Noname manuscript No.
(will be inserted by the editor)

Josselin Garnier · George Papanicolaou · Tzu-Wei Yang

Mean Field Model for Collective Motion Bistability

Abstract We consider the Czirók model for collective motion of locusts along a one-dimensional torus. In the model, each agent's velocity locally interacts with other agents' velocities in the system, and there is also exogenous randomness to each agent's velocity. The interaction tends to create the alignment of collective motion. By analyzing the associated nonlinear Fokker-Planck equation, we obtain the condition for the existence of stationary order states and the conditions for their linear stability. These conditions depend both on the noise level, which should be strong enough, and on the interaction between the agent's velocities, which should be neither too small, nor too strong. We carry out the fluctuation analysis of the interacting system and describe the large deviation principle to calculate the transition probability from one order state to the other. Extensive numerical simulations confirm our analytical findings.

Keywords Czirók model · collective motion · mean field · interacting random processes

Mathematics Subject Classification (2000) 92D25, 35Q84, 60K35

Josselin Garnier
Centre de Mathématiques Appliquées, Ecole Polytechnique, 91128 Palaiseau Cedex, France
E-mail: josselin.garnier@polytechnique.edu

George Papanicolaou
Department of Mathematics, Stanford University, Stanford, CA 94305, USA
E-mail: papanicolaou@stanford.edu

Tzu-Wei Yang
School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA
E-mail: yangx953@umn.edu

1 Introduction

Collective motion has become an emerging topic in biology and statistical physics and has a large potential for the applications in other fields, for example, opinion dynamics in social science [19] and autonomous robotic networks in engineering [23]. Collective motion describes the systemic behavior of a large group of animals, insects or bacteria. In a collective motion model, the system has a large number of individuals, expressed as their locations and velocities, (x_i, u_i), $i = 1, \ldots, N$. The movement of each individual is affected by the movements of other individuals in the system, mostly attractive interactions. Most biologically realistic and mathematically interesting models suppose *local* attractive interactions, that is, any individual is affected only by its neighbors within a limited range. In addition, a collective motion model generally adds independent, internal or external randomness to the individuals' movements.

Two approaches are widely used to model collective motion: self-propelled particle (SPP) models and coarse-grained (CG) models. In an SPP model (also known as an individual-based model), the system is viewed as a large number of coupled stochastic processes describing the evolutions of the individuals in the system. In an SPP model we can easily build an elaborated interaction between an individual and its neighbors to match actual biological observations. However, an SPP model consists of numerous nonlinear equations and is difficult to analyze. To obtain the analytical tractability, further simplifications are usually required [29,14]; otherwise the system behavior has to be verified by numerical simulations. On the other hand, a CG model (also known as a macroscopic or mean-field model) assumes infinitely many individuals in the system and considers the evolution of the distribution of the individuals. Most CG models in the literature lead to hydrodynamic equations so that the system behavior can be analyzed. CG models have some limitations. Although some CG models [3,4,18] are indeed constructed from microscopic dynamics, it is generally difficult to relate the models to actual individuals' properties [1]. Another limitation of CG models is the lack of stochasticity. The hydrodynamic equations only describe the deterministic evolution and cannot capture and analyze some important stochastic events like transitions from one order state to another order state observed for instance in [6].

In the recent years, there is a huge amount of literature concerning the modeling of collective motion so we mention only a few papers guiding our own work. Vicsek *et al.* [27] propose the scalar noise model that is currently the most well-known SPP model. The Czirók model [9] is a one-dimensional model on a torus. It is a simplified version of the Vicsek model and is used to describe the collective motion of locusts for instance (see [28,1] and references therein). Buhl *et al.* [6] experimentally and numerically study the switching of the alignments of locusts and find that the probability of switching is exponentially decaying with the number of locusts. In [29,14] the authors assume global and uniform interactions in the Czirók model and they derive the nonlinear Fokker-Planck equations for the average velocity and consider overall stochastic behavior. Some interesting features, like the existences of

leaders or predators are considered in [8] and [21], respectively. Other various individual-based models related to the Czirók model are discussed in [20, 7, 22, 2], and [24, 13, 25, 3, 4, 18, 17, 26] construct and analyze CG models of collective motion.

One of the most important goals of the collective motion modeling is to mathematically reproduce and analyze the *ordering phenomenon* where the individuals show a coherence in their movements, such as swarms of locusts, flocks of birds, or schools of fish. Such an ordering phenomenon is called an order state of the system and there are generally multiple order states that represent different directions and rotations. The counterpart of the order states is the disorder state, where the system does not show such a pattern. From the perspective of statistical physics, in this paper we address several interesting questions that are also considered in many of the literature:

1. Existence and number of order states.
2. Quantities that control the phase transition between the disorder state and the order states.
3. Transition probability from one order state to another order state.

Our contributions in this paper are the following. We interpret the Czirók model as an interacting particle system in the McKean-Vlasov framework [10, 11]. In this sense, we characterize the state of the system by the empirical probability measure of the locations and velocities of all the individuals, and as the number of individual goes to infinity, the random, empirical measure converges to a deterministic, probability measure whose density is the solution of a nonlinear Fokker-Planck equation. Therefore, when the system is large, we can consider the Fokker-Planck equation instead of all the individuals. In this way, we are also able to provide a clear connection between the SPP model and the CG model. This procedure was proved to be very efficient to study the stationary states and their stability properties in the case of opinion dynamics models [15]. It turns out that it is still efficient to address collective motion models, although the framework is different: In the opinion dynamics model addressed in [15], the Fokker-Planck equation is non-degenerate while it is degenerate in the collective motion model that we addresse here, as the external noise affects only the velocities and not the locations. As a result noise plays a very different role which strongly affects the conditions for the existence and stability of the stationary states. For instance, the noise has to be large enough to ensure the existence of the non-trivial stationary states in the opinion dynamics model, while we will see that the existence of the order states in the collective motion model does not depend on the noise level and their stability depends in a complicated way of the noise level. Noise has to be large enough to ensure the linear stability of the order states, but large noise also triggers frequent transitions from one order state to another one.

By analyzing the Fokker-Planck equation obtained in the Czirók model, we find that the system may have one disorder state and two order states (clockwise and counterclockwise rotations at constant velocity). We find the necessary and sufficient condition for the existence of the order states. This existence condition quantitatively describes the phase transition between the disorder state and the order states. We then perform the linear stability

analysis to examine the stabilities of the order and disorder states, and we find that the order states are stable when the external noises are sufficiently strong. Such a phenomenon that the randomness of movement improves the alignment of collective motion coincides with the observation in [29]. The linear stability also requires that the interaction between the particles' velocities is neither too small, nor too strong. Finally, when the size of the system is large but finite, the empirical measure is still stochastic and has a small probability to transit from one order state to the other one. This small probability can be described by the large deviation principle. It increases with the noise level. It is exponentially decaying as a function of the number of individuals; this fact is also observed in [6,29]. We confirm our analysis by extensive numerical simulations.

The paper is organized as follows. The interacting particle model and its mean-field limit is presented in section 2. The equilibria of the nonlinear Fokker-Planck equation are analyzed in section 3. In section 4, the linearized stability analysis of the nonlinear Fokker-Planck equation is performed. We provide the fluctuation analysis and the large deviation principle in sections 5 and 6, respectively. In section 7, we verify our analytical findings by numerical simulations. We briefly consider the nonsymmetric Czirók model with normalized influence functions in section 8. We end with a brief summary and conclusions in section 9.

2 Model and Mean Field Limit

We consider the Czirók model, a system of N agents moving along the torus $[0, L]$. The position x_i and velocity u_i of particle i satisfy the following stochastic differential equations: for $i = 1, \dots, N$,

$$\begin{aligned} dx_i &= u_i dt, \\ du_i &= \left[G(\langle u \rangle_i) - u_i\right] dt + \sigma dW_i(t), \end{aligned} \tag{1}$$

where $\{W_i(t)\}_{i=1}^N$ are independent Brownian motions and $\langle u \rangle_i$ is a weighted average of the velocities $\{u_j\}_{j=1}^N$:

$$\langle u \rangle_i = \frac{1}{N} \sum_{j=1}^{N} u_j \phi(|x_j - x_i|), \tag{2}$$

with the weights depending on the distance (on the torus) between the position x_i and the positions of the other agents. Here $\phi(x)$ is a nonnegative influence function normalized so that

$$\frac{1}{L} \int_0^L \phi(|x|) dx = 1, \tag{3}$$

and $G(u)$ is an odd and smooth function. Note that there is no loss of generality in assuming (3) since we can always rescale G to ensure that this condition is satisfied. As we will see in this paper the interesting configuration is when $u \mapsto u - G(u)$ derives from a double-well potential, with two

symmetric global minima and one saddle point at zero. For instance, we can think at

$$G(u) = 2\tanh(u) \tag{4}$$

and then $u - G(u)$ derives from the double-well potential $u^2/2 - 2\log(\cosh(u))$, or

$$G(u) = 2u - u^3 \tag{5}$$

and then $u - G(u)$ derives from the double-well potential $u^4/4 - u^2/2$.

We define the empirical probability measure:

$$\mu_N(t, dx, du) = \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i(t), u_i(t))}(dx, du). \tag{6}$$

Let us assume that $\mu_N(0, dx, du)$ converges to a deterministic measure $\bar{\rho}(x, u)dxdu$ as $N \to \infty$. This happens in particular when the positions and velocities of the agents at the inital time $t = 0$ are independent and identically distributed with the distribution with density $\bar{\rho}(x, u)$. Then, as $N \to \infty$, $\mu_N(t, dx, du)$ converges to the deterministic measure $\rho(t, x, u)dxdu$ whose density is the solution of the nonlinear Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = -u\frac{\partial \rho}{\partial x} - \frac{\partial}{\partial u}\left\{\left[G\left(\iint u'\phi(|x'|)\rho(t, x - x', u')du'dx'\right) - u\right]\rho\right\} + \frac{1}{2}\sigma^2\frac{\partial^2 \rho}{\partial u^2}, \tag{7}$$

starting from $\rho(t = 0, x, u) = \bar{\rho}(x, u)$. Note that the Fokker-Planck equation (7) is degenerate because there is no diffusion in x. The convergence proof is standard and can be found in [10,16].

3 Stationary Analysis for Order and Disorder States

We look for a stationary equilibrium, that is to say, a probability density function stationary in time:

$$-u\frac{\partial \rho}{\partial x} - \frac{\partial}{\partial u}\left\{\left[G\left(\iint u'\phi(|x'|)\rho(x - x', u')du'dx'\right) - u\right]\rho\right\} + \frac{1}{2}\sigma^2\frac{\partial^2 \rho}{\partial u^2} = 0. \tag{8}$$

Proposition 1 *The probability-density-valued solutions $\rho(x, u)$ to (8) have the following form:*

$$\rho_\xi(x, u) = \frac{1}{L}F_\xi(u), \quad F_\xi(u) = \frac{1}{\sqrt{\pi\sigma^2}}\exp\Big(-\frac{(u - \xi)^2}{\sigma^2}\Big). \tag{9}$$

They are uniform in space, Gaussian in velocity, and their mean velocity ξ satisfies the compatibility condition:

$$\xi = G(\xi). \tag{10}$$

There are, therefore, as many stationary equilibria as there are solutions to the compatibility equation (10).

Proof It is straightforward to check that (9) is a solution of (8). Reciprocally, let ρ be a solution of (8). After integrating (8) from $u=-\infty$ to $u=+\infty$, we obtain

$$-\frac{\partial}{\partial x}\int_{-\infty}^{\infty} u\rho(x,u)du = 0,$$

which shows that $\int_{-\infty}^{\infty} u\rho(x,u)du = \xi/L$ is a constant independent of x. Then Eq. (8) can be written as a linear equation

$$-u\frac{\partial \rho}{\partial x} - \frac{\partial}{\partial u}\{[G(\xi)-u]\rho\} + \frac{1}{2}\sigma^2\frac{\partial^2\rho}{\partial u^2} = 0. \tag{11}$$

Let $F(u)=\int_0^L \rho(x,u)dx$. Then $F(u)$ satisfies

$$-\frac{\partial}{\partial u}\{[G(\xi)-u]F(u)\} + \frac{1}{2}\sigma^2\frac{\partial^2 F(u)}{\partial u^2} = 0, \tag{12}$$

which shows that $F(u)$ is a Gaussian density function:

$$F(u) = \frac{1}{\sqrt{\pi\sigma^2}}\exp\Big(-\frac{1}{\sigma^2}(u-G(\xi))^2\Big). \tag{13}$$

The condition $\xi = L\int u\rho(x,u)du$ requires that

$$\xi = \frac{1}{L}\int_0^L \xi dx = \int_0^L\int u\rho(x,u)dudx = \int uF(u)du = G(\xi). \tag{14}$$

To show that $\rho(x,u)$ is uniform in x, we first note that $\rho(x,u)$ is periodic in x. We expand it as:

$$\rho(x,u) = \sum_{k=-\infty}^{\infty}\rho_k(u)e^{-i2\pi kx/L}, \quad \rho_k(u) = \frac{1}{L}\int_0^L \rho(x,u)e^{i2\pi kx/L}dx. \tag{15}$$

From (11) and (10), ρ_k satisfies

$$\frac{i2\pi k}{L}u\rho_k - \frac{\partial}{\partial u}[(\xi-u)\rho_k] + \frac{1}{2}\sigma^2\frac{\partial^2\rho_k}{\partial u^2} = 0. \tag{16}$$

We take a Fourier transform in u:

$$\hat{\rho}_k(\eta) = \int_{-\infty}^{\infty}\rho_k(u)e^{-i\eta u}du. \tag{17}$$

From (16), $\hat{\rho}_k$ satisfies the ordinary differential equation

$$-\Big(\frac{2\pi k}{L}+\eta\Big)\frac{\partial\hat{\rho}_k}{\partial\eta} = \Big(i\xi\eta + \frac{1}{2}\sigma^2\eta^2\Big)\hat{\rho}_k. \tag{18}$$

This equation can be solved. Let $k<0$. For any $\eta\in(-\infty,-2\pi k/L)$ we have

$$|\hat{\rho}_k(\eta)| = |\hat{\rho}_k(0)|\left|1+\frac{L\eta}{2\pi k}\right|^{-\frac{2\pi^2k^2\sigma^2}{L^2}}\exp\Big(-\frac{\sigma^2\eta^2}{4}+\frac{\pi k\sigma^2\eta}{L}\Big),$$

which goes to $+\infty$ as $\eta \nearrow -2\pi k/L$ if $\hat{\rho}_k(0) \neq 0$. However $|\hat{\rho}_k(\eta)| \leq \int |\rho_k(u)|du \leq (1/L)\int\int_0^L \rho(x,u)dxdu = 1/L$ is uniformly bounded, which imposes $\hat{\rho}_k(0) = 0$ and therefore $\hat{\rho}_k(\eta) = 0$ for any $\eta \in (-\infty, -2\pi k/L)$. Similarly, for any $\eta \in (-2\pi k/L, \infty)$ we have

$$|\hat{\rho}_k(\eta)| = |\hat{\rho}_k(-4\pi k/L)|\Big|1+\frac{L\eta}{2\pi k}\Big|^{-\frac{2\pi^2 k^2\sigma^2}{L^2}} \exp\Big(\big(-\frac{\sigma^2\eta}{4}+\frac{2\pi k\sigma^2}{L}\big)\big(\eta+\frac{4\pi k}{L}\big)\Big),$$

which goes to $+\infty$ as $\eta \searrow -2\pi k/L$ if $\hat{\rho}_k(-4\pi k/L) \neq 0$. The boundedness of $\hat{\rho}_k$ imposes $\hat{\rho}_k(-4\pi k/L) = 0$ and therefore $\hat{\rho}_k(\eta) = 0$ for any $\eta \in (-2\pi k/L, \infty)$. Moreover, $\hat{\rho}_k$ is continuous by the Riemann-Lebesgue lemma and therefore $\hat{\rho}_k(\eta) = 0$ for any $\eta \in \mathbb{R}$ and $k < 0$. Since $\hat{\rho}_{-k}(-\eta) = \overline{\hat{\rho}_k(\eta)}$, we also find that $\hat{\rho}_k(\eta) = 0$ for any $\eta \in \mathbb{R}$ and $k > 0$. Therefore $\rho_k(u) = 0$ for any $u \in \mathbb{R}$ and $k \neq 0$ and the stationary solution is equal to $\rho(x,u) = \rho_0(u)$. □

When G is such that $u - G(u)$ derives from a double-well potential, such as the two examples (4) and (5), there are three ξ satisfying the compatibility condition (10): 0 and $\pm\xi_e$, with $\xi_e > 0$. The equilibrium $\rho_0(x,u)$ is the disorder state and $\rho_{\pm\xi_e}(x,u)$ are the two order states of the locusts marching on the torus, the clockwise and counterclockwise rotations. The existence of the nonzero solution $\pm\xi_e$ and hence the stationary states $\rho_{\pm\xi_e}(x,u)$ is independent of the value of the noise strength σ. This observation is unusual and in fact inconsistent with the conclusion in many models in statistical physics and opinion dynamics [10,15]. We will see, however, that the noise strength plays a crucial role in the stability of the stationary states.

4 Linear Stability Analysis

In this section, we assume that the compatibility condition (10) has solutions and we use the linear stability analysis to study the stability of the stationary states. Let ξ be such that $G(\xi) = \xi$ and consider

$$\rho(t,x,u) = \rho_\xi(x,u) + \rho^{(1)}(t,x,u) = \frac{1}{L}F_\xi(u) + \rho^{(1)}(t,x,u), \tag{19}$$

for small perturbation $\rho^{(1)}$. By linearizing the nonlinear Fokker-Planck equation (7) we find that $\rho^{(1)}$ satisfies

$$\begin{aligned}\frac{\partial\rho^{(1)}}{\partial t} &= -u\frac{\partial\rho^{(1)}}{\partial x} - \frac{\partial}{\partial u}\Big[(\xi-u)\rho^{(1)}\Big] \\ &\quad - \frac{1}{L}G'(\xi)\left[\iint u'\phi(|x'|)\rho^{(1)}(t,x-x',u')du'dx'\right]F'_\xi(u) + \frac{1}{2}\sigma^2\frac{\partial^2\rho^{(1)}}{\partial u^2}.\end{aligned} \tag{20}$$

Our goal is to clarify under which circumstances the linearized system is stable.

4.1 Expressions of the linearized modes

The perturbation $\rho^{(1)}$ is periodic in x. We expand it as:

$$\rho^{(1)}(t,x,u) = \sum_{k=-\infty}^{\infty} \rho_k^{(1)}(t,u) e^{-i2\pi kx/L}, \quad \rho_k^{(1)}(t,u) = \frac{1}{L}\int_0^L \rho^{(1)}(t,x,u) e^{i2\pi kx/L} dx. \tag{21}$$

The mode $\rho_k^{(1)}$ satisfies

$$\begin{aligned} \frac{\partial \rho_k^{(1)}}{\partial t} &= \frac{i2\pi k}{L} u \rho_k^{(1)} - \frac{\partial}{\partial u}\left[(\xi - u)\rho_k^{(1)}\right] \\ &\quad - G'(\xi)\phi_k \left[\int u' \rho_k^{(1)}(t,u') du'\right] F_\xi'(u) + \frac{1}{2}\sigma^2 \frac{\partial^2 \rho_k^{(1)}}{\partial u^2}, \end{aligned} \tag{22}$$

where the discrete Fourier coefficients

$$\phi_k = \frac{1}{L}\int_0^L \phi(|x|) e^{i2\pi kx/L} dx \tag{23}$$

are real-valued (because $|L - x| = |x|$) and bounded by one (because $|\phi_k| \le \int_0^L \phi(|x|)dx/L = 1$). We note that the equations are uncoupled in k and the system is linearly stable if all modes are stable. We then take a Fourier transform in u:

$$\hat{\rho}_k(t,\eta) = \int_{-\infty}^{\infty} \rho_k^{(1)}(t,u) e^{-i\eta u} du. \tag{24}$$

From (22), $\hat{\rho}_k$ satisfies the following first-order partial differential equation:

$$\frac{\partial}{\partial t}\hat{\rho}_k + \Big(\frac{2\pi k}{L} + \eta\Big)\frac{\partial}{\partial \eta}\hat{\rho}_k = \Big(-i\xi\eta - \frac{1}{2}\sigma^2\eta^2\Big)\hat{\rho}_k + G'(\xi)\phi_k\Big[\frac{\partial}{\partial \eta}\hat{\rho}_k(t,0)\Big]\eta\hat{F}_\xi, \tag{25}$$

starting from an initial condition that is assumed to satisfy

$$\|\hat{\rho}_k(0,\cdot)\|_1 + \|\hat{\rho}_k(0,\cdot)\|_\infty + \|\partial_\eta \hat{\rho}_k(0,\cdot)\|_\infty < \infty.$$

We say that the kth-order mode is stable if

$$\sup_{t\ge 0}\left\{\|\hat{\rho}_k(t,\cdot)\|_1 + \|\hat{\rho}_k(t,\cdot)\|_\infty + \|\partial_\eta \hat{\rho}_k(t,\cdot)\|_\infty\right\} < \infty.$$

By the method of characteristics, we can obtain the following implicit expression for $\hat{\rho}_k(t,\eta)$ that is useful for the stability analysis:

Lemma 1 *Let $w_k(t) = \partial_\eta \hat{\rho}_k(t,0)$. $\hat{\rho}_k(t,\eta)$ has the following expression:*

$$\begin{aligned} \hat{\rho}_k(t,\eta) &= e^{-g_k(t,\eta)}\hat{\rho}_k\Big(0, e^{-t}\eta - D_k(1-e^{-t})\Big) \\ &\quad + G'(\xi)\phi_k \int_0^t e^{-g_k(t-s,\eta)} w_k(s) H_\xi\Big(e^{-(t-s)}\eta - D_k(1-e^{-(t-s)})\Big) ds, \end{aligned} \tag{26}$$

where $D_k = 2\pi k/L$, $H_\xi(\eta) = \eta\hat{F}_\xi(\eta)$ *and*

$$g_k(t,\eta) = \frac{\sigma^2}{4}(\eta + D_k)^2(1 - e^{-2t}) + (i\xi - \sigma^2 D_k)(\eta + D_k)(1 - e^{-t}) \quad (27)$$
$$+ \Big(- i\xi D_k + \frac{\sigma^2}{2} D_k^2 \Big) t.$$

Proof The function $\hat{\rho}_k(t,\eta)$ satisfies (25). We consider the mapping $\eta(t) = \eta_0 e^t - 2\pi k/L$ so that $\eta'(t) = \eta(t) + 2\pi k/L$. By the method of characteristics,

$$\frac{d}{dt}\hat{\rho}_k(t,\eta(t)) = \Big(- i\xi\eta - \frac{1}{2}\sigma^2\eta^2 \Big)\hat{\rho}_k(t,\eta) + G'(\xi)\phi_k \Big[\frac{\partial}{\partial\eta}\hat{\rho}_k(t,0) \Big] \eta\hat{F}_\xi(\eta).$$

Then by using the method of integrating factors, we have

$$\hat{\rho}_k(t,\eta(t)) = e^{-\int_0^t g(s)ds}\hat{\rho}_k(0,\eta_0 - 2\pi k/L)$$
$$+ G'(\xi)\phi_k \int_0^t e^{-\int_s^t g(w)dw} \Big[\frac{\partial}{\partial\eta}\hat{\rho}_k(s,0) \Big] \eta(s)\hat{F}_\xi(\eta(s))ds,$$

where $g(t) = i\xi\eta(t) + \frac{1}{2}\sigma^2\eta^2(t)$. We note that $\int_s^t g(w)dw = g_k(t-s,\eta(t))$, where g_k is defined by (27) and $D_k = 2\pi k/L$. By letting $H_\xi(\eta) = \eta\hat{F}_\xi(\eta)$ and $w_k(s) = \partial_\eta\hat{\rho}_k(s,0)$, $\hat{\rho}_k(t,\eta)$ can be written as (26). □

4.2 Stability of the 0th-order mode

In this section we find a necessary and sufficient condition for the stability of the mode $k = 0$. For $k = 0$, $\rho_0^{(1)}$ satisfies

$$\frac{\partial\rho_0^{(1)}}{\partial t} = -\frac{\partial}{\partial u}\Big[(\xi - u)\rho_0^{(1)}\Big] - G'(\xi)\left[\int u'\rho_0^{(1)}(t,u')du'\right] F'_\xi(u) + \frac{1}{2}\sigma^2\frac{\partial^2\rho_0^{(1)}}{\partial u^2}. \quad (28)$$

We use the method of moments. We let $m_{0,0}^{(1)}$, $m_{0,1}^{(1)}$ denote the zeroth and first moments of $\rho_0^{(1)}$, respectively:

$$m_{0,0}^{(1)}(t) = \int \rho_0^{(1)}(t,u)du, \quad m_{0,1}^{(1)}(t) = \int u\rho_0^{(1)}(t,u)du. \quad (29)$$

From (28), we get

$$\frac{\partial}{\partial t}m_{0,0}^{(1)} = 0, \quad \frac{\partial}{\partial t}m_{0,1}^{(1)} = \xi m_{0,0}^{(1)} - m_{0,1}^{(1)} + G'(\xi)m_{0,1}^{(1)}.$$

The first moment $m_{0,1}^{(1)}$ is stable if and only if $G'(\xi) \le 1$. More exactly, if $G'(\xi) = 1$ then the first moment grows linearly in time. If $G'(\xi) < 1$ then the first moment is bounded. Once $m_{0,1}^{(1)}$ has been shown to bounded, then the solution $\rho_0^{(1)}$ of (28) can be seen as the solution of a linear parabolic equation and it can be shown to be smooth and bounded as explained in the following proposition.

Proposition 2 *The 0th-order mode is stable if and only if* $G'(\xi) < 1$.

Proof We know that $m_{0,1}^{(1)}(t)$ is uniformly bounded in t if and only if $G'(\xi) < 1$. From (26) in Lemma 1 and noting that $w_0(t) = \partial_\eta \hat{\rho}_0(t,0) = -im_{0,1}^{(1)}(t)$, we have the following expression for $\hat{\rho}_0(t,\eta)$:

$$\hat{\rho}_0(t,\eta) = \hat{\rho}_0(0,\eta e^{-t}) \exp\Big(- i\xi\eta(1-e^{-t}) - \frac{1}{4}\sigma^2\eta^2(1-e^{-2t})\Big) \tag{30}$$
$$- i\int_0^t F_0(\eta e^{-s}) m_{0,1}^{(1)}(t-s) \exp\Big(- i\xi\eta(1-e^{-s}) - \frac{1}{4}\sigma^2\eta^2(1-e^{-2s})\Big) ds,$$

with $F_0(\eta) = G'(\xi)\eta\hat{F}_\xi(\eta)$. We find that, for any $t \geq 1$,

$$\begin{aligned}
|\hat{\rho}_0(t,\eta)| &\leq \|\hat{\rho}_0(0,\cdot)\|_\infty e^{-\frac{1}{4}\sigma^2\eta^2(1-e^{-2})} \\
&\quad + \|m_{0,1}^{(1)}\|_\infty \int_0^t |F_0(\eta e^{-s})| \exp\Big(- \frac{1}{4}\sigma^2\eta^2(1-e^{-2s})\Big) ds \\
&\leq \|\hat{\rho}_0(0,\cdot)\|_\infty e^{-\frac{1}{4}\sigma^2\eta^2(1-e^{-2})} + \|m_{0,1}^{(1)}\|_\infty \int_0^1 |F_0(\eta e^{-s})| ds \\
&\quad + \|m_{0,1}^{(1)}\|_\infty |G'(\xi)| \|\hat{F}_\xi\|_\infty \int_1^t |\eta| e^{-s} \exp\Big(- \frac{1}{4}\sigma^2\eta^2(1-e^{-2})\Big) ds \\
&\leq \|\hat{\rho}_0(0,\cdot)\|_\infty e^{-\frac{1}{4}\sigma^2\eta^2(1-e^{-2})} + \|m_{0,1}^{(1)}\|_\infty \int_0^1 |F_0(\eta e^{-s})| ds \\
&\quad + \|m_{0,1}^{(1)}\|_\infty |G'(\xi)| \|\hat{F}_\xi\|_\infty |\eta| e^{-\frac{1}{4}\sigma^2\eta^2(1-e^{-2})}.
\end{aligned}$$

Therefore, there exists a constant C that depends only on G, ξ and σ such that

$$\|\hat{\rho}_0(t,\cdot)\|_1 + \|\hat{\rho}_0(t,\cdot)\|_\infty \leq C.$$

We can proceed in the same way after differentiating with respect to η the expression (30) of $\hat{\rho}_0$ to show that $\|\partial_\eta \hat{\rho}_0(t,\cdot)\|_\infty$ is bounded uniformly in t. □

Remark 1 When $u - G(u)$ derives from a double-well potential, such as the examples (4) and (5), then, as soon as there exist nonzero solutions $\pm\xi_e$ to the compatibility equation $\xi = G(\xi)$, we have $G'(\pm\xi_e) < 1$. Therefore, from Proposition 1 and Proposition 2, the existence of the order states $\rho_{\pm\xi_e}$ in (7) is equivalent to the stability of the 0-th mode $\rho_0^{(1)}$ in (28). In addition, the condition that the equation $\xi = G(\xi)$ has the nonzero solutions $\pm\xi_e$ implies $G'(0) > 1$, and therefore the disorder state ρ_0 is an unstable equilibrium to the nonlinear Fokker-Planck equation (7) when the order states $\rho_{\pm\xi_e}$ exist.

4.3 Sufficient condition for the stability of the kth-order mode

In this section we find a sufficient condition for the stability of the kth-order modes for $k \neq 0$. We show that all the nonzero modes are stable for

sufficiently large σ. From (26) in Lemma 1, $w_k(t) = \partial_\eta \hat{\rho}_k(t,0)$ satisfies

$$\begin{aligned} w_k(t) = &\left(-\partial_\eta g_k(t,0)\hat{\rho}_k(0, D_k(e^{-t}-1)) + \partial_\eta \hat{\rho}_k(0, D_k(e^{-t}-1))e^{-t}\right) e^{-g_k(t,0)} \\ &+ G'(\xi)\phi_k \int_0^t \Big[-\partial_\eta g_k(t-s,0) H_\xi(D_k(e^{-(t-s)}-1)) \\ &\quad + H'_\xi(D_k(e^{-(t-s)}-1))e^{-(t-s)}\Big] e^{-g_k(t-s,0)} w_k(s) ds. \end{aligned}$$

If we let

$$\beta_k(t) = D_k(1-e^{-t}), \tag{31}$$

$$\psi_k(t) = \left(-\partial_\eta g_k(t,0)\hat{\rho}_k(0,-\beta_k) + \partial_\eta \hat{\rho}_k(0,-\beta_k)e^{-t}\right) e^{-g_k(t,0)}, \tag{32}$$

$$R_k(t) = G'(\xi)\phi_k \left[-\partial_\eta g_k(t,0) H_\xi(-\beta_k) + H'_\xi(-\beta_k)e^{-t}\right] e^{-g_k(t,0)}, \tag{33}$$

then $w_k(t)$ can be rewritten as

$$w_k(t) = \psi_k(t) + \int_0^t R_k(t-s) w_k(s) ds. \tag{34}$$

The strategy is to show that $\psi_k \in L^\infty(0,\infty)$ and that $\int_0^\infty |R_k(t)|dt < 1$, and thus $w_k \in L^\infty[0,\infty)$ by (34). Once it is known that w_k is uniformly bounded, it is not difficult to show that the kth mode is stable by the inspection of the formula (26).

Lemma 2 *ψ_k defined by (32) belongs to $L^\infty[0,\infty)$.*

Proof We have

$$\begin{aligned} g_k(t,0) &= \frac{\sigma^2}{4} D_k^2 (1-e^{-2t}) + (i\xi - \sigma^2 D_k) D_k (1-e^{-t}) + (-i\xi D_k + \frac{\sigma^2}{2} D_k^2) t \\ &\simeq (-i\xi D_k + \frac{\sigma^2}{2} D_k^2) t, \quad \text{as } t \to \infty, \end{aligned}$$

$$\partial_\eta g_k(t,0) = \frac{\sigma^2}{2} D_k (1-e^{-2t}) + (i\xi - \sigma^2 D_k)(1-e^{-t}) \simeq (i\xi - \frac{\sigma^2}{2} D_k), \quad \text{as } t \to \infty.$$

Since $\hat{\rho}_k(0,\cdot)$ and $\partial_\eta \hat{\rho}_k(0,\cdot)$ are bounded we find that $\psi_k(t)$ decays exponential in time as $\exp(-\sigma^2 D_k^2 t/2)$, which gives the desired result. □

Lemma 3 *Let $R_k(t)$ be defined by (33). $\int_0^\infty |R_k(t)|dt < 1$ if*

$$|G'(\xi)\phi_k| < \left(1 + \frac{3\sqrt{2\pi}}{\sigma|D_k|} + \frac{3|\xi|}{\frac{\sigma^2}{2}|D_k|} + \frac{e^{-1}}{1+\frac{\sigma^2}{2}D_k^2}\right)^{-1}. \tag{35}$$

Proof We can show that $R_k(t) \in L^1[0,\infty)$ because it is the product of a bounded function with $e^{-g_k(t,0)}$ which decays exponentially as $t \to \infty$. Note that it is important that $k \neq 0$ and $\sigma > 0$ to ensure the decay. Recall that

$D_k = 2\pi k/L$ and $\beta_k = D_k(1-e^{-t})$. Then we can write $g_k(t,0)$, $\partial_\eta g_k(t,0)$, $H_\xi(-\beta_k)$, and $H'_\xi(-\beta_k)$ in terms of D_k and β_k:

$$\begin{aligned}
g_k(t,0) &= -\frac{1}{4}\sigma^2\beta_k^2 - \frac{1}{2}\sigma^2 D_k\beta_k + i\xi\beta_k - i\xi D_k t + \frac{1}{2}\sigma^2 D_k^2 t,\\
\partial_\eta g_k(t,0) &= \frac{1}{2}\sigma^2(e^{-t}-1)\beta_k + i\xi(1-e^{-t}),\\
H_\xi(-\beta_k) &= -\beta_k \hat{F}_\xi(-\beta_k) = -\beta_k \exp\Big(i\xi\beta_k - \frac{1}{4}\sigma^2\beta_k^2\Big),\\
H'_\xi(-\beta_k) &= \Big(1 + i\xi\beta_k - \frac{1}{2}\sigma^2\beta_k^2\Big)\exp\Big(i\xi\beta_k - \frac{1}{4}\sigma^2\beta_k^2\Big).
\end{aligned}$$

Therefore,

$$-\partial_\eta g_k(t,0)H_\xi(-\beta_k) + e^{-t}H'_\xi(-\beta_k) = \Big(-\frac{1}{2}\sigma^2\beta_k^2 + i\xi\beta_k + e^{-t}\Big)\exp\Big(i\xi\beta_k - \frac{1}{4}\sigma^2\beta_k^2\Big),$$

and we can also write $R_k(t)$ in terms of D_k and β_k:

$$\begin{aligned}
R_k(t) &= G'(\xi)\phi_k\left[-\partial_\eta g_k(t,0)H_\xi(-\beta_k) + e^{-t}H'_\xi(-\beta_k)\right]e^{-g_k(t,0)}\\
&= G'(\xi)\phi_k\Big(-\frac{1}{2}\sigma^2\beta_k^2 + i\xi\beta_k + e^{-t}\Big)\exp\Big[\frac{\sigma^2}{2}D_k\beta_k + i\xi D_k t - \frac{\sigma^2}{2}D_k^2 t\Big].
\end{aligned}$$

The L^1-norm of R_k can be written as

$$\int_0^\infty |R_k(t)|dt = |G'(\xi)\phi_k|\int_0^\infty \Big|-\frac{1}{2}\sigma^2\beta_k^2 + i\xi\beta_k + e^{-t}\Big|\exp\Big[\frac{\sigma^2}{2}D_k\beta_k - \frac{\sigma^2}{2}D_k^2 t\Big]dt. \tag{36}$$

Because $\beta_k = D_k(1-e^{-t})$, we use the following bounds for (36):

$$\begin{aligned}
\Big|-\frac{1}{2}\sigma^2\beta_k^2 + i\xi\beta_k + e^{-t}\Big| &\le \frac{\sigma^2}{2}D_k^2(1-e^{-t})^2 + |\xi||D_k|(1-e^{-t}) + e^{-t}\\
&\le \begin{cases} \frac{\sigma^2}{2}D_k^2 t^2 + |\xi||D_k|t + 1, & 0\le t\le 1,\\ \frac{\sigma^2}{2}D_k^2 + |\xi||D_k| + e^{-t}, & 1 < t,\end{cases}
\end{aligned}$$

$$\exp\Big[\frac{\sigma^2}{2}D_k\beta_k - \frac{\sigma^2}{2}D_k^2 t\Big] = \exp\Big[\frac{\sigma^2}{2}D_k^2(1-e^{-t}-t)\Big] \le \begin{cases} e^{\frac{\sigma^2}{2}D_k^2(-\frac{1}{4}t^2)}, & 0\le t\le 1,\\ e^{\frac{\sigma^2}{2}D_k^2(1-t)}, & 1 < t.\end{cases}$$

We have the following estimates:

$$\int_0^1 \Big(\frac{\sigma^2}{2}D_k^2 t^2 + |\xi||D_k|t + 1\Big)e^{\frac{\sigma^2}{2}D_k^2(-\frac{1}{4}t^2)}dt \le \frac{\sqrt{2\pi}}{\frac{\sigma}{2}|D_k|}\Big(1 + \sqrt{\frac{2}{\pi}}\frac{|\xi|}{\sigma} + \frac{1}{2}\Big),$$

$$\int_1^\infty \Big(\frac{\sigma^2}{2}D_k^2 + |\xi||D_k| + e^{-t}\Big)e^{\frac{\sigma^2}{2}D_k^2(1-t)}dt = 1 + \frac{|\xi|}{\frac{\sigma^2}{2}|D_k|} + \frac{e^{-1}}{1+\frac{\sigma^2}{2}D_k^2}.$$

Then (35) is a sufficient condition to ensure that $\int_0^\infty |R_k(t)|dt < 1$ in (36). □

By the fact that $|D_k| = 2\pi|k|/L$ is increasing in $|k|$, we can have a simplified sufficient condition from (35).

Lemma 4 *Let $R_k(t)$ be defined by (33). $\int_0^\infty |R_k(t)|dt < 1$ if*

$$|G'(\xi)\phi_k| < \left(1 + \frac{3\sqrt{2\pi}}{\sigma D_1} + \frac{3|\xi|}{\frac{\sigma^2}{2}D_1} + \frac{e^{-1}}{1+\frac{\sigma^2}{2}D_1^2}\right)^{-1}. \tag{37}$$

We can now show that the kth mode is stable for $k \neq 0$.

Proposition 3 *Let $k \neq 0$. If the condition (35) or (37) is satisfied, then the kth-order mode is stable.*

Proof First we write $\hat{\rho}_k(t,\eta)$ as the form (26). Because (35) or (37) is satisfied, $w_k(t) \in L^\infty[0,\infty)$ and thus we can show that $\|\hat{\rho}_k(t,\cdot)\|_1 + \|\hat{\rho}_k(t,\cdot)\|_\infty + \|\partial_\eta\hat{\rho}_k(t,\cdot)\|_\infty$ is uniformly bounded by the estimate similar to the proof of Proposition 2. □

4.4 Sufficient condition for the instability of the kth-order mode

The linear stability analysis that we have presented reveals that the stability of the k-th mode ($k \neq 0$) is determined by the behavior of the function w_k solution of (34). The k-th mode is stable if and only if the function w_k is bounded. We have shown that the condition $\int_0^\infty |R_k(t)|dt < 1$ is a sufficient condition for the stability of the k-th mode, which is not, however, necessary as we will see in the numerical section. It is also possible to give a sufficient condition for the instability of the k-th mode. This condition is that

$$C_k = \left\{\gamma \in \mathbb{C} \text{ s.t. } \mathrm{Re}(\gamma) > 0 \text{ and } \int_0^\infty R_k(t)e^{-\gamma t}dt = 1\right\} \tag{38}$$

is non-empty. If this happens, then the Fourier-Laplace transform of w_k:

$$\mathcal{W}_k(\gamma) = \int_0^\infty w_k(t)e^{-\gamma t}dt$$

must blow up when the complex frequency γ goes to an element in C_k, so we can conclude that $|w_k(t)|$ grows exponentially with the growth rate $\gamma_r(k) = \sup\{\mathrm{Re}(\gamma),\ \gamma \in C_k\}$. If there exists a unique $\gamma(k) \in C_k$ with $\mathrm{Re}(\gamma(k)) = \gamma_r(k)$, then we can predict that the k-th mode grows like $\exp(\gamma(k)t)$. The real part $\gamma_r(k)$ is the growth rate of the k-th mode, the imaginary part $\gamma_i(k)$ describes the dynamical behavior of the unstable mode as follows.

The mode with the largest growth rate is the one that drives the instability. If we denote $k_{\max} = \mathrm{argmax}_k(\gamma_r(k))$, then the instability has the form (up to a multiplicative constant):

$$\exp[-i(2\pi k_{\max}x/L - \gamma_i(k_{\max})t)]\exp[\gamma_r(k_{\max})t]. \tag{39}$$

For a fixed time, its spatial profile has the form of a periodic modulation with the spatial period $L/k_{\max}$. This modulation moves in time at the velocity $\gamma_i(k_{\max})L/(2\pi k_{\max})$ and grows exponentially in time with the rate $\gamma_r(k_{\max})$. This shows that the initial uniform distribution of the positions is not stable and that spatial clustering appears.

4.5 Conclusion

If $\xi = G(\xi)$ and $G'(\xi) < 1$, then the 0th-order mode of the stationary state ρ_ξ is stable. If, additionally, $|G'(\xi)\phi_k| < 1$ for all $k \neq 0$, then the stability of all the kth-order modes for $k \neq 0$ is ensured by Eq. (37) for a sufficiently large σ. Under these conditions the stationary state ρ_ξ is stable. When G is such that $u - G(u)$ derives from a double-well potential, such as the two examples (4) and (5), this means that the order states

$$\rho_{\pm\xi_e}(x,u) = \frac{1}{L}\frac{1}{\sqrt{\pi\sigma^2}}\exp\Big(-\frac{(u\mp\xi_e)^2}{\sigma^2}\Big), \quad \xi_e = G(\xi_e) > 0,$$

are stable equilibria to the nonlinear Fokker-Planck equation (7), while the disorder state

$$\rho_0(x,u) = \frac{1}{L}\frac{1}{\sqrt{\pi\sigma^2}}\exp\Big(-\frac{u^2}{\sigma^2}\Big) \tag{40}$$

is an unstable equilibrium. This shows that the noise strength σ improves the stability of the order states $\rho_{\pm\xi_e}(x,u)$. Such a phenomenon that the randomness of the movements establishes the alignment of collective motion is also found in [29] both numerically and experimentally.

Note, however, an interesting phenomenon: if $G'(\xi) < 1$, then the 0th-order mode is stable, but for some $k \neq 0$ the kth-order mode may be unstable. This instability is very different from the instability that results from violating the condition $G'(\xi) < 1$, because it manifests itself in an instability in the distribution of the positions of the particles which cannot stay uniform. This situation happens when $u - G(u)$ derives from a double-well potential and the wells are very deep, as we will see later in the numerical section. The prediction that the order states become unstable when the depths of the wells are too deep seems counterintuitive and indeed this never happens when the interaction is uniform $\phi \equiv 1$, but it happens when the interaction is local.

5 Fluctuations Analysis

In this section, we consider the fluctuation analysis for the interacting system of agents around the stationary state. Recall that μ_N is the empirical probability measure of the locations and velocities:

$$\mu_N(t,dx,du) = \frac{1}{N}\sum_{i=1}^{N}\delta_{(x_i(t),u_i(t))}(dx,du). \tag{41}$$

Here we consider a solution ξ to the compatibility equation $\xi = G(\xi)$ and we assume that $\mu_N(0,dx,du)$ converges to the stationary state $\rho_\xi(x,u)dxdu$ in (9) as $N \to \infty$. Therefore, $\mu_N(t,dx,du)$ converges to $\rho_\xi(x,u)dxdu$ for all t as $N \to \infty$ by (7). We define the fluctuation process around $\rho_\xi(x,u)dxdu$:

$$X_N(t,dx,du) = \sqrt{N}[\mu_N(t,dx,du) - \rho_\xi(x,u)dxdu],$$

and study its dynamics as $N \to \infty$.

We define, for any test function $f(x,u)$ and any measure $X(dx,du)$ (in the space of tempered Schwartz distributions):

$$\langle X, f\rangle = \int_0^L \int_{-\infty}^{\infty} f(x,u) X(dx,du).$$

We have

$$\langle \mu_N(t), f\rangle = \frac{1}{N}\sum_{i=1}^{N} f(x_i(t), u_i(t)).$$

First, at time 0, we assume that the initial positions and velocities of the agents are independent and identically distributed with the distribution with density $\rho_\xi(x,u)$. This ensures that $\mu_N(0,dx,du)$ converges to the stationary state $\rho_\xi(x,u)dxdu$, and also that $X_N(t=0,dx,du)$ converges weakly as $N \to \infty$ to a Gaussian measure $X(t=0,dx,du)$ with mean zero and covariance

$$\mathbb{E}\big[\langle X(t=0), f_1\rangle \langle X(t=0), f_2\rangle\big] = \int_{\mathbb{R}^2} f_1(x,u) f_2(x,u) \rho_\xi(x,u) dxdu - \int_{\mathbb{R}^2} f_1(x,u)\rho_\xi(x,u)dxdu \int_{\mathbb{R}^2} f_2(x,u)\rho_\xi(x,u)dxdu, \quad (42)$$

for any test functions f_1 and f_2.

Second, by Itô's formula,

$$\begin{aligned} d\langle \mu_N, f\rangle &= \frac{1}{N}\sum_{i=1}^{N} df(x_i,u_i) \\ &= \langle \mu_N, u\frac{\partial f}{\partial x}\rangle dt + \langle \mu_N, \big[G\big(\langle \mu_N(t,dx',du'), u'\phi(|x'-x|)\rangle\big) - u\big]\frac{\partial f}{\partial u}\rangle dt \\ &\quad + \langle \mu_N, \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial u^2}\rangle dt + \langle \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i,u_i)}(dx,du) dW_i(t), \sigma\frac{\partial f}{\partial u}\rangle. \end{aligned}$$

By integration by parts, we can write

$$\begin{aligned} d\mu_N &= -u\frac{\partial \mu_N}{\partial x}dt - \frac{\partial}{\partial u}\left\{[G(\langle \mu_N(t,dx',du'), u'\phi(|x'-x|)\rangle) - u]\mu_N\right\} dt \\ &\quad + \frac{1}{2}\sigma^2 \frac{\partial^2 \mu_N}{\partial u^2}dt - \sigma\frac{\partial}{\partial u}\Big[\frac{1}{N}\sum_{i=1}^{N}\delta_{(x_i,u_i)}(dx,du)dW_i(t)\Big]. \end{aligned}$$

Therefore X_N satisfies

$$\begin{aligned} dX_N &= \sqrt{N} d[\mu_N(t,dx,du) - \rho_\xi(x,u)dxdu] \\ &= -u\frac{\partial X_N}{\partial x}dt - \frac{\partial}{\partial u}\Big\{[G(\langle u'\phi(|x'-x|), \mu_N(t,dx',du')\rangle) - u]\sqrt{N}\mu_N\Big\} dt \\ &\quad + \frac{\partial}{\partial u}[(G(\xi)-u)\sqrt{N}\rho_\xi]dt + \frac{1}{2}\sigma^2\frac{\partial^2 X_N}{\partial u^2}dt \\ &\quad - \frac{\sigma}{\sqrt{N}}\sum_{i=1}^{N}\frac{\partial}{\partial u}\delta_{(x_i,u_i)}(dx,du)dW_i(t), \end{aligned}$$

that is to say

$$dX_N = -u\frac{\partial X_N}{\partial x}dt - \frac{\partial}{\partial u}\left\{\left[G\big(\langle u'\phi(|x'-x|), \mu_N(t,dx',du')\rangle\big) - u\right]X_N\right\}dt$$
$$+ \frac{\partial}{\partial u}\left\{\sqrt{N}\left[G(\xi) - G\big(\langle u'\phi(|x'-x|), \mu_N(t,dx',du')\rangle\big)\right]\rho_\xi\right\}dt$$
$$+ \frac{1}{2}\sigma^2\frac{\partial^2 X_N}{\partial u^2}dt - \sigma dW^N(t), \tag{43}$$

where $W^N(t)$ is the measure-valued stochastic process

$$W^N(t) = -\int_0^t \frac{1}{\sqrt{N}}\sum_{i=1}^N \frac{\partial}{\partial u}\delta_{(x_i,u_i)}(dx,du)dW_i(s).$$

The measure-valued stochastic process $W^N(t)$ is such that, for any test function $f(x,u)$,

$$\langle W^N(t), f\rangle = -\int_0^t\int_0^L\int_{-\infty}^{\infty} f(x,u)\frac{1}{\sqrt{N}}\sum_{i=1}^N \frac{\partial}{\partial u}\delta_{(x_i,u_i)}(dx,du)dW_i(s)$$
$$= \int_0^t \frac{1}{\sqrt{N}}\sum_{i=1}^N \frac{\partial f}{\partial u}(x_i(s),u_i(s))dW_i(s).$$

It is a zero-mean martingale whose quadratic variation is, for any test functions f_1 and f_2,

$$\big[\langle W^N, f_1\rangle, \langle W^N, f_2\rangle\big]_t = \int_0^t \frac{1}{N}\sum_{i=1}^N \frac{\partial f_1}{\partial u}(x_i(s),u_i(s))\frac{\partial f_2}{\partial u}(x_i(s),u_i(s))ds$$
$$= \int_0^t \langle \frac{\partial f_1}{\partial u}\frac{\partial f_2}{\partial u}, \mu_N(s,dx,du)\rangle ds,$$

which converges to the deterministic process

$$\big[\langle f_1, W^N\rangle, \langle f_2, W^N\rangle\big]_t \overset{N\to\infty}{\longrightarrow} t\langle \frac{\partial f_1}{\partial u}\frac{\partial f_2}{\partial u}, \rho_\xi(x,u)dxdu\rangle.$$

Consequently, as $N\to\infty$, W^N converges to a Brownian field, and this implies that X_N converges to an Ornstein-Uhlenbeck process, as explained in the following proposition.

Proposition 4 *Let $\mu_N(t,dx,du) = \frac{1}{N}\sum_{i=1}^N \delta_{(x_i(t),u_i(t))}(dx,du)$ and ξ be a solution to $\xi = G(\xi)$. If $\mu_N(0,dx,du)$ converges to the stationary state $\rho_\xi(x,u)dxdu$ as $N\to\infty$ and $X_N(0,dx,du)$ converges to the Gaussian measure-valued process $X(0,dx,du)$ with mean zero and covariance (42), then as*

$N \to \infty$, $X_N(t, dx, du)$ converges to the Gaussian measure-valued process $X(t, dx, du)$ satisfying

$$dX = -u\frac{\partial X}{\partial x}dt - \frac{\partial}{\partial u}[(G(\xi) - u)X]dt \tag{44}$$
$$- G'(\xi)\frac{\partial \rho_\xi}{\partial u}\left[\int_0^L \int_{-\infty}^{\infty} u'\phi(|x' - x|)X(t, dx', du')\right] dt + \frac{1}{2}\sigma^2\frac{\partial^2 X}{\partial u^2}dt + \sigma dW_\xi,$$

where $W_\xi(t, x, u)$ is a space-time Gaussian random noise with mean zero and covariance

$$\mathbf{Cov}\left(\int_0^L \int_{-\infty}^{\infty} W_\xi(t, x, u)f_1(x, u)dxdu, \int_0^L \int_{-\infty}^{\infty} W_\xi(t', x, u)f_2(x, u)dxdu\right)$$
$$= \min(t, t')\int_0^L \int_{-\infty}^{\infty} \frac{\partial f_1}{\partial u}(x, u)\frac{\partial f_2}{\partial u}(x, u)\rho_\xi(x, u)dxdu, \tag{45}$$

for any test functions f_1 and f_2.

Proof We only provide a formal derivation of Proposition 4. The rigorous proof is lengthy but standard. We first note that the solution X_N of (43) is the solution of a martingale problem. To prove the existence, we need to prove that the set $\{X_N\}_{N=1}^{\infty}$ solution of (43) is weakly compact and thus the sequence $\{X_N, N = 1, 2, 3, \ldots\}$ has limits. To prove the uniqueness, we need to prove that the infinitesimal generator of X_N converges to the infinitesimal generator of X as $N \to \infty$, and then the uniqueness of the limit is ensured by the uniqueness to the solution of the martingale problem. Note that, because

$$G(\xi) = G\big(\langle \rho_\xi dxdu, u'\phi(|x' - x|)\rangle\big),$$

we have

$$\big[G\big(\langle \mu_N(t, dx', du'), u'\phi(|x' - x|)\rangle\big) - u\big]X_N \overset{N\to\infty}{\longrightarrow} \big[G(\xi) - u\big]X$$

and

$$\sqrt{N}[G(\xi) - G(\langle \mu_N(t, dx', du'), u'\phi(|x' - x|)\rangle)] \overset{N\to\infty}{\longrightarrow} -G'(\xi)\langle X(t, dx'du'), u'\phi(|x' - x|)\rangle,$$

when X_N weakly converges to X. For the details of a rigorous proof for such a type, readers can see the classical work [10] by Dawson. □

The equation (44) for the fluctuation process X involves the same linearized operator as in (20), which is not a surprise. So the linear stability analysis carried out in the previous section can be invoked to claim that the system (1) is linearly stable when the initial empirical distribution is initially close to ρ_ξ, where ξ is such that $G(\xi) = \xi$, $G'(\xi) < 1$, $|\phi_k G'(\xi)| < 1$ for all $k \neq 0$, and σ is large enough. When G is such that $u - G(u)$ derives from a double-well potential, such as the two examples (4) and (5), this means that the order states $\rho_{\pm\xi_e}$ are stable equilibria, in the sense that if the initial empirical distribution is close to one of them, then the empirical distribution will remain close to it, with small normal fluctuations, while the disorder state ρ_0 is an unstable equilibrium, in the sense that if the initial empirical distribution is close to it, then the empirical distribution will quickly move away from it.

6 Large Deviations Analysis

Here we assume that two order states $\rho_{\pm\xi_e}$ exist and are stable and that the initial empirical measure $\mu_N(t=0,dx,du)$ converges to ρ_{ξ_e}. By the mean-field theory, the empirical measure

$$\mu_N(t,dx,du) = \frac{1}{N}\sum_{i=1}^{N}\delta_{(x_i(t),u_i(t))}(dx,du),$$

converges to $\rho_{\xi_e}(x,u)dxdu$ as $N\to\infty$ and the stability analysis ensures that $\mu_N(t,dx,du)$ is approximately ρ_{ξ_e} over the time interval $[0,T]$. However, when N is large but finite, the empirical measure μ_N is still stochastic and there is a very small probability that, eventhough $\mu_N \approx \rho_{\xi_e}$ at time $t=0$, $\mu_N \approx \rho_{-\xi_e}$ at time $t=T$. Here we use the large deviation principle (LDP) to formally write down the asymptotic probability of a transition from one stable order state to the other one. The empirical measure μ_N satisfies the large deviation principle in the space of continuous probability-measure-valued processes as the following: for a set A of paths of probability measures $\mu=(\mu(t,dx,du))_{t\in[0,T]}$, we have

$$\begin{aligned}-\inf_{\mu\in\mathring{A}} I(\mu) &\le \liminf_{N\to\infty}\frac{1}{N}\log\mathbf{P}(\mu_N\in A)\\ &\le \limsup_{N\to\infty}\frac{1}{N}\log\mathbf{P}(\mu_N\in A) \le -\inf_{\mu\in\bar{A}} I(\mu),\end{aligned}$$

where $\mathring{A}$ and $\bar{A}$ are the interior and closure of A, respectively, with an appropriate topology. The functional I is called the rate function:

$$I(\mu)=\frac{1}{2\sigma^2}\int_0^T \sup_{f(x,u):\langle\mu(t,\cdot,\cdot),(\frac{\partial f}{\partial u})^2\rangle\neq 0}\frac{\langle\frac{\partial\mu}{\partial t}(t,\cdot,\cdot)-\mathcal{L}^*_{\mu(t,\cdot,\cdot)}\mu(t,\cdot,\cdot),f\rangle^2}{\langle\mu(t,\cdot,\cdot),(\frac{\partial f}{\partial u})^2\rangle}dt, \tag{46}$$

where $\mathcal{L}^*_\nu$ is the differential operator associated to the Fokker-Planck equation:

$$\mathcal{L}^*_\nu\mu=-u\frac{\partial\mu}{\partial x}-\frac{\partial}{\partial u}\left[\left(G\left(\iint u'\phi(|x'|)\nu(x-x',u')du'dx'\right)-u\right)\mu\right]+\frac{1}{2}\sigma^2\frac{\partial^2\mu}{\partial u^2}.$$

We are interested in the rare event A which corresponds to the transition from one order state $\rho_{\xi_e}(x,u)dxdu$ to the other one $\rho_{-\xi_e}(x,u)dxdu$. We therefore assume that the initial conditions correspond to independent agents with the distribution $\rho_{\xi_e}(x,u)dxdu$ and we consider the event

$$A=\left\{(\mu(t,dx,du))_{t\in[0,T]} \,:\, \|\mu(T,dx,du)-\rho_{-\xi_c}(x,u)dxdu\|\le\delta\right\}, \tag{47}$$

for some small δ, where $\|\cdot\|$ stands for the metric of the space.

Roughly speaking, for large but finite N, the probability of transitions from one stable state to the other is an exponential function of N whose exponential decay rate is governed by the rate function I:

$$\mathbf{P}(\mu_N\in A)\overset{N\gg 1}{\approx}\exp\left(-N\inf_{\mu\in A} I(\mu)\right). \tag{48}$$

The transition probability (48) of the switching of the alignments of locusts is experimentally and numerically observed in [6].

Remark 2 In this section we have only provided a formal large deviation description, because the classical large deviation result [11] requires that the noise must be non-degenerate. Therefore, the result of [11] can not be directly applied to our model and we are not able to well define the space for our case until the rigorous large deviation is proven. However, a rigorous large deviation principle could be possibly constructed by the technique in [5].

7 Numerical Simulations

We use the numerical simulations to verify our theoretical analysis. The spatial domain is the torus $[0, L] = [0, 10]$, the influence function is

$$\phi(|x|) = 5 \times \mathbf{I}_{[0,L/10]}(|x|),$$

which is such that $\int_0^L \phi(|x|)dx/L = 1$ and $\phi_k = \mathrm{sinc}(\pi k/5)$, and we let

$$G(u) = \frac{h+1}{5}u - \frac{h}{125}u^3,$$

which is such that $u - G(u)$ derives from the potential $\frac{4-h}{10}u^2 + \frac{h}{500}u^4$, that is a double-well potential as soon as $h > h_c := 4$. The parameter h allows us to quantify the depths of the two wells of the double-well potential when $h > h_c$.

We discretize the time domain: $t_n = n\Delta t$, $n = 0, 1, 2, 3, \ldots$ and use the Euler method to simulate (1):

$$\begin{aligned} x_i^{n+1} - x_i^n &= u_i^n \Delta t, \\ u_i^{n+1} - u_i^n &= [G(\langle u \rangle_i^n) - u_i^n]\, \Delta t + \sigma \Delta W_i^n, \end{aligned}$$

where $\{\Delta W_i^n\}_{i,n}$ are independent Gaussian random variables with mean 0 and variance Δt, and $\langle u \rangle_i^n$ is a weighted average of the velocities with respect to u_i^n:

$$\langle u \rangle_i^n = \frac{1}{N} \sum_{j=1}^{N} u_j^n \phi(|x_j^n - x_i^n|). \tag{49}$$

7.1 Existence of the Order States

From Proposition 1, the number of stationary states (7) is equal to the number of solutions ξ for the compatibility condition (10). When $h \le h_c = 4$, the function $u - G(u)$ is increasing and 0 is the unique solution of the compatibility equation (10). When $h > h_c$, the function $u - G(u)$ derives from a double well-potential and there are three solutions $0, \pm\xi_e$ to the compatibility equation (10), with

$$\xi_e = 5\sqrt{\frac{h-4}{h}}.$$

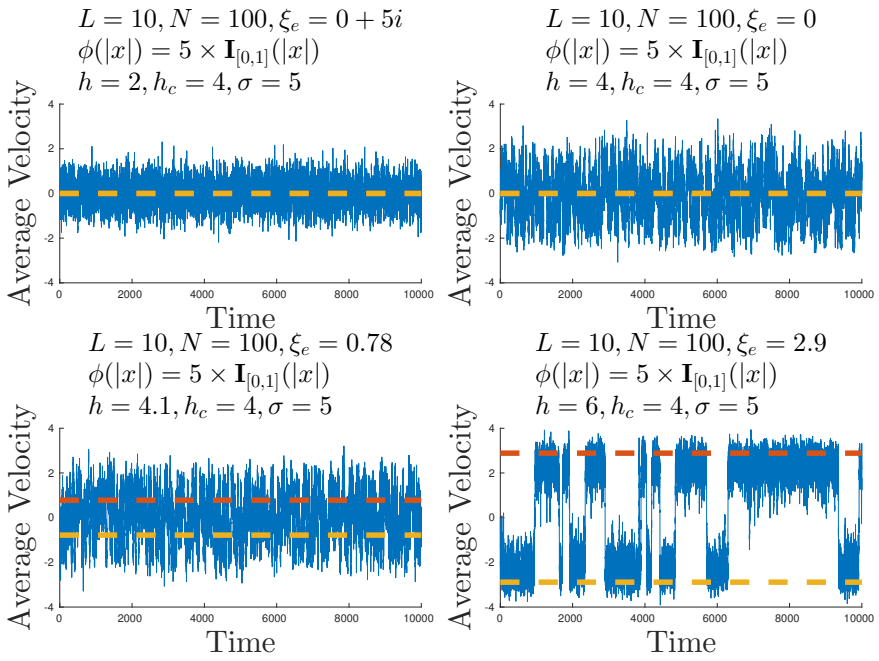

Fig. 1 The empirical average velocity $\bar{u}^n = \frac{1}{N}\sum_{i=1}^N u_i^n$ at each time step t_n for $h = 2$, 4, 4.1, and 6. The critical value to have the nonzero equilibria is $h_c = 4$. The red and yellow thick dashed lines are the theoretical average velocities ξ_e and $-\xi_e$, respectively. When $h \le h_c$ (top figures), the empirical average velocity oscillates around zero and when $h > h_c$ (bottom figures), it oscillates around either ξ_e or $-\xi_e$.

We test the cases of $h = 2$, 4, 4.1, and 6 and show the empirical average velocity $\bar{u}^n = \frac{1}{N}\sum_{i=1}^N u_i^n$ in Figure 1. The initial locations and velocities $\{(x_i^0, u_i^0)\}_{i=1}^N$ are sampled from the distribution $\rho_0(x,u)dxdu$, the stationary state in (9) with $\xi = 0$. That is, $\{x_i^0\}_{i=1}^N$ are uniformly sampled over $[0, L]$ and $\{u_i^0\}_{i=1}^N$ are sampled from the Gaussian distribution with mean 0 and variance $\sigma^2/2$. The parameters used for Figure 1 are $\Delta t = 0.5$, $N = 100$, and $\sigma = 5$.

When $h < h_c = 4$, the only solution of $\xi = G(\xi)$ is $\xi = 0$, and $G'(0) = (h+1)/5 < 1$ which means that ρ_0 is stable. Indeed, we can see in Figure 1 that the empirical average velocity $\bar{u}^n$ is oscillating around zero.

When $h > h_c$, the order states $\rho_{\pm\xi_e}$ exist. In addition, $G'(\xi_e) = (13 - 2h)/5 < (13 - 2h_c)/5 = 1$ and $G'(0) = (h+1)/5 > (h_c+1)/5 = 1$. Therefore, ρ_0 is an unstable state while $\rho_{\pm\xi_e}$ are stable states in the point of view of their zeroth-order modes, and therefore, in the point of view of their average velocities. Under these circumstances $\bar{u}^n$ oscillates around either ξ_e or $-\xi_e$ as can be seen in Figure 1.

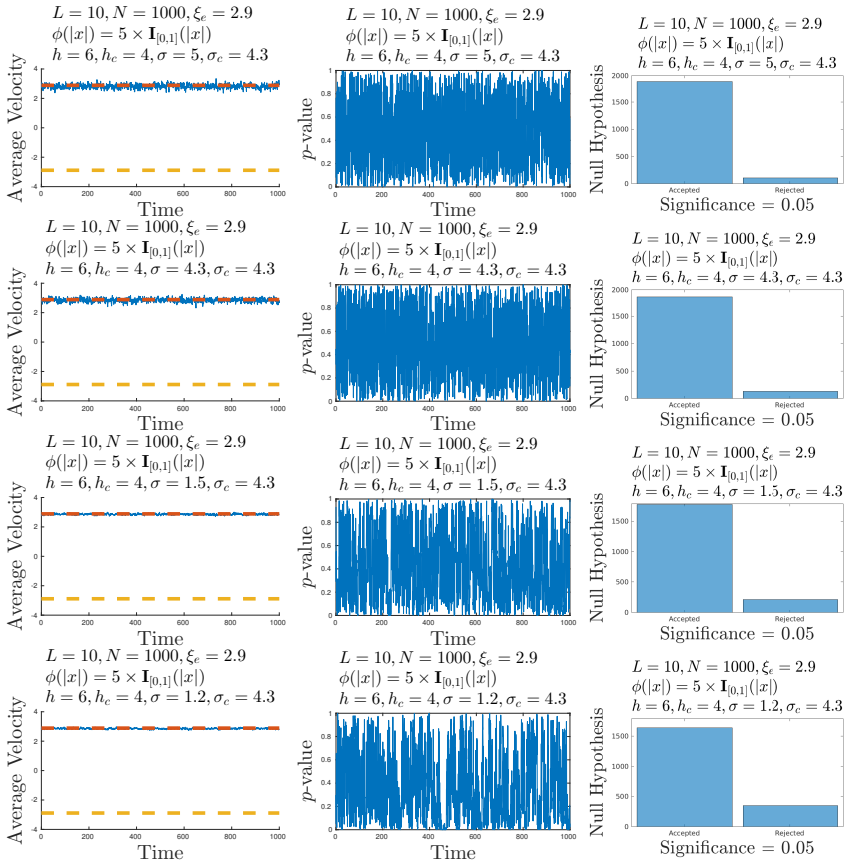

Fig. 2 The Kolmogorov-Smirnov test for the null hypothesis that $\{x_i^n\}_{i=1}^N$ is uniformly distributed on $[0, L]$ for each time step t_n. We plot the empirical average velocity $\bar{u}^n = \frac{1}{N}\sum_{i=1}^N u_i^n$ (left), the p-value (center) and the counts of the null hypothesis test with the significance level 0.05 (right) for $\sigma = 5$, 4.3, 1.5, and 1.2 (from top to bottom). When $\sigma \geq 1.2$ we are not able to reject the null hypothesis for most t_n.

7.2 Linear Stability

Here we assume that $h > h_c = 4$ so that there are three stationary states ρ_0, $\rho_{\pm\xi_e}$. Then we have $G'(0) = (h+1)/5 > 1$, $G'(\pm\xi_e) = (13-2h)/5 < 1$, and thus the 0th order mode of ρ_0 is unstable while the 0th-order modes of $\rho_{\pm\xi_e}$ are stable. An immediate conclusion is that ρ_0 is an unstable state for the nonlinear Fokker-Planck equation (7), and the empirical distribution of $\{(x_i^n, u_i^n)\}_{i=1}^N$ cannot stay close to the distribution ρ_0.

The zeroth-order mode of $\rho_{\pm\xi_e}$ is stable. However, the stabilities of $\rho_{\pm\xi_e}$ requires that all the nonzeroth-order modes of $\rho_{\pm\xi_e}$ are also stable. From Proposition 3, we know that, for $|G'(\xi_e)\phi_k| < 1$ for all $k \neq 0$, that is to say, for $h < h^C = 13/2 + 5/[2\mathrm{sinc}(\pi/5)] \simeq 9.2$, and for a sufficiently large σ, all the nonzeroth-order modes are stable, and the conditions (35) and (37)

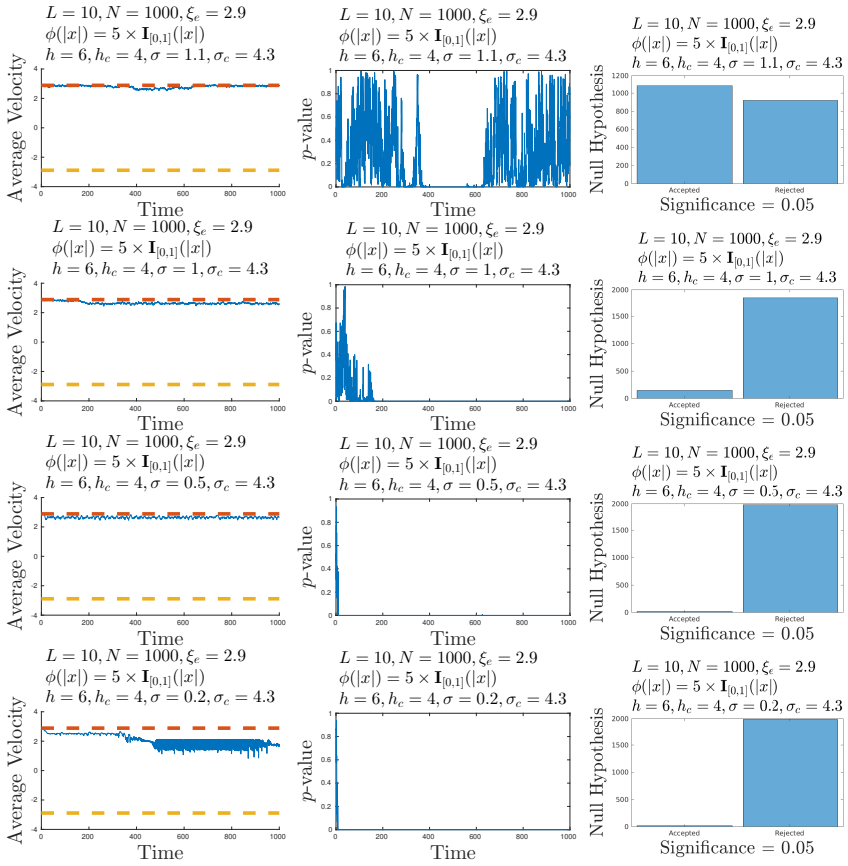

Fig. 3 Same as in Figure 2. We test the cases $\sigma = 1.1$, 1, 0.5 and 0.2 (from top to bottom). When $\sigma = 1.1$, we are able to reject roughly half of the null hypotheses and when $\sigma \leq 1$, we are able to reject the null hypothesis for most t_n.

suggest how to find the critical value of σ for the stability. We note that ρ_{ξ_e} and $\rho_{-\xi_e}$ are symmetric and we only need to analyze ρ_{ξ_e}.

In order to analyze the impact of σ to the stability of ρ_{ξ_e} for fixed h, we consider a sufficient condition slightly stricter than (35) so that we can obtain the algebraic tractability:

$$|G'(\xi_e)\phi_k| < \left(1 + \frac{3\sqrt{2\pi}}{\sigma|D_k|} + \frac{3|\xi_e|}{\frac{\sigma^2}{2}|D_k|} + \frac{e^{-1}}{\frac{\sigma^2}{2}D_k^2}\right)^{-1}, \tag{50}$$

where $D_k = 2\pi k/L$. To satisfy (50), h must be between 4 and 9.2 (so that $|G'(\xi)\phi_k|$ is less than one for all $k \neq 0$) and σ should be greater than the

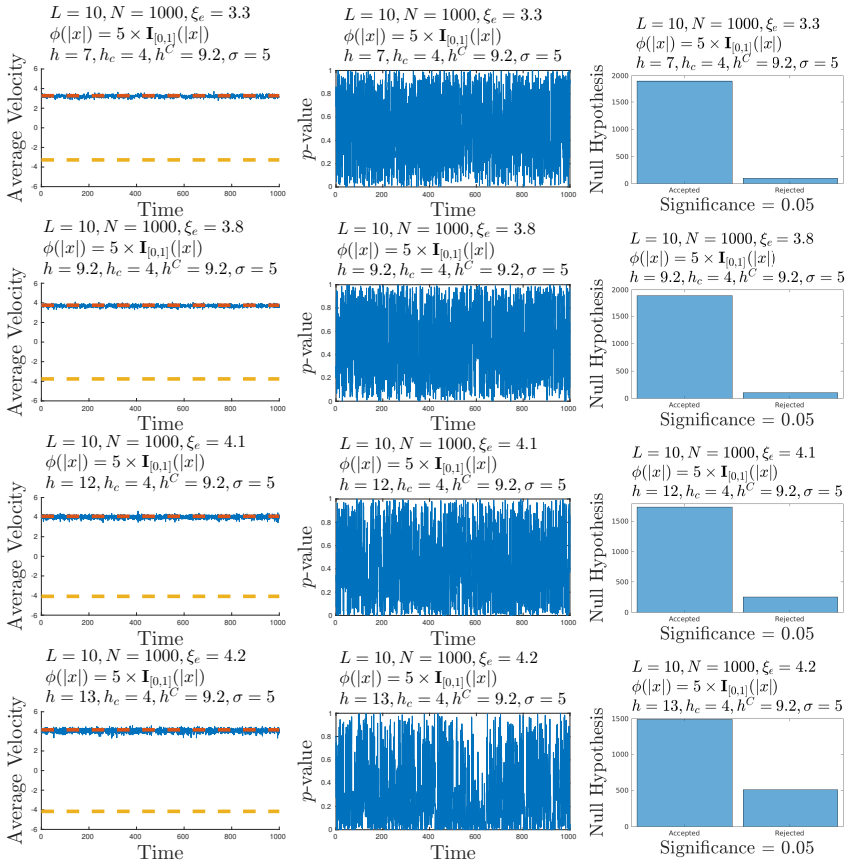

Fig. 4 The Kolmogorov-Smirnov test for the null hypothesis that $\{x_i^n\}_{i=1}^N$ is uniformly distributed on $[0, L]$ for each time step t_n. We plot the empirical average velocity $\bar{u}^n = \frac{1}{N}\sum_{i=1}^N u_i^n$ (left), the p-value (center) and the counts of the null hypothesis test with the significance level 0.05 (right) for $h = 7$, 9.2, 12 and 13 (from top to bottom). When $h \leq 13$ we are not able to reject the null hypothesis for the majority of t_n.

critical value σ_c:

$$\sigma > \sigma_c = \max_{k\neq 0} \frac{3\sqrt{2\pi} + \sqrt{(3\sqrt{2\pi})^2 + 4(|G'(\xi_e)\phi_k|^{-1} - 1)(6|\xi_e||D_k| + 2e^{-1})}}{2|D_k|(|G'(\xi_e)\phi_k|^{-1} - 1)}. \tag{51}$$

By the fact that $|\phi_k| \to 0$ and $|D_k| \to \infty$ as $|k| \to \infty$, the condition (50) becomes easier to satisfy for large $|k|$, and thus it is sufficient to find the maximum over $0 < |k| \leq K$ for an appropriate K. In our case, if $h = 6$, then $\sigma_c = 4.3$. We find from our simulations that ρ_{ξ_e} is stable when $\sigma > \sigma_c$ and becomes unstable when σ is small. We note that $\sigma_c = 4.3$ is derived from the sufficient condition (50), which is not a necessary condition. Therefore,

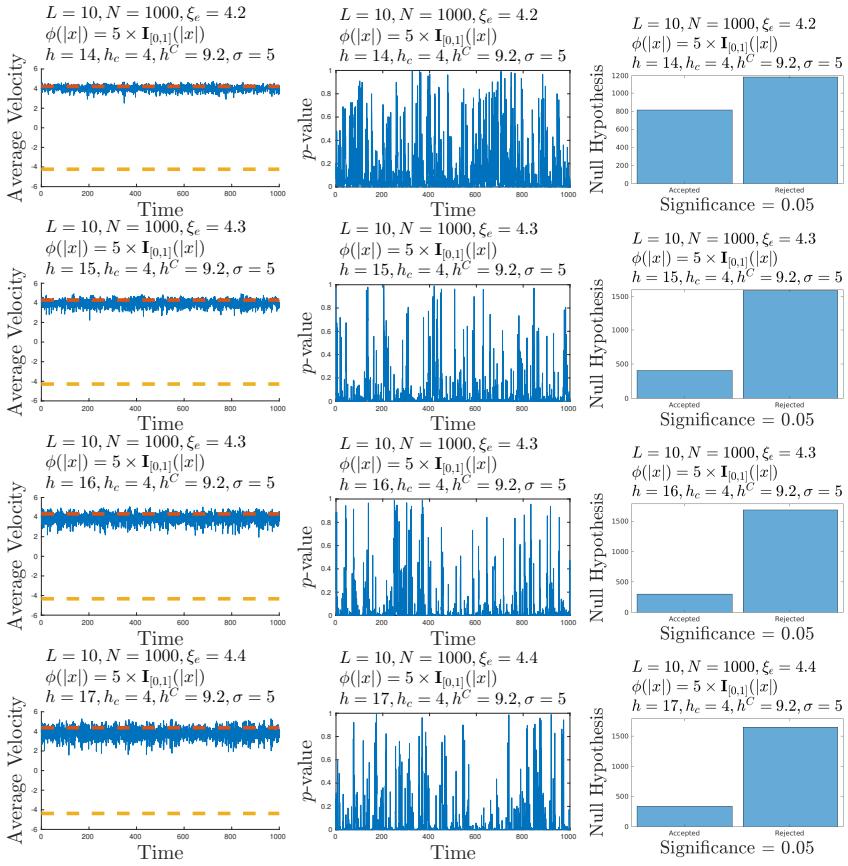

Fig. 5 Same as in Figure 4 for $h = 14$, 15, 16, and 17 (from top to bottom). When $h = 14$, we are able to reject roughly half of the null hypotheses and when $h \geq 15$, we are able to reject the null hypothesis for the majority of t_n.

$\sigma > \sigma_c = 4.3$ ensures the stability of $\rho_{\xi_e}(x, u)$, but when $\sigma \leq \sigma_c = 4.3$, it is still possible that $\rho_{\xi_e}(x, u)$ is stable.

To see the instability of ρ_{ξ_e}, it is not enough to see the empirical average velocity $\bar{u}^n = \frac{1}{N}\sum_{i=1}^N u_i^n$, because the marginal distribution $\int_0^L \rho_{\xi_e}(x, u)dx$, which is the zeroth-order mode of ρ_{ξ_e}, is always stable as soon as $h > h_c$. Instead, we observe that if ρ_{ξ_e} is not a stable state, then $\{(x_i^n, u_i^n)\}_{i=1}^N$ will not follow the distribution ρ_{ξ_e}. In particular, $\{x_i^n\}_{i=1}^N$ will not be uniformly distributed on $[0, L]$, and then we can use the Kolmogorov-Smirnov test to quantify the instability of ρ_{ξ_e}. We use the Kolmogorov-Smirnov test to compute the p-value and to perform the hypothesis testing. For each time step t_n, we will reject the null hypothesis that $\{x_i^n\}_{i=1}^N$ is uniformly distributed on $[0, L]$ if the p-value is lower than the significance level $\alpha = 0.05$. For the details of hypothesis testings and the Kolmogorov-Smirnov test, readers can see [12], for example. With $h = 6$, we have $\sigma_c = 4.3$ in (51). We test the cases $\sigma > \sigma_c$, $\sigma = \sigma_c$, and $\sigma < \sigma_c$. We compute the empirical average velocities

$\bar{u}^n$ at each time step t_n, the p-value from the Kolmogorov-Smirnov test at each time step t_n, and the counts of the null hypothesis tests. The initial locations and velocities $\{(x_i^0, u_i^0)\}_{i=1}^N$ are sampled from ρ_{ξ_e}. The parameters used for the simulations are $\Delta t = 0.5$, $N = 1000$ and $h = 6$. Here we use a larger N to achieve a better performance of the Kolmogorov-Smirnov test. From Figures 2 and 3, we find that when $\sigma \geq 1.2$, we are not able to reject the null hypothesis that $\{x_i^n\}_{i=1}^N$ are uniformly distributed on $[0, L]$ for most t_n. $\sigma = 1.1$ is close to the critical case and roughly half of time we can reject the null hypothesis with significance level 0.05. When $\sigma \leq 1$, the p-value is almost 0 and we are sure with a very high probability that $\{x_i^n\}_{i=1}^N$ are not uniformly distributed on $[0, L]$ and therefore ρ_{ξ_e} is unstable in this case. Here we see that $\sigma_c = 4.3$ is obtained from a sufficient condition and is larger than the actual critical value ($\sigma \approx 1.1$) in our simulations.

To see the impact of h to the stability of ρ_{ξ_e}, for fixed σ, we first observe that if $h \in (h_c, h^C) \simeq (4, 9.2)$, then $|G'(\xi_e)\phi_k| < 1$ for all $k \neq 0$ and the sufficient conditions (35) or (37) for stability are satisfied if σ is large enough. We therefore use the Kolmogorov-Smirnov test to perform the hypothesis testing that $\{x_i^n\}_{i=1}^N$ is uniformly distributed on $[0, L]$ for $h < h^C$, $h = h^C$ and $h > h^C$. In Figures 4 and 5, we see that when $h \leq 13$, we cannot reject the null hypothesis that $\{x_i^n\}_{i=1}^N$ are uniformly distributed on $[0, L]$ for the majority of t_n, when $h = 14$ we can reject the null hypothesis for roughly one half of $\{x_i^n\}_{i=1}^N$, and when $h \geq 15$, we can reject the null hypothesis for the majority of $\{x_i^n\}_{i=1}^N$. When $h > 17$, the empirical average $\bar{u}^n$ begins to frequently switch between ξ_e and $-\xi_e$ so that the results cannot be used to compare. This illustrates out theoretical finding that the order states become unstable when the depths of the wells become too large. We can see that the theoretical threshold value $h^C = 9.2$ has been obtained from a sufficient condition for stability and is smaller than the critical value ($h \approx 14$) in our simulations.

Finally, we study the sufficient condition for instability studied in Subsection 4.4. This allows us to clarify the instability mechanism: the average velocity of the particles is stable, but the uniform distribution of the positions of the particles is not stable and a modulation grows that gives rise to one (or several) cluster. If $h = 6$, then the theory in Subsection 4.4 predicts that the sufficient condition for the instability of the system is that σ should be smaller than 0.8 and that the most unstable mode is the first one $k_{\max} = 1$. Note that Figure 3 indeed shows that the uniform distribution of the positions is not stable when σ is smaller than 1. Figure 6 plots the predicted growth rate $\gamma_r(1)$ of the first mode and the coefficient $\gamma_i(1)$. For $\sigma = 0.5$ for instance, the formula predicts that the growth rate is 0.07, the spatial period of the modulation is $L = 10$, which means that there should be one growing moving cluster. The apparent velocity for the cluster is predicted to be approximately equal to $2.15 \times 10/(2\pi) \simeq 3.4$ by (39), which is close to, although slightly larger than, the average velocity of the particles $\xi_e = 2.9$. There is no contradiction as the apparent velocity of the cluster can be interpreted as a "phase velocity" while the average velocity of the particles can be interpreted as a "group velocity". The growth of the first mode giving rise to the moving cluster can be observed in the simulations (see Figure 7). Moreover,

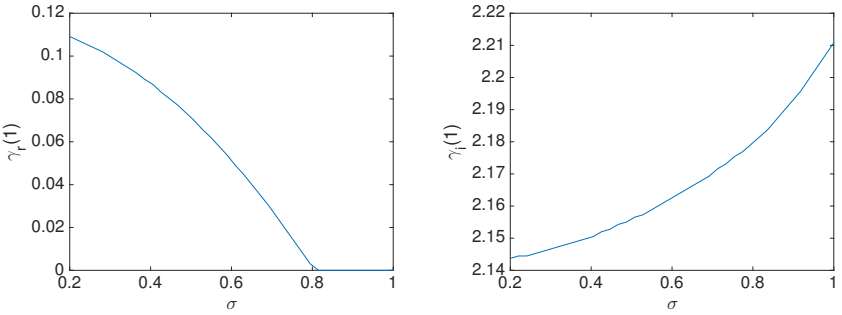

Fig. 6 The real and imaginary parts of the complex growth rate of the first mode (the most unstable mode) as a function of σ. Here $\phi(|x|) = 5 \times \mathbf{1}_{[0,1]}(|x|)$, $L = 10$, $h = 6$.

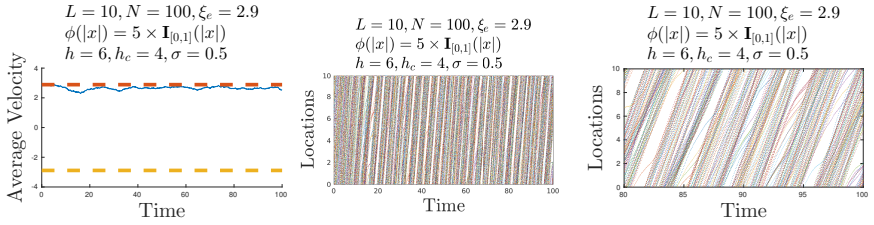

Fig. 7 The empirical average velocity (left), the trajectories of the particles for $t \in [0, 100]$ (center) and for $t \in [80, 100]$ (right). Here $\phi(|x|) = 5 \times \mathbf{1}_{[0,1]}(|x|)$, $L = 10$, $N = 100$, $h = 6$, $\sigma = 0.5$.

we can see in the right picture of Figure 7 that the particles in the tail of the moving cluster are transfered to the front of the "next" cluster (which is the same one, by periodicity), which indeed allows the cluster to apparently move faster than the average velocity of the particles: in the simulations, the average velocity of the particles is about 2.8, while the velocity of the cluster is about 3.6, very close to the predicted value.

7.3 Large Deviations

In this subsection we assume that the conditions are fulfilled so that ρ_{ξ_e} and $\rho_{-\xi_e}$ exist and are stable. We consider the transition probability from one stable state to the other:

$$\mathbf{P}(\mu_N \in A) \overset{N \gg 1}{\approx} \exp\left(-N \inf_{\mu \in A} I(\mu)\right),$$

where the rare event A is the set (47) of all possible transitions from ρ_{ξ_e} at $t = 0$ to $\rho_{-\xi_e}$ at $t = T$, and the rate function I is defined in (46).

In Figures 8, 9 and 10, we qualitatively examine the effects of N, h and σ to the transition probability $\mathbf{P}(\mu_N \in A)$ by checking the frequencies of transitions of the empirical average velocity $\bar{u}^n = \frac{1}{N}\sum_{i=1}^{N} u_i^n$ between the two stable order states $\rho_{\pm\xi_e}$. In Figure 8, we can observe that the system

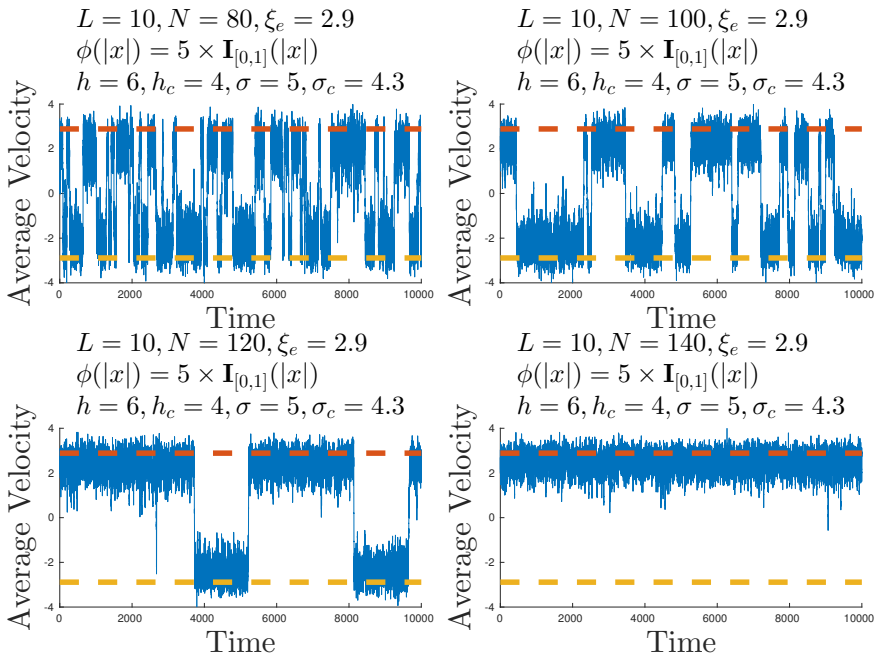

Fig. 8 The empirical average velocity $\bar{u}^n = \frac{1}{N}\sum_{i=1}^N u_i^n$ at each time step t_n for $N = 80$, 100, 120 and 140. The frequencies of the transitions between the two stable order states decay when N increases. The system has less transitions with a higher number of agents.

has less transitions as N becomes larger. In other words, the probability of transition $\mathbf{P}(\mu_N \in A)$ becomes smaller as N becomes larger; this is consistent with the fact that the large deviation principle predicts that $\mathbf{P}(\mu_N \in A)$ is an exponential decay function of N. In Figure 9, we can observe that the system experiences more transitions as σ becomes larger. Therefore, $\mathbf{P}(\mu_N \in A)$ becomes larger as σ becomes larger and this is confirmed by the fact that the rate function is approximately proportional to $1/\sigma^2$ (note, however, that σ is in the operator $\mathcal{L}^*$). In Figure 10, we find a non-trivial behavior. The probability of transition first decreases as h increases and then increases again. This shows that the rate function I is not a monotone function of h. We note that the critical value σ_c defined by (51), which is related to the linear stability of the system, is also a non-monotone function of h.

8 The Nonsymmetric Case

Let us briefly revisit the previous analysis when the average velocity has the form

$$\langle u \rangle_i = \begin{cases} \frac{1}{N_i}\sum_{j=1}^N u_j \phi(|x_j - x_i|), & \text{if } N_i > 0, \\ 0, & \text{if } N_i = 0, \end{cases} \tag{52}$$

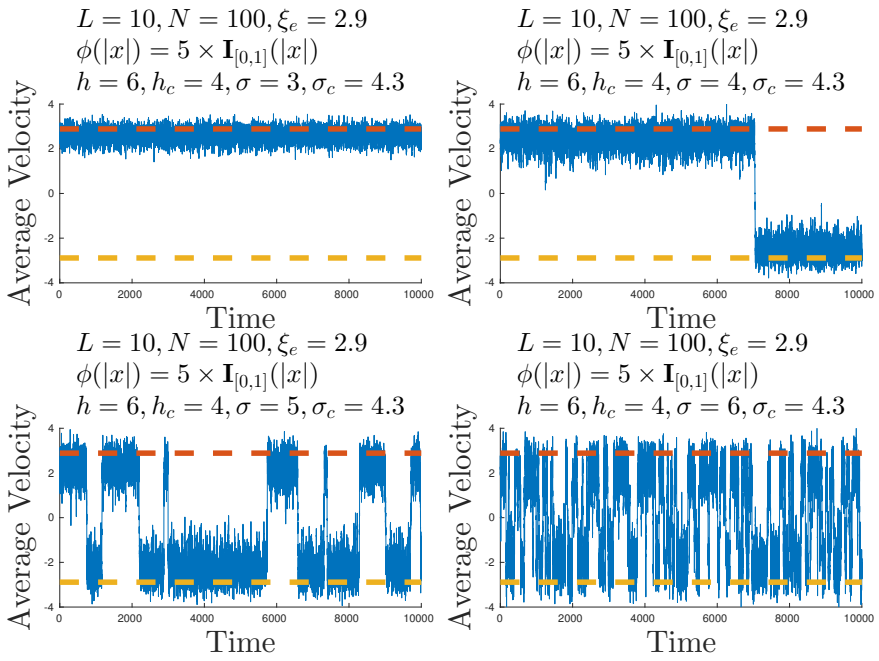

Fig. 9 The empirical average velocity $\bar{u}^n = \frac{1}{N}\sum_{i=1}^{N} u_i^n$ at each time step t_n for $\sigma = 3$, 4, 5, and 6. The frequencies of the transitions between the two stable order states increases when σ increases. The system has more transitions with a higher σ.

where N_i is the weighted number of agents in the neighborhood of the ith agent:

$$N_i = \sum_{j=1}^{N} \phi(|x_j - x_i|). \tag{53}$$

If $\mu_N(0, dx, du)$ converges to a deterministic measure $\bar{\rho}(x, u)dxdu$, then $\mu_N(t, dx, du)$ converges to the deterministic measure $\rho(t, x, u)dxdu$ whose density is the solution of the nonlinear Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = -u\frac{\partial \rho}{\partial x} - \frac{\partial}{\partial u}\left[\left(G\left(\frac{\iint u'\phi(|x'|)\rho(t, x - x', u')du'dx'}{\iint \phi(|x'|)\rho(t, x - x', u')du'dx'}\right) - u\right)\rho\right] + \frac{1}{2}\sigma^2\frac{\partial^2 \rho}{\partial u^2}, \tag{54}$$

starting from $\rho(t = 0, x, u) = \bar{\rho}(x, u)$.

The nonlinear Fokker-Planck equation (54) has stationary states that we describe in the following proposition.

Proposition 5 *Let Ξ be the set of solutions of the compatibility condition equation*

$$\xi = G(\xi). \tag{55}$$

For any $\xi \in \Xi$, the state

$$\rho_\xi(x, u) = \frac{1}{L}F_\xi(u), \quad F_\xi(u) = \frac{1}{\sqrt{\pi\sigma^2}}\exp\left(-\frac{(u - \xi)^2}{\sigma^2}\right) \tag{56}$$

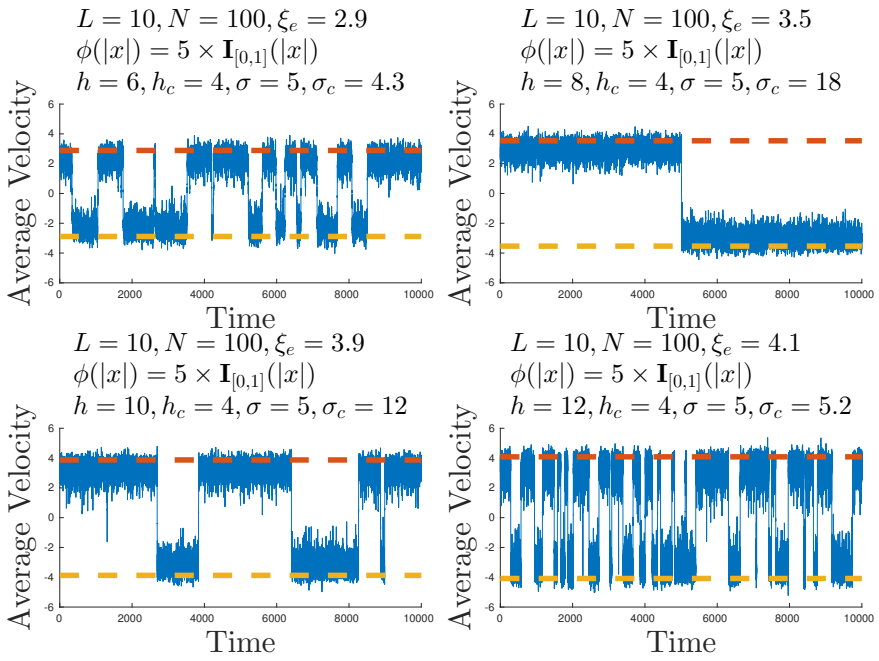

Fig. 10 The empirical average velocity $\bar{u}^n = \frac{1}{N}\sum_{i=1}^N u_i^n$ at each time step t_n for $h = 6$, 8, 10, and 12. The frequencies of the transitions between the two stable order states first decays and then increases with h. There is an optimal well depth to enhance stability in a bistable situation.

is a stationary solution of (54).

Note, however, that we could not prove that any stationary state is of the form (56). It is possible to prove that any stationary state that is uniform in space is of the form (56), but we could not prove that a stationary state must be uniform in space.

The linear stability analysis of the stationary states listed in Proposition 5 follows the same lines as in the symmetric case (2). Let $\xi \in \Xi$ and consider

$$\rho(t,x,u) = \rho_\xi(x,u) + \rho^{(1)}(t,x,u) = \frac{1}{L}F_\xi(u) + \rho^{(1)}(t,x,u), \tag{57}$$

for small perturbation $\rho^{(1)}$. We linearize the nonlinear Fokker-Planck equation (54)

$$\begin{aligned}\frac{\partial \rho^{(1)}}{\partial t} &= -u\frac{\partial \rho^{(1)}}{\partial x} - \frac{\partial}{\partial u}\left[(\xi - u)\rho^{(1)}\right] \\ &+ \frac{1}{L}G'(\xi)\left[\iint (\xi - u')\phi(|x'|)\rho^{(1)}(t, x - x', u')du'dx'\right]F_\xi'(u) + \frac{1}{2}\sigma^2\frac{\partial^2\rho^{(1)}}{\partial u^2}.\end{aligned} \tag{58}$$

The perturbation $\rho^{(1)}$ is periodic in x and can be expanded as (21). The Fourier coefficients $\hat{\rho}_k^{(1)}$ satisfy uncoupled equations

$$\frac{\partial \rho_k^{(1)}}{\partial t} = \frac{i2\pi k}{L} u\rho_k^{(1)} - \frac{\partial}{\partial u}\left[(\xi - u)\rho_k^{(1)}\right] + G'(\xi)\phi_k \left[\int (\xi - u')\rho_k^{(1)}(t, u')du'\right] F'_\xi(u) + \frac{1}{2}\sigma^2 \frac{\partial^2 \rho_k^{(1)}}{\partial u^2}. \quad (59)$$

The necessary and sufficient condition for the linear stability of the 0th order mode is

$$G'(\xi) < 1. \quad (60)$$

The sufficient condition for the linear stability of the other modes is that $|G'(\xi)\phi_k| < 1$ for all $k \neq 0$ and the noise level σ should be larger than a threshold value σ_c, which is, however different from the threshold value of the symmetric case (1-2).

9 Conclusion

We have analyzed the Czirók model for the collective motion of locusts. The mean-field theory is used to obtain a nonlinear Fokker-Planck equation as the number of agents tends to infinity. We analyze the phase transition between the disorder and order states by the existence condition for the order states. We then perform the linear stability analysis on the stationary states, and prove that the order states are stable for a sufficiently large noise level and when the interaction between the velocities of the particles is neither too small nor too strong. We provide the formal result of the fluctuation analysis and the large deviations principle. For a large but finite system we calculate the asymptotic, exponentially small transition probability from one order state to the other, which is known experimentally and numerically. Our analytical findings are verified by the extensive numerically simulations and are found in agreement with the experimental observations.

References

1. Ariel, G., Ayali, A.: Locust collective motion and its modeling. PLoS Comput Biol **11**(12), 1–25 (2015). DOI 10.1371/journal.pcbi.1004522. URL http://dx.doi.org/10.1371%2Fjournal.pcbi.1004522 2
2. Ariel, G., Ophir, Y., Levi, S., Ben-Jacob, E., Ayali, A.: Individual pause-and-go motion is instrumental to the formation and maintenance of swarms of marching locust nymphs. PLoS ONE **9**(7), 1–15 (2014). DOI 10.1371/journal.pone.0101636. URL http://dx.doi.org/10.1371%2Fjournal.pone.0101636 3
3. Bertin, E., Droz, M., Grégoire, G.: Boltzmann and hydrodynamic description for self-propelled particles. Phys. Rev. E **74**, 022,101 (2006). DOI 10.1103/PhysRevE.74.022101. URL http://link.aps.org/doi/10.1103/PhysRevE.74.022101 2, 3

4. Bertin, E., Droz, M., Grégoire, G.: Hydrodynamic equations for self-propelled particles: microscopic derivation and stability analysis. Journal of Physics A: Mathematical and Theoretical **42**(44), 445,001 (2009). URL http://stacks.iop.org/1751-8121/42/i=44/a=445001 2, 3
5. Budhiraja, A., Dupuis, P., Fischer, M.: Large deviation properties of weakly interacting processes via weak convergence methods. Ann. Probab. **40**(1), 74–102 (2012). DOI 10.1214/10-AOP616 19
6. Buhl, J., Sumpter, D.J.T., Couzin, I.D., Hale, J.J., Despland, E., Miller, E.R., Simpson, S.J.: From disorder to order in marching locusts. Science **312**(5778), 1402–1406 (2006). DOI 10.1126/science.1125142. URL http://science.sciencemag.org/content/312/5778/1402 2, 4, 19
7. Chaté, H., Ginelli, F., Grégoire, G., Peruani, F., Raynaud, F.: Modeling collective motion: variations on the vicsek model. The European Physical Journal B **64**(3), 451–456 (2008). DOI 10.1140/epjb/e2008-00275-9. URL http://dx.doi.org/10.1140/epjb/e2008-00275-9 3
8. Couzin, I.D., Krause, J., Franks, N.R., Levin, S.A.: Effective leadership and decision-making in animal groups on the move. Nature **433**(7025), 513–516 (2005). DOI 10.1038/nature03236. URL http://dx.doi.org/10.1038/nature03236 3
9. Czirók, A., Barabási, A.L., Vicsek, T.: Collective motion of self-propelled particles: Kinetic phase transition in one dimension. Phys. Rev. Lett. **82**, 209–212 (1999). DOI 10.1103/PhysRevLett.82.209. URL http://link.aps.org/doi/10.1103/PhysRevLett.82.209 2
10. Dawson, D.A.: Critical dynamics and fluctuations for a mean-field model of cooperative behavior. J. Statist. Phys. **31**(1), 29–85 (1983) 3, 5, 7, 17
11. Dawson, D.A., Gärtner, J.: Large deviations from the Mckean-Vlasov limit for weakly interacting diffusions. Stochastics **20**(4), 247–308 (1987). DOI 10.1080/17442508708833446. URL http://dx.doi.org/10.1080/17442508708833446 3, 19
12. DeGroot, M.H., Schervish, M.J.: Probability and Statistics. Addison-Wesley (2012). URL https://books.google.com/books?id=4TlEPgAACAAJ 24
13. Edelstein-Keshet, L., Watmough, J., Grunbaum, D.: Do travelling band solutions describe cohesive swarms? an investigation for migratory locusts. Journal of Mathematical Biology **36**(6), 515–549 (1998). DOI 10.1007/s002850050112. URL http://dx.doi.org/10.1007/s002850050112 3
14. Escudero, C., Yates, C.A., Buhl, J., Couzin, I.D., Erban, R., Kevrekidis, I.G., Maini, P.K.: Ergodic directional switching in mobile insect groups. Phys. Rev. E **82**, 011,926 (2010). DOI 10.1103/PhysRevE.82.011926. URL http://link.aps.org/doi/10.1103/PhysRevE.82.011926 2
15. Garnier, J., Papanicolaou, G., Yang, T.W.: Consensus convergence with stochastic effects. Vietnam Journal of Mathematics pp. 1–25 (2016). DOI 10.1007/s10013-016-0190-2. URL http://dx.doi.org/10.1007/s10013-016-0190-2 3, 7
16. Gärtner, J.: On the McKean-Vlasov limit for interacting diffusions. Math. Nachr. **137**, 197–248 (1988). DOI 10.1002/mana.19881370116 5
17. Ihle, T.: Kinetic theory of flocking: Derivation of hydrodynamic equations. Phys. Rev. E **83**, 030,901 (2011). DOI 10.1103/PhysRevE.83.030901. URL http://link.aps.org/doi/10.1103/PhysRevE.83.030901 3
18. Mishra, S., Baskaran, A., Marchetti, M.C.: Fluctuations and pattern formation in self-propelled particles. Phys. Rev. E **81**, 061,916 (2010). DOI 10.1103/PhysRevE.81.061916. URL http://link.aps.org/doi/10.1103/PhysRevE.81.061916 2, 3
19. Motsch, S., Tadmor, E.: Heterophilious dynamics enhances consensus. SIAM Review **56**(4), 577–621 (2014). DOI 10.1137/120901866. URL http://dx.doi.org/10.1137/120901866 2
20. O'Loan, O.J., Evans, M.R.: Alternating steady state in one-dimensional flocking. Journal of Physics A: Mathematical and General **32**(8), L99 (1999). URL http://stacks.iop.org/0305-4470/32/i=8/a=002 3
21. Reynolds, A.M., Sword, G.A., Simpson, S.J., Reynolds, D.R.: Predator percolation, insect outbreaks, and phase polyphenism. Current Biology **19**(1), 20

– 24 (2009). DOI http://dx.doi.org/10.1016/j.cub.2008.10.070. URL http://www.sciencedirect.com/science/article/pii/S0960982208015479 3
22. Romanczuk, P., Bär, M., Ebeling, W., Lindner, B., Schimansky-Geier, L.: Active brownian particles. The European Physical Journal Special Topics **202**(1), 1–162 (2012). DOI 10.1140/epjst/e2012-01529-y. URL http://dx.doi.org/10.1140/epjst/e2012-01529-y 3
23. Savkin, A.V.: Coordinated collective motion of groups of autonomous mobile robots: analysis of vicsek's model. In: Decision and Control, 2003. Proceedings. 42nd IEEE Conference on, vol. 3, pp. 3076–3077 Vol.3 (2003). DOI 10.1109/CDC.2003.1273095 2
24. Toner, J., Tu, Y.: Long-range order in a two-dimensional dynamical XY model: How birds fly together. Phys. Rev. Lett. **75**, 4326–4329 (1995). DOI 10.1103/PhysRevLett.75.4326. URL http://link.aps.org/doi/10.1103/PhysRevLett.75.4326 3
25. Toner, J., Tu, Y.: Flocks, herds, and schools: A quantitative theory of flocking. Phys. Rev. E **58**, 4828–4858 (1998). DOI 10.1103/PhysRevE.58.4828. URL http://link.aps.org/doi/10.1103/PhysRevE.58.4828 3
26. Topaz, C.M., D'Orsogna, M.R., Edelstein-Keshet, L., Bernoff, A.J.: Locust dynamics: Behavioral phase change and swarming. PLoS Comput Biol **8**(8), 1–11 (2012). DOI 10.1371/journal.pcbi.1002642. URL http://dx.doi.org/10.1371%2Fjournal.pcbi.1002642 3
27. Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., Shochet, O.: Novel type of phase transition in a system of self-driven particles. Phys. Rev. Lett. **75**, 1226–1229 (1995). DOI 10.1103/PhysRevLett.75.1226. URL http://link.aps.org/doi/10.1103/PhysRevLett.75.1226 2
28. Vicsek, T., Zafeiris, A.: Collective motion. Physics Reports **517**, 71 – 140 (2012). DOI http://dx.doi.org/10.1016/j.physrep.2012.03.004. URL http://www.sciencedirect.com/science/article/pii/S0370157312000968. Collective motion 2
29. Yates, C.A., Erban, R., Escudero, C., Couzin, I.D., Buhl, J., Kevrekidis, I.G., Maini, P.K., Sumpter, D.J.T.: Inherent noise can facilitate coherence in collective swarm motion. Proceedings of the National Academy of Sciences **106**(14), 5464–5469 (2009). DOI 10.1073/pnas.0811195106. URL http://www.pnas.org/content/106/14/5464.abstract 2, 4, 14