EMERGENCE OF TIME-ASYMPTOTIC FLOCKING IN A STOCHASTIC CUCKER-SMALE SYSTEM[*]

SEUNG-YEAL HA[†], KISEOP LEE[‡], AND DORON LEVY[§]

Abstract. We study a stochastic Cucker-Smale flocking system in which particles interact with the environment through white noise. We provide the definition of flocking for the stochastic system, and show that when the communication rate is constant, the system exhibits a flocking behavior independent of the initial configurations. For the case of a radially symmetric communication rate with a positive lower bound, we show that the relative fluctuations of the particle velocity around the mean velocity have a uniformly bounded variance in time. We conclude with numerical simulations that validate our analytical results.

1. Introduction

Flocking, or '*coordinated motion*', arises from the interaction between self-propelled mobile agents that follow simple local rules of interaction. Flocking phenomena are ubiquitous in many biological and physical systems such as bacteria, bees, flies, birds, fish, humans, and interacting oscillators, to name a few [1, 4, 5, 6, 9, 11, 12, 13, 14, 15, 17, 18, 19]. Recently several mathematical models for flocking were introduced and analyzed [2, 3, 4, 5, 6, 16]. Among them, our main interest in this paper is the work of Cucker and Smale [2, 3]. The Cucker-Smale (C-S) particle model is a simple relaxation-type model for an N-particle system that reveals a phase transition depending on the level of communication rates (see [2, 3, 8, 10, 16] for details). The C-S particle model can be described as follows: we assume N autonomous agents and let $(\boldsymbol{x}_i, \boldsymbol{v}_i) \in \mathbb{R}^{2d}$ denote the phase-space coordinate of the i-th agent. Then the C-S model reads as

$$\begin{cases} \dfrac{d\boldsymbol{x}_i}{dt} = \boldsymbol{v}_i, \\[2mm] \dfrac{d\boldsymbol{v}_i}{dt} = \dfrac{\lambda}{N} \sum_{j=1}^{N} \psi(\boldsymbol{x}_j, \boldsymbol{x}_i)(\boldsymbol{v}_j - \boldsymbol{v}_i), \end{cases} \qquad 1 \le i \le N. \tag{1.1}$$

Here, λ is a nonnegative coupling strength. The communication rate, ψ, is a nonnegative measurable function, and it satisfies the symmetry condition and translation invariance:

$$\psi(\boldsymbol{x}_i, \boldsymbol{x}_j) = \psi(\boldsymbol{x}_j, \boldsymbol{x}_i), \qquad 1 \le i, j \le N, \tag{1.2}$$

$$\psi(\boldsymbol{x}_j, \boldsymbol{x}_i) = \psi(\boldsymbol{x}_j + \mathbf{C}, \boldsymbol{x}_i + \mathbf{C}), \qquad \forall C \in \mathbb{R}^d. \tag{1.3}$$

In [2, 3] Cucker and Smale showed that if the communication has a long-range interaction, then a global unconditional flocking occurs, which means that the velocities of all agents converge to the same asymptotic velocity as long as the initial configurations are compactly supported in the phase space. In contrast, when the communication has a short range, then conditional flocking occurs for the restricted class of initial configurations. For more details we refer to [2, 3, 8, 9, 10, 12, 16].

The C-S model (1.1) does not take into account any interactions between the particle system and the environment. One possible way of modeling such interactions is to add noise terms to the deterministic dynamical system. Indeed, in this paper we introduce such a stochastic particle model, i.e., a stochastically perturbed Cucker-Smale system (in short: *the SCS system*).

Our starting point is to revisit the definition of flocking. Once a proper definition in the stochastic context is established, we can then ask whether flocking can be expected at all in a stochastically driven particle system, and if this is the case, under what conditions does it emerge? The purpose of this paper is to study the time-asymptotic flocking for the new SCS system.

The rest of the paper is organized as follows. In Sec. 2 we present a stochastic counterpart of the Cucker-Smale model. This system is then decomposed into two parts: one system that follows the macroscopic (coarse-scale) dynamics, and a second system that follows the microscopic (fine-scale) fluctuations. The system is analyzed in Sec. 3. For a constant communication rate, we explicitly obtain all of the statistical quantities for the random velocity process which leads to time-asymptotic strong flocking. When the mutual communication rate depends on the distance between particles, we provide a sufficient a priori condition that leads to strong flocking. Numerical simulations for the cases studied in this work are shown and discussed in Sec. 4. We end at Sec. 5 with concluding remarks.

2. A stochastic Cucker-Smale model for flocking

In this section we present a new stochastic Cucker-Smale (SCS) flocking model in which we assume that particles interact with the environment via stochastic noise. We model these mutual interactions between the particle system and their environment by multi-dimensional white noise processes, and illustrate several concepts of stochastic flocking.

Throughout the paper we use superscripts to denote the components of vectors and subscripts to denote sequences of vectors.

Consider a stochastic particle system consisting of N-identical self-propelled autonomous agents (or particles) of unit mass in the presence of white noise of uniform strength $\sqrt{D}$. Let $(\boldsymbol{x}_i(t), \boldsymbol{v}_i(t)) \in \mathbb{R}^{2d}$ be the position and velocity of particle i. The stochastic C-S dynamics is then governed by the following Ornstein-Uhlenbeck process:

$$\begin{cases} d\boldsymbol{x}_i = \boldsymbol{v}_i dt, \\ \\ d\boldsymbol{v}_i = \dfrac{\lambda}{N} \displaystyle\sum_{j=1}^{N} \psi(\boldsymbol{x}_j, \boldsymbol{x}_i)(\boldsymbol{v}_j - \boldsymbol{v}_i) dt + \sqrt{D} d\boldsymbol{W}_i, \end{cases} \qquad 1 \leq i \leq N, \qquad (2.1)$$

subject to "*deterministic*" initial data

$$(\boldsymbol{x}_i, \boldsymbol{v}_i)(0) = (\boldsymbol{x}_{i0}, \boldsymbol{v}_{i0}), \qquad i = 1, \ldots, N.$$

Similar to the deterministic system (1.1), λ is a positive coupling strength and ψ is the bi-mutual communication rate between particles. The noise term $d\boldsymbol{W}_i$ is an i.i.d.

d-dimensional white noise characterized by mean zero and the following covariance relations: for $1 \leq \alpha, \beta \leq d$, $1 \leq i, j \leq N$,

$$\langle d\boldsymbol{W}_i^\alpha(t) \rangle = 0, \quad \left\langle d\boldsymbol{W}_i^\alpha(t) d\boldsymbol{W}_j^\beta(t_*) \right\rangle = \delta_{\alpha\beta} \delta_{ij} \delta(t - t_*). \tag{2.2}$$

The notation $\langle \cdot \rangle$ denotes an ensemble average. We next define several concepts on the asymptotic flocking for the SCS system (2.1).

DEFINITION 2.1. *The system* (2.1) *has a (time-asymptotic) flocking if and only if the solutions* $\{\boldsymbol{x}_i, \boldsymbol{v}_i\}, i = 1, \ldots, N$ *to* (1.1) *satisfy the following two conditions: for* $1 \leq i, j \leq N$,

  1. *The differences of expectations in all velocity process go to zero time-asymptotically (velocity alignment):*

$$\lim_{t \to +\infty} |\langle \boldsymbol{v}_i(t) \rangle - \langle \boldsymbol{v}_j(t) \rangle| = 0.$$

  2. *The mean diameter of a group is uniformly bounded in time t (group formation):*

$$\sup_{0 \leq t < \infty} |\langle \boldsymbol{x}_i(t) \rangle - \langle \boldsymbol{x}_j(t) \rangle| < \infty.$$

REMARK 2.2. When noise is turned off, the above definition coincides with flocking in the deterministic case (see [2, 3, 8, 10]).

2.1. A macro-micro decomposition. We introduce a macro-micro decomposition [8, 9] which decomposes the system into two parts: one system that describes the macroscopic (coarse-scale) dynamics and a second system that describes the microscopic (fine-scale) dynamics. We start by setting the macroscopic part for the solution as the center of mass system $(\boldsymbol{x}_c, \boldsymbol{v}_c)$:

$$\boldsymbol{x}_c \equiv \frac{1}{N} \sum_{i=1}^N \boldsymbol{x}_i, \qquad \boldsymbol{v}_c \equiv \frac{1}{N} \sum_{i=1}^N \boldsymbol{v}_i. \tag{2.3}$$

The microscopic variables are then taken as the fluctuations with respect to the center of mass, both in location and in velocity:

$$\hat{\boldsymbol{x}}_i \equiv \boldsymbol{x}_i - \boldsymbol{x}_c, \quad \hat{\boldsymbol{v}}_i \equiv \boldsymbol{v}_i - \boldsymbol{v}_c. \tag{2.4}$$

Under Assumption (1.3), the macroscopic and microscopic parts of (2.1) can be written as

$$\text{(Macro)} \quad \begin{cases} d\boldsymbol{x}_c = \boldsymbol{v}_c dt, \\ d\boldsymbol{v}_c = \dfrac{\sqrt{D}}{N} \sum_{i=1}^N d\boldsymbol{W}_i, \end{cases} \tag{2.5}$$

and

$$\text{(Micro)} \quad \begin{cases} d\hat{\boldsymbol{x}}_i = \hat{\boldsymbol{v}}_i dt, \\ d\hat{\boldsymbol{v}}_i = \dfrac{\lambda}{N} \sum_{j=1}^N \psi(\hat{\boldsymbol{x}}_j, \hat{\boldsymbol{x}}_i)(\hat{\boldsymbol{v}}_j - \hat{\boldsymbol{v}}_i) dt + \sqrt{D}\left(1 - \dfrac{1}{N}\right) d\boldsymbol{W}_i \\ \qquad\quad - \dfrac{\sqrt{D}}{N} \sum_{j \neq i} d\boldsymbol{W}_j. \end{cases} \tag{2.6}$$

REMARK 2.3.

(i) It is easy to see that the initial data for the microscopic part must satisfy

$$\sum_{i=1}^{N} \hat{\boldsymbol{x}}_i = \sum_{i=1}^{N} \hat{\boldsymbol{v}}_i = 0. \tag{2.7}$$

(ii) In the noiseless case ($D=0$) the macroscopic part (2.5) and the microscopic part (2.6) are completely decoupled and the dynamics of the microscopic part is identical to the dynamics of the original system except for the constraint (2.7) on the initial data.

2.2. The dynamics of the macroscopic variables.   We now study the dynamics of the macroscopic variables $(\boldsymbol{x}_c, \boldsymbol{v}_c)$ with deterministic initial data $(\boldsymbol{x}_c(0), \boldsymbol{v}_c(0))$. In this case, the macroscopic system (2.5) is exactly solvable:

$$\begin{aligned}
\boldsymbol{v}_c(t) &= \boldsymbol{v}_c(0) + \frac{\sqrt{D}}{N} \sum_{i=1}^{N} \boldsymbol{W}_i(t), \\
\boldsymbol{x}_c(t) &= \boldsymbol{x}_c(0) + t\boldsymbol{v}_c(0) + \frac{\sqrt{D}}{N} \sum_{i=1}^{N} \int_0^t \boldsymbol{W}_i(s)ds,
\end{aligned} \tag{2.8}$$

where the initial data $(\boldsymbol{x}_c(0), \boldsymbol{v}_c(0))$ are deterministic, i.e.,

$$\langle \boldsymbol{x}_c(0) \rangle = \boldsymbol{x}_c(0), \quad \langle \boldsymbol{v}_c(0) \rangle = \boldsymbol{v}_c(0).$$

PROPOSITION 2.4. *Let* $(\boldsymbol{x}_c, \boldsymbol{v}_c)$ *be the solution of* (2.5) *given by the formula* (2.8). *Then for* $\alpha, \beta \in \{1, \ldots, d\}$, *we have*

$$(i) \ \langle \boldsymbol{v}_c(t) \rangle = \boldsymbol{v}_c(0), \quad var[\boldsymbol{v}_c^{\alpha}(t)] = \frac{Dt}{N},$$

$$(ii) \ \langle \boldsymbol{x}_c(t) \rangle = \boldsymbol{x}_c(0) + t\boldsymbol{v}_c(0), \quad var[\boldsymbol{x}_c^{\alpha}(t)] = \frac{Dt^2}{2N}.$$

Proof.

(i) We use the basic properties of Brownian motion

$$\langle \boldsymbol{W}_i(t) \rangle = 0, \qquad \left\langle \boldsymbol{W}_i^{\alpha}(t) \boldsymbol{W}_j^{\beta}(t) \right\rangle = \delta_{\alpha\beta} \delta_{i,j} t,$$

to obtain the variance estimate

$$var[\boldsymbol{v}_c^{\alpha}(t)] \equiv \left\langle |\boldsymbol{v}_c^{\alpha}(t) - \boldsymbol{v}_c^{\alpha}(0)|^2 \right\rangle = \frac{D}{N^2} \sum_{1 \le i,j \le N} \langle \boldsymbol{W}_i^{\alpha}(t) \boldsymbol{W}_j^{\alpha}(t) \rangle = \frac{Dt}{N}.$$

(ii) As in (i), we only need to consider the estimate for the variance

$$\begin{aligned}
var[\boldsymbol{x}_c^{\alpha}(t)] &\equiv \left\langle |\boldsymbol{x}_c^{\alpha}(t) - \boldsymbol{x}_c^{\alpha}(0) - t\boldsymbol{v}_c^{\alpha}(0)|^2 \right\rangle \\
&= \frac{D}{N^2} \sum_{1 \le i,j \le N} \int_0^t \langle \boldsymbol{W}_i^{\alpha}(s) \boldsymbol{W}_j^{\alpha}(s) \rangle ds = \frac{Dt^2}{2N}.
\end{aligned}$$

□

REMARK 2.5. In the mean-field limit $(N \to \infty)$, the dynamics of the macroscopic variables is governed by the pure translation motion as in the noiseless case [2, 3, 8, 9]. The statistical estimates in Proposition 2.4 imply the strong law of large numbers.

COROLLARY 2.1. *For any fixed $t \in [0, \infty)$, the macroscopic velocity $\boldsymbol{v}_c(t) - \boldsymbol{v}_c(0)$ converges to 0 almost surely as $N \to \infty$, i.e.,*

$$\boldsymbol{v}_c(t) - \boldsymbol{v}_c(0) \to 0, \quad a.s.$$

Proof. It suffices to show that for $\alpha \in \{1, \dots, d\}$,

$$v_c^\alpha(t) - v_c^\alpha(0) \to 0 \quad \text{a.s} \quad \text{as } N \to \infty.$$

Note that

$$\boldsymbol{v}_c^\alpha(t) - \boldsymbol{v}_c^\alpha(0) = \frac{\sqrt{D}}{N} \sum_{i=1}^N W_i^\alpha(t).$$

Since for each finite t, W_i^α is an i.i.d. random variable with mean zero and variance 1, by the strong law of large numbers (see [7]), we have

$$\frac{1}{N} \sum_{i=1}^N W_i^\alpha(t) \to 0 \quad \text{a.s.}, \quad \text{as } N \to \infty. \qquad \square$$

3. The dynamics of the microscopic variables

In this section, we present several estimates on the statistical fluctuations of the microscopic variables $(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{v}}_i)$. To simplify the notations we drop the hat notation in the microscopic variables and use $(\boldsymbol{x}_i, \boldsymbol{v}_i)$ instead of $(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{v}}_i)$.

Consider the dynamics of the microscopic variables given by the system (2.6) which we rewrite without the hat notation as:

$$\begin{cases} d\boldsymbol{x}_i = \boldsymbol{v}_i dt, \\ d\boldsymbol{v}_i = \dfrac{\lambda}{N} \sum_{j=1}^N \psi(\boldsymbol{x}_j, \boldsymbol{x}_i)(\boldsymbol{v}_j - \boldsymbol{v}_i)dt + \sqrt{D}\left(1 - \dfrac{1}{N}\right)d\boldsymbol{W}_i - \dfrac{\sqrt{D}}{N} \sum_{j \neq i} d\boldsymbol{W}_j. \end{cases} \tag{3.1}$$

The system (3.1) is augmented with the initial data

$$\boldsymbol{x}_i(0) = \boldsymbol{x}_{i0}, \quad \boldsymbol{v}_i(0) = \boldsymbol{v}_{i0}, \quad i = 1, \dots, N. \tag{3.2}$$

Since by (2.7) the microscopic initial data is constrained to satisfy

$$\sum_{i=1}^N \boldsymbol{x}_{i0} = 0 \quad \text{and} \quad \sum_{i=1}^N \boldsymbol{v}_{i0} = 0,$$

and the system (3.1) conserves the momentum, we have $\forall t \geq 0$

$$\sum_{i=1}^N \boldsymbol{x}_i(t) = 0 \quad \text{and} \quad \sum_{i=1}^N \boldsymbol{v}_i(t) = 0. \tag{3.3}$$

We also note that the strong solutions to (3.1) satisfy Itô's integral representations:

$$
\begin{cases}
\boldsymbol{x}_i(t) = \boldsymbol{x}_{i0} + \displaystyle\int_0^t \boldsymbol{v}_i(s)ds, \\[2ex]
\boldsymbol{v}_i(t) = \boldsymbol{v}_{i0} + \dfrac{\lambda}{N}\displaystyle\sum_{j=1}^N \int_0^t \psi(\boldsymbol{x}_j(s),\boldsymbol{x}_i(s))(\boldsymbol{v}_j(s)-\boldsymbol{v}_i(s))ds \\[2ex]
\qquad\qquad + \sqrt{D}\left(1-\dfrac{1}{N}\right)\boldsymbol{W}_i - \dfrac{\sqrt{D}}{N}\displaystyle\sum_{j\neq i}\boldsymbol{W}_j.
\end{cases}
\tag{3.4}
$$

Before we analyze the nonlinear system (3.4), we first consider the linear case.

3.1. The linear case.   In this case we assume a constant communication rate, $\psi = 1$. Under such a condition, the dynamics of $\boldsymbol{v}_i$ is decoupled from the dynamics of $\boldsymbol{x}_i$. The equation for $\boldsymbol{v}_i$ then becomes

$$
d\boldsymbol{v}_i = \frac{\lambda}{N}\sum_{j=1}^N (\boldsymbol{v}_j - \boldsymbol{v}_i)dt + \sqrt{D}\left(1-\frac{1}{N}\right)d\boldsymbol{W}_i - \frac{\sqrt{D}}{N}\sum_{j\neq i}d\boldsymbol{W}_j.
\tag{3.5}
$$

We now use property (3.3) to simplify the mean-field interaction term

$$
\frac{\lambda}{N}\sum_{j=1}^N (\boldsymbol{v}_j - \boldsymbol{v}_i) = \frac{\lambda}{N}\sum_{j=1}^N \boldsymbol{v}_j - \lambda\boldsymbol{v}_i = -\lambda\boldsymbol{v}_i,
$$

so that (3.5) becomes

$$
d\boldsymbol{v}_i = -\lambda\boldsymbol{v}_i dt + \sqrt{D}\left(1-\frac{1}{N}\right)d\boldsymbol{W}_i - \frac{\sqrt{D}}{N}\sum_{j\neq i}d\boldsymbol{W}_j.
\tag{3.6}
$$

Hence the solution $\boldsymbol{v}_i$ satisfies

$$
\boldsymbol{v}_i(t) = e^{-\lambda t}\boldsymbol{v}_{i0} + \int_0^t e^{-\lambda(t-s)}\left[\sqrt{D}\left(1-\frac{1}{N}\right)d\boldsymbol{W}_i(s) - \frac{\sqrt{D}}{N}\sum_{j\neq i}d\boldsymbol{W}_j(s)\right].
\tag{3.7}
$$

We now use the explicit representation (3.7) for $\boldsymbol{v}_i$ to derive several statistical estimates.

PROPOSITION 3.1. *The velocity random process $(\boldsymbol{v}_i(t))$ given by (3.7) satisfies the following estimates. For $1\leq i,j\leq N$ and $1\leq\alpha\leq d$,*

(i) $\langle\boldsymbol{v}_i^\alpha(t)\rangle = e^{-\lambda t}\boldsymbol{v}_i^\alpha(0),$

(ii) $var[\boldsymbol{v}_i^\alpha(t)] = \dfrac{D}{2\lambda}\left(1-\dfrac{1}{N}\right)(1-e^{-2\lambda t}),$

(iii) $\left|\langle\boldsymbol{v}_i^\alpha(t)\rangle - \langle\boldsymbol{v}_j^\alpha(t)\rangle\right| = e^{-\lambda t}\left|\langle\boldsymbol{v}_i^\alpha(0)\rangle - \langle\boldsymbol{v}_j^\alpha(0)\rangle\right|,$

(iv) $\left\langle\left|\boldsymbol{v}_i^\alpha(t) - \boldsymbol{v}_j^\alpha(t)\right|^2\right\rangle = e^{-2\lambda t}\left|\boldsymbol{v}_i^\alpha(0) - \boldsymbol{v}_j^\alpha(0)\right|^2 + \dfrac{D}{\lambda}(1-e^{-2\lambda t}), \quad \forall\ i\neq j.$

Proof. Since (i) and (iii) are straightforward, we only consider (ii) and (iv).
(ii) We use (3.7) and the result (i) to find

$$\boldsymbol{v}_i^\alpha(t) - \langle \boldsymbol{v}_i^\alpha(t) \rangle = \int_0^t e^{-\lambda(t-s)} \left(\sqrt{D} d\boldsymbol{W}_i^\alpha(s) - \frac{\sqrt{D}}{N} \sum_{k=1}^N d\boldsymbol{W}_k^\alpha(s) \right).$$

This yields

$$(\boldsymbol{v}_i^\alpha(t) - \langle \boldsymbol{v}_i^\alpha(t) \rangle)^2$$
$$= \int_0^t \int_0^t e^{-\lambda(2t-s_1-s_2)} \left(\sqrt{D} d\boldsymbol{W}_i^\alpha(s_1) - \frac{\sqrt{D}}{N} \sum_{k=1}^N d\boldsymbol{W}_k^\alpha(s_1) \right)$$
$$\times \left(\sqrt{D} d\boldsymbol{W}_i^\alpha(s_2) - \frac{\sqrt{D}}{N} \sum_{k=1}^N d\boldsymbol{W}_k^\alpha(s_2) \right),$$
$$= D \int_0^t \int_0^t e^{-\lambda(2t-s_1-s_2)} \left(d\boldsymbol{W}_i^\alpha(s_1) d\boldsymbol{W}_i^\alpha(s_2) - \frac{1}{N} \sum_{k=1}^N d\boldsymbol{W}_k^\alpha(s_1) d\boldsymbol{W}_i^\alpha(s_2) \right.$$
$$\left. - \frac{1}{N} \sum_{k=1}^N d\boldsymbol{W}_k^\alpha(s_2) d\boldsymbol{W}_i^\alpha(s_1) + \frac{1}{N^2} \sum_{k,l=1}^N d\boldsymbol{W}_k^\alpha(s_1) d\boldsymbol{W}_l^\alpha(s_2) \right).$$

We now take an ensemble average to find the variance of $\boldsymbol{v}_i^\alpha$:

$$\mathrm{var}[\boldsymbol{v}_i^\alpha(t)] = \left\langle (\boldsymbol{v}_i^\alpha(t) - \langle \boldsymbol{v}_i^\alpha(t) \rangle)^2 \right\rangle,$$
$$= D \left(1 - \frac{1}{N} \right) \int_0^t \int_0^t e^{-\lambda(2t-s_1-s_2)} \delta(s_1 - s_2) ds_2 ds_1,$$
$$= D \left(1 - \frac{1}{N} \right) \int_0^t e^{-2\lambda(t-s)} ds,$$
$$= \frac{D}{2\lambda} \left(1 - \frac{1}{N} \right) \left(1 - e^{-2\lambda t} \right).$$

(iv) Note that $\boldsymbol{v}_i^\alpha - \boldsymbol{v}_j^\alpha$ satisfies

$$\boldsymbol{v}_i^\alpha(t) - \boldsymbol{v}_j^\alpha(t) = e^{-\lambda t} (\boldsymbol{v}_i^\alpha(0) - \boldsymbol{v}_j^\alpha(0)) + \sqrt{D} \int_0^t e^{-\lambda(t-s)} (d\boldsymbol{W}_i^\alpha(s) - d\boldsymbol{W}_j^\alpha(s)). \quad (3.8)$$

Then it is easy to see that when $i \neq j$

$$\left\langle \left| \boldsymbol{v}_i^\alpha(t) - \boldsymbol{v}_j^\alpha(t) \right|^2 \right\rangle = e^{-2\lambda t} \left| \boldsymbol{v}_i^\alpha(0) - \boldsymbol{v}_j^\alpha(0) \right|^2 + 2D \left\langle \left(\int_0^t e^{-\lambda(t-s)} d\boldsymbol{W}_i^\alpha(s) \right)^2 \right\rangle,$$
$$= e^{-2\lambda t} \left| \boldsymbol{v}_i^\alpha(0) - \boldsymbol{v}_j^\alpha(0) \right|^2 + \frac{D}{\lambda} \left(1 - e^{-2\lambda t} \right).$$

$\square$

REMARK 3.2.
(i) Note that as $t \to \infty$, the variance of $\boldsymbol{v}_i^\alpha$ does not go to zero. Instead, the variance of $\boldsymbol{v}_i^\alpha$ converges to $\frac{D}{2\lambda} \left(1 - \frac{1}{N} \right)$, independent of α. This is a contrasted difference from the noiseless case.

(ii) In (3.7), since the first term $e^{-\lambda t}\boldsymbol{v}_{i0}$ is a deterministic process and the second (integral) term follows the normal distribution $N\left(0,\frac{D}{2\lambda}\left(1-e^{-2\lambda t}\right)\right)$, we can conclude that $\boldsymbol{v}_i^\alpha$ also follows the normal distribution:

$$N\left(e^{-\lambda t}\boldsymbol{v}_{i0}^\alpha, \frac{D}{2\lambda}\left(1-e^{-2\lambda t}\right)\right).$$

In the limit as $t\to\infty$,

$$\boldsymbol{v}_i^\alpha \sim N\left(0,\frac{D}{2\lambda}\right).$$

In the next theorem, we show that the linear system (3.5) satisfies the flocking estimate in the sense of Definition 2.1.

THEOREM 3.3. *Let* $(\boldsymbol{x}_i(t),\boldsymbol{v}_i(t))$ *be the solution to* (3.5). *Then for any* $\varepsilon>0$, *we have*

$$(i)\ \lim_{t\to\infty}|\langle\boldsymbol{v}_i(t)\rangle-\langle\boldsymbol{v}_j(t)\rangle|=0, \quad \sup_{0\leq t<\infty}|\langle\boldsymbol{x}_i(t)\rangle-\langle\boldsymbol{x}_j(t)\rangle|<\infty,$$

$$(ii)\ \lim_{t\to\infty}\mathbb{P}\left(|\boldsymbol{v}_i(t)-\boldsymbol{v}_j(t)|^2>\varepsilon\right)\leq\frac{D}{\lambda\varepsilon}.$$

Proof.
(i) The estimate for the first relation directly follows from Proposition 3.1 (iii). Recall that $\boldsymbol{x}_i-\boldsymbol{x}_j$ satisfies

$$\boldsymbol{x}_i(t)-\boldsymbol{x}_j(t)=\boldsymbol{x}_{i0}-\boldsymbol{x}_{j0}+\int_0^t(\boldsymbol{v}_i(s)-\boldsymbol{v}_j(s))ds.$$

We now use the estimate (iii) of Proposition 3.1 to obtain

$$|\langle\boldsymbol{x}_i(t)\rangle-\langle\boldsymbol{x}_j(t)\rangle|\leq|\langle\boldsymbol{x}_{i0}\rangle-\langle\boldsymbol{x}_{j0}\rangle|+\int_0^t|\langle\boldsymbol{v}_i(s)\rangle-\langle\boldsymbol{v}_j(s)\rangle|ds$$
$$\leq|\langle\boldsymbol{x}_{i0}\rangle-\langle\boldsymbol{x}_{j0}\rangle|+\frac{C}{\lambda}(1-e^{-\lambda t})|\langle\boldsymbol{v}_{i0}\rangle-\langle\boldsymbol{v}_{j0}\rangle|.$$

This yields the desired result.

(ii) We again use Markov's inequality together with Proposition 3.1(iii) to obtain

$$\mathbb{P}(|\boldsymbol{v}_i^\alpha(t)-\boldsymbol{v}_j^\alpha(t)|^2\geq\varepsilon)$$
$$\leq\frac{1}{\varepsilon}\left\langle|\boldsymbol{v}_i^\alpha(t)-\boldsymbol{v}_j^\alpha(t)|^2\right\rangle$$
$$\leq\frac{1}{\varepsilon}\left[e^{-2\lambda t}|\boldsymbol{v}_i^\alpha(0)-\boldsymbol{v}_j^\alpha(0)|^2+\frac{D}{\lambda}(1-e^{-2\lambda t})\right]\to\frac{D}{\lambda\varepsilon}, \qquad \text{as } t\to\infty.$$

$\square$

3.2. Radially symmetric communication rate function. In this part, we study the dynamics of the SCS system for which we assume that the communication rate ψ depends on the distance between the particles and is non-increasing in its argument, i.e.,

$$\psi=\bar{\psi}(|\boldsymbol{x}_j-\boldsymbol{x}_i|^2), \tag{3.9}$$

where $\bar{\psi}$ is a non-increasing function. We thus consider the microscopic dynamics:

$$\begin{cases} d\boldsymbol{x}_i = \boldsymbol{v}_i dt, \\ d\boldsymbol{v}_i = \dfrac{\lambda}{N} \sum_{j=1}^{N} \bar{\psi}(|\boldsymbol{x}_j - \boldsymbol{x}_i|^2)(\boldsymbol{v}_j - \boldsymbol{v}_i)dt + \sqrt{D}\left(1 - \dfrac{1}{N}\right)d\boldsymbol{W}_i - \dfrac{\sqrt{D}}{N}\sum_{j \neq i} d\boldsymbol{W}_j. \end{cases} \qquad (3.10)$$

3.2.1. Special case: two particles on a line In the special case when there are only two particles, the system can be solved explicitly. In this case, the system reads as

$$\begin{cases} dx_1 = v_1 dt, \quad dx_2 = v_2 dt, \\ dv_1 = \dfrac{\lambda}{2}\bar{\psi}(|x_2 - x_1|^2)(v_2 - v_1)dt + \dfrac{\sqrt{D}}{2}(dW_1 - dW_2), \\ dv_2 = \dfrac{\lambda}{2}\bar{\psi}(|x_1 - x_2|^2)(v_1 - v_2)dt + \dfrac{\sqrt{D}}{2}(dW_2 - dW_1), \end{cases} \qquad (3.11)$$

subject to the initial data

$$(x_i(0), v_i(0)) = (x_{i0}, v_{i0}), \qquad i = 1, 2.$$

We set

$$x(t) \equiv x_2(t) - x_1(t), \quad v(t) \equiv v_2(t) - v_1(t),$$

which enables us to rewrite (3.11) as

$$\begin{cases} dx = vdt, \\ dv = -\lambda v \bar{\psi}(|x|^2)dt + \sqrt{D}(dW_2 - dW_1), \\ (x(0), v(0)) = (x_0, v_0). \end{cases} \qquad (3.12)$$

The system (3.12) can be reduced to a single equation:

$$dv = -\lambda \bar{\psi}(|x|^2)dx + \sqrt{D}(dW_2 - dW_1), \qquad v|_{x=x_0} = v_0,$$

which we integrate to get the explicit solution:

$$v = v_0 - \lambda \int_{x_0}^{x} \bar{\psi}(|\xi|^2)d\xi + \sqrt{D}(W_2 - W_1). \qquad (3.13)$$

3.2.2. Uniform bound for the variance of fluctuations. In this part, we discuss the uniform boundedness of the variance of fluctuations. Consider the more general case of N particles. Given the phase-space coordinates of N particles, we let

$$\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N), \qquad \boldsymbol{v} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N),$$

and denote the l_2 norm by $\|\cdot\|$. We introduce two auxiliary functionals:

$$\mathcal{X}(t) \equiv \sum_{i=1}^{N} \|\boldsymbol{x}_i\|^2, \qquad \mathcal{V}(t) \equiv \sum_{i=1}^{N} \|\boldsymbol{v}_i\|^2.$$

Then it is easy to see that the ensemble average of the functional $\mathcal{V}$ bounds the variance of a fluctuation $\boldsymbol{v}_i$, i.e.,

$$\text{var}[\boldsymbol{v}_i^{\alpha}] = \langle |\boldsymbol{v}_i^{\alpha}|^2 \rangle - \langle \boldsymbol{v}_i^{\alpha} \rangle^2 \leq \langle \mathcal{V}(t) \rangle. \qquad (3.14)$$

We next derive a system of stochastic differential inequalities from the microscopic dynamics (3.10).

LEMMA 3.4. *Let $(\boldsymbol{x}, \boldsymbol{v})$ be the solution to the system* (3.10). *Then* $(\mathcal{X}(t), \mathcal{V}(t))$ *satisfies*

$$
\begin{cases}
d\mathcal{X} \le 2\sqrt{\mathcal{X}}\sqrt{\mathcal{V}}dt, \\[2mm]
d\mathcal{V} \le -2N\bar{\psi}(2\mathcal{X})\mathcal{V}dt + dD\left(1 - \dfrac{1}{N}\right)dt \\[2mm]
\quad + \sqrt{D}\displaystyle\sum_{i=1}^{N}\boldsymbol{v}_i\cdot\left[\left(1 - \dfrac{1}{N}\right)d\boldsymbol{W}_i - \dfrac{\sqrt{D}}{N}\sum_{j\ne i}d\boldsymbol{W}_j\right].
\end{cases}
\tag{3.15}
$$

Proof.
(i) We use Itô's formula, $dt \cdot dt = 0$ and the Cauchy-Schwartz inequality to obtain

$$
d\mathcal{X} = 2\sum_{i=1}^{N}\boldsymbol{x}_i\cdot d\boldsymbol{x}_i + \sum_{i=1}^{N}d\boldsymbol{x}_i\cdot d\boldsymbol{x}_i \le 2\sqrt{\mathcal{X}}\sqrt{\mathcal{V}}dt.
$$

(ii) Since $|\boldsymbol{x}_j - \boldsymbol{x}_i|^2 \le 2\mathcal{X}$ and $\bar{\psi}$ is non-increasing, we have

$$
\bar{\psi}(|\boldsymbol{x}_j - \boldsymbol{x}_i|) \ge \bar{\psi}(2\mathcal{X}).
$$

We again use Itô's formula to find

$$
\begin{aligned}
d\mathcal{V} &= 2\sum_{i=1}^{N}\boldsymbol{v}_i\cdot d\boldsymbol{v}_i + \sum_{i=1}^{N}d\boldsymbol{v}_i\cdot d\boldsymbol{v}_i \\
&= 2\sum_{i=1}^{N}\boldsymbol{v}_i\cdot\left[\sum_{j}\bar{\psi}(|\boldsymbol{x}_j - \boldsymbol{x}_i|)(\boldsymbol{v}_j - \boldsymbol{v}_i)dt + \sqrt{D}\left(1 - \frac{1}{N}\right)d\boldsymbol{W}_i - \frac{\sqrt{D}}{N}\sum_{j\ne i}d\boldsymbol{W}_j\right] \\
&\quad + dD\left(1 - \frac{1}{N}\right)dt, \\
&= 2\sum_{i=1}^{N}\boldsymbol{v}_i\cdot\left[\sum_{j}\bar{\psi}(|\boldsymbol{x}_j - \boldsymbol{x}_i|)(\boldsymbol{v}_j - \boldsymbol{v}_i)dt\right] \\
&\quad + \sqrt{D}\sum_{i=1}^{N}\boldsymbol{v}_i\cdot\left[\left(1 - \frac{1}{N}\right)d\boldsymbol{W}_i - \frac{1}{N}\sum_{j\ne i}d\boldsymbol{W}_j\right] + dD\left(1 - \frac{1}{N}\right)dt.
\end{aligned}
\tag{3.16}
$$

Here we used Ito's rules $dt\cdot dt = dt\cdot d\boldsymbol{W}_i^{\alpha} = 0$ and $d\boldsymbol{W}_i^{\alpha}d\boldsymbol{W}_j^{\beta} = \delta_{ij}\delta_{\alpha\beta}dt$ to find

$$
\begin{aligned}
&d\boldsymbol{v}_i\cdot d\boldsymbol{v}_i \\
&= \left[\sqrt{D}\left(1 - \frac{1}{N}\right)d\boldsymbol{W}_i - \frac{\sqrt{D}}{N}\sum_{j\ne i}d\boldsymbol{W}_j\right]\cdot\left[\sqrt{D}\left(1 - \frac{1}{N}\right)d\boldsymbol{W}_i - \frac{\sqrt{D}}{N}\sum_{j\ne i}d\boldsymbol{W}_j\right] \\
&= dD\left(1 - \frac{1}{N}\right)dt.
\end{aligned}
$$

The first term on the right hand side of (3.16) can be treated by changing $i \leftrightarrow j$.

$$2\sum_{i=1}^{N} \boldsymbol{v}_i \cdot \left[\sum_j \bar{\psi}(|\boldsymbol{x}_j - \boldsymbol{x}_i|)(\boldsymbol{v}_j - \boldsymbol{v}_i)dt\right]$$

$$= -2\sum_{i=1}^{N} \boldsymbol{v}_j \cdot \left[\sum_j \bar{\psi}(|\boldsymbol{x}_j - \boldsymbol{x}_i|)(\boldsymbol{v}_j - \boldsymbol{v}_i)dt\right]$$

$$= -\sum_{i,j} \bar{\psi}(|\boldsymbol{x}_j - \boldsymbol{x}_i|)||\boldsymbol{v}_j - \boldsymbol{v}_i||^2 dt$$

$$\leq -\bar{\psi}(2\mathcal{X})\sum_{i,j} ||\boldsymbol{v}_j - \boldsymbol{v}_i||^2 dt$$

$$\leq -2N\bar{\psi}(2\mathcal{X})\mathcal{V}dt,$$

where we used (3.3) and the identity

$$\sum_{i,j} ||\boldsymbol{v}_j - \boldsymbol{v}_i||^2 = 2N\sum_i ||\boldsymbol{v}_i||^2 = 2N\mathcal{V}.$$

Combining this with the second term on the right hand side of (3.16) concludes the proof.   □

THEOREM 3.5. *Let $(\boldsymbol{x}_i, \boldsymbol{v}_i)$ be the solution of system* (3.10) *with a bounded initial configuration $(\boldsymbol{x}_{i0}, \boldsymbol{v}_{i0})$. Suppose the communication rate satisfies* (3.9) *and the additional lower bound condition*

$$\min_{s \in \mathbb{R}_+} \bar{\psi}(s) \geq \bar{\psi}_* > 0, \quad \textit{for some positive constant } \bar{\psi}_*.$$

Then the ensemble average of $\mathcal{V}$ is uniformly bounded in t. More precisely,

$$\langle \mathcal{V}(t) \rangle \leq \mathcal{V}(0)e^{-2N\bar{\psi}_* t} + \frac{dD}{2N\bar{\psi}_*}\left(1 - \frac{1}{N}\right)(1 - e^{-2N\bar{\psi}_* t}).$$

Proof. We use the above positive lower bound condition of $\bar{\psi}$, and the equation for $\mathcal{V}$ becomes

$$d\mathcal{V} \leq -2N\bar{\psi}_* \mathcal{V}dt + dD\left(1 - \frac{1}{N}\right)dt$$

$$+ \sqrt{D}\sum_{i=1}^{N} \boldsymbol{v}_i \cdot \left[\left(1 - \frac{1}{N}\right)d\boldsymbol{W}_i - \frac{1}{N}\sum_{j \neq i} d\boldsymbol{W}_j\right].$$

We now apply Ito's formula to the function $e^{2N\bar{\psi}_* t}\mathcal{V}$ to find

$$d(e^{2N\bar{\psi}_* t}\mathcal{V}) = 2N\bar{\psi}_* e^{2N\bar{\psi}_* t}\mathcal{V}dt + e^{2N\bar{\psi}_* t}d\mathcal{V}$$

$$\leq dD e^{2N\bar{\psi}_* t}\left(1 - \frac{1}{N}\right)dt$$

$$+ \sqrt{D}e^{2N\bar{\psi}_* t}\sum_{i=1}^{N} \boldsymbol{v}_i \cdot \left[\left(1 - \frac{1}{N}\right)d\boldsymbol{W}_i - \frac{1}{N}\sum_{j \neq i} d\boldsymbol{W}_j\right].$$

We now integrate the above stochastic differential inequality in t to find

$$\mathcal{V}(t) \leq e^{-2N\bar{\psi}_* t}\mathcal{V}(0) + \frac{dD}{2N\bar{\psi}_*}\left(1 - \frac{1}{N}\right)(1 - e^{-2N\bar{\psi}_* t})$$
$$+ \sqrt{D}\int_0^t e^{-2N\bar{\psi}_*(t-s)}\sum_{i=1}^N \boldsymbol{v}_i(s) \cdot \left[\left(1 - \frac{1}{N}\right)d\boldsymbol{W}_i(s) - \frac{1}{N}\sum_{j\neq i}d\boldsymbol{W}_j(s)\right]ds.$$

We now take an ensemble average to get the desired result.   □

REMARK 3.6.

(i) As a direct application of Theorem 3.5 and relation (3.14), we have a uniform bound for the variance of fluctuations

$$\lim_{t\to\infty}\text{Var}[\boldsymbol{v}_i^\alpha] \leq \frac{dD}{2N\bar{\psi}_*}.$$

(ii) Note that as $D \to 0$, we can see that the variances of fluctuation vanish.

(iii) Since $\sum_{i,j}|\boldsymbol{v}_i - \boldsymbol{v}_j|^2 = 2N\mathcal{V}$, we have

$$\sup_{t\in[0,\infty)}\sum_{i,j}\langle|\boldsymbol{v}_i - \boldsymbol{v}_j|^2\rangle = 2N\langle\mathcal{V}\rangle \leq C < \infty,$$
$$\sup_{t\in[0,\infty)}\langle|\boldsymbol{v}_i - \boldsymbol{v}_j|\rangle \leq \sqrt{2N\langle\mathcal{V}\rangle} \leq C.$$

We note that this is not the flocking estimate in the sense of Definition 2.1, i.e., $\lim_{t\to\infty}|\langle\boldsymbol{v}_i - \boldsymbol{v}_j\rangle| = 0$.

(iv) Note that Theorem 3.5 does not apply to the Cucker-Smale communication rate, however we study this issue numerically in the next section and show that even for the Cucker-Smale's communication rate, the variances of fluctuations are uniformly bounded in time.

(v) For the constant communication rate case, $\psi = 1$, we have an explicit closed formula for v_i itself (see (3.7)). Hence all estimates in Proposition 3.1 do not involve the number of particles N. However, for the general radially symmetric case, we do not have a formula for v_i itself. Hence we instead considered the quantities $(\mathcal{X}, \mathcal{V})$ that are globally defined in the sense that all x_i and v_i are involved. That is why the estimates in Theorem 3.5 depend on N. In fact, these estimates hold for finite N. Of course, if we could consider one single quantity v_i (which may not be possible for the general nonlinear case), there will be no N-dependence.

4. Numerical examples

In this section we present results of numerical simulations of the microscopic system (2.6) for the two types of communication rates that were studied in Sec. 3. The first example is the case of a constant communication rate, $\psi = 1$. We solve the system (2.6) for $N = 100$ particles in $\boldsymbol{x}_i, \boldsymbol{v}_i \in \mathbb{R}^2$. The initial locations and velocities for 99 particles are randomly distributed in the interval $[-50, 50]$. The location and the velocity of the last particle are set as to satisfy the constraint (2.7). The parameters are $\lambda = 10$ and $D = 10$. The system (3.10) is solved up to time $T = 0.3$ using Euler's method. We compute 250 realizations of the solution with the same initial data. The results are shown in figures 4.1 & 4.2.
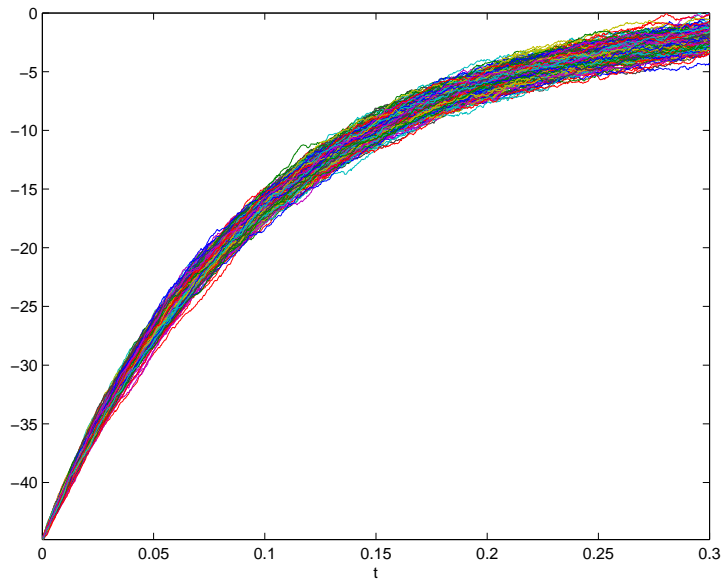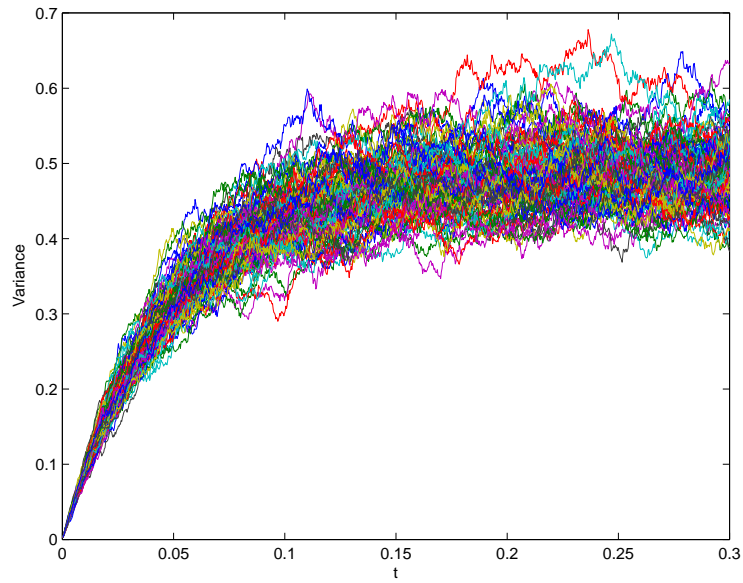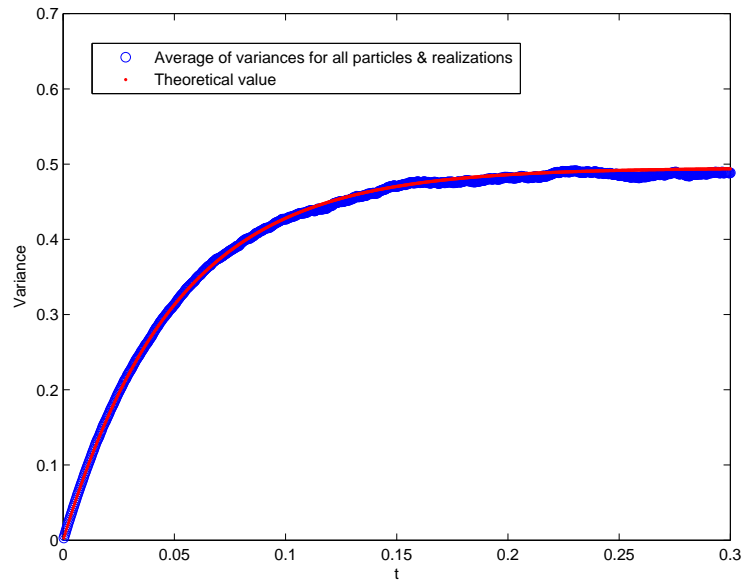
(a)
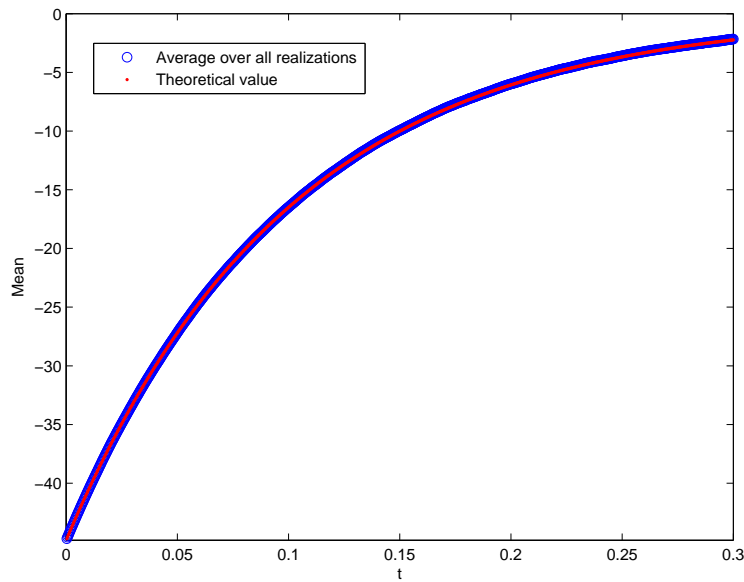
(b)

FIG. 4.1. *Example 1: Constant communication rate* $\psi = 1$. *(a) All the realizations of the trajectories of* $\boldsymbol{v}_1^1$ *(the first component of the velocity v of particle #1). (b) The average of all the realizations shown in (a) and the theoretical value from Proposition 3.1(i).*

Figure 4.1(a) shows the realizations for all the trajectories of $\boldsymbol{v}_1^1$ (the first component of the velocity of particle #1). The average over all these realizations is shown in figure 4.1(b) and is shown to agree with the theoretical value given by Proposi-

(a)

(b)

FIG. 4.2. *Example 1: Constant communication rate* $\psi = 1$. *(a) Each curve shows the variance of one of the 100 particles in the x component of the velocity v over all 250 realizations. (b) The average of all the realizations shown in (a) and the theoretical value from Proposition 3.1(ii).*

tion 3.1(i).

Figure 4.2(a) shows the variances for all 100 particles as computed for each particle for all 250 realizations. The average over all these variances is shown in figure 4.2 and
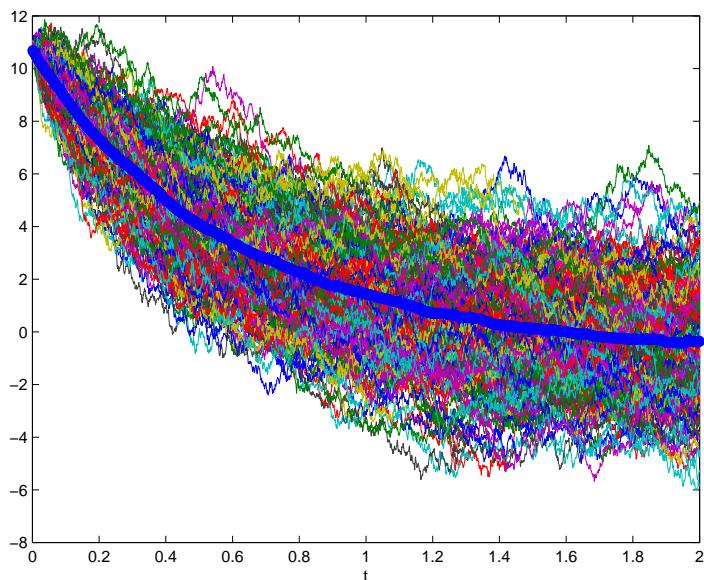
FIG. 4.3. *Example 2: A radially symmetric communication rate function (4.1). All 250 realizations of the trajectories of $\boldsymbol{v}_1^1$ (the first component of the velocity v of particle #1). The average over all these realizations of the trajectories is the solid line.*
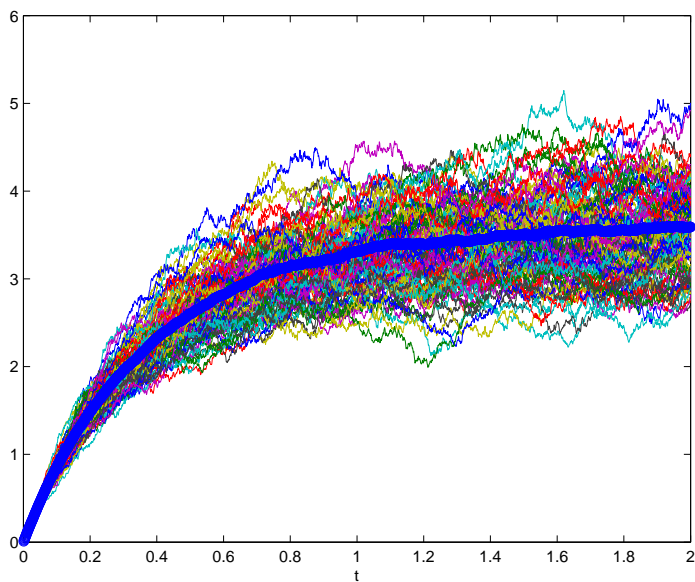

FIG. 4.4. *Example 2: A radially symmetric communication rate function (4.1). Each curve shows the variance of one of the 100 particles in the x component of the velocity v over all 250 realizations. The average of the variances over all these realizations is the solid line.*

is shown to agree with the theoretical value given by Proposition 3.1(ii).

In the second example we set the communication function as

$$\psi(\boldsymbol{x}_j, \boldsymbol{x}_j) = \frac{1}{(1 + |\boldsymbol{x}_j - \boldsymbol{x}_i|^2)^{\frac{1}{4}}}. \tag{4.1}$$

We select initial data for $N = 100$ particles in the same way as in the first example. We obtain 250 realizations of the system that is solved until time $T = 2$ using Euler's method. The other parameters are set as before, $\lambda = 10$ and $D = 10$. Figure 4.3 shows all the realizations of the trajectories of $\boldsymbol{v}_1^1$.

The average over all these realizations is the solid line, and it is shown to converge exponentially fast to 0. In figure 4.4 we see the variances of all 100 particles when computed over all 250 realizations. The average of the variances over all these realizations is shown in the solid line. While we do not have a theoretical formula for this average, it is clear from the figure that it is bounded.

5. Conclusion

In this paper we presented a new stochastic version of the Cucker-Smale model (the SCS system), defined flocking in a stochastic system, and studied flocking in the SCS system.

Flocking in the SCS system was studied in two setups: a constant communication rate and a radially symmetric communication rate. When the communication rate between the particles was assumed to be constant, we showed that the system exhibits a flocking behavior that is independent of the initial configuration. In the radially symmetric communication rate case we added a lower bound assumption and showed that the relative fluctuations of the particle velocity around the mean velocity have a uniformly bounded variance in time.

Our study concluded with several numerical results. The first example was for the case of a constant communication rate case. For this case we demonstrated some of the results of Proposition 3.1. In the second numerical example we studied a non-constant radially symmetric communication rate. In this case, for which we have no analytical results, we showed that the average velocity goes to zero exponentially fast, while the average variance remains bounded.

Clearly, it is of interest to see if one can derive explicit variance bound of fluctuations for more general communication rates at least in some special nonlinear cases. In addition, it would also be of great interest to obtain results that are valid for individual realizations (e.g. statements about an almost sure convergence). The numerical simulations suggest that such results could possibly be obtained. These issues are beyond the scope of this paper and are left for future research.

REFERENCES

[1] I. Aoki, *A simulation study on the schooling mechanism in fish*, Bulletin of the Japan Society of Scientific Fisheries., 48, 1081–1088, 1982.
[2] F. Cucker and S. Smale, *On the mathematics of emergence*, Japan. J. Math., 2, 197–227, 2007.
[3] F. Cucker and S. Smale, *Emergent behavior in flocks*, IEEE Trans. Automat. Control, 52, 852–862, 2007.
[4] P. Degond and S. Motsch, *Macroscopic limit of self-driven particles with orientation interaction*, C.R. Math. Acad. Sci. Paris, 345, 555–560, 2007.
[5] P. Degond and S. Motsch, *Large-scale dynamics of the Persistent Turing Walker model of fish behavior*, J. Stat. Phys., 131, 989–1022, 2008.
[6] P. Degond and S. Motsch, *Continuum limit of self-driven particles with orientation interaction*, Math. Models Methods Appl. Sci., suppl., 18, 1193–1215, 2008.

[7] R. Durret, *Probability: Theory and Examples*, Duxbury Press, 2nd edition 1996.
[8] S.Y. Ha and E. Tadmor, *From particle to kinetic and hydrodynamic description of flocking*, Kinetic and Related Models., 1, 415–435, 2008.
[9] S.Y. Ha and D. Levy, *Particle, kinetic and fluid models for phototaxis*, Disc. Cont. Dyn. Sys. B, to appear.
[10] S.Y. Ha and J.G. Liu, A simple proof of Cucker-Smale flocking dynamics and mean field limit, Commun. Math. Sci., 7(2), 297–325, 2009.
[11] H. Levine and W.J. Rappel, *Self-organization in systems of self-propelled particles*, Phys. Re. E, 63, 017101, 2000.
[12] P. Milewski and X. Yang, *A simple model for biological aggregation with asymmetric sensing*, Commun. Math. Sci., 6, 397–416, 2008.
[13] J.K. Parrish, S.V. Viscido and D. Grünbaum, *Self-organized fish schools: an examination of emergent properties*, The Biological Bulletin, 202, 296–305, 2002.
[14] B.L. Partridge, *The structure and function of fish schools*, Sci. Amer., 246, 114, 1982.
[15] E. Shaw, *Schooling fishes*, American Scientist, 66, 116, 1978.
[16] J. Shen, *Cucker-Smale flocking under hierarchical leadership*, SIAM J. Appl. Math., 68, 694–719, 2007.
[17] J. Toner and Y. Tu, *Flocks, herds, and schools: a quantitative theory of flocking*, Physical Review E., 58, 4828–4858, 1998.
[18] C.M. Topaz and A.L. Bertozzi, *Swarming patterns in a two-dimensional kinematic model for biological groups*, SIAM J. Appl. Math., 65, 152–174, 2004.
[19] T. Vicsek, Czirók, E. Ben-Jacob, I. Cohen and O. Schochet, *Novel type of phase transition in a system of self-driven particles*, Phys. ReV. Lett., 75, 1226–1229, 1995.