term of equation (2.6)).

Pairwise copying interactions: The pairwise copying interaction rate (r_2), surprisingly, does not appear in the deterministic term of the group-level dynamics. This is because the pairwise interactions exhibit no bias in the directionality of state-change and thus, on an average, cause equal numbers of individuals to switch states from 1 to 2 and 2 to 1.

However, sampling errors while individuals choose copying partners can cause substantial variation around this zero mean effect. Its effect is larger for smaller groups. When the group is at the disordered state ($m = 0$), the sampling error can only cause the degree of consensus to increase and hence the strength of noise is maximum when $m = 0$. On the other hand, copying (and associated sampling errors) will have least effect at/near the ordered state ($m = \pm 1$), where nearly all individuals are in the same state. Therefore, the net effect of sampling errors due to copying is captured by the state-dependent or multiplicative noise term $(1 - m^2)r_2$ in equation (2.6). This simple structure of the noise pushes the system away from $m = 0$ and when the group has high order it resides there longer owing to low levels of noise. Thus, the non-monotonic structure of group-level noise, driven by pairwise copying interactions, pushes the system away from disorder ($m = 0$) and towards group consensus ($m = \pm 1$).

Ternary interactions: Moving onto ternary interactions (with rate r_3), we note that they cause the minority of the three interacting partners to switch its state towards the majority. Consequently, its mean effect creates an ordered state and hence appears in the deterministic term of equation (2.8). The residual stochasticity is exactly like the pairwise interactions. When $r_3 > 4r_1$, the mean or the deterministic effect alon pushes the system away from disorder towards an ordered state; the role of noise is not important to the collective dynamics in this model. However, for $r_3 < 4r_1$, the model behaves qualitatively similarly to the pairwise model with a cubic deterministic term.

In summary, all individual-level probabilistic interactions contribute to the noise at the group level. However, interactions whose mean effect is zero at the group level do not contribute to the deterministic dynamics.

Table 1. Scaling from individual stochastic interaction rates to group-level dynamics (deterministic and stochastic terms).

model	stochastic interaction rates	deterministic term $f(m)$	stochastic term $g(m)$
pairwise copying model	r_1: spontaneous switching rate	$-2r_1m$	$\sqrt{\frac{2}{N}}\sqrt{2r_1 + r_2(1-m^2)}$
	r_2: pairwise copying rate	depends on switching but not on pairwise copying	depends on both rates
ternary interaction model	r_1 and r_2: same as above	$-2r_1m + \frac{1}{2}r_3m(1-m^2)$	$\sqrt{\frac{2}{N}}\sqrt{2r_1 + (r_2 + r_3/2)(1-m^2)}$
	r_3: ternary interaction rate	depends on switching and ternary but not on pairwise copying	depends on all three rates

Ternary interactions model: In this model, individuals continue to exhibit a spontaneous switching between states at a rate r_1 and a pairwise copying interaction at a rate r_2, exactly as in equations (2.4) and (2.5). In addition, individuals exhibit a ternary interaction given by the following reactions:

$$2X_1 + X_2 \xrightarrow{r_3} 3X_1 \tag{2.7a}$$

and

$$X_1 + 2X_2 \xrightarrow{r_3} 3X_2. \tag{2.7b}$$

Here, interactions can happen between three individuals at a time. In an interacting triad, the individual who is in a minority switches his/her state to those of majority, at a rate r_3 [38].

The mesoscopic dynamics of m for this model are given in [11,18]:

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -2r_1m + \frac{r_3}{2}m(1-m^2) + \frac{1}{\sqrt{N}}\sqrt{2r_1 + (r_2 + \frac{r_3}{2})(1-m^2)}\ \eta(t), \tag{2.8}$$

where $\eta(t)$ again represents the uncorrelated Gaussian noise.

The functional form of the stochastic term here is similar to that of the pairwise copying model but with an additional term associated with the ternary interaction rate (r_3). However, in contrast to the pairwise copying model, the

deterministic term here is a cubic function arising solely from the ternary interaction rate r_3. Focusing on the limit of large group sizes $N \to \infty$ and thus ignoring the stochastic term, the dynamics of order are determined by the cubic equation $\dot{m} = -2r_1 m + (r_3/2)\, m(1 - m^2)$. Here, when $r_3 > 4r_1$, the system has two additional fixed points at $|m^*| > 0$ and are stable. Furthermore, the $m^* = 0$ fixed point becomes unstable. In other words, the ternary interaction model exhibits order or group consensus primarily via the deterministic term. Therefore, the order is present even in the large group size limit ($N \to \infty$).

Summary of model results: In mathematical terms, the consensus in the ternary model is driven by deterministic terms and hence is realized in the large group size limit. This is a significant contrast to the pairwise model, which shows consensus only when N is less than a threshold value and does not admit consensus in the $N \to \infty$ limit. Therefore, the mechanism causing order in the ternary interaction model is fundamentally different from the pairwise copying model, thus providing a useful contrast.

Intriguingly, the stochastic terms are of the same form in both models, yet the importance and the role of noise is different in both models. It is also worth emphasizing that an additive noise of the form shown in equation (2.2) of box 1 does not produce any nontrivial ordering effects. Therefore, an interplay of deterministic and stochastic terms is necessary to produce noise-induced order. We refer the readers to box 2 and table 1 therein for an intuitive discussion of how individual-level interactions scale to mesoscale dynamical terms.

Many previous studies on collective behaviour have investigated the role of fluctuations in collectives via adding an additive noise term to the deterministic term of the ternary interaction model [39–41]. In [16], the authors begin from microscopic interactions and show that the fluctuations (*temperature*) depend on the group polarization (*velocity field*). This implies that the state-dependent (multiplicative noise) can emerge even though the individual rules have no such multiplicative noise—which is consistent with the models we have described here [11]. However, these conclusions are derived considering the infinite-size limit or they do not account for finite system sizes explicitly. Therefore, we re-emphasize here that our framework accounts explicitly for fluctuations arising from finite system sizes.

3. Characterizing noise-induced states from data

With this background, we now turn our attention to the inverse question, which is also the main goal of this manuscript: given time-series data of a collective state (or order parameter), we ask is it possible to infer if the order is noise-induced?

To address this question, we perform stochastic simulations of both pairwise and ternary interaction models using the Gillespie algorithm [42,43]. In figure 2*a*,*b*, we display the time series of the degree of consensus (denoted by M) among 50 individuals using the pairwise and ternary copying models, respectively. We denote the order parameter obtained by simulations by the capital letter M.

We observe that in both systems the degree of consensus does not reach an equilibrium value but shows dynamic patterns, sometimes reaching a consensus ($M = \pm 1$) but repeatedly switching back and forth between two consensus values (i.e. $M = 1$ or -1).

In figure 2*c*,*d*, we display the graphs of the probability density functions of M. These show that the most likely state in both models is a high degree of consensus ($M \approx \pm 1$).

We recall that there are fundamental differences between the nature of collective dynamics in these models; while the collective order in the pairwise copying model is driven by stochasticity (i.e. noise-induced), the order in the ternary copying system is entirely driven by the deterministic term. Yet, visual inspection reveals no qualitative features that distinguish the two model outcomes in figure 2*a–d*— in terms of either dynamics or the most likely states.

However, as shown in the previous section, the SDEs that govern the dynamics of the consensus in two models are indeed different. Therefore, if we can use the time-series data shown in figure 2*a*,*b* to construct SDEs of the form (see box 1)

$$\dot{m} = F(m) + G(m)\eta(t), \tag{3.1}$$

we may decipher the role of stochasticity in each of the datasets. Here, as before, η is a Gaussian white noise with mean zero and unit variance, $F(m)$ represents the deterministic term (also called *drift* coefficient) and $G(m)$ is the stochastic term (with $G^2(m)$ called the *diffusion* coefficient) driving the dynamics.

Note that we have used capital letters to denote simulated data (M) and the data-constructed functional forms of deterministic ($F(m)$) and stochastic terms ($G(m)$). While the simulated M is necessarily discrete owing to finite numbers of individuals N in simulations, the order parameter in SDEs is assumed/approximated as a continuous-order parameter; hence in the functional form we keep the m notation, resulting in a composite notation such as $F(m)$ and $G(m)$. This notation also helps to distinguish from analytically derived formal equations such as equations (2.6) and (2.8).

(a) Method for constructing stochastic differential equations from data

Following [15,20,22,25,44], the deterministic component (or the drift coefficient) can be approximately obtained by the *first jump-moment*, defined as

$$F(m) = \left\langle \frac{M(t + \Delta t) - M(t)}{\Delta t} \right\rangle \Bigg|_{M(t) \in [m, m+\epsilon]}, \tag{3.2}$$

where the angular brackets denote an average over all instances in the time series where $M(t)$ is close to a given m. In either real or simulated time series, the observable will never (or rarely) be exactly equal to a given m; hence the average is obtained considering all $M \in [m,\ m + \epsilon]$, where ϵ is a small value (we choose $\epsilon = 0.01$). In other words, the deterministic part $f(m)$ is the *average or expected change per unit time* in the observable quantity when it is at (or near) the value m.

Likewise, the stochastic term (or the diffusion coefficient) can be approximately computed via the *second jump-moment*, defined as

$$G^2(m) = \left\langle \frac{R^2(m)}{\delta t} \right\rangle \tag{3.3}$$

where, $R(m) = (M(t + \delta t) - M(t))\,|_{M(t) \in [m, m+\epsilon]} - F(m)\,\delta t$.

Here too, the averaging is done over the entire time series, and in the vicinity of m as described above.

To obtain intuition concerning this formula, we decompose the *residual* term R into two parts. The first part is the term $M(t + \delta t) - M(t)$, representing the actual change in the observable over a time δt from t. The second part is $F(m)\delta t$,

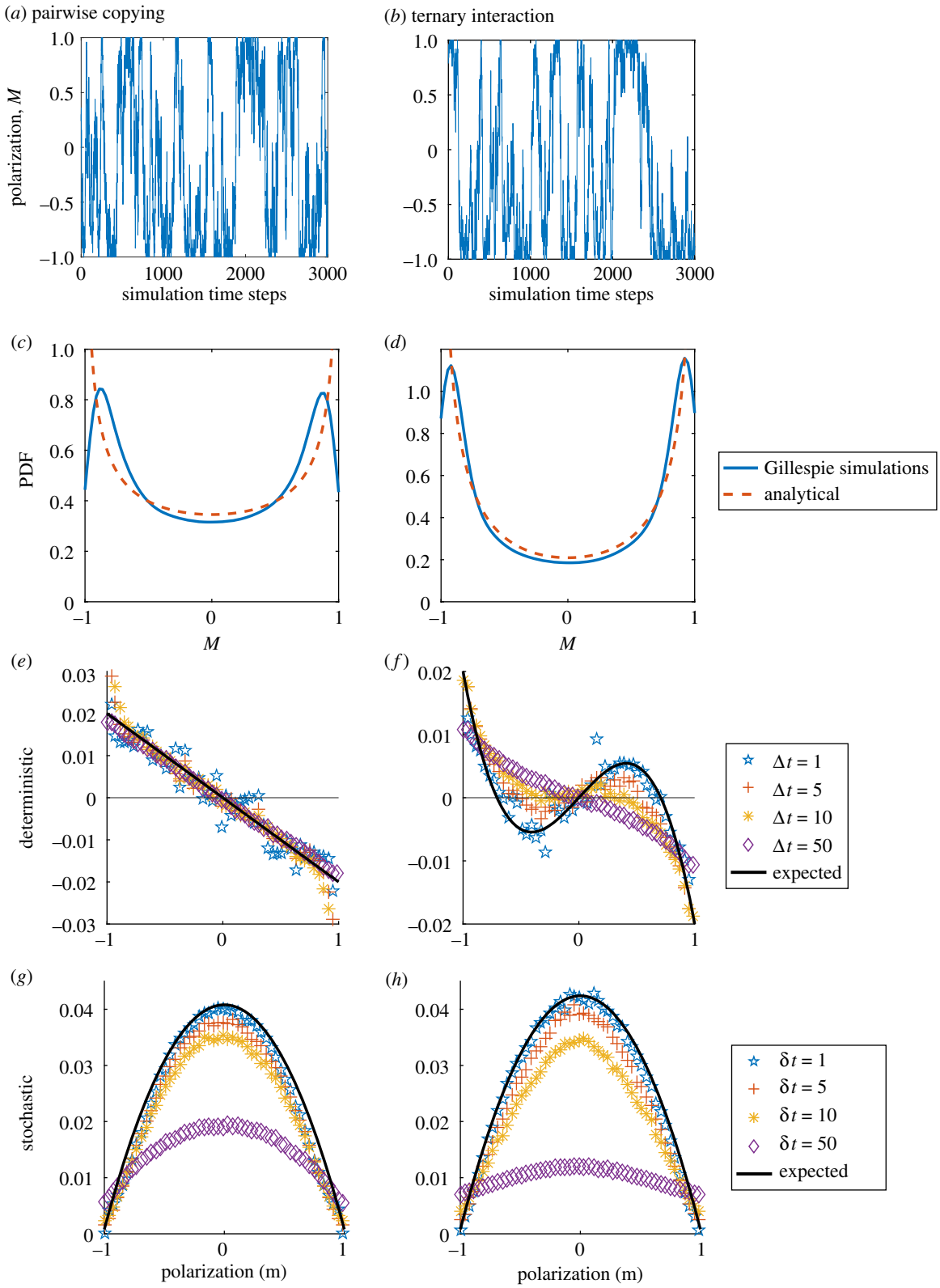

Figure 2. Characterizing mesoscopic dynamics of collective behaviour from data. (*a*,*b*) Representative time series of $M(t)$ for the pairwise and the ternary interaction models show that the system does not reach an equilibrium value. (*c*,*d*) Probability density function (PDF) of the data showing two modes corresponding to ordered states for both the pairwise and the ternary interaction model. The red dashed line represents the PDF derived using the analytical expression (see [11]). (*e*,*f*) Data-derived deterministic terms match the expected functional forms when $\Delta t = 50$ for the pairwise and $\Delta t \approx 5$ for the ternary model. (*g*,*h*) Data-derived stochastic terms for both the models are similar and match the expected function when $\delta t = 1$. The black solid line in (*e*–*h*) represents the analytical forms. Parameters: (pairwise model) $r_1 = 0.01$, $r_2 = 1$, $N = 50$; (ternary model) $r_1 = 0.01$, $r_2 = 1$, $r_2 = 0.08$, $N = 50$. (Online version in colour.)

which is the expected change in the observable based on the deterministic term alone. Therefore, for any given value of m, the term R in the numerator is basically the difference between the observed change and the expected change from the deterministic term. Considering squaring of this difference and then the averaging,, we may readily recognize the numerator as the second moment and it hence captures the stochasticity in the dynamics of the state variable m.

Although this method of constructing equations from data has been used earlier [15,26], a fundamental issue of choosing the right time scales to compute the deterministic and stochastic terms is overlooked. Note that we have deliberately chosen

Box 3. Time scale to construct the mesoscopic dynamics.

Time scale to construct deterministic term Δt: The time scale to compute the deterministic component must be comparable to (but less than) the autocorrelation time—the time difference above which two measurements of the observable M become essentially uncorrelated (see electronic supplementary material, §S1.1 for formal definition)—of the time series (denoted by τ_c). At very fine time scales ($\Delta t \ll \tau_c$), stochasticity of individual-level interactions will cause constant perturbations to the system away from the deterministic. The time scale over which these perturbations decay and the system relaxes back to deterministic stable states is typically given by τ_c. If we choose a $\Delta t \gg \tau_c$, we are likely to miss the relaxation dynamics of perturbations. Therefore, to capture the dynamics driven by the deterministic forces, we conjecture that a time scale Δt comparable to τ_c is most appropriate.

Time scale to construct stochastic term δt: The time scale to compute the stochastic component should be much smaller in order to capture stochastic effects provided the number of probabilistic events in that time window follow a Gaussian distribution. Equivalently, the residuals $R(m)$ for any m must follow a Gaussian distribution. This expectation is based on the key assumption of the mesoscopic SDE description where the noise $\eta(t)$ is uncorrelated and follows a Gaussian distribution. For larger time windows, although these assumptions could still hold true, the stochastic effects would average out. Therefore, we expect that δt is much smaller than Δt.

different notations for time steps Δt and δt in the formulae to compute the deterministic and stochastic terms, respectively. A naive choice could be that both time steps must be equal to the smallest time step, i.e. at the finest resolution in which data are available. However, that is not the case. Here, we conjecture and later confirm via simulations that appropriate time scales for constructing the deterministic and stochastic forces are not the same. More specifically, while Δt must be comparable to the autocorrelation time of the time-series data, δt must be much smaller provided that the Gaussian approximation of noise in the mesoscopic SDE is still valid (see box 3).

(b) Model parameters

Parameter values: We apply the above described method on the data generated from simulation of the models. For generating this data, we vary different parameters (N, r_1, r_2, r_3) in the models.

The pairwise model is a special case of the ternary interaction model with $r_3 = 0$. Without loss of generality, we set $r_2 = 1$. We then choose $r_1 = 0.01$, which as per the analytical results of [12] sets the critical system size to $N_c = 100$; that is, if $N < N_c$, the group exhibits noise-induced order ($|m| > 0$), whereas the group exhibits disorder ($m \approx 0$) if $N > N_c$. Therefore, for demonstration of the method, we choose three different values of $N = 50$, 100 and 200 corresponding to the so-called subcritical, critical and super-critical regimes, respectively.

For the ternary model, when the ternary copying rate is sufficiently high (specifically, $r_3 > 4r_1$), the system exhibits deterministically driven ordered dynamics ($m > 0$). Keeping the rest of the parameters the same as for the pairwise model, we arbitrarily set $r_3 = 0.08$ for our analyses. We note that the dynamics of this model are qualitatively similar to the pairwise model in the subcritical regime of this model ($r_3 < 4r_1$) and hence we do not consider those values for our demonstration.

Correlation time: The correlation time (τ_c) of a time series is the time difference above which two measurements of the observable become essentially uncorrelated (see electronic supplementary material, §S1.1 for a formal definition and also figure S1 for how autocorrelation changes with time). We find that τ_c does not change with N for the pairwise model, but it increases with N for the ternary model. For both the models, τ_c decreases with the spontaneous rate r_1 (electronic supplementary material, figure S2). Hence, to understand the role of correlation time in the reconstruction of the SDE, we vary r_1 for the pairwise model and N for the ternary model.

Distance between the expected and derived function: We measure the distance between the derived and the expected functions by calculating the normalized root mean square distance metric (D), defined as

$$D(F, f) = \frac{\|A - a\|_2}{\|a\|_2}, \tag{3.4}$$

where $\|f\|_2 = \sqrt{\sum_m f^2(m)}$ defines the ℓ^2-norm for a function f. A represents the derived function and a represents the expected functional form from the analytic derivations of the mesoscopic SDE.

(c) Stochastic differential equations constructed from data reveal the role of stochasticity in collective dynamics

We now demonstrate the method of SDE construction by using the data generated by individual-based collective behaviour models described in §2a (see figure 2*a*,*b* for representative graphs of time series).

To construct the deterministic term, we apply equation (3.2) to time series of M for both models. For the pairwise copying model, we find that the deterministic term is a linear function of m (figure 2*e*). Analysis of the time series of the ternary interaction model reveals a deterministic term that is a cubic function of m (figure 2*e*). Reassuringly, the functions thus constructed for both models match remarkably well with the analytically expected deterministic terms of equations (2.6) and (2.8).

In the above data-driven construction of deterministic dynamics, we considered a range of values of Δt. The results for some Δt are shown in figure 2*e*,*f*. The smallest time step ($\Delta t = 1$) yields a noisy pattern around the analytically expected functions for both the models. However, the constructed functions become closer to the analytical expectations (table 1) for larger values of Δt. When Δt is around an order of magnitude less than the autocorrelation time of the time series, we find that the fit is most accurate, i.e. the distance between the analytically expected and the data-constructed functions reaches a minimum value (figure 3*a*,*b*). Furthermore, we find a strong positive relationship between the optimum value of Δt and the autocorrelation time (τ_c) (figure 3*c*,*d*); τ_c changes as model parameters vary (see §3b; electronic supplementary material, S1.2)).

We now turn our attention to constructing the stochastic term, by applying equation (3.3a) to time-series data from

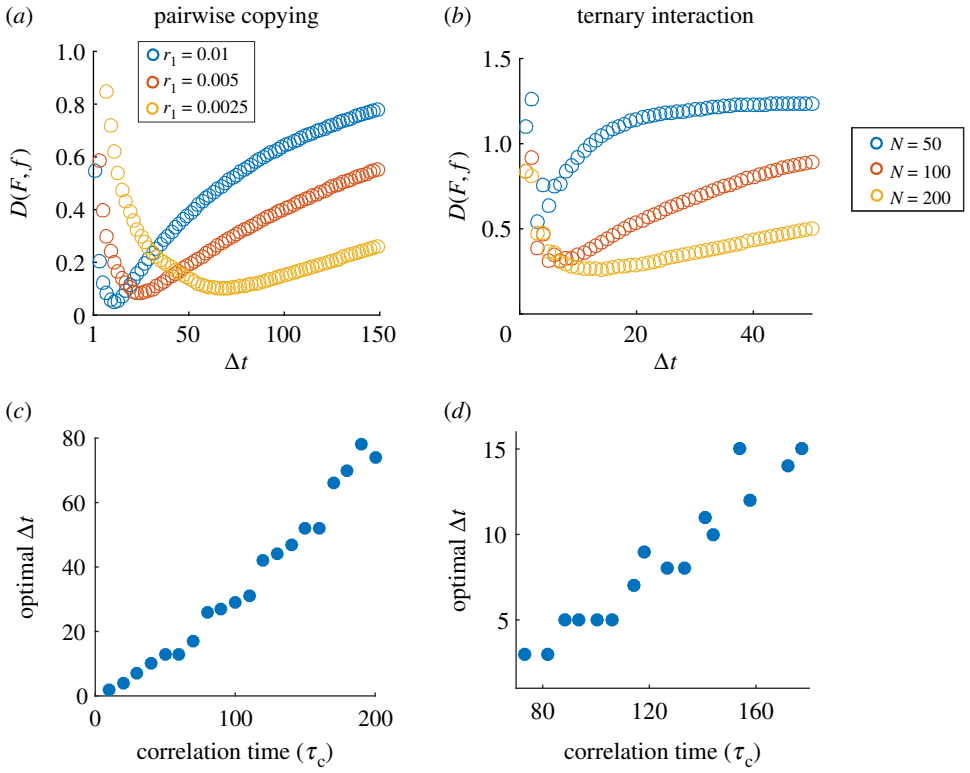

Figure 3. Optimum time scale to extract the deterministic term. Distance between the data-derived and the expected form for the deterministic term, $D(F, f) = ||F - f||_2/||f||_2$, is used to find the optimum time scale (Δt_{opt}) for extraction (see equation (3.4)). (*a*) This distance as a function of Δt for the pairwise model for three values of spontaneous reaction rate (r_1) and (*b*) three values of system size (N) for the ternary model; *note that parameters of the model that influence correlation time in that model are chosen for these analysis (see §3b)*. The Δt_{opt} corresponding to the minima of the above plots is the optimum time scale to derive the deterministic component. (*c*,*d*) Δt_{opt} as a function of correlation time (τ_c) follows the same pattern for both models, suggesting a possible universal rule that Δt_{opt} is roughly an order of magnitude less than τ_c. (Online version in colour.)

both models, for a range of values of δt. Here too, for both models, we are able to obtain the analytically expected functional form of an inverted parabolic function to a remarkable accuracy. Interestingly, the smallest δt yields the most accurate stochastic force function. This match becomes rapidly worse with increasing the time step (δt), a pattern exactly opposite to that of constructing the deterministic term (figure 2*g–h*; see electronic supplementary material §S2 and figure S3 for derivation with different parameter values).

We explore the role of δt further, by first generating very-high-resolution data of both individual-based models and then analysing the noise ($\eta(t)$) structure. Interestingly, we find that for very-high-resolution data the noise distribution is not perfectly Gaussian (figure 4*a*,*c*), thus violating the assumption of SDE (equation (2.1)). We find the distribution tends to Gaussian for coarser time scales (figure 4*b*,*d* also see electronic supplementary material, §S3 and figure S4 for $\delta t = 100$). We also analyse the correlations in noise by plotting its autocorrelation function. For all time scales in our analysis, we find the noise in the data to be uncorrelated (figure 4*i–l*), satisfying a key assumption of our framework (§2). Both these checks, i.e. whether $\eta(t)$ is Gaussian as well as uncorrelated in data, are important to validate key underlying assumptions (figure 5). Furthermore, though the distribution of the residuals/noise ($R(m)$) is Gaussian above a certain time scale in the data, we conjecture and confirm that there is an optimum of δt beyond which the extraction of the stochastic term becomes inaccurate (figure 4*i*,*k*).

However, if the data were available at a coarser resolution, we might not find the optimum δt as depicted by figure 4*i*,*k*. For example, if the time interval between consecutive data points is more than the optimum δt, we find that the smallest time step of such coarser time series yields minimum error in the reconstruction (figure 4*j*,*l*). Indeed, this is the reason why we find that the best reconstruction of the stochastic term in figure 2 corresponds to $\delta t = 1$.

Finally, we confirm that the method of construction of SDE from simulated data of the individual-based models is consistent for different parameter values in models (electronic supplementary material, §S2). Reassuringly, in all the cases, the data-derived deterministic and stochastic functions match not only the qualitative features of the analytically expected functions but also quantitatively.

(d) Consistency of the method and model

We ask whether our proposed method is *self-consistent*? To do this, we simulate the data-derived SDE to generate high-resolution time series of the collective state variable m. We then apply the same set of steps to reconstruct the SDE from this time series. We find that such a procedure yields comparable values of time scales (Δt and δt) and qualitatively similar forms of the deterministic and stochastic terms of the SDE (electronic supplementary material, §S3, §S4 and figure S5), thus demonstrating the consistency of the method.

Next, we ask whether the data-derived SDE model produces dynamical features consistent with the experimental data; the latter in our case corresponds to Gillespie simulation of the microscopic interaction rules. We recall that the construction of SDEs relied only on the first and second jump-moments, together with the autocorrelation time, of the time series of the state variable. We consider two functions—(i) autocorrelation

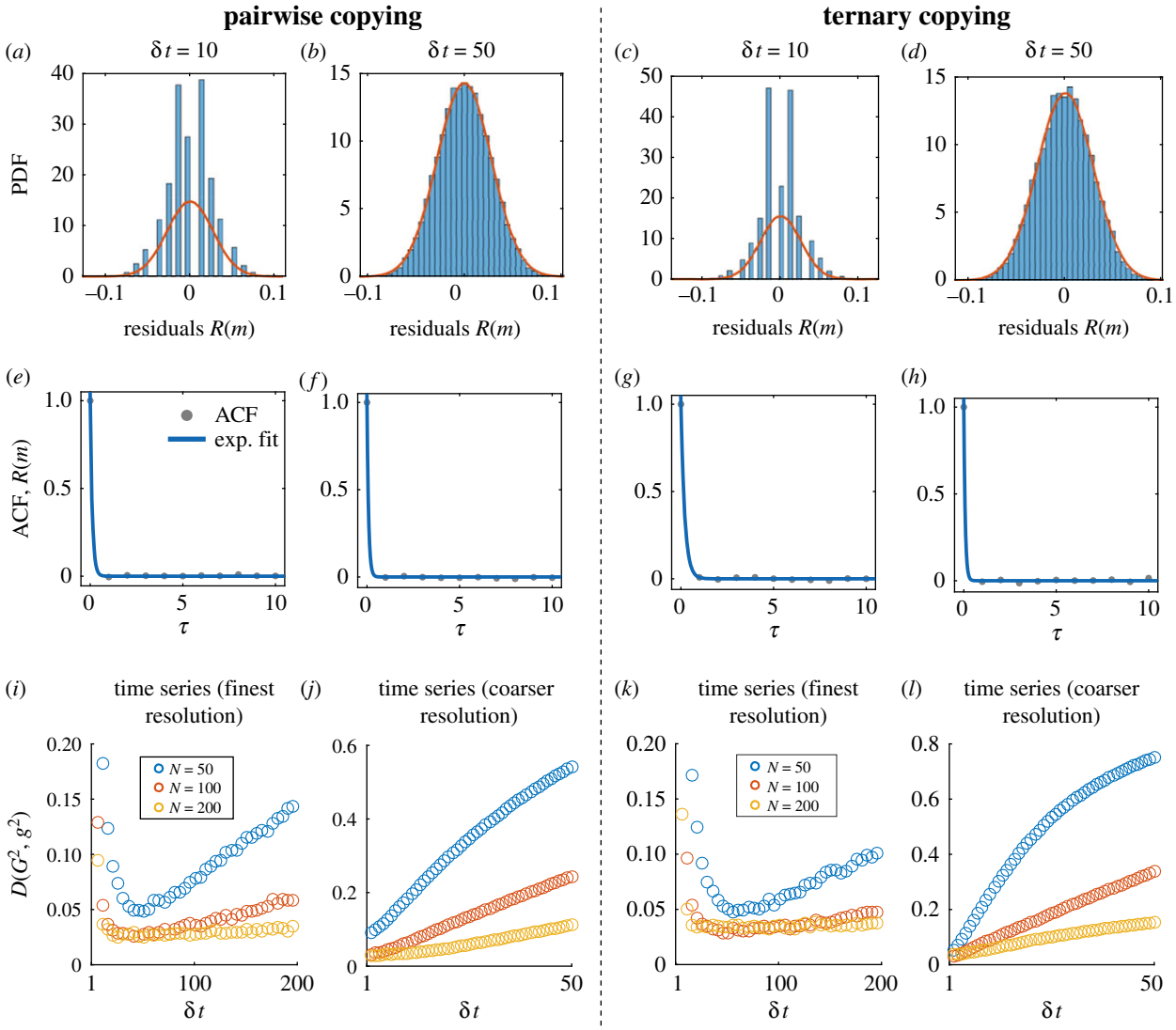

Figure 4. Optimum time scale to extract the stochastic term. (Top row) Distribution of residuals $R(m)$ for $M = 0$ for different δt using the fine time series for both models. For both the models, the distribution is not perfectly Gaussian for smaller δt (*a*,*c*) but increasing δt improves the Gaussian approximation (*b*,*d*). (Middle row) (*e*–*h*) The autocorrelation function (ACF) of the residuals ($R(m)$) for $M = 0$ shows that noise is uncorrelated in our simulated datasets. (Circles, ACF; line exponential fit.) (Bottom row) Distance, $D(G, g) = ||(G^2 - g^2)||_2/||g^2||_2$ (see equation (3.4)), plotted as a function of δt for the pairwise copying model and the ternary interaction model for three values of system size (N). In (*i*,*k*), the time series is much finer and $\delta t_{\text{opt}} > 1$. In (*j*,*l*), the time series is coarse and hence we find $\delta t_{\text{opt}} = 1$. Therefore, the optimum (δt_{opt}) corresponds to the smallest δt at which the Gaussian structure of noise is preserved. (Online version in colour.)

function of m and (ii) probability density function of m—neither of which were directly used to construct the model. We find that the autocorrelation function of the time series of the SDE model is in excellent qualitative agreement with that of the Gillespie simulations of the microscopic model (electronic supplementary material, figure S1). Likewise, the probability density function of m computed from the SDE and the Gillespie simulations of the microscopic model also show qualitative agreement across a range of parameter values for both models (figure 2*c*,*d*; electronic supplementary material, figures S4 and S5 do not appear to have been cited). Put together, these analyses suggest self-consistency of both the method of construction of SDE that we proposed and the model we have derived.

4. Discussion

We demonstrate a method to characterize the dynamics of collective behaviour that accounts for intrinsic stochastic effects in groups. Such noise arises owing to small sizes of groups and probabilistic interactions among group members. Specifically, given a high-resolution time-series dataset of collective behaviour, we characterized the dynamics via a stochastic differential equation (SDE) which accounts for both deterministic and stochastic drivers. Our key contribution lies in finding optimum time scales over which to compute the deterministic and stochastic terms of an SDE. With this characterization, we highlight the potential of intrinsic noise in producing group order even though the deterministic limit does not predict order.

(a) Novelty and applicability of the method

Strikingly, this method can help us distinguish whether the observed collective order is due to deterministic or stochastic drivers. To demonstrate this, we use two well-studied toy models of collective behaviour. For these models, we know the exact forms of the mesoscopic-scale SDEs from previous analytical studies [12,18,45]. Specifically, while the pairwise interaction model

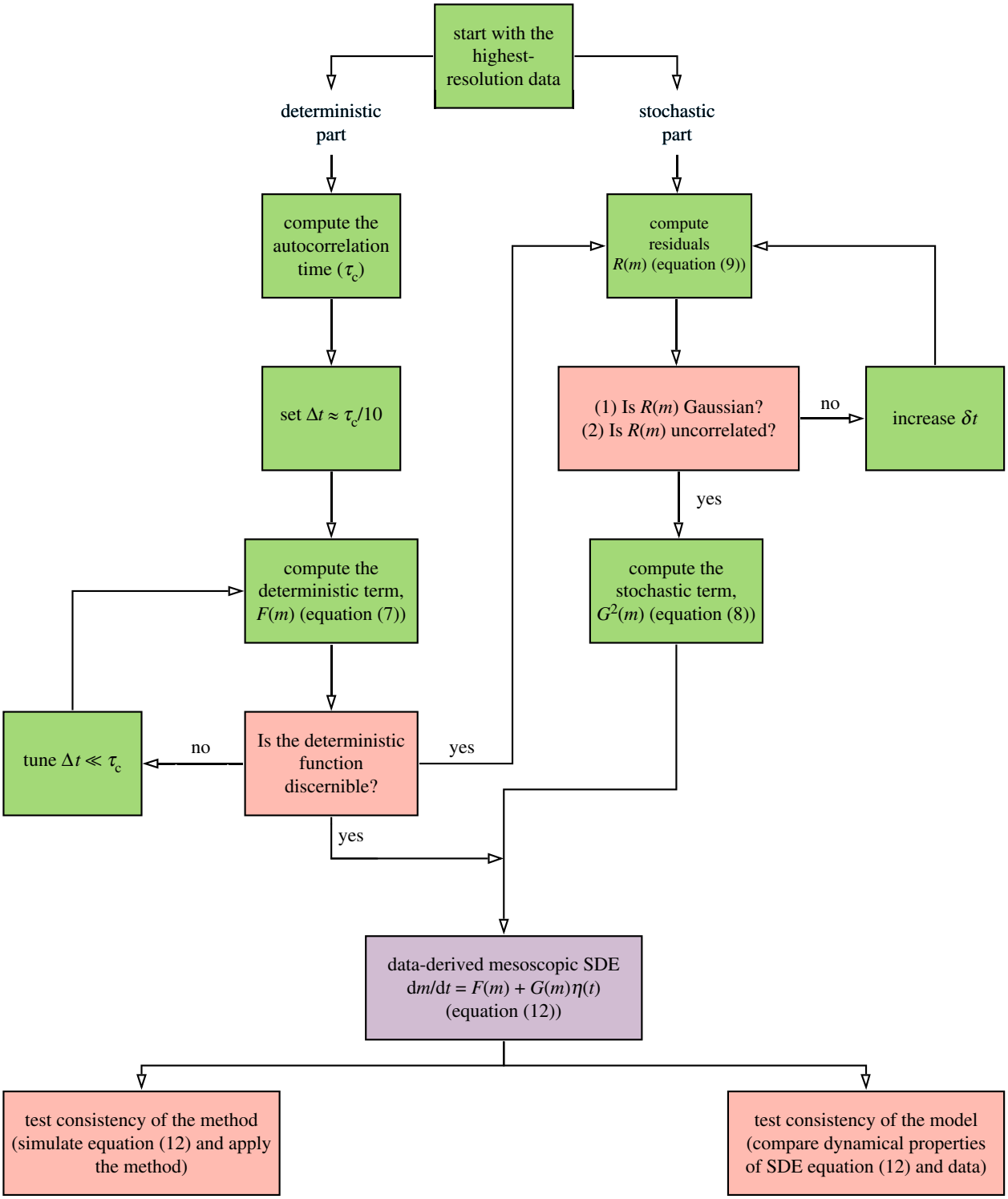

Figure 5. A flowchart summarizing the method to derive the stochastic differential equation (3.1) from time-series data. (Online version in colour.)

exhibits a noise-induced order, the ternary interaction dominated system exhibits order driven by deterministic terms. The qualitative features of the time series of collective behaviour for these models are similar (figure 2*a–d*). From the same high-resolution temporal data of the group order, we are able to confirm that the described method faithfully characterizes the mesoscopic SDEs. This is reassuring and therefore instils confidence that we can employ the method in more complex scenarios including real data.

Although simple and elegant, this method has rarely been used in the biology literature (but see [14,15,25,26]). One possible reason might be the lack of clarity on methods of constructing the deterministic and stochastic terms. A key finding from our study is that the construction of both the parts needs to be done at different time scales in the data. To construct the deterministic term, a time scale slightly smaller than the autocorrelation time of the data seems optimal. By contrast, the stochastic term needs to be constructed at much finer time scales. In figure 5, we provide a flow chart of the procedure.

In the context of collective behaviour, a study on marching locusts applied the same method and found evidence for multiplicative noise that is of a similar form to the pairwise and ternary interaction models [15]. However, in their system, the deterministic term was cubic, like in the ternary model. Hence, the deterministic term alone could explain the order. A recent study on fish schools (*Etroplus suratensis*) of small to intermediate group sizes showed that the highly aligned

motion is a noise-induced effect, best explained by the simple pairwise alignment interaction model we discussed above [14]. The method has also been applied to study single-cell migrations and find that movement in normal and cancerous cell types differs qualitatively. While cancerous cells show migration that is driven purely by deterministic effects, normal cells are driven only by stochastic factors [19]. These examples show that the method not only offers a rigorous quantitative description of collective dynamics (or even single organismal behaviour), but may also offer insights on the individual-level processes.

We have demonstrated this method for two non-spatial models of collective behaviour where each individual could be in one of the two discrete decision states. Despite the simplicity of the model framework we used, the applicability of the method to construct mesoscopic dynamics is wider. For example, as in the locust study [15], two states could be interpreted as two directions of movement in an annulus, and hence the group order may correspond to the degree of alignment of collective motion. For swarming systems in a continuous two- (or three-) dimensional world, the number of states is infinitely large. However, we may define the group order using a vectorial representation and construct a mesoscopic SDE by a fairly trivial extension of the method proposed here [14].

(b) Limitations and future directions

The method to construct mesoscopic SDEs requires further exploration in several contexts. A key assumption of the SDE equation (2.1) is that the noise ($\eta(t)$) is Gaussian and uncorrelated. However, for time series (finest resolution) in many systems, noise can be temporally correlated [46] and may exhibit large fluctuations [47]. Therefore, it is important to test these assumptions while applying the method to real datasets (figure 4*e–h*). Furthermore, in the wild, the dynamics of animal groups are likely influenced by external stimuli as individuals respond to threats, food availability, mates and so on. Consequently, the resulting group dynamics may often be non-stationary. However, our method is applicable only to stationary time series. Thus, it would be interesting to generalize the method of SDE construction in such complex scenarios.

The dynamics of other collective state variables such as rotational and dilational order are also of interest in many biological and physical contexts. Our method can be readily applied to these order parameters as well. However, in many contexts the state of a collective may not be fully described by a single coarse-grained variable but rather requires coupled dynamical variables (e.g. density together with the order parameter). Future work may extend the method of characterization of noise to such scenarios.

In this study, we considered only simple non-spatial descriptions. While neglecting space might be justified for small groups, in larger groups explicitly accounting for space is crucial to analyse the role of fluctuations. Recent analytical work suggests that we can derive stochastic partial differential equations to account for finite-group dynamics [48,49] where group order can be described as a function of both time and space. We note that promising efforts have been made in the context of Navier–Stokes equations of physical systems [50]. These approaches may shed light on extending our approaches to develop data-driven hydrodynamic descriptions of collective motion [39,41,51].

Finally, we comment on the potentially interesting biological consequences of noise-induced order. A system exhibiting (intrinisic) noise-induced states is characterized by the existence of a critical group size, at which key collective properties qualitatively change. This is particularly important for the design of experiments on collective behaviour, with the implication that conclusions based on small (or large) group sizes cannot be trivially applied to other group sizes. Another important question that merits attention is whether the noise-induced collective behaviour is adaptive. Natural selection is strongest at the level of individuals; therefore, it is likely that the observed mesoscopic noise-induced order is a simple consequence of selection for the microscopic interactions (e.g. pairwise copying versus higher-order copying); however, we may expect nontrivial feedback from the emergent properties of the group. It would be interesting to explore, both theoretically and experimentally, the adaptive significance of noise-induced collective behaviour.

(c) Concluding remarks

Our study highlights a much neglected but an elegant concept of noise-induced states in empirical studies on collective behaviour. The method we described to infer the role of stochasticity can be readily applied to data on collective motion across animal species. Stochastic interactions as well as finite (and small) sizes are inescapable features of the biological world. Therefore, we expect that noise-induced states are likely to occur in larger classes of biological phenomena. We hope that our study inspires further studies on the role of noise in collective motion, especially with a focus on functional aspects of collective behaviour.

Data accessibility. The codes used for analysis in the manuscript and to analyse real data can be downloaded from https://doi.org/10.5281/zenodo.3893644.

Authors' contributions. J.J. and V.G. conceived the project. J.J. performed the analysis. V.G. and J.J. wrote the manuscript.

Competing interests. We declare we have no competing interest.

Funding. No funding has been received for this article.

Acknowledgements. We thank Richard G. Morris, M. Danny Raj, Aakanksha Rathore, Vivek Jadhav and Ayan Das for discussions and critical comments on the manuscript. We thank Ashwin Karichannavar for his help with the Git repository. V.G. acknowledges support from the DBT-IISc partnership program, SERB (DST) and infrastructure support from DST-FIST. J.J. acknowledges support from CSIR-India for a research scholarship.

References

1. Ballerini M *et al.* 2008 Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *Proc. Natl Acad. Sci. USA* **105**, 1232–1237. (doi:10.1073/pnas.0711437105)
2. Lukeman R, Li Y-X, Edelstein-Keshet L. 2010 Inferring individual rules from collective behavior. *Proc. Natl Acad. Sci. USA* **107**, 12 576–12 580. (doi:10.1073/pnas.1001763107)
3. Gautrais J, Ginelli F, Fournier R, Blanco S, Soria M, Chaté H, Theraulaz G. 2012 Deciphering interactions in moving animal groups. *PLoS Comput. Biol.* **8**, e1002678. (doi:10.1371/journal.pcbi.1002678)

4. Herbert-Read JE. 2016 Understanding how animal groups achieve coordinated movement. *J. Exp. Biol.* **219**, 2971–2983. (doi:10.1242/jeb.129411)
5. Katz Y, Tunstrøm K, Ioannou CC, Huepe C, Couzin ID. 2011 Inferring the structure and dynamics of interactions in schooling fish. *Proc. Natl Acad. Sci. USA* **108**, 18 720–18 725. (doi:10.1073/pnas.1107583108)
6. Herbert-Read JE, Perna A, Mann RP, Schaerf TM, Sumpter DJT, Ward AJW. 2011 Inferring the rules of interaction of shoaling fish. *Proc. Natl Acad. Sci. USA* **108**, 18 726–18 731. (doi:10.1073/pnas.1109355108)
7. Hinz RC, de Polavieja GG. 2017 Ontogeny of collective behavior reveals a simple attraction rule. *Proc. Natl Acad. Sci. USA* **114**, 2295–2300. (doi:10.1073/pnas.1616926114)
8. Horsthemke W, Lefever R. 1984 *Noise-induced transitions: theory and applications in physics, chemistry and biology*. Berlin, Germany: Springer.
9. Ridolfi L, D'Odorico P, Laio F. 2011 *Noise-induced phenomena in the environmental sciences*. Cambridge, UK: Cambridge University Press.
10. Boettiger C. 2018 From noise to knowledge: how randomness generates novel phenomena and reveals information. *Ecol. Lett.* **21**, 1255–1267. (doi:10.1111/ele.13085)
11. Jhawar J, Morris RG, Guttal V. 2019 Deriving mesoscopic models of collective behavior for finite populations. In *Integrated population biology and modeling* (*Handbook of Statistics,* vol. 40), part B, pp. 551–594. Amsterdam, The Netherlands: Elsevier.
12. Biancalani T, Dyson L, McKane AJ. 2014 Noise-induced bistable states and their mean switching time in foraging colonies. *Phys. Rev. Lett.* **112**, 038101. (doi:10.1103/PhysRevLett.112.038101)
13. Kolpas A 2008 Coarse-grained analysis of collective motion in animal groups. PhD dissertation, University of California, Santa Barbara.
14. Jhawar J, Morris RG, Amith-Kumar UR, Raj MD, Rogers T, Rajendran H, Guttal V. 2020 Noise-induced schooling of fish. *Nat. Phys.* **16**, 488–493. (doi:10.1038/s41567-020-0787-y)
15. Yates CA, Erban R, Escudero C, Couzin ID, Buhl J, Kevrekidis IG, Maini PK, Sumpter DJT. 2009 Inherent noise can facilitate coherence in collective swarm motion. *Proc. Natl Acad. Sci. USA* **106**, 5464–5469. (doi:10.1073/pnas.0811195106)
16. Romanczuk P, Schimansky-Geier L. 2012 Mean-field theory of collective motion due to velocity alignment. *Ecol. Complex.* **10**, 83–92. (doi:10.1016/j.ecocom.2011.07.008)
17. Bertin E, Chaté H, Ginelli F, Mishra S, Peshkov A, Ramaswamy S. 2013 Mesoscopic theory for fluctuating active nematics. *New J. Phys.* **15**, 085032. (doi:10.1088/1367-2630/15/8/085032)
18. Dyson L, Yates CA, Buhl J, McKane AJ. 2015 Onset of collective motion in locusts is captured by a minimal model. *Phys. Rev. E* **92**, 052708. (doi:10.1103/PhysRevE.92.052708)
19. Brückner DB, Fink A, Schreiber C, Röttgermann PJF, Rädler JO, Broedersz CP. 2019 Stochastic nonlinear dynamics of confined cell migration in two-state systems. *Nat. Phys.* **15**, 595–601. (doi:10.1038/s41567-019-0445-4)
20. Van Kampen NG. 1981 Stochastic variables. In *Stochastic processes in chemistry and physics*, pp. 120–127. Amsterdam, The Netherlands: North Holland.
21. Gardiner C. 2009 *Stochastic methods*, vol. 4. Berlin, Germany: Springer.
22. Gradišek J, Siegert S, Friedrich R, Grabec I. 2000 Analysis of time series from stochastic processes. *Phys. Rev. E* **62**, 3146–3155. (doi:10.1103/PhysRevE.62.3146)
23. Honisch C, Friedrich R. 2011 Estimation of Kramers-Moyal coefficients at low sampling rates. *Physical Review E* **83**, 066701.
24. Ghasemi F, Peinke J, Tabar MRR, Sahimi M. 2006 Statistical properties of the interbeat interval cascade in human hearts. *Int. J. Mod. Phys. C* **17**, 571–580. (doi:10.1142/S0129183106008704)
25. Kolpas A, Moehlis J, Kevrekidis IG. 2007 Coarse-grained analysis of stochasticity-induced switching between collective motion states. *Proc. Natl Acad. Sci. USA* **104**, 5931–5935. (doi:10.1073/pnas.0608270104)
26. Boedeker HU, Beta C, Frank TD, Bodenschatz E. 2010 Quantitative analysis of random ameboid motion. *Europhys. Lett.* **90**, 28005. (doi:10.1209/0295-5075/90/28005)
27. Carpenter SR, Arani BM, Hanson PC, Scheffer M, Stanley EH, Van Nes E. 2020 Stochastic dynamics of Cyanobacteria in long-term high-frequency observations of a eutrophic lake. *Limnology and Oceanography Letters*.
28. Seeley TD, Visscher PK, Schlegel T, Hogan PM, Franks NR, Marshall JAR. 2012 Stop signals provide cross inhibition in collective decision-making by honeybee swarms. *Science* **335**, 108–111. (doi:10.1126/science.1210361)
29. Marshall JAR, Reina A, Bose T. 2019 Multiscale modelling tool: mathematical modelling of collective behaviour without the maths. *PLoS ONE* **14**, e0222906. (doi:10.1371/journal.pone.0222906)
30. Altschuler SJ, Angenent SB, Wang Y, Wu LF. 2008 On the spontaneous emergence of cell polarity. *Nature* **454**, 886–889. (doi:10.1038/nature07119)
31. Kirman A. 1993 Ants, rationality, and recruitment. *Q. J. Econ.* **108**, 137–156. (doi:10.2307/2118498)
32. Alfarano S, Lux T, Wagner F. 2008 Time variation of higher moments in a financial market with heterogeneous agents: an analytical approach. *J. Econ. Dyn. Control* **32**, 101–136. (doi:10.1016/j.jedc.2006.12.014)
33. Cox JT, Griffeath D. 1986 Diffusive clustering in the two dimensional voter model. *Ann. Probab.* **14**, 347–370. (doi:10.1214/aop/1176992521)
34. Cox JT. 1989 Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$. *Ann. Probab.* **17**, 1333–1366. (doi:10.1214/aop/1176991158)
35. Granovsky BL, Madras N. 1995 The noisy voter model. *Stoch. Process. Appl.* **55**, 23–43. (doi:10.1016/0304-4149(94)00035-R)
36. Dall'Asta L, Castellano C. 2007 Effective surface-tension in the noise-reduced voter model. *Europhys. Lett.* **77**, 60005. (doi:10.1209/0295-5075/77/60005)
37. Liggett TM. 2013 *Stochastic interacting systems: contact, voter and exclusion processes*. Berlin, Germany: Springer Science & Business Media.
38. Schulze C, Stauffer D. 2005 Monte Carlo simulation of the rise and the fall of languages. *Int. J. Mod. Phys. C* **16**, 781–787. (doi:10.1142/S0129183105007479)
39. Toner J, Tu Y. 1995 Long-range order in a two-dimensional dynamical XY model: how birds fly together. *Phys. Rev. Lett.* **75**, 4326. (doi:10.1103/PhysRevLett.75.4326)
40. Solon AP, Chaté H, Tailleur J. 2015 From phase to microphase separation in flocking models: the essential role of nonequilibrium fluctuations. *Phys. Rev. Lett.* **114**, 068101. (doi:10.1103/PhysRevLett.114.068101)
41. Ramaswamy S. 2010 The mechanics and statistics of active matter. *Annu. Rev. Condens. Matter Phys.* **1**, 323–345. (doi:10.1146/annurev-conmatphys-070909-104101)
42. Gillespie DT. 1976 A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–434. (doi:10.1016/0021-9991(76)90041-3)
43. Gillespie DT. 1977 Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361. (doi:10.1021/j100540a008)
44. Friedrich R, Peinke J, Sahimi M, Tabar MRR. 2011 Approaching complexity by stochastic methods: from biological systems to turbulence. *Phys. Rep.* **506**, 87–162. (doi:10.1016/j.physrep.2011.05.003)
45. McKane AJ, Newman TJ. 2004 Stochastic models in population biology and their deterministic analogs. *Phys. Rev. E* **70**, 041902. (doi:10.1103/PhysRevE.70.041902)
46. Hänggi P. 1978 Correlation functions and masterequations of generalized (non-Markovian) Langevin equations. *Z. Phys. B Condens. Matter* **31**, 407–416. (doi:10.1007/BF01351552)
47. Murakami H, Niizato T, Tomaru T, Nishiyama Y, Gunji Y-P. 2015 Inherent noise appears as a Lévy walk in fish schools. *Scient. Rep.* **5**, 10605. (doi:10.1038/srep10605)
48. Laighléis EÓ, Evans MR, Blythe RA. 2018 Minimal stochastic field equations for one-dimensional flocking. *Phys. Rev. E* **98**, 062127. (doi:10.1103/PhysRevE.98.062127)
49. Chatterjee P, Goldenfeld N. 2019 Three-body interactions drive the transition to polar order in a simple flocking model. *Phys. Rev. E* **100**, 040602. (doi:10.1103/PhysRevE.100.040602)
50. Rudy SH, Brunton SL, Proctor JL, Kutz JN. 2017 Data-driven discovery of partial differential equations. *Sci. Adv.* **3**, e1602614. (doi:10.1126/sciadv.1602614)
51. Grossmann R, Schimansky-Geier L, Romanczuk P. 2013 Self-propelled particles with selective attraction–repulsion interaction: from microscopic dynamics to coarse-grained theories. *New J. Phys.* **15**, 085014. (doi:10.1088/1367-2630/15/8/085014)