CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.SO-2023-0013

Vol. **5**, No. 1, pp. 73-109
March 2024

Flocking Behaviors of a Body Attitude Coordination Model with Velocity Alignment

Zhengyang Qiao, Yicheng Liu* and Xiao Wang

Department of Mathematics, National University of Defense Technology, Changsha 410073, P.R. China.

Received 23 March 2023; Accepted 28 October 2023

Abstract. Body attitude coordination plays an important role in multi-airplane synchronization. In this paper, we study the flocking dynamics of a modified model for body attitude coordination. In contrast to the original body attitude alignment models in Degond *et al.* (Math. Models Methods Appl. Sci., 27(6):1005–1049, 2017) and Ha *et al.* (Discrete Contin. Dyn. Syst., 40(4):2037–2060, 2020), we introduce the velocity alignment term and assume the velocity of each agent is variable. More precisely, the adjoint coefficient will vary with the linked individual changes. In this case, synchronization would include the body attitude alignment and velocity alignment. It will generate a new collective behaviour which is called body attitude flocking. As results, we present two sufficient frameworks leading to the body attitude flocking by technique estimates. Also, we show the finite-in-time stability of the system which is valid on any finite time interval. In addition, we formally derive a kinetic model of the model for body attitude coordination using the BBGKY hierarchy. We prove the well-posedness of the kinetic equation and show a rigorous justification for the mean-field limit of our model. Moreover, we present a sufficient condition for asymptotic flocking in the kinetic model. Finally, we also give the numerical simulations to verify our analysis results.

AMS subject classifications: 34D06, 70F45, 70G60, 82C22

Key words: Body attitude coordination, flocking, stability, measure valued solutions, mean-field limit.

1 Introduction

Collective behavior is ubiquitous in the nature world: schools of fish, flocks of birds, herds of animals, colonies, pedestrian dynamics, etc. Explaining the emergence of these

*Corresponding author. *Email addresses:* `qiaozhengyang1@163.com` (Z. Qiao), `liuyc2001@hotmail.com` (Y. Liu), `wxiao_98@nudt.edu.cn` (X. Wang)

collective behaviors in terms of microscopic decisions of each member is a hot topic of research in the natural sciences. Recently, many mathematical models on the phenomena of flocking have appeared, such as the Vicsek model [29], the Cucker-Smale model [6], the Kromoto model [21], the Lohe model [24], etc, and these models have been extensively studied in literature [3–5,10,11,14,16–20,23,26,31], more literature can be found in [1,25].

These models presented above are built with mass point particles as objects, when collective behavior requires body attitude synchronization, such as synchronization of satellite attitudes [22] and camera pose averaging [28], the mass point models fall short. In this paper we are mainly interested in the model of self-propelled particles with body attitude coordination.

Degond *et al.* [9] proposed an agent-based model for alignment of body attitudes where the states of agents are described by the positions of their center of mass and body attitudes. Specifically, agents move with the same speed and the direction of motions are determined by body attitudes and agents try to adjust their body attitudes with their neighboring agents toward average orientation. For simplicity of modeling, other detailed internal structures are ignored at the level of modeling. The body attitude model is as follows:

$$\begin{cases} dx_i = v_0 A_i \mathbf{e}_1 dt, \\ dA_i = \nu P_{T_{A_i}} \circ \Big[\big(\mathrm{PD}(\mathbb{M}_i)\cdot A_i\big)\mathrm{PD}(\mathbb{M}_i)dt + 2\sqrt{D}dW_t^k \Big], \\ \mathbb{M}_i := \dfrac{1}{N}\displaystyle\sum_{k=1}^{N} K(|x_i - x_k|)A_k, \quad \big(x_i(0),A_i(0)\big) \in \mathbb{R}^3 \times \mathrm{SO}(3), \end{cases} \tag{1.1}$$

where ν is a constant, $K(x)$ is the communication function, $x_i \in \mathbb{R}^3$ is the position of the i-th agent and $A_i \in \mathrm{SO}(3)$ is the body attitude of the i-th agent, $\mathrm{PD}(\mathbb{M}_i)$ denotes the orthogonal matrix which comes from the polar decomposition of $\mathbb{M}_i, W_t^k$ is a noise term and $P_{T_{A_i}}$ is the projection operator on the tangent space $T_{A_i}\mathrm{SO}(3)$,

$$P_{T_A}(B) = \frac{1}{2}(B - AB^T A), \quad A \in \mathrm{SO}(3), \quad B \in M(\mathbb{R},3).$$

The term $A_i\mathbf{e}_1$ describes the direction of movement for the i-th agent where v_0 denotes a constant common speed of the agents and $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ denotes the canonical basis. The authors of [9] formally derived system (1.1) corresponding kinetic and hydrodynamic model, and studied the hydrodynamic limit for body attitude coordination. Based on [9], Ha *et al.* [15] studied the following orientation flocking model (OFM):

$$\begin{cases} \dot{x}_i = v_0 A_i \mathbf{e}_1, \quad t>0, \quad i=1,\dots,N, \\ \dot{A}_i A_i^{-1} = H_i + \dfrac{\kappa}{N}\displaystyle\sum_{k=1}^{N} \psi(|x_i - x_k|)\big(A_k A_i^{-1} - A_i A_k^{-1}\big), \quad \big(x_i(0),A_i(0)\big) = \big(x_i^0, A_i^0\big), \end{cases} \tag{1.2}$$

where $(x_i,A_i) \in \mathbb{R}^d \times \mathrm{SO}(d)$, the skew-symmetric matrix H_i is a generalized frequency-like matrix, κ denotes the coupling strength between the agents. $\psi(r)$ is the communication function and satisfies

$$0 \le \psi_m \le \psi(r) \le \psi_M, \quad r \ge 0, \tag{1.3}$$

where ψ_m and ψ_M are positive constants. In the system (1.2), the dynamics of A_i obey the Lohe matrix model in $\mathrm{SO}(d)$. The authors of [15] presented orientation flocking estimates and showed the system (1.2) is stable with respect to the initial data in any finite time interval. For more studies of body attitude coordination models, refer to [7, 8, 12, 13].

The models (1.1) and (1.2) are regarded as the Vicsek-type flocking model, i.e. all agents travel at a constant speed in given directions. The states of agents are described by the positions and body attitudes. However, in many collective phenomena, each agent has a different speed and is constantly changing. Therefore, we would like to address the following question:

- Can we design a body attitude coordination model with velocity alignment? Such as every agent adjusts its velocity by the velocity of the other agents. If so, under what conditions on system parameters and initial data can the proposed system exhibit asymptotic flocking and body attitude alignment?

We obtained the following body attitude coordination model by considering the addition of velocities in the rotating coordinate system (modeling details see Section 2). The model is as follows:

$$\begin{cases} \dot{x}_i = v_i, \quad t>0, \quad 1 \le i \le N, \\ \dot{v}_i = \dfrac{k_1}{N} \displaystyle\sum_{j=1}^{N} \varphi_{ik} \big((v_k - v_i) + \dot{A}_i A_i^{-1} (x_k - x_i) \big), \\ \dot{A}_i A_i^{-1} = \dfrac{k_2}{N} \displaystyle\sum_{j=1}^{N} \psi_{ik} \big(A_k A_i^{-1} - A_i A_k^{-1} \big), \quad \big(x_i(0), v_i(0), A_i(0) \big) = \big(x_i^0, v_i^0, A_i^0 \big), \end{cases} \tag{1.4}$$

where $k_1, k_2 > 0$ are the coupling strength, $(x_i^0, v_i^0, A_i^0) \in \mathbb{R}^{2d} \times \mathrm{SO}(d)$. The communication weight φ and ψ are bounded, positive, non increasing and Lipschitz continuous on $\mathbb{R}$,

$$\psi_{ik} = \psi(|x_i - x_k|), \quad \varphi_{ik} = \varphi(|x_i - x_k|).$$

We consider the channel capacity in the Cucker-Smale model whose form

$$\varphi(r) = (1 + r^2)^{-\beta}.$$

And $\psi(r)$ follows the form in [15] to satisfy (1.3), i.e.

$$\psi_m \le \psi(r) \le \psi_M.$$

But we consider the case of $\psi(r)$ without a positive lower bound in Theorem 3.2. Comparing models (1.1) and (1.2), the speed of every agent in our model is variable and the direction of velocity is not determined by body attitude.

The model (1.4) is a particle description to collective behaviors. When the number of agents is sufficiently large $N\to\infty$, we are forced to study the mean field limit of the system (1.4). Hence, we present the kinetic model of (1.4). Let $f=f(t,x,v,A)$ be the one-particle distribution function of the system (1.4) infinite ensemble at the phase space point (x,v,A) at time t. By the BBGKY hierarchy, we formally derive the following kinetic equation (see Appendix B):

$$\begin{aligned}&\partial_t f+v\cdot\nabla_x f+\nabla_v\cdot(L[f]f)+\nabla_A\cdot(Q[f]f)=0,\\ &L[f](t,x,v,A)=k_1\int_\Omega \varphi(|x-y|)\big[(v_*-v)+k_2\psi(|x-y|)K(A,A^*)(x-y)\big]\\ &\qquad\qquad\qquad\times f(t,y,v_*,A_*)dydv_*dA_*,\\ &Q[f](t,x,v,A)=k_2\int_\Omega \psi(|x-y|)(A_*-AA_*^{-1}A)f(t,y,v_*,A_*)dydv_*dA_*,\end{aligned}\tag{1.5}$$

where $K(A,A^*)=A_*A^{-1}-AA_*^{-1}$ and $\Omega=\mathbb{R}^{2d}\times\mathrm{SO}(d)$, ∇_A is divergence in $\mathrm{SO}(d)$. For a detailed calculation of ∇_A and integration with respect to A can refer to [9].

In this paper, our main results can be summarized as follows. The first result is that we present two sufficient frameworks leading to the body attitude flocking (see Definition 2.1). In the first sufficient framework, we require the communication function $\psi(r)$ to satisfy (1.3) and obtain a body attitude flocking estimate that is independent of the number of agents N. In the second framework, we remove the requirement for a positive lower bound on $\psi(r)$ but place a restriction on its growth rate, and we present a body attitude flocking estimate associated with N. In addition, we present the finite-in-time stability estimate which is valid on any finite time interval.

The second result of this paper is that we prove the well-posedness of the kinetic equation based on the framework of [3] and we show also the convergence of particle systems (1.4) to their corresponding kinetic equations (1.5), i.e. convergence of the particle method towards a measure solution of the kinetic equation (1.5). Moreover, we present a sufficient condition for asymptotic flocking in the kinetic system (1.5).

The rest of the paper is organized as follows. In Section 2, we discuss physical derivation of our model and present definitions and lemmas needed in the later section. In Section 3, we give two sufficient frameworks leading to the body attitude flocking and show that the system (1.4) is stable with respect to the initial data in any finite time interval. In Section 4, we prove the existence and uniqueness of measure solutions to the kinetic equation (1.5) and present the stability estimate of the measure-valued solutions in W_1-distance. In Section 5, we provide several numerical simulations consistent with the theoretical analysis obtained in Section 3.

Notation. For $x_i=(x_i^{(1)},\cdots,x_i^{(d)})$ and a matrix $A=(a_{ij})\in\mathrm{SO}(d)$

$$\|x_i\|_p=\left(\sum_{j=1}^d |x_i^j|^p\right)^{\frac{1}{p}},\quad \|A\|=\big(\mathrm{tr}(AA^T)\big)^{\frac{1}{2}},$$

and $|x_i|$ denotes ℓ_2 norm $\|x_i\|_2$.

2 Model formulation and preliminaries

In this section, we present physical derivation of our system, the definition of the body attitude flocking and review some background on Wasserstein distances.

2.1 Model formulation

The Cucker-Smale-type flocking model postulates every agent adjusts its velocity by a weighted average of the differences of its velocity with those of the other agents. We can design our model based on this postulate. Consider the following system:

$$\begin{cases} \dot{x}_i(t)=v_i(t), \quad t>0, \quad 1\le i\le N, \\ \dot{v}_i(t)=\dfrac{\kappa}{N}\displaystyle\sum_{k=1}^{N}\varphi(|x_j-x_i|)V_{ji}, \end{cases} \tag{2.1}$$

where V_{ji} denotes the relative velocity of the j-th agent and the i-th agent.

In order to obtain the expression for V_{ji}, we need to consider motions in moving coordinate systems. The inertial system is denoted by L (laboratory system) and the moving system by M_i (moving system). O is the origin of L, o_i is the origin of M_i (o_i can be regarded as the center of mass of the i-th agent). According to [2], a motion of M_i relative to L is a map smoothly depending on t

$$D_i^t : M_i \to L,$$

and every motion D_t^i can be uniquely written as the composition of a rotation $A_i^t : M_i \to L$ and a translation $C_i^t : M_i \to L$, i.e.

$$D_t^i = C_t^i \circ A_t^i.$$

Let $x_i(t)$ be the displacement vector of o_i relative to O, $x_j(t)$ be the displacement vector of o_j relative to O and $Q_{ji}(t)$ be the displacement vector of o_j relative to o_i in the moving system M_i. Then we have

$$x_j(t)=D_tQ_{ji}(t)=A_t^iQ_{ji}(t)+x_i(t).$$

Taking the derivative of the above equation with respect to t we can obtain

$$\dot{x}_j(t)=\dot{A}_t^iQ_{ji}(t)+A_t^i\dot{Q}_{ji}(t)+\dot{x}_i(t).$$

We set $v_i(t)=\dot{x}_j(t)$ is the absolute velocity of $o_i, V_{ji}(t)=A_t^i\dot{Q}_{ji}(t)\in L$ is the relative velocity. Note that $V_{ji}(t)\in L$ and $\dot{Q}_{ji}(t)\in M_i$ are different. $\dot{Q}_{ji}(t)$ is the velocity of o_j relative to o_i in the M_i reference system, while $V_{ji}(t)$ is the velocity of o_j relative to x_i in the L reference system. Since

$$Q_{ji}(t)=(A_t^i)^{-1}\big(x_j(t)-x_i(t)\big).$$

We can obtain the following expression in the inertial system L:

$$\begin{aligned} v_j(t) &= v_i(t)+V_{ji}(t)+\dot{A}_t^i\left(A_t^i\right)^{-1}\left(x_j(t)-x_i(t)\right) \\ &= v_i(t)+V_{ji}(t)+\omega_i(t)\times\left(x_j(t)-x_i(t)\right), \end{aligned}$$

where $\omega_i(t)$ is the angular velocity of the i-th agent, $\omega_i(t)\times(x_j(t)-x_i(t))$ is the transferred velocity of rotation. Let $A_t^i=A_i(t)\in\mathrm{SO}(3)$ denote the rotation of the i-th agent, $A_i(t)(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)=(\mathbf{e}_1^{'},\mathbf{e}_2^{'},\mathbf{e}_3^{'})$ (see Fig. 1). Hence, we can obtain

$$V_{ji}(t)=v_j(t)-v_i(t)+\dot{A}_i(t)A_i^{-1}(t)\left(x_i(t)-x_j(t)\right).$$

The above equation also holds in the d-dimension. Clearly, the system (2.1) degenerates to the Cucker-Smale model when agents are not rotating, i.e. $\dot{A}_i=0$.

Next, we need to give the dynamic system of A_i. Similar to [9,15], we set the dynamics of A_i is the gradient flow of the weighted average $\mathbb{M}_i$

$$\begin{aligned} \dot{A}_i(t) &= 2k_2\nabla_A(\mathbb{M}_i\cdot A)|_{A=A_i}=P_{T_{A_i}}(\mathbb{M}_i) \\ &= \frac{k_2}{N}\sum_{j=1}^N\psi(|x_j-x_i|)\left(A_j-A_iA_j^{-1}A_i\right). \end{aligned} \tag{2.2}$$

The Eq. (2.2) means that agents try to coordinate their body attitude with other agents. In summary, we obtain the following body attitude coordination model:

$$\begin{cases} \dot{x}_i=v_i, \quad t>0, \quad 1\le i\le N, \\ \dot{v}_i=\dfrac{k_1}{N}\displaystyle\sum_{j=1}^N\varphi_{ik}\left((v_k-v_i)+\dot{A}_iA_i^{-1}(x_k-x_i)\right), \\ \dot{A}_iA_i^{-1}=\dfrac{k_2}{N}\displaystyle\sum_{j=1}^N\psi_{ik}\left(A_kA_i^{-1}-A_iA_k^{-1}\right), \quad (x_i(0),v_i(0),A_i(0))=\left(x_i^0,v_i^0,A_i^0\right). \end{cases}$$

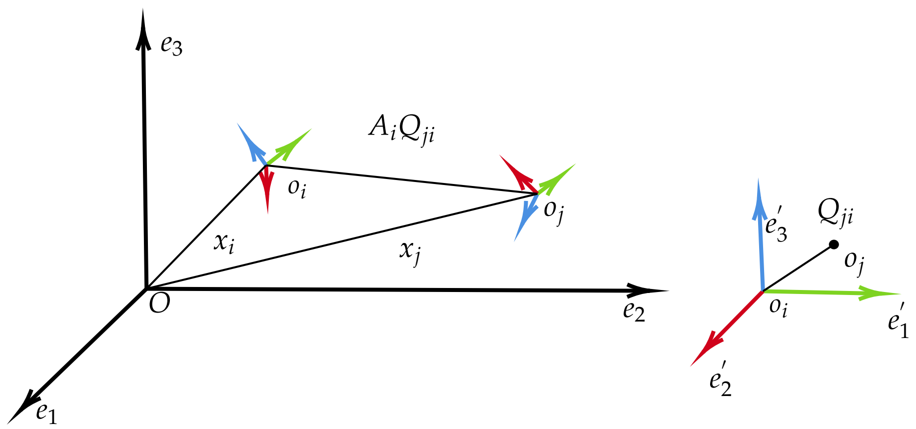

Figure 1: Addition of velocities.

2.2 Definitions and lemmas

We introduce the definition of the body attitude flocking and the finite-in-time stability in the following. Before we proceed further, we give the following notation:

$$X=(x_1,\cdots,x_N),\quad V=(v_1,\cdots,v_N),\quad \mathcal{A}=(A_1,\cdots,A_N),$$

and

$$\begin{aligned}
D\big(X(t)\big)&=\max_{1\le i,j\le N}|x_i-x_j|,\\
D\big(V(t)\big)&=\max_{1\le i,j\le N}|v_i-v_j|,\\
D\big(\mathcal{A}(t)\big)&=\max_{1\le i,j\le N}\|A_i-A_j\|.
\end{aligned}$$

We set $\mathcal{Z}=(X,V,\mathcal{A})$ and define the $\ell_{2,\infty}$-norm as follow:

$$\|\mathcal{Z}\|_{2,\infty}=\max_{1\le i\le N}|x_i|+\max_{1\le i\le N}|v_i|+\max_{1\le i\le N}\|A_i\|. \tag{2.3}$$

Definition 2.1. *Let (X,V,A) be a global solution to* (1.4). *We say that system* (1.4) *exhibits body attitude flocking if and only if the solution $\{x_i,v_i,A_i\}_{i=1}^N$ to* (1.4) *satisfies the following three conditions:*

1. *(Group formation). The position fluctuation is uniformly bounded in time t*

$$\sup_{0\le t<\infty}\max_{1\le i,j\le N}|x_i-x_j|<\infty.$$

2. *(Velocity alignment). The velocity fluctuation goes to zero time-asymptotically*

$$\lim_{t\to\infty}\max_{1\le i,j\le N}|v_i-v_j|=0.$$

3. *(Body attitude alignment). The body attitude fluctuation converges to zero as time goes to infinity*

$$\lim_{t\to\infty}\max_{1\le i,j\le N}\|A_i-A_j\|=0.$$

Definition 2.2 ([15])**.** *For any two global solutions $\mathcal{Z}$ and $\tilde{\mathcal{Z}}$, there exists a positive constant $G=G(T),T\in(0,\infty)$, which is depending on the initial data and independent of time t and the number of agents N. If the following estimate holds, then we say the system* (1.4) *is finite-in-time stable with respect to the initial data:*

$$\sup_{t\in[0,T)}\|\mathcal{Z}-\tilde{\mathcal{Z}}\|_{2,\infty}\le G(T)\|\mathcal{Z}^0-\tilde{\mathcal{Z}}^0\|_{2,\infty}. \tag{2.4}$$

In what follows, we present a selection of lemmas devoid of accompanying proofs.

Lemma 2.1. *Let A,B be $d\times d$ matrices and x be a d-dimensional vector. Then the following relations hold:*

$$\|AB\|\le\|A\|\|B\|,\quad |A\mathbf{e}|\le\|A\|,\quad |Ax|\le\sqrt{d}\|A\||x|,$$

where $\mathbf{e}$ is an any unit vector in $\mathbb{R}^d$. If $A,B\in\mathrm{SO}(d)$, the following equation holds:

$$\|A-B\|=2d-2\mathrm{tr}(A^TB).$$

Remark 2.1. In Lemma 2.1, the third inequality can be easily obtained. In fact, for any $x=(x_1,\cdots,x_d)\in\mathbb{R}^d$,

$$|Ax|=|x_1A\mathbf{e}_1+\cdots+x_dA\mathbf{e}_d|\le\|A\|\sum_{i=1}^d|x_i|\le\sqrt{d}\|A\||x|.$$

Lemma 2.2 ([15]). *Let $(X,\mathcal{A})$ be a global solution to* (1.2). *Then the following estimate holds for all $t>0$:*

$$\frac{d}{dt}D(\mathcal{A})\le -k_2(3\psi_m-\psi_M)D(\mathcal{A})+2k_2\psi_MD(\mathcal{A})^2.$$

Proof. The details of the proof are similar to [15, Lemma 2.4]. □

Lemma 2.3 ([15]). *Let $y=y(t)$ denote a nonnegative function in the class of C^1, which satisfies the following differential inequality:*

$$\begin{cases}\dot{y}\le -py+qy^2, & t>0,\\ y(0)=y_0,\end{cases}$$

where p,q are positive constants. Then we have

$$y(t)\le\left(\left(\frac{1}{y(0)}-\frac{q}{p}\right)e^{pt}+\frac{q}{p}\right)^{-1},\quad t>0.$$

2.3 Wasserstein distances

In this subsection, let us recall some notations and known results about mass transportation that we will use in the Section 4. For a more detailed approach, the reader can refer to [30]. We set $\Omega=\mathbb{R}^{2d}\times\mathrm{SO}(d)$. Let Ω equip with the norm $\|\cdot\|_\Omega$,

$$\|z\|_\Omega=|x|+|v|+\|A\|,\quad z=(x,v,A)\in\Omega.$$

The Frobenius norm of $\mathrm{SO}(d)\subset\mathbb{R}^{d^2}$ corresponds to the ℓ_2-norm of $\mathbb{R}^{d^2}$. $A\in\mathrm{SO(d)}$ is considered as a vector in $\mathbb{R}^{d^2}$. Thus, $(\Omega,\|\cdot\|_\Omega)$ is a Polish space. We denote by $\mathcal{P}_1(\Omega)$ the set of Borel probability measures with finite first moment. Let $f,g\in\mathcal{P}_1(\Omega)$, then the

Wasserstein metric W_1 (or Monge-Kantorovich-Rubinstein distance) between f and g is given by

$$W_1(f,g)=\sup\left\{\left\|\int_\Omega \varphi(z)\big(f(z)-g(z)\big)dz\right\|_\Omega, \varphi\in \mathrm{Lip}(\Omega), \mathrm{Lip}(\varphi)\le 1\right\},$$

where $\mathrm{Lip}(\Omega)$ denotes the set of Lipschitz functions on Ω and $\mathrm{Lip}(\varphi)$ is the Lipschitz constant of the function φ. Let $\Pi(f,g)$ is the set of transport plans between f and g, i.e. the set of elements in $\mathcal{P}_1(\Omega\times\Omega)$ with first and second marginals f and g, respectively. By Kantorovich duality we have

$$W_1(f,g)=\inf_{\pi\in\Pi(f,g)}\left\{\int_{\Omega\times\Omega}|z_1-z_2|d\pi(z_1,z_2)\right\}.$$

$(\mathcal{P}_1(\Omega),W_1)$ is a Polish space. In the following proposition, we recall some of its properties.

Proposition 2.1. *The following properties of the distance W_1 hold:*

(i) The infimum in the definition of the distance W_1 can be achieved. If the probability measure π_* satisfying

$$W_1(f,g)=\int_{\Omega\times\Omega}|z_1-z_2|d\pi_*(z_1,z_2),$$

then π_* is called an optimal transference plan.

(ii) Given $\{f_k\}_{k\ge1}$ and f in $\mathcal{P}_1(\Omega)$, the following assertions are equivalent:

(a) $W_1(f_k,f)$ tends to 0 as k goes to infinity.

(b) f_k tends to f weakly- as measures and*

$$\int_\Omega \|z\|_\Omega f_k(z)dz \;\to\; \int_\Omega \|z\|_\Omega f(z)dz \quad as \quad k\;\to\;+\infty.$$

For $T>0$ and a function $f:[0,T]\times\mathbb{R}^n\to\mathbb{R}^m$, we set $\mathrm{Lip}_R(f)$ to denote the Lipschitz constant of f in the closed ball $B_R\subset\mathbb{R}^n$ with center 0 and radius $R>0$. That is

$$|f(t,x_1)-f(t,x_2)|\le \mathrm{Lip}_R(f)|x_1-x_2|$$

for all $x_1,x_2\in B_R, t\in[0,T)$.

3 Emergence of body attitude flocking

In this section, we present two body attitude flocking estimates for (1.4) using different methods. In the first estimate, it is required that $\psi(r)$ satisfies (1.3), and we use Lemmas 2.2 and 2.3 to obtain an estimate that is independent of the number of agents N. In the second estimate, we remove the requirement for a positive lower bound on $\psi(r)$ and we obtain a body attitude flocking estimate associated with N. In addition, we show that the system (1.4) is finite-in-time stable.

3.1 The body attitude flocking estimates

3.1.1 Flocking estimates independent of the number of agents N

Compared to the model (1.2), the body attitude evolution equation in our model (1.4) is a special case of model (1.2), i.e. $H_i \equiv 0$. Therefore, we directly quote the conclusion about the diameter functional $D(\mathcal{A})$ in Lemma 2.2. By Lemmas 2.2 and 2.3, if ψ satisfies (1.3), we can obtain

$$D(\mathcal{A}) \leq \left(\left(\frac{1}{D(\mathcal{A}_0)} - \frac{2\psi_M}{3\psi_m - \psi_M} \right) e^{k_2(3\psi_m - \psi_M)t} + \frac{2\psi_M}{3\psi_m - \psi_M} \right)^{-1} \leq c_0 \mathrm{e}^{-at}, \tag{3.1}$$

where

$$c_0 = \left(\frac{1}{D(\mathcal{A}_0)} - \frac{2\psi_M}{3\psi_m - \psi_M} \right)^{-1}, \quad a = k_2(3\psi_m - \psi_M).$$

Next, we show our first result on the emergence of body attitude flocking of system (1.4).

Theorem 3.1. *Let $(X,V,\mathcal{A})$ be a global solution to* (1.4). *If the system parameters and initial data satisfy $0 \leq \beta < 1/2$,*

$$D\big(\mathcal{A}(0)\big) < \frac{a}{2k_2\psi_M}, \quad 0 < a < \frac{2k_1}{\big(1 + D\big(X(0)\big)^2\big)^{\beta}},$$

and ψ satisfies (1.3), *i.e. $0 < \psi_m \leq \psi(r) \leq \psi_M$. Then, system* (1.4) *admits a body attitude flocking in the sense of Definition* 2.1. *Furthermore, there exist positive constants C, M, c such that*

$$\begin{aligned} &D\big(X(t)\big) \leq C D\big(X(0)\big) + D\big(V(0)\big), \\ &D\big(V(t)\big) \leq M \mathrm{e}^{-\frac{c}{2}t} D\big(V(0)\big), \\ &D\big(\mathcal{A}(t)\big) \leq c_0 \mathrm{e}^{-at}. \end{aligned}$$

Proof. Note that there exist at most countable number of increasing times t_k such that we can choose indices i and j such that $D(V(t)) = |v_i(t) - v_j(t)|$ on any time interval (t_k, t_{k+1}) since the number of agents is finite and continuity of the velocity trajectories. This allows us to estimate the time evolution of $D(V(t))$ as

$$\begin{aligned} \frac{1}{2}\frac{d}{dt} D\big(V(t)\big)^2 &= \frac{1}{2}\frac{d|v_i - v_j|^2}{dt} = \langle \dot{v}_i - \dot{v}_j, v_i - v_j \rangle \\ &= \left\langle \frac{k_1}{N}\sum_{k=1}^N \varphi_{ik}(v_k - v_i) - \frac{k_1}{N}\sum_{k=1}^N \varphi_{jk}(v_k - v_j), v_i - v_j \right\rangle \\ &\quad + \left\langle \frac{k_1}{N}\sum_{k=1}^N \varphi_{ik}\dot{A}_i A_i^{-1}(x_i - x_k) - \frac{k_1}{N}\sum_{k=1}^N \varphi_{jk}\dot{A}_j A_j^{-1}(x_j - x_k), v_i - v_j \right\rangle. \end{aligned}$$

Because

$$\langle v_i(t) - v_j(t), v_k(t) - v_i(t) \rangle \leq 0, \quad \langle v_i(t) - v_j(t), v_k(t) - v_j(t) \rangle \geq 0,$$

we can obtain

$$\left\langle \frac{k_1}{N}\sum_{k=1}^{N}\varphi_{ik}(v_k-v_i)-\frac{k_1}{N}\sum_{k=1}^{N}\varphi_{ik}(v_k-v_j),v_i-v_j\right\rangle \le -k_1\varphi_m(t)|v_i(t)-v_j(t)|^2,$$

where $\varphi_m(t)=\varphi(D(X(t)))$. By Lemma 2.1,

$$\begin{aligned}\frac{d}{dt}D\big(V(t)\big)^2 &\le -2k_1\varphi_m(t)D\big(V(t)\big)^2\\ &\quad +2\left|\frac{k_1}{N}\sum_{k=1}^{N}\varphi_{ik}\dot{A}_iA_i^{-1}(x_i-x_k)-\frac{k_1}{N}\sum_{k=1}^{N}\varphi_{jk}\dot{A}_jA_j^{-1}(x_j-x_k)\right|D\big(V(t)\big)\\ &\le -2k_1\varphi_m(t)D\big(V(t)\big)^2+4k_1\sqrt{d}\big\|\dot{A}_iA_j^{-1}\big\|D\big(X(t)\big)D\big(V(t)\big).\end{aligned}$$

According to (3.1), we know

$$D(\mathcal{A})=\big\|A_iA_j^{-1}-I\big\|\le c_0\mathrm{e}^{-at}.$$

Thus,

$$\big\|\dot{A}_iA_j^{-1}\big\|\le\frac{k_2}{N}\sum_{k=1}^{N}\psi_M\big\|A_kA_j^{-1}-A_jA_k^{-1}\big\|\le 2k_2\psi_MD(\mathcal{A})\le 2k_2\psi_Mc_0\mathrm{e}^{-at}.$$

We can get the following estimation:

$$\begin{aligned}\frac{d}{dt}D\big(V(t)\big)&\le -2k_1\varphi_m(t)D\big(V(t)\big)+4k_1k_2\sqrt{d}\psi_MD\big(\mathcal{A}(t)\big)D\big(X(t)\big)\\ &\le -2k_1\varphi_m(t)D\big(V(t)\big)+8k_1k_2\sqrt{d}\psi_Mc_0\mathrm{e}^{-at}D\big(X(t)\big)\\ &\le -2k_1\varphi_m(t)D\big(V(t)\big)+c_1\mathrm{e}^{-at}D\big(X(t)\big),\end{aligned}\tag{3.2}$$

where $c_1=8k_1k_2\sqrt{d}\psi_Mc_0$. By (1.4) and the definitions of $D(X(t)),D(V(t))$ and $D(X(t))$ has a bound

$$D\big(X(t)\big)\le D\big(X(0)\big)+\int_0^t D\big(X(s)\big)ds.$$

We substitute this into (3.2) to get

$$\frac{d}{dt}D\big(V(t)\big)\le c_1\mathrm{e}^{-at}D\big(X(t)\big)\le c_1\mathrm{e}^{-at}\left(D\big(X(0)\big)+\int_0^t D\big(V(s)\big)ds\right).$$

Then we can obtain the following estimation for $D(V)$:

$$D\big(V(t)\big)\le D\big(V(0)\big)+c_1\int_0^t\mathrm{e}^{-at}\left(\big(D(X(0)\big)+\int_0^s D\big(V(u)\big)du\right).$$

For all $\tau \le t$, we consider the maximum of $D(V(\tau))$ in the interval $[0,t]$

$$
\begin{aligned}
\max_{\tau\in[0,t]} D\big(V(\tau)\big) &\le D\big(V(0)\big)+c_1\int_0^t e^{-at}\left(D\big(X(0)\big)+\int_0^s D\big(V(u)\big)du\right)ds \\
&\le D\big(V(0)\big)+c_1\int_0^t e^{-at}\left(D\big(X(0)\big)+\max_{\tau\in[0,s]} D\big(V(\tau)\big)s\right)ds \\
&\le D\big(V(0)\big)+\frac{c_1}{a}D\big(X(0)\big)+c_1\int_0^t e^{-at}\max_{\tau\in[0,s]} D\big(V(\tau)\big)sds \\
&\le \left(D\big(V(0)\big)+\frac{c_1}{a}D\big(X(0)\big)\right)e^{\frac{c_1}{a^2}}.
\end{aligned}
$$

The last inequality is obtained by Gronwall's lemma. Because the above equation is independent of t. For all $t>0, D(V(t))$ has a bound

$$
D\big(V(t)\big) \le \left(D\big(V(0)\big)+\frac{c_1}{a}D\big(X(0)\big)\right)e^{\frac{c_1}{a^2}} = c_2.
$$

Thus,

$$
D\big(X(t)\big) \le D\big(X(0)\big)+c_2t \le c_3(1+t),
$$

where $c_3=\max\{c_2,D(X(0))\}$. We can get a estimation of $\varphi_m(t)$

$$
\varphi_m(t)=\varphi\big(D\big(X(t)\big)\big) \le \varphi\big(c_3(1+t)\big)=\frac{1}{\big(1+c_3^2(1+t)^2\big)^{\beta}}.
$$

We return to the estimation for $D(V)$. We set $2k_1/(1+c_3^2(1+t)^2)^{\beta}=g(t)$, then

$$
\frac{d}{dt}D\big(V(t)\big) \le -g(t)D\big(V(t)\big)+c_1e^{-at}D\big(X(t)\big).
$$

The differential inequality can be integrated to get

$$
\begin{aligned}
D\big(V(t)\big) \le\, & D\big(V(0)\big)\exp\left(-\int_0^t g(s)ds\right) \\
&+c_1\exp\left(-\int_0^t g(s)ds\right)\int_0^t e^{-as}\exp\left(\int_0^s g(u)du\right)c_3(1+s)ds \\
\le\, & D\big(V(0)\big)\exp\left(-\int_0^t g(s)ds\right) \\
&+c_1\exp\left(-\int_0^t g(s)ds\right)\int_0^t e^{-as}\exp\left(\frac{k_1}{(1+c_3^2)^{\beta}}s\right)c_3(1+s)ds \\
\le\, & D\big(V(0)\big)\exp\left(-\int_0^t g(s)ds\right)+c_1\exp\left(-\int_0^t g(s)ds\right)\int_0^t e^{-(a-c_4)s}c_3(1+s)ds \\
\le\, & \exp\left(-\int_0^t g(s)ds\right)\left(D\big(V(0)\big)+\frac{c_1c_3}{a-c_4}+\frac{c_1c_3}{(a-c_4)^2}\right),
\end{aligned}
\tag{3.3}
$$

where

$$c_4=\frac{2k_1}{(1+c_3^2)^\beta},\quad a>\frac{2k_1}{\left(1+D\big(X(0)\big)^2\right)^\beta}\geq c_4.$$

When $0\leq\beta<1/2$, it is trivial to check that $\lim_{t\to\infty}(1+t)g(t)=\infty$. Thus, for any $K>0$, there exists $T>0$ such that if $t>T$, then $(1+t)g(t)\geq K$, and

$$\int_0^t g(s)ds\geq\int_0^T g(s)ds+K\ln(1+t)-K\ln(1+T),$$

then

$$\exp\left(-\int_0^t g(s)ds\right)\leq\exp\left(-\int_0^T g(s)ds+K\ln(1+T)\right)(1+t)^{-K}=F(T)(1+t)^{-K},$$

where

$$F(T)=\exp\left(-\int_0^T g(s)ds+K\ln(1+T)\right).$$

We can derive the estimate

$$\begin{aligned}D\big(X(t)\big)&\leq D\big(X(0)\big)+\int_0^t D\big(V(s)\big)ds\\&\leq D\big(X(0)\big)+F(T)\left(D\big(V(0)\big)+\frac{c_1c_3}{a-c_4}+\frac{c_1c_3}{(a-c_4)^2}\right)\int_0^t(1+s)^{-K}ds.\end{aligned}\tag{3.4}$$

Because the constant K can be chosen arbitrarily large by choosing T big enough, we can know the integral is bounded, and that there exists x_∞ such that $D(X(t))\leq x_\infty$. Hence, (3.2) can be writed

$$\frac{d}{dt}D\big(V(t)\big)\leq-2k_1\varphi(x_\infty)D\big(V(t)\big)+c_1\mathrm{e}^{-at}D\big(X(t)\big).$$

Because $a>c_4$, we can obtain that

$$\begin{aligned}D\big(V(t)\big)&\leq\mathrm{e}^{-2k_1\varphi(x_\infty)t}D\big(V(0)\big)+c_1\mathrm{e}^{-2k_1\varphi(x_\infty)t}\int_0^t\mathrm{e}^{(2k_1\varphi(x_\infty)-a)s}c_3s ds\\&\leq\mathrm{e}^{-c_5t}D\big(V(0)\big)+c_1\mathrm{e}^{-c_5t}D\big(V(0)\big)c_3s^2,\end{aligned}$$

where $c_5=2k_1\varphi(x_\infty)$. We take the constants $M_1=8c_1c_3/c_5^2$. Then we have

$$\begin{aligned}&D\big(V(t)\big)\leq(1+M_1)e^{-\frac{c_5}{2}t}D\big(V(0)\big)=Me^{-\frac{c_5}{2}t}D\big(V(0)\big),\\&D\big(X(t)\big)\leq\frac{2M}{c_5}D\big(V(0)\big)+D\big(X(0)\big),\end{aligned}$$

where $M=1+M_1$. $\square$

3.1.2 Flocking estimates without positive lower bound of ψ

In Theorem 3.1, we require $0<\psi_m<\psi(r)<\psi_M$. This assumption is reasonable, and we obtain an estimate of the flocking convergence rate independent of N. This is crucial in the later analysis of finite-time stability. Next, we give another body attitude flocking estimate for (1.4), which removes the requirement for a positive lower bound on $\psi(r)$.

Lemma 3.1. *Let* $(X,V,\mathcal{A})$ *be a global solution to* (1.4)*. If* $D(\mathcal{A}(0))<\sqrt{2}$*, then the following estimate holds for all* $t>0$*:*

$$D\big(\mathcal{A}(t)\big)\leq \exp\left(-\eta\int_0^t \psi_m(s)ds\right)D\big(\mathcal{A}(0)\big),$$

where

$$\eta=\frac{4}{N}\left(1-\frac{1}{2}D\big(\mathcal{A}(0)\big)^2\right).$$

Proof. See Appendix A. □

We make the following assumptions on the communication function $\psi(r)$:

$(\mathbf{H})$: $\varphi(r)=(1+r^2)^{\beta}$, $\beta<1/2$ and $\psi(r)$ is not increased, positive, Lipschitz continuous on $\mathbb{R}$ and satisfies

$$e^{\Phi_1(t)-\Psi_1(t)}(\xi_3+t)^{\xi_4}\leq 1,\quad e^{\Phi_2(t)-\Psi_2(t)}(\xi_3+t)^{\xi_4}\leq 1,\quad \xi_3>0,\quad \xi_4>2, \tag{3.5}$$

where $\Phi_i(s)$ and $\Psi_i(s), i=1,2$, denote that

$$\begin{aligned}
&\Phi_1(s)=2k_1\int_0^t \varphi\big(\xi_1(s+1)e^{\xi_2 s^2}\big)ds, \qquad \Psi_1(s)=\eta\int_0^t \psi\big(\xi_1(s+1)e^{\xi_2 s^2}\big)ds,\\
&\Phi_2(s)=2k_1\int_0^t \varphi\big(D\big(X(0)\big)+C_4 s\big)ds, \quad \Psi_2(s)=\eta\int_0^t \psi\big(D\big(X(0)\big)+C_4 s\big)ds,\\
&C_2=\int_0^\infty \frac{1}{(\xi_3+s)^{\xi_4}}, \quad C_3=4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(0)\big)D\big(x(0)\big)\int_0^t \frac{s}{(\xi_3+s)^{\xi_4}}ds,\\
&C_4=\Big(D\big(\mathcal{V}(0)\big)+4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(0)\big)D\big(X(0)\big)C_2\Big)e^{C_3}.
\end{aligned}$$

Here ξ_1,ξ_2 satisfy

$$\max\big\{D\big(X(0)\big),D\big(V(0)\big)\big\}\leq \xi_1,\quad 4k_1k_2d\psi_M\leq \xi_2.$$

Remark 3.1. The assumption $(\mathbf{H})$ can be satisfied, we can choose

$$\psi(t)=\frac{2k_1}{\eta}\varphi(t)+\phi(t).$$

Then (3.5) holds is equivalent to the following inequality holds:

$$\int_0^t \eta\phi\big(\theta(s)\big)ds\geq \xi_4\ln(\xi_3+t), \tag{3.6}$$

where $\theta(t)=\xi_1(t+1)\mathrm{e}^{\xi_2 t^2}$. We can choose

$$\phi(t)=\frac{\xi_5}{\xi_3+\ln\big(\ln(t+\mathrm{e})\big)}.$$

We can compare the derivatives of both sides of the above inequality

$$\frac{d}{dt}\int_0^t \eta\phi\big(\xi_1(s+1)\mathrm{e}^{\xi_2 s^2}\big)ds=\frac{\eta\xi_5}{\xi_3+\ln\big(\ln\big(\xi_1(t+1)\mathrm{e}^{\xi_2 t^2}+\mathrm{e}\big)\big)},$$

and

$$\lim_{t\to\infty}\frac{\eta\xi_5\big(\xi 3+\ln\big(\ln\big(\xi_1(t+1)\mathrm{e}^{\xi_2 t^2}+\mathrm{e}\big)\big)\big)^{-1}}{\xi_4(\xi_3+t)^{-1}}=+\infty.$$

Therefore, the left side of inequality (3.6) grows faster than the right side. Therefore, choosing the appropriate ξ_5 can make (3.6) hold. Note that when $\theta(t)=D(X(0))+C_4 t$, as long as ξ_5 is chosen large enough (3.6) still holds.

Theorem 3.2. *Let* $(X,V,\mathcal{A})$ *be a global solution to* (1.4)*. If the system parameters and initial data satisfy* $D(\mathcal{A}(0))<\sqrt{2}$ *and the assumption* $(\mathbf{H})$ *holds. Then, system* (1.4) *exhibits a body attitude flocking in the sense of Definition* 2.1*.*

Proof. Firstly, we roughly estimate the upper bound of $D(X(t))$

$$\frac{d}{dt}D\big(X(t)\big)\le D\big(X(0)\big)+\int_0^t D\big(V(s)\big)ds.$$

Because

$$D\big(V(t)\big)\le D\big(V(0)\big)+\int_0^t 4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(s)\big)D\big(X(s)\big)ds$$

and $D(\mathcal{A}(t))\le 2\sqrt{d}$,

$$\frac{d}{dt}D\big(X(t)\big)\le D\big(X(0)\big)+D\big(V(0)\big)t+\int_0^t\int_0^s 8k_1k_2 d\psi_M D\big(X(u)\big)duds.$$

Similarly to the estimate in the proof of Theorem 3.1, we have

$$D\big(X(t)\big)\le\big(D\big(X(0)\big)+D\big(V(0)\big)t\big)\mathrm{e}^{4k_1k_2d\psi_M t^2}\le\xi_1(t+1)\mathrm{e}^{\xi_2 t^2}.$$

By Lemma 3.1, we can obtain

$$\begin{aligned}\frac{d}{dt}D\big(V(t)\big)&\le-2k_1\varphi_m(t)D\big(V(t)\big)+4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(t)\big)D\big(X(t)\big)\\&\le-2k_1\varphi_m(t)D\big(V(t)\big)+4k_1k_2\sqrt{d}\psi_M\exp\left(-\eta\int_0^t\psi_m(s)ds\right)D\big(\mathcal{A}(0)\big)D\big(X(t)\big).\end{aligned}$$

According to Gronwall's lemma and the assumption $(\mathbf{H})$, we can get

$$D\big(V(t)\big)\leq \mathrm{e}^{-\Phi_1(s)}D\big(\mathcal{V}(0)\big)+4k_1k_2\sqrt{d}\psi_M\mathrm{e}^{-\Phi_1(s)}\int_0^t \mathrm{e}^{\Phi_1(s)-\Psi_1(s)}D\big(\mathcal{A}(0)\big)D\big(X(t)\big)ds.$$

Then we have

$$\begin{aligned}D\big(V(t)\big)&\leq D\big(\mathcal{V}(0)\big)+4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(0)\big)\int_0^t \mathrm{e}^{\Phi_1(s)-\Psi_1(s)}\Big(D\big(X(0)\big)+\max_{\tau\in[0,s]}D\big(V(\tau)\big)s\Big)ds\\&\leq D\big(\mathcal{V}(0)\big)+4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(0)\big)\int_0^t \bigg(\frac{1}{(\xi_3+s)^{\xi_4}}D\big(X(0)\big)+\max_{\tau\in[0,s]}D\big(V(\tau)\big)s\bigg)ds\\&\leq D\big(\mathcal{V}(0)\big)+4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(0)\big)D\big(x(0)\big)C_2\\&\quad+4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(0)\big)D\big(x(0)\big)\int_0^t\frac{s}{(\xi_3+s)^{\xi_4}}\max_{\tau\in[0,s]}D\big(V(\tau)\big)ds\\&\leq\Big(D\big(\mathcal{V}(0)\big)+4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(0)\big)D\big(X(0)\big)C_2\Big)\mathrm{e}^{C_3}=C_4.\end{aligned}\tag{3.7}$$

Thus, we can get $D(X(t))\leq D(X(0))+C_4t$. We set

$$g_1(t)=2k_1\Big(1+\big(C_4t+D\big(X(0)\big)\big)^2\Big)^{-\beta}.$$

Obviously,

$$2k_1\varphi_m(t)\leq 2k_1\Big(1+\big(C_4t+D\big(X(0)\big)\big)^2\Big)^{-\beta}=g_1(t).$$

Hence,

$$\begin{aligned}D\big(V(t)\big)\leq D\big(V(0)\big)&\exp\Big(-\int_0^t g_1(s)ds\Big)\\&+4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(0)\big)\exp\Big(-\int_0^t g_1(s)ds\Big)\\&\times\int_0^t \mathrm{e}^{\Phi_2(s)-\Psi_2(s)}\big(D\big(X(0)\big)+C_4t\big)ds.\end{aligned}$$

Because $\mathrm{e}^{\Phi_2(s)-\Psi_2(s)}(\xi_3+t)^{\xi_4}\leq 1$ and $\xi_4>2$, we know

$$\int_0^t \mathrm{e}^{\Phi_2(s)-\Psi_2(s)}\big(D\big(X(0)\big)+C_4t\big)ds=C_5<\infty.$$

Then we can get a estimate similar to (3.3)

$$D\big(V(t)\big)\leq\Big(D\big(V(0)\big)+4k_1k_2\sqrt{d}\psi_M D\big(\mathcal{A}(0)\big)C_5\Big)\exp\Big(-\int_0^t g_1(s)ds\Big).$$

We come back to the same situation as for (3.4). Therefore, the system (1.4) has a body attitude flocking. ◻

Remark 3.2. If $0<\psi_m\leq\psi(r)\leq\psi_M$, then the assumption $(\mathbf{H})$ is easily satisfied. But this does not mean that Theorem 3.1 is a special case of Theorem 3.2. Both the conditions and the conclusions in Theorem 3.1 are independent of the number of agents N. But in Theorem 3.2, η is dependent on N. The number of agents become larger, then the flocking converges rate will become smaller. Thus, the result of Theorem 3.2 is not extendable to the macroscopic or kinetic case.

3.2 Finite-in-time stability

In this subsection, we study the finite-in-time stability of system (1.4) with respect to initial data.

Theorem 3.3. *Suppose $\mathcal{Z}$ and $\tilde{\mathcal{Z}}$ are two global solutions to system* (1.4), *the system parameters satisfy $\beta<1/2, 0<a=k_2(3\psi_m-\psi_M), 0<\psi_m<\psi(r)<\psi_M$,*

$$\max\left\{D\big(\mathcal{A}(0)\big),D\big(\tilde{\mathcal{A}}(0)\big)\right\}<\frac{a}{2k_2\psi_M},$$

$$\left(1+\left(\max\left\{D\big(\mathcal{X}(0)\big),D\big(\tilde{\mathcal{X}}(0)\big)\right\}\right)^2\right)^\beta<\frac{2k_1}{a}.$$

Then, for any $T>0$, there exists a positive constant $G=G(T)$ such that

$$\sup_{t\in[0,T)}\|\mathcal{Z}-\tilde{\mathcal{Z}}\|_{2,\infty}\leq G\|\mathcal{Z}^0-\tilde{\mathcal{Z}}^0\|_{2,\infty}.$$

Proof. We firstly define some symbols for concise representation. Let $\mathcal{Z}=(X,V,\mathcal{A})$ and $\tilde{\mathcal{Z}}=(\tilde{X},\tilde{V},\tilde{\mathcal{A}})$ are two solutions to (1.4). The initial data $\mathcal{Z}^0,\tilde{\mathcal{Z}}^0$ satisfy the conditions of Theorem 3.1. We set

$$\mathcal{X}=\max_{1\leq i\leq N}|x_i-\tilde{x}_i|,\quad \mathcal{V}=\max_{1\leq i\leq N}|v_i-\tilde{v}_i|,\quad \Lambda=\max_{1\leq i\leq N}\|A_i-\tilde{A}_i\|.$$

Step A. (Estimate of $\Lambda(t)$). For the estimate of Λ we directly use the following conclusion in [15]:

$$\frac{d}{dt}\Lambda\leq 4k_2(1+d)\Big(\psi_M\max_{i,k}\|A_i-A_k\|+\psi_M\Big)\Lambda. \tag{3.8}$$

The proof of (3.8) can be referred to [15, Theorem 4.4]. Because

$$\max_{i,k}\|A_i-A_k\|\leq 2\max_i\|A_i\|=2\sqrt{d},$$

we have

$$\frac{d}{dt}\Lambda(t)\leq\delta\Lambda(t), \tag{3.9}$$

where $\delta=4k_2(1+d)\psi_M(2\sqrt{d}+1)$. Thus, $\Lambda\leq e^{\delta t}\Lambda(0)$.

Step B. (Estimate of $\mathcal{V}(t)$). There exist at most countable number of increasing times t_k such that we can choose index i such that $\mathcal{V}=|v_i(t)-\tilde{v}_i(t)|$ on any time interval (t_k,t_{k+1})

$$\begin{aligned}\frac{d}{dt}|v_i-\tilde{v}_i|^2&=\langle\dot{v}_i-\dot{\tilde{v}}_i,v_i-\tilde{v}_i\rangle\\&=\frac{k_1}{N}\sum_{k=1}^{N}\big\langle\varphi_{ik}(v_k-v_i)-\tilde{\varphi}_{ik}(\tilde{v}_k-\tilde{v}_i),v_i-\tilde{v}_i\big\rangle\\&\quad+\frac{k_1}{N}\sum_{k=1}^{N}\Big\langle\varphi_{ik}\dot{A}_iA_i^{-1}(x_i-x_k)-\tilde{\varphi}_{ik}\dot{\tilde{A}}_i\tilde{A}_i^{-1}(\tilde{x}_i-\tilde{x}_k),v_i-\tilde{v}_i\Big\rangle\\&=I_1+I_2,\end{aligned}$$

$$\begin{aligned}I_1&=\frac{k_1}{N}\sum_{k}^{N}\langle\varphi_{ik}(v_k-v_i)-\varphi_{ik}(\tilde{v}_k-\tilde{v}_i)+\varphi_{ik}(\tilde{v}_k-\tilde{v}_i)-\tilde{\varphi}_{ik}(\tilde{v}_k-\tilde{v}_i),v_i-\tilde{v}_i\rangle\\&=\frac{k_1}{N}\sum_{k}^{N}\langle\varphi_{ik}(v_k-v_i-\tilde{v}_k+\tilde{v}_i),v_i-\tilde{v}_i\rangle+\frac{k_1}{N}\sum_{k}^{N}\langle(\varphi_{ik}-\tilde{\varphi}_{ik})(\tilde{v}_k-\tilde{v}_i),v_i-\tilde{v}_i\rangle\\&=I_{11}+I_{12}.\end{aligned}$$

Because $|v_i-\tilde{v}_i|=\max\limits_j|v_j-\tilde{v}_j|$,

$$\langle(v_k-v_i-\tilde{v}_k+\tilde{v}_i),v_i-\tilde{v}_i\rangle=\langle v_k-\tilde{v}_k,v_i-\tilde{v}_i\rangle-|v_i-\tilde{v}_i|\le 0.$$

Hence, we can get $I_{11}\le 0$. According to Theorem 3.1, we can get

$$\begin{aligned}I_{12}&\le\frac{k_1}{N}\sum_{k}^{N}\big|\varphi(|x_i-x_k|)-\tilde{\varphi}(|\tilde{x}_i-\tilde{x}_k|)\big||\tilde{v}_k-\tilde{v}_i||v_i-\tilde{v}_i|\\&\le\frac{k_1}{N}M\mathrm{e}^{-\frac{c}{2}t}D\big(V(0)\big)\sum_{k}^{N}\mathrm{Lip}(\varphi)\big||x_i-x_k|-|\tilde{x}_i-\tilde{x}_k|\big||v_i-\tilde{v}_i|\\&\le\frac{k_1}{N}MD\big(V(0)\big)\mathrm{Lip}(\varphi)\mathrm{e}^{-\frac{c}{2}t}\sum_{k}^{N}(|x_i-\tilde{x}_i|+|x_k-\tilde{x}_k|)|v_i-\tilde{v}_i|\\&\le 2k_1M\mathrm{Lip}(\varphi)D\big(V(0)\big)\mathrm{e}^{-\frac{c}{2}t}\mathcal{V}(t)\mathcal{X}(t),\end{aligned}$$

where M and c are defined in Theorem 3.1. For I_2, we have the following estimate:

$$\begin{aligned}I_2&=\frac{k_1}{N}\sum_{k}^{N}\Big\langle\varphi_{ik}\dot{A}_iA_i^{-1}(x_i-x_k)-\tilde{\varphi}_{ik}\dot{\tilde{A}}_i\tilde{A}_i^{-1}(\tilde{x}_i-\tilde{x}_k),v_i-\tilde{v}_i\Big\rangle\\&=\frac{k_1}{N}\sum_{k}^{N}\Big\langle\varphi_{ik}\big(\dot{A}_iA_i^{-1}(x_i-x_k)-\dot{\tilde{A}}_i\tilde{A}_i^{-1}(\tilde{x}_i-\tilde{x}_k)\big),v_i-\tilde{v}_i\Big\rangle\\&\quad+\frac{k_1}{N}\sum_{k}^{N}\Big\langle(\varphi_{ik}-\tilde{\varphi}_{ik})\dot{\tilde{A}}_i\tilde{A}_i^{-1}(\tilde{x}_i-\tilde{x}_k),v_i-\tilde{v}_i\Big\rangle=I_{21}+I_{22},\end{aligned}$$

$$\begin{aligned}I_{21}=\frac{k_1}{N}\sum_{k}^{N}\Big\langle\varphi_{ik}\Big[&\dot{\tilde{A}}_i\tilde{A}_i^{-1}(x_i-\tilde{x}_i)-\dot{\tilde{A}}_i\tilde{A}_i^{-1}(x_k-\tilde{x}_k)\\&+\big(\dot{A}_iA_i^{-1}-\dot{\tilde{A}}_i\tilde{A}_i^{-1}\big)(x_i-x_k)\Big],v_i-\tilde{v}_i\Big\rangle.\end{aligned}$$

Because

$$\begin{aligned}\big\|\dot{A}_iA_i^{-1}-\dot{\tilde{A}}_i\tilde{A}_i^{-1}\big\|&\le\frac{k_2}{N}\sum_{k=1}^{N}\Big\|\psi_{ik}\big(A_kA_i^{-1}-A_iA_k^{-1}\big)-\tilde{\psi}_{ik}\big(\tilde{A}_k\tilde{A}_i^{-1}-\tilde{A}_i\tilde{A}_k^{-1}\big)\Big\|\\&\le\frac{k_2}{N}\sum_{k=1}^{N}\Big(\psi_{ik}\big\|A_kA_i^{-1}-\tilde{A}_k\tilde{A}_i^{-1}-A_iA_k^{-1}+\tilde{A}_i\tilde{A}_k^{-1}\big\|\\&\qquad+\big\|(\psi_{ik}-\tilde{\psi}_{ik})\big(\tilde{A}_k\tilde{A}_i^{-1}-\tilde{A}_i\tilde{A}_k^{-1}\big)\big\|\Big)\\&=I_{211}+I_{212}.\end{aligned}$$

Because $\|A_k\|=\sqrt{d}$ and $\max_{1\le i\le N}\|A_i-\tilde{A}_i\|=\Lambda(t)$,

$$\begin{aligned}I_{211}=\frac{k_2}{N}\sum_{k=1}^{N}\psi_{ik}\Big\|&\big(A_k\big(A_i^{-1}-\tilde{A}_i^{-1}\big)+(A_k-\tilde{A}_k)\tilde{A}_i^{-1}\\&-A_i\big(A_k^{-1}-\tilde{A}_k^{-1}\big)-(A_i-\tilde{A}_i)\tilde{A}_k^{-1}\big)\Big\|\\\le 4\sqrt{d}k_2\psi_M\Lambda(t).&\end{aligned}$$

The last inequality sign is due to

$$\big\|A_i^{-1}-\tilde{A}_i^{-1}\big\|=\big\|A_i-\tilde{A}_i\big\|$$

for $A_1,\tilde{A}_i\in\mathrm{SO}(d)$. And

$$I_{212}\le\frac{2k_2}{N}\mathrm{Lip}(\psi)\sum_{k=1}^{N}\big(|x_i-\tilde{x}_i|+|x_k-\tilde{x}_k|\big)D\big(\tilde{\mathcal{A}}(t)\big)\le 4k_2\mathrm{Lip}(\psi)\tilde{c}_0\mathrm{e}^{-at}\mathcal{X}(t),$$

where

$$\tilde{c}_0=\left(\frac{1}{D(\tilde{\mathcal{A}}_0)}-\frac{2\psi_M}{3\psi_m-\psi_M}\right)^{-1}.$$

Hence,

$$\big\|\dot{A}_iA_i^{-1}-\dot{\tilde{A}}_i\tilde{A}_i^{-1}\big\|\le 4\sqrt{d}k_2\psi_M\Lambda(t)+4k_2\mathrm{Lip}(\psi)\tilde{c}_0\mathrm{e}^{-at}\mathcal{X}(t).$$

Then we can obtain that

$$\begin{aligned}I_{21}\le&\frac{k_1\sqrt{d}\psi_M}{N}\sum_{k}^{N}\Big(\big\|\dot{\tilde{A}}_i\tilde{A}_i^{-1}\big\||x_i-\tilde{x}_i|+\big\|\dot{\tilde{A}}_i\tilde{A}_i^{-1}\big\||x_k-\tilde{x}_k|\Big)\mathcal{V}(t)\\&+\frac{k_1\sqrt{d}\psi_M}{N}\sum_{k}^{N}\big\|\dot{A}_iA_i^{-1}-\dot{\tilde{A}}_i\tilde{A}_i^{-1}\big\||x_i-x_k|\mathcal{V}(t)\end{aligned}$$

$$
\begin{aligned}
&\le \Big(4k_1k_2\sqrt{d}\psi_M D\big(\tilde{\mathcal{A}}(t)\big)\mathcal{X}(t)+k_1\psi_M\sqrt{d}\big\|\dot{A}_iA_i^{-1}-\dot{\tilde{A}}_i\tilde{A}_i^{-1}\big\|x_\infty\Big)\mathcal{V}(t)\\
&\le \Big(4k_1k_2\sqrt{d}\psi_M\tilde{c}_0\mathrm{e}^{-\alpha t}\mathcal{X}(t)+4k_1dk_2\psi_M^2x_\infty\Lambda(t)\Big)\mathcal{V}(t)\\
&\quad +4k_1k_2\sqrt{d}\mathrm{Lip}(\psi)\tilde{c}_0x_\infty\mathrm{e}^{-at}\mathcal{X}(t)\mathcal{V}(t).
\end{aligned}
$$

Similarly, we can obtain an estimate of I_{22}

$$
\begin{aligned}
I_{22}&\le\frac{k_1\sqrt{d}}{N}\sum_k^N\mathrm{Lip}(\varphi)\big(|x_i-\tilde{x}_i|+|x_k-\tilde{x}_k|\big)\big\|\dot{\tilde{A}}_i\tilde{A}_i^{-1}\big\||\tilde{x}_i-\tilde{x}_k||v_i-\tilde{v}_i|\\
&\le 4k_1k_2\sqrt{d}\mathrm{Lip}(\varphi)\tilde{c}_0\tilde{x}_\infty\mathrm{e}^{-\alpha t}\mathcal{X}(t)\mathcal{V}(t).
\end{aligned}
$$

In summary, we can obtain

$$
\begin{aligned}
\frac{d}{dt}\mathcal{V}(t)^2\le{}&2k_1M\mathrm{Lip}(\varphi)\mathrm{e}^{\frac{-c}{2}t}D\big(V(0)\big)\mathcal{X}(t)\mathcal{V}(t)+4k_1k_2\sqrt{d}\mathrm{Lip}(\varphi)\tilde{c}_0\tilde{x}_\infty\mathrm{e}^{-\alpha t}\mathcal{X}(t)\mathcal{V}(t)\\
&+4k_1k_2\sqrt{d}\psi_M\Big(\tilde{c}_0\mathrm{e}^{-\alpha t}\mathcal{X}(t)+\sqrt{d}x_\infty\Lambda(t)+\mathrm{Lip}(\psi)\tilde{c}_0x_\infty\mathrm{e}^{-at}\mathcal{X}(t)\Big)\mathcal{V}(t).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{d}{dt}\mathcal{V}(t)\le{}&k_1M\mathrm{Lip}(\varphi)\mathrm{e}^{\frac{-c}{2}t}D\big(V(0)\big)\mathcal{X}(t)+2k_1k_2\sqrt{d}\mathrm{Lip}(\varphi)\tilde{c}_0\tilde{x}_\infty\mathrm{e}^{-\alpha t}\mathcal{X}(t)\\
&+2k_1k_2\sqrt{d}\psi_M\Big(\tilde{c}_0\mathrm{e}^{-\alpha t}\mathcal{X}(t)+\sqrt{d}\psi_Mx_\infty\Lambda(t)+\mathrm{Lip}(\psi)\tilde{c}_0x_\infty\mathrm{e}^{-at}\mathcal{X}(t)\Big)\\
={}&\Xi_1\mathrm{e}^{-\gamma t}\mathcal{X}(t)+\Xi_2\Lambda(t),
\end{aligned}
$$

where Ξ_1 and Ξ_2 represent the coefficients and $\gamma=\min\{c/2,a\}$. Note that Ξ_1 and Ξ_2 are independent of t and N. Since $\Lambda(t)\le\mathrm{e}^{\delta t}\Lambda(0)$,

$$
\frac{d}{dt}\mathcal{V}(t)\le\Xi_1\mathrm{e}^{-\gamma t}\left(\mathcal{X}(0)+\int_0^t\mathcal{V}(s)ds\right)+\Xi_2\Lambda(t).
$$

Integrating this differential inequality yields

$$
\begin{aligned}
\mathcal{V}(t)&\le\mathcal{V}(0)+\Xi_1\int_0^t\mathrm{e}^{-\gamma s}\left(\mathcal{X}(0)+\int_0^s\mathcal{V}(u)du\right)ds+\frac{\Xi_2}{\delta}\mathrm{e}^{\delta t}\Lambda(0),\\
\max_{\tau\in[0,t]}\mathcal{V}(\tau)&\le\mathcal{V}(0)+\Xi_1\int_0^t\mathrm{e}^{-\gamma s}\Big(\mathcal{X}(0)+\max_{\tau\in[0,t]}\mathcal{V}(\tau)s\Big)ds+\frac{\Xi_2}{\delta}\mathrm{e}^{\delta t}\Lambda(0)\\
&\le\left(\mathcal{V}(0)+\frac{\Xi_1}{\gamma}\mathcal{X}(0)+\frac{\Xi_2}{\delta}\mathrm{e}^{\delta t}\Lambda(0)\right)\mathrm{e}^{\frac{\Xi_1}{\gamma^2}}.
\end{aligned}
$$

Then, for $t\in[0,T)$,

$$
\mathcal{V}(t)\le\left(\mathcal{V}(0)+\frac{\Xi_1}{\gamma}\mathcal{X}(0)+\frac{\Xi_2}{\delta}\mathrm{e}^{\delta T}\Lambda(0)\right)\mathrm{e}^{\frac{\Xi_1}{\gamma^2}}.
$$

Step C. (Estimate of $\mathcal{X}(t)$).

$$\mathcal{X}(t) \le \left(\mathcal{V}(0) + \frac{\Xi_1}{\gamma}\mathcal{X}(0) + \frac{\Xi_2}{\delta}e^{\delta T}\Lambda(0)\right)e^{\frac{\Xi_1}{\gamma^2}T}.$$

To sum up, we can get

$$\begin{aligned}\sup_{t\in[0,T)} \|\mathcal{Z}-\tilde{\mathcal{Z}}\|_{2,\infty} &= \mathcal{X}(t)+\mathcal{V}(t)+\Lambda(t)\\ &\le e^{\frac{\Xi_1}{\gamma^2}}(T+1)\mathcal{V}(0) + \frac{\Xi_1}{\gamma}e^{\frac{\Xi_1}{\gamma^2}}(T+1)\mathcal{X}(0)\\ &\quad + e^{\delta T}\left(\frac{\Xi_2}{\delta}e^{\frac{\Xi_1}{\gamma^2}}(1+T)+1\right)\Lambda(0).\end{aligned}$$

The positive constant $G(T)$ is given by the following explicit form:

$$G(T) = \max\left\{e^{\frac{\Xi_1}{\gamma^2}}(T+1), \frac{\Xi_1}{\gamma}e^{\frac{\Xi_1}{\gamma^2}}(T+1), e^{\delta T}\left(\frac{\Xi_2}{\delta}e^{\frac{\Xi_1}{\gamma^2}}(1+T)+1\right)\right\}.$$

The proof is complete. □

4 The well-posedness of the kinetic equation and mean field limit

In this section, we consider the well-posedness of the kinetic equation (1.5) and give a rigorous proof for the mean-field limit. Firstly, we define some notations

$$\mathcal{E} = \mathcal{C}\big([0,T];\mathcal{P}_c(\Omega)\big),$$

where $\Omega = \mathbb{R}^{2d}\times \mathrm{SO}(d)$, $\mathcal{P}_c(\Omega)$ denotes the space of probability measures with compact support in Ω and endowed with the 1-Wasserstein distance W_1. B_R is the closed ball with center 0 and radius $R>0$ in $\mathbb{R}^{2d}$.

Lemma 4.1. *Take any $R_0>0$ and $f,g\in\mathcal{E}$ such that*

$$\mathrm{supp}(f_t)\cup\mathrm{supp}(g_t)\subseteq B_{R_0}\times \mathrm{SO}(d), \quad \forall t\in[0,T].$$

Then for any $B_R\times\mathrm{SO}(d)\subset\Omega$, there exist a constant $C=C(R,R_0)$ such that

$$\max_{t\in[0,T]}\mathrm{Lip}_R(L[f])\le C, \quad \max_{t\in[0,T]}\mathrm{Lip}_R(Q[f])\le C, \tag{4.1}$$

and

$$\begin{aligned}&\sup_{t\in[0,T]}\left\{\|Q[f]-Q[g]\|_{L^\infty(B_R\times\mathrm{SO}(d))}\right\}\le CW_1(f,g),\\ &\sup_{t\in[0,T]}\left\{\|L[f]-L[g]\|_{L^\infty(B_R\times\mathrm{SO}(d))}\right\}\le CW_1(f,g).\end{aligned} \tag{4.2}$$

Proof. We first prove that (4.1). Fixing v and A,

$$\begin{aligned}
&\big|L[f](t,x_1,v,A)-L[f](t,x_2,v,A)\big| \\
\le& k_1\int_\Omega \big|\varphi(|x_1-y|)(v_*-v)-\varphi(|x_2-y|)(v_*-v)\big|f(t,y,v_*,A_*)dydv_*dA_* \\
&+\bigg|k_1k_2\int_\Omega \varphi(|x_1-y|)\psi(|x_1-y|)K(A,A_*)(x_1-y)f(t,y,v_*,A_*)dydv_*dA_* \\
&\quad -k_1k_2\int_\Omega \varphi(|x_2-y|)\psi(|x_2-y|)K(A,A_*)(x_2-y)f(t,y,v_*,A_*)dydv_*dA_*\bigg| \\
=&I_1+I_2.
\end{aligned}$$

Because the compactness of the rotation group and $\mathrm{supp}(f_t)\cup\mathrm{supp}(g_t)\subseteq B_{R_0}\times\mathrm{SO}(d)$ and $|v|<R$. This is equivalent to $|v_*-v|\le R+R_0=c(R,R_0)$

$$\begin{aligned}
I_1\le& k_1\int_\Omega \mathrm{Lip}(\varphi)|x_1-x_2||v_*-v|f(t,y,v_*,A_*)dydv_*dA_* \\
\le& k_1\mathrm{Lip}(\varphi)c(R,R_0)|x_1-x_2|.
\end{aligned}$$

We define $h(r)=\varphi(r)\psi(r)r$. We can get $h(r)$ is locally Lipschitz. $\mathrm{Lip}_R(h)$ is the Lipschitz constant in B_R

$$\begin{aligned}
I_2\le& \sqrt{d}k_1k_2\int_\Omega \mathrm{Lip}_R(h)|x_1-x_2|\,\|K(A,A_*)\|f(t,y,v_*,A_*)dydv_*dA_* \\
\le& 2\sqrt{d}k_1k_2\int_\Omega \mathrm{Lip}_R(h)|x_1-x_2|\,\|A^*A^{-1}\|f(t,y,v_*,A_*)dydv_*dA_* \\
\le& 2dk_1k_2\mathrm{Lip}_R(h)|x_1-x_2|.
\end{aligned}$$

The uniform Lipschitz continuity of $L[f]$ with respect to v is obvious

$$\big|L[f](t,x,v_1,A)-L[f](t,x,v_2,A)\big|\le|v_1-v_2|.$$

Fixing x and v,

$$\begin{aligned}
&\big|L[f](t,x,v,A_1)-L[f](t,x,v,A_2)\big| \\
\le& k_1k_2\sqrt{d}\int_\Omega \psi_M|x_1-x_2|\big\|A_*\big(A_1^{-1}-A_2^{-1}\big)+\big(A_1-A_2\big)A_*^T\big\|f(t,y,v_*,A_*)dydv_*dA_* \\
\le& 2dc(R,R_0)\psi_M\|A_1-A_2\|.
\end{aligned}$$

Thus, we can obtain

$$\mathrm{Lip}_R(L[f])\le\max\big\{k_1\mathrm{Lip}(\psi)c(R,R_0)+2dk_1k_2\mathrm{Lip}_R(h),1,2dc(R,R_0)\big\}=\omega_1(R,R_0).$$

The uniform Lipschitz continuity of $Q[f]$ with respect to v is obvious. Fixing v and A,

$$\begin{aligned}
&\|Q[f](t,x_1,v,A)-Q[f](t,x_2,v,A)\| \\
\le& k_2\int_0^t|\psi(|x_1-y|)-\psi(|x_2-y|)|\,\|A_*-AA_*A^{-1}\|dydv_*dA_* \\
\le& 2\sqrt{d}k_2\mathrm{Lip}(\psi)|x_1-x_2|.
\end{aligned}$$

By

$$\left\|A_1A_*^{-1}A_1-A_2A_*^{-1}A_2\right\|=\left\|(A_1-A_2)A_*^{-1}A_1+A_2A_*^{-1}(A_1-A_2)\right\|\le 2\sqrt{d}\left\|A_1-A_2\right\|,$$

we have

$$\begin{aligned}&\|Q[f](t,x,v,A_1)-Q[f](t,x,v,A_2)\|\\ =&k_2\int_\Omega \psi_M\left\|A_1A_*^{-1}A_1-A_2A_*^{-1}A_2\right\|f(t,y,v_*,A_*)dydv_*dA_*\\ \le& 2\sqrt{d}k_2\|A_1-A_2\|.\end{aligned}$$

Hence, $Q[f]$ is uniform Lipschitz continuous

$$\mathrm{Lip}_R(Q[f])\le\max\{2\sqrt{d}k_2,2\sqrt{d}k_2\mathrm{Lip}(\psi)\}=\omega_2(R,R_0).$$

Next we prove that (4.2). By Proposition 2.1, let π be an optimal transportation plan between the measures f and g. The π has marginals f and g and the support of π is contained in $(B_{R_0}\times\mathrm{SO}(d))\times(B_{R_0}\times\mathrm{SO}(d))$

$$\begin{aligned}&\|Q[f](t,x,v,A)-Q[g](t,x,v,A)\|\\ =&k_2\int_\Omega \psi(|x-y_1|)\big(A_1-AA_1^{-1}A\big)-\psi(|x-y_2|)\big(A_2-AA_2^{-1}A\big)d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\ \le&k_2\int_\Omega\psi_M\left\|A_1-AA_1^{-1}A-A_2+AA_2^{-1}A\right\|d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\ &\quad+k_2\int_\Omega\mathrm{Lip}(\psi)|y_1-y_2|\left\|A_2-AA_2^{-1}A\right\|d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\ \le&k_2\int_\Omega\psi_M(\sqrt{d}+1)\|A_1-A_2\|d\pi(y_1,v_1,A_1,y_2,v_2,A_2)+2\sqrt{d}k_2\mathrm{Lip}(\psi)W_1(f,g)\\ \le&2k_2\big(\sqrt{d}\mathrm{Lip}(\psi)+(\sqrt{d}+1)\psi_M\big)W_1(f,g)=\omega_3(R,R_0)W_1(f,g).\end{aligned}$$

For $L[f]$, we can obtain the following estimate:

$$\begin{aligned}&\|L[f](t,x,v,A)-L[g](t,x,v,A)\|\\ =&k_1\int_\Omega\big(\varphi(|x-y_1|)(v_1-v)-\varphi(|x-y_2|)(v_2-v)\big)\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\ &\quad+k_1k_2\int_\Omega\varphi(|x-y_1|)\psi(|x-y_1|)K(A,A_1)d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\ &\quad-k_1k_2\int_\Omega\varphi(|x-y_2|)\psi(|x-y_2|)K(A,A_2)d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\ =&I_3+I_4.\end{aligned}$$

We can obtain that

$$\begin{aligned}I_3=&k_1\int_\Omega\varphi(|y_1-x|)-\varphi(|y_2-x|)(v_1-v)\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\ &+k_1\int_\Omega\varphi(|x-y_2|)(v_1-v_2)\pi(y_1,v_1,A_1,y_2,v_2,A_2)\end{aligned}$$

$$
\begin{aligned}
&\le k_1\int_\Omega \mathrm{Lip}(\varphi)|y_1-y_2|(v_1-v)+\varphi(|x-y_2|)(v_1-v_2)\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\
&\le \big(k_1\mathrm{Lip}(\varphi)c(R_0,R)+k_1\big)W_1(f,g),
\end{aligned}
$$

$$
\begin{aligned}
I_4=&k_1k_2\int_\Omega \varphi(|x-y_1|)\psi(|x-y_1|)(x-y_1)K(A,A_1)d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\
&-k_1k_2\int_\Omega \varphi(|x-y_1|)\psi(|x-y_1|)(x-y_1)K(A,A_2)d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\
&+k_1k_2\int_\Omega \varphi(|x-y_1|)\psi(|x-y_1|)(x-y_1)K(A,A_2)d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\
&-k_1k_2\int_\Omega \varphi(|x-y_2|)\psi(|x-y_2|)(x-y_2)K(A,A_2)d\pi(y_1,v_1,A_1,y_2,v_2,A_2).
\end{aligned}
$$

For I_4,

$$
\begin{aligned}
I_4\le& k_1k_2\int_\Omega \psi_M\sqrt{d}c(R_0,R)\|A_1A^{-1}-AA_1^{-1}-A_2A^{-1}+AA_2^{-1}\|d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\
&+k_1k_2\int_\Omega \mathrm{Lip}_R(h)|y_1-y_2|\|K(A,A_2)\|d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\
\le& k_1k_2\int_\Omega 2d\psi_M c(R_0,R)\|A_1-A_2\|+2d\mathrm{Lip}_R(h)|y_1-y_2|d\pi(y_1,v_1,A_1,y_2,v_2,A_2)\\
\le& k_1k_2\big(2d\psi_M c(R_0,R)+2d\mathrm{Lip}_R(h)\big)W_1(f,g).
\end{aligned}
$$

Thus, we can obtain

$$
\|L[f](t,x,v,A)-L[g](t,x,v,A)\|\le\omega_4(R,R_0)W_1(f,g),
$$

where

$$
\omega_4(R,R_0)=k_1k_2\big(2d\psi_M c(R_0,R)+2d\mathrm{Lip}_R(h)\big)+k_1\mathrm{Lip}(\varphi)c(R_0,R)+k_1.
$$

In summary we can set $C(R,R_0)=\max\{\omega_i(R,R_0)\}_{i=1}^4$. $\square$

The associated characteristic system of (1.5) is

$$
\begin{cases}
\dot{x}=v,\\
\dot{v}=L[f](t,x,v,A),\\
\dot{A}=Q[f](t,x,v,A),
\end{cases}
\tag{4.3}
$$

where $(x,v,A)\in\Omega$. The system (4.3) can be conveniently written as

$$
\frac{d}{dt}\mathcal{Q}=\Psi_f(t,\mathcal{Q}),
$$

where $\mathcal{Q}=(x,v,A)$ and $\Psi_f:[0,T]\times\Omega\to\Omega$ is the right-hand side of (4.3). We denote a flow at $t\in[0,T)$ of (4.3), i.e.

$$
\mathcal{I}_f^t(x_0,v_0,A_0)=\big(x(t),v(t),A(t)\big),\quad (x_0,v_0,A_0)\in\Omega.
$$

Next, we give the definition of weak solution of system (1.5) (see [3]).

Definition 4.1 (Weak Solution)**.** *Take a measure $f_0 \in \mathcal{P}_1(\Omega)$. We say that a function $f:[0,T] \to \mathcal{P}_1(\Omega)$ is a weak solution to* (1.5) *with initial condition f_0, if it satisfies that*

$$f_t = \mathcal{I}_f^t \# f_0, \quad \forall t \in [0,T),$$

where $\mathcal{I}^t \# f_0$ is defined as

$$\int_\Omega \zeta(x) d\big(\mathcal{I}_f^t \# f_0\big)(x) = \int_\Omega \zeta\big(\mathcal{I}_f^t(x)\big) d f_0(x)$$

for every measurable function $\zeta : \Omega \to \mathbb{R}$.

We give some basic regularity results for the (4.3).

Lemma 4.2. *Consider the closed ball $B_R \subseteq \mathbb{R}^{2d}, R > 0$ and $T > 0$*

(1) Ψ_f is bounded in compact sets. For $\mathcal{Q} \in B_R \times \mathrm{SO}(d)$ and $t \in [0,T]$

$$\|\Psi_f(t,\mathcal{Q})\|_\Omega \leq C_\Psi,$$

where C_{Ψ_f} which depends only on R.

(2) Ψ_f is locally Lipschitz with respect to x,v,A. For all $\mathcal{Q}_1, \mathcal{Q}_2 \in B_R \times \mathrm{SO}(d)$ and $t \in [0,T]$,

$$\|\Psi_f(t,\mathcal{Q}_1) - \Psi_f(t,\mathcal{Q}_2)\| \leq \mathrm{Lip}_R(\Psi_f) \|\mathcal{Q}_1 - \mathcal{Q}_2\|.$$

Proof. It is obvious for (1). And (2) can be obtained by Lemma 4.1. □

Based on Lemmas 4.1, 4.2 and the framework developed in [3], we can provide existence and uniqueness of a compactly supported global weak solution to Eq. (1.5) in a compactly supported measure initial condition, see [3, Theorem 4.11] for details. We give the following results about existence, uniqueness and stability of weak solutions to the kinetic equation (1.5).

Theorem 4.1. *Let f_0 be a measure on Ω with compact support. There exists a solution f_t on $[0,+\infty)$ to* (1.5) *with initial condition f_0. Furthermore*

$$f_t \in \mathcal{C}\big([0,+\infty); \mathcal{P}_c(\Omega)\big), \tag{4.4}$$

and there is an increasing function $R = R(T)$ such that for all $T > 0$,

$$\operatorname{supp} f_t \subseteq B_{R(T)} \times \mathrm{SO}(d) \subseteq \Omega, \quad \forall t \in [0,T]. \tag{4.5}$$

This solution is unique among the family of solutions satisfying (4.4) *and* (4.5)*. Moreover, given any other initial data $g_0 \in \mathcal{P}_c(\Omega)$ and g its corresponding solution. Then there exists a strictly increasing function $r(t) > 0$ with $r(0) = 1$ the size of the support of f_0 and g_0 such that*

$$W_1(f_t, g_t) \leq r(t) W_1(f_0, g_0), \quad t \geq 0. \tag{4.6}$$

Proof. According to Lemmas 4.1 and 4.2, the proof can be done using the framework developed in [3]. We briefly state the main ideas here. Suppose $f_0 \in \mathcal{P}_c(\Omega)$ with the support contained in B_{R_0}. We define a set as follows:

$$\mathcal{F} = \big\{ f \in C\big([0,T];\mathcal{P}_c(\Omega)\big), \mathrm{supp} f \subseteq B_R \big\},$$

where $R = 2R_0$. For $f \in \mathcal{F}$, we define

$$\Gamma[f](t) = \mathcal{I}_f^t \# f_0.$$

For any $\mathcal{Q} \in B_{R_0} \times \mathrm{SO}(d)$, because $\|\Psi(t,\mathcal{Q})\|_\Omega \le C_{\Psi_f}$ we have

$$\mathcal{I}_f^t(\mathcal{Q}) \le C_{\Psi_f} T.$$

As long as T is small enough we can obtain $\mathrm{supp}(\mathcal{I}_f^t \# f_0) \subseteq B_R \times \mathrm{SO}(d)$. The continuity of $\mathcal{I}_f^t \# f_0$ can be referred to [3, Theorem 3.12]. Hence, we have $\Gamma : \mathcal{F} \to \mathcal{F}$. For $f,g \in \mathcal{F}$ and $\mathcal{Q}_0 \in B_{R_0} \times \mathrm{SO}(d)$, we set

$$\mathcal{Q}_1(t) = \mathcal{I}_f^t(\mathcal{Q}_0), \quad \mathcal{Q}_2(t) = \mathcal{I}_g^t(\mathcal{Q}_0).$$

Then we have

$$\begin{aligned}
\|\mathcal{Q}_1(t) - \mathcal{Q}_2(t)\|_\Omega &\le \int_0^t \big\|\Psi_f\big(s,\mathcal{Q}_1(s)\big) - \Psi_g\big(s,\mathcal{Q}_2(s)\big)\big\|_\Omega ds \\
&\le \int_0^t \big\|\Psi_f\big(s,\mathcal{Q}_1(s)\big) - \Psi_f\big(s,\mathcal{Q}_2(s)\big)\big\|_\Omega ds \\
&\quad + \int_0^t \big\|\Psi_f\big(s,\mathcal{Q}_2(s)\big) - \Psi_g\big(s,\mathcal{Q}_2(s)\big)\big\|_\Omega ds \\
&\le \mathrm{Lip}_R(\Psi_f) \int_0^t \|\mathcal{Q}_1 - \mathcal{Q}_2\|_\Omega ds \\
&\quad + \int_0^t \Big(\|Q[f] - Q[g]\|_{L^\infty(B_R \times \mathrm{SO}(d))} + \|L[f] - L[g]\|_{L^\infty(B_R \times \mathrm{SO}(d))} \Big) ds.
\end{aligned}$$

By Gronwall's lemma,

$$\begin{aligned}
\|\mathcal{I}_f^t - \mathcal{I}_g^t\|_\Omega \le & C_1 \sup_{t\in[0,T]} \big\{ \|Q[f] - Q[g]\|_{L^\infty(R \times \mathrm{SO}(d))} \big\} \\
& + C_1 \sup_{t \in [0,T]} \big\{ \|L[f] - L[g]\|_{L^\infty(B_R \times \mathrm{SO}(d))} \big\},
\end{aligned}$$

where $C_1 = (e^{\mathrm{Lip}_R(\Psi_f)T} - 1)/\mathrm{Lip}_R(\Psi_f)$. By [3, Theorem 3.11] and Lemma 4.1,

$$\begin{aligned}
W_1\big(\mathcal{I}_f^t \# f_0, \mathcal{I}_g^t \# f_0\big) &\le C_1 \sup_{t\in[0,T]} \big\{ \|Q[f] - Q[g]\|_{L^\infty(R \times \mathrm{SO}(d))} \big\} \\
&\quad + C_1 \sup_{t\in[0,T]} \big\{ \|L[f] - L[g]\|_{L^\infty(B_R \times \mathrm{SO}(d))} \big\} \\
&\le 2CC_1 W_1(f,g).
\end{aligned}$$

We can in addition choose T small enough so that $2CC_1<1$. Thus, Γ is contractive, and this proves that there is a unique fixed point in $\mathcal{F}$. We obtained local existence and uniqueness of solutions. For initial value $\mathcal{Q}_0$, by (4.3) we can get

$$\|\mathcal{Q}\|_\Omega \leq e^{C_2 t}, \tag{4.7}$$

where C_2 depending on $\|\mathcal{Q}_0\|_\Omega, T, k_1, k_2$. According to (4.7), we can estimate the support set $B_R \times \mathrm{SO}(d)$ for any $T>0$. Hence, we can extend this solution as long as the support of the solution remains compact. The stability result (4.6) is similar to the proof of [3, Theorem 3.16]. ☐

Next we investigate the mean-field approximation, i.e. the approximation of a continuum measure by empirical measures. Firstly, we introduce the empirical measure

$$f_0^N = \sum_{i=1}^N m_i \delta_{(x_i^0, v_i^0, A_i^0)}, \quad \sum_{i=1}^N m_i = 1, \tag{4.8}$$

and denote f_t^N is the weak solution to Eq. (1.5) with initial value f_0^N on $t\in[0,T)$. By [3, Lemma 5.1], we know f_t^N is the empirical measure with trajectories $(x_i(t), v_i(t), A_i(t))$

$$f_t^N = \sum_{i=1}^N m_i \delta_{(x_i(t), v_i(t), A_i(t))}, \quad t\in[0,T),$$

where $(x_i(t), v_i(t), A_i(t))$ satisfies the following equation:

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = \sum_{i=1}^N m_i L\left[f_t^N\right](t, x_i, v_i, A_i), \\ \dot{A}_i = \sum_{i=1}^N m_i Q\left[f_t^N\right](t, x_i, v_i, A_i), \quad x_i(0) = x_i^0, \quad v_i(0) = v_i^0, \quad A_i(0) = A_i^0. \end{cases} \tag{4.9}$$

When $m_i = 1/N, 1\leq i\leq N$, obviously the above equation is system (1.4) by the properties of Dirac measure. The mean-field limit of the system (1.4) is given by the following corollary.

Corollary 4.1 (The Mean-Field Limit). *Let f_0^N be a sequence of empirical measure with the form* (4.8) *and $f_0\in\mathcal{P}_c(\Omega)$ with compact support. f_0^N and f_0 satisfy*

$$\lim_{N\to\infty} W_1\left(f_0, f_0^N\right) = 0.$$

For $T>0, f_t$ and f_t^N are the unique weak solutions to the kinetic equation (1.5) *with initial value f_0 and f_0^N, respectively. Then, there exists $T_*\in(0,T)$ such that*

$$\lim_{N\to\infty} W_1\left(f_t, f_t^N\right) = 0, \quad \forall t\in(0,T^*]. \tag{4.10}$$

Proof. By (4.6), we can find a strictly increasing function $r(t)$ such that

$$W_1\big(f_t,f_t^N\big)\le r(t)W_{f_0,f_0^N}$$

for all $t\in[0,T^*)$ and $N\in\mathbb{N}_+$. Because $r(t)$ is bounded on $[0,T^*)$, when $N\to\infty$, it is clear that (4.10) holds. □

Remark 4.1. By the stability result (4.6), we can extend the result of Theorem 3.1 to the kinetic system (1.5). We define

$$\begin{aligned}
&D_x[f]=\operatorname{diam}(\operatorname{supp}_x f),\\
&D_v[f]=\operatorname{diam}(\operatorname{supp}_v f),\\
&D_A[f]=\operatorname{diam}(\operatorname{supp}_A f),
\end{aligned}$$

where $\operatorname{supp}_x f$ denotes the x-projection of $\operatorname{supp} f$ and similarly for $\operatorname{supp}_v f$ and $\operatorname{supp}_A f$. Suppose $f_t\in C([0,T],\mathcal{P}_c(\Omega))$ is a weak solution of (1.5) subject to a compactly supported initial datum $f_0\in\mathcal{P}_c(\Omega)$. And we assume that $0\le\beta<1/2$, ψ satisfies (1.3),

$$D_A[f_0]<\frac{a}{2k_2\psi_M},\quad 0<a<\frac{2k_1}{(1+D_x[f_0]^2)^\beta}.$$

Because $(\mathcal{P}_c(\Omega),W_1)$ is a Polish space, we can find a N-particle approximations of f_0 and denoted as $\{f_0^N\}_{N\in\mathbb{N}}$, i.e.

$$f_0^N=\frac{1}{N}\sum_{i=1}^N\delta\big(x-x_i(0)\big)\otimes\delta\big(v-v_i(0)\big)\otimes\delta\big(A-A_i(0)\big),\quad t\in[0,T),$$

and

$$\lim_{N\to\infty}W_1\big(f_0^N,f_0\big)=0.$$

By Theorem 3.1, we have

$$D_v\big[f_t^N\big]\le MD_v\big[f_0^N\big]e^{-\frac{c}{2}t},\quad D_A\big[f_t^N\big]\le c_0e^{-at},\quad t\in[0,T),$$

where f_t^N is a weak solution of (1.5) subject to the initial datum f_0^N and

$$c_0=\left(\frac{1}{D_A[f_0^N]}-\frac{2\psi_M}{3\psi_m-\psi_M}\right)^{-1}.$$

According to (4.6),

$$W_1\big(f_t,f_t^N\big)\le r(T)W_1\big(f_0,f_0^N\big),\quad t\in[0,T].$$

Because $r(T),M,c,c_0,a$ is independent of N. Fixing $T>0$, letting $N\to\infty$, then $D_v[f_t^N]=D_v[f_t]$ and $D_A[f_t^N]=D_A[f_t]$, i.e.

$$D_v[f_t]\le MD_v[f_0]e^{-\frac{c}{2}t},\quad D_A[f_t]\le c_0e^{-at},\quad t\in[0,T).$$

And $\sup_{t\in[0,T]}D_x[f_t]$ is obvious.

5 Numerical simulations

In this section, we show the numerical simulations for the emergent dynamics of the system (1.4). Let $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ denote the canonical basic, where $\{\mathbf{e}_i\}_{i=1}^3$ are standard unit vectors. We use $(\mathbf{e}_1^i,\mathbf{e}_2^i,\mathbf{e}_3^i)$ to denote the body attitude of the i-th agent

$$(\mathbf{e}_1^i,\mathbf{e}_2^i,\mathbf{e}_3^i)=(A_i\mathbf{e}_1,A_i\mathbf{e}_2,A_i\mathbf{e}_3).$$

Display $\mathbf{e}_1^i,\mathbf{e}_2^i,\mathbf{e}_3^i$ axes in colors red, green, blue, respectively as in Fig. 2.

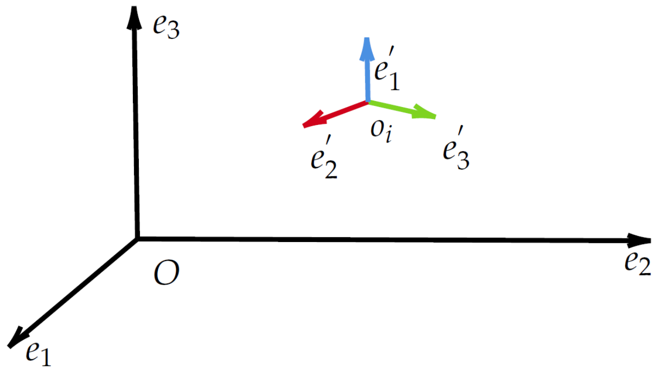

Figure 2: Illustration of the body attitude.

In Figs. 3(a)-3(b), we can find the asymptotic synchronization of the body attitude of agents. And in Figs. 4(a) and 4(b) illustrate the asymptotic alignment of the velocity fluctuation and the body attitude fluctuation, the position fluctuation is bounded. This implies the existence of asymptotic body attitude flocking for the system (1.4).

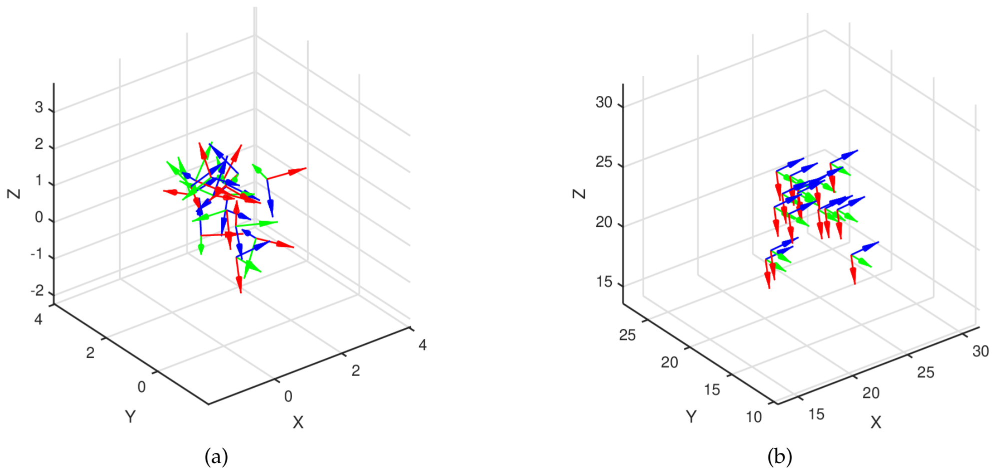

Figure 3: (a): Agents with no coordinated body attitude at the initial moment. (b): Alignment of body-orientations at the final moment.

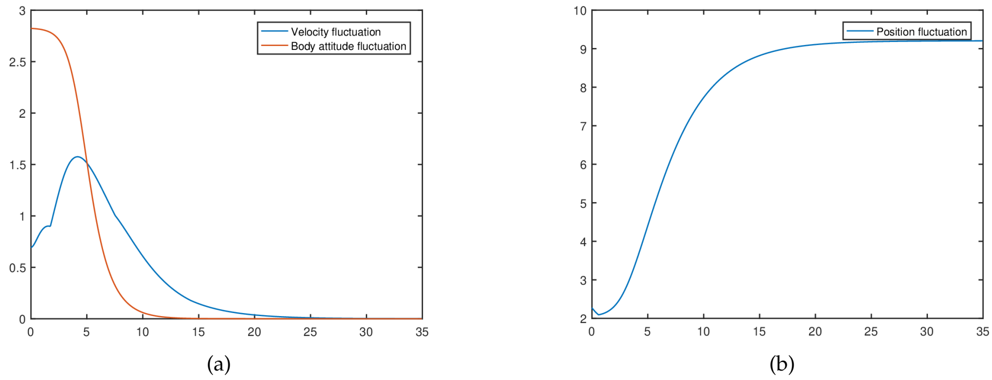

Figure 4: (a): Velocity fluctuation and body attitude fluctuation. (b): Position fluctuation.

6 Conclusion

In this paper, we have studied body attitude flocking behaviors of a new model for body attitude coordination. Unlike the body attitude alignment models in [9, 15], the velocity of each agent is constantly adjusted according to the velocity of other agents. In this case, flocking would include the body attitude alignment and velocity alignment. It will generate a new collective behaviour which is called body attitude flocking (see Definition 2.1). We present two sufficient frameworks leading to the body attitude flocking. The first framework requires a positive lower bound on communication function $\psi(r)$, which yields flocking estimates that are independent of the number of agents N. The second framework removes the assumption that the communication function $\psi(r)$ has a positive lower bound, but obtains flocking estimates related to the number of agents. Based on the sufficient framework, we present the finite-in-time stability estimate which is valid on any finite time interval. In addition, we formally derive a kinetic model of the model for body attitude coordination using the BBGKY hierarchy. We prove the well-posedness of the kinetic equation based on the framework of [3] and give a rigorous proof for the mean-field limit of the system (1.4). We extend the result of Theorem 3.1 to the kinetic system (1.5) and present a sufficient condition for asymptotic flocking in the kinetic system (1.5). Of course, there are still lots of interesting open questions such as the extension of stability estimate to the whole time interval and uniform-in-time asymptotic flocking dynamics of the kinetic model. These issues will be addressed in future works.

Appendix A. Proof for Lemma 3.1

In this appendix, we provide the proof of Lemma 3.1.

Proof. Firstly we estimate $d(A_i^T A_j)/dt$. Although the estimate is similar to [12, Theorem 5.11], we provide the details here for the readers' convenience

$$
\begin{aligned}
\frac{d}{dt}\big(A_i^T A_j\big) &= \dot{A}_i^T A_j + A_i^T A_j \\
&= \frac{1}{N}\sum_{k=1}^{N}\Big(\psi_{ik}\big(A_k^T A_j - A_i^T A_k A_i^T A_j\big) + \psi_{jk}\big(A_i^T A_k - A_i^T A_j A_k^T A_j\big)\Big).
\end{aligned}
$$

We make the following transformations:

$$
\begin{aligned}
A_k^T A_j - A_i^T A_k A_i^T A_j &= 2A_k^T A_j - \big(I + A_i^T A_k A_i^T A_k\big) A_k^T A_j \\
&= 2A_k^T A_j - 2A_i^T A_j - \big(I - 2A_i^T A_k + A_i^T A_i A_i^T A_k\big) A_k^T A_j \\
&= 2(A_k - A_i)^T A_j - \big(I - A_i^T A_k\big)^2 A_k^T A_j.
\end{aligned}
$$

Taking the trace of the above equation yields

$$
\begin{aligned}
&\operatorname{tr}\big(A_k^T A_j - A_i^T A_k A_i^T A_j\big) \\
={}& 2\operatorname{tr}\big((A_k - A_i)^T A_j\big) - \operatorname{tr}\big((I - A_i^T A_k)^2 A_k^T A_j\big) \\
={}& 2\operatorname{tr}\big((A_k - A_i)^T A_j\big) - \operatorname{tr}\big(A_i^T (A_i - A_k)\big(A_k^T - A_i^T\big) A_j\big) \\
={}& 2\operatorname{tr}\big((A_k - A_i)^T A_j\big) - \operatorname{tr}\big((A_i - A_k)\big(A_k^T - A_i^T\big) A_j A_i^T\big) \\
={}& 2\operatorname{tr}\big((A_k - A_i)^T A_j\big) - \frac{1}{2}\operatorname{tr}\big((A_i - A_k)\big(A_k^T - A_i^T\big)\big(A_j A_i^T + A_i A_j^T\big)\big) \\
={}& 2\operatorname{tr}\big((A_k - A_i)^T A_j\big) - \frac{1}{2}\operatorname{tr}\big((A_i - A_k)\big(A_k^T - A_i^T\big)\big(2I - (A_i - A_j)(A_i - A_j)^T\big)\big) \\
={}& 2\operatorname{tr}\big((A_k - A_i)^T A_j\big) + \operatorname{tr}\big((A_i - A_k)(A_i - A_k)^T\big) \\
&\quad - \frac{1}{2}\operatorname{tr}\big((A_i - A_k)(A_i - A_k)^T (A_i - A_j)(A_i - A_j)^T\big) \\
={}& 2\operatorname{tr}\big((A_k - A_i)^T A_j\big) + \|A_i - A_k\|^2 - \frac{1}{2}\big\|(A_i - A_k)^T (A_i - A_j)\big\|^2.
\end{aligned}
$$

Then we can get

$$
\begin{aligned}
&\operatorname{tr}\big(A_k^T A_j - A_i^T A_i A_i^T A_j\big) \\
\geq{}& 2\Big(\operatorname{tr}\big(A_k^T A_j\big) + \operatorname{tr}\big(A_i^T A_j\big)\Big) + \|A_i - A_k\|^2 \left(1 - \frac{1}{2}\|A_i - A_j\|^2\right) \\
\geq{}& 2\Big(\operatorname{tr}\big(A_k^T A_j\big) + \operatorname{tr}\big(A_i^T A_j\big)\Big) + \|A_i - A_k\|^2 \Big(\operatorname{tr}\big(R_i^T R_j\big) - d + 1\Big).
\end{aligned}
$$

Repeating the above estimation for $(A_i^T A_k - A_i^T A_j A_k^T A_j)$ yields

$$
\begin{aligned}
\frac{d}{dt} A_i^T A_j \geq{}& \frac{2}{N}\sum_{k=1}^{N}\Big(\psi_{ik}\Big(\operatorname{tr}\big(A_k^T A_j\big) - \operatorname{tr}\big(A_i^T A_j\big)\Big) + \psi_{jk}\Big(\operatorname{tr}\big(A_k^T A_i\big) - \operatorname{tr}\big(A_i^T A_j\big)\Big)\Big) \\
&+ \frac{1}{N}\Big(\operatorname{tr}\big(A_i^T A_j\big) - d + 1\Big)\sum_{k=1}^{N}\big(\psi_{ik}\|A_i - A_k\|^2 + \psi_{jk}\|A_j - A_k\|^2\big). \qquad \text{(A.1)}
\end{aligned}
$$

There exist at most countable number of increasing times t_m such that we can choose indices $1\le i,j\le N$ such that $\|A_i-A_j\|=D(\mathcal{A})$ for $t\in(t_m,t_{m+1})$. Because $\|A_i-A_j\|^2=2d-2\mathrm{tr}(A_i^TA_j)$,

$$\mathrm{tr}\big(A_i^TA_j\big)=\min_{1\le k,l\le N}\mathrm{tr}\big(A_k^TA_l\big).$$

By (A.1), we can obtain

$$\begin{aligned}\frac{d}{dt}\mathrm{tr}\big(A_i^TA_j\big)&\ge\frac{1}{N}\big(\mathrm{tr}\big(A_i^TA_j\big)-d+1\big)\sum_{k=1}^N\big(\psi_{ik}\|A_i-A_k\|^2+\psi_{jk}\|A_j-A_k\|^2\big)\\&\ge\frac{1}{N}\big(\mathrm{tr}\big(A_i^TA_j\big)-d+1\big)\big(\psi_{ij}\|A_i-A_j\|^2+\psi_{ji}\|A_j-A_i\|^2\big)\\&\ge\frac{2}{N}\psi_m(t)\big(\mathrm{tr}\big(A_i^TA_j\big)-d+1\big)\big(2d-2\mathrm{tr}\big(A_i^TA_j\big)\big)\\&\ge\frac{4}{N}\psi_m(t)\big(\mathrm{tr}\big(A_i^TA_j\big)-d+1\big)\big(d-\mathrm{tr}\big(A_i^TA_j\big)\big),\end{aligned}\tag{A.2}$$

where $\psi_m(t)=\psi(D(X)(t))$. Since $D(\mathcal{A}(0))<\sqrt{2}$ and $A_i^TA_j\le d$, we have

$$d-1<\mathrm{tr}\Big(\big(A_i^0\big)^TA_j^0\Big)\le d.$$

Hence,

$$\begin{aligned}&d-\frac{1}{2}D\big(\mathcal{A}(0)\big)^2<\mathrm{tr}\big(A_i^TA_j\big)\le d,\quad\forall t\in[0,\infty),\\&\frac{d}{dt}\big(d-\mathrm{tr}\big(A_i^TA_j\big)\big)\le\frac{4}{N}\psi_m(t)\big(\mathrm{tr}\big(A_i^TA_j\big)-d+1\big)\big(\mathrm{tr}\big(A_i^TA_j\big)-d\big).\end{aligned}$$

Then we can obtain

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|A_i-A_j\|^2\le-\frac{4}{N}\psi_m(t)\left(1-\frac{1}{2}D\big(\mathcal{A}(0)\big)^2\right)\|A_i-A_j\|^2,\\&\frac{d}{dt}D\big(\mathcal{A}(t)\big)\le-\frac{4}{N}\psi_m(t)\left(1-\frac{1}{2}D\big(\mathcal{A}(0)\big)^2\right)D\big(\mathcal{A}(t)\big)\le-\eta\psi_m(t)D\big(\mathcal{A}(t)\big),\end{aligned}$$

where

$$\eta=\frac{4}{N}\left(1-\frac{1}{2}D\big(\mathcal{A}(0)\big)^2\right).$$

The above differential inequality can be integrated to get

$$D\big(\mathcal{A}(t)\big)\le\exp\left(-\eta\int_0^t\psi_m(s)ds\right)D\big(\mathcal{A}(0)\big).$$

The proof is complete. $\square$

Appendix B. Derivation of a mean-field model

In this appendix, we use the standard BBGKY hierarchy to derive the mean-field kinetic model of the system (1.4) formally (see [14,20,27] for related results). Let

$$f^N = f^N(t,x_1,v_1,A_1,x_2,v_2,A_2,\cdots,x_N,v_N,A_N)$$

be the N-particle distribution function. Since particles are indistinguishable, f^N is symmetric

$$f^N(t,\cdots,x_i,v_i,A_i,\cdots,x_j,v_j,A_j,\cdots) = f^N(t,\cdots,x_j,v_j,A_j,\cdots,x_i,v_i,A_i,\cdots).$$

Based on the Liouville equation, we have

$$\partial_t f^N + \sum_{i=1}^N \mathrm{div}_{x_i}\left(\dot{x}_i f^N\right) + \sum_{i=1}^N \mathrm{div}_{v_i}\left(\dot{v}_i f^N\right) + \sum_{i=1}^N \mathrm{div}_{A_i}\left(\dot{A}_i f^N\right) = 0. \tag{B.1}$$

Let $f^{N:1}(x_1,v_1,A_1,t)$ denote the marginal distribution

$$f^{N:1}(t,x_1,v_1,A_1) = \int_{\Omega^{N-1}} f^N(t,x_1,v_1,A_1,x_2,v_2,A_2,\cdots,x_N,v_N,A_N)dx_-dv_-dA_-,$$

where

$$\Omega = \mathbb{R}^{2d}\times\mathrm{SO}(d), (x_-,v_-,A_-) = (x_2,v_2,A_2,\cdots,x_N,v_N.A_N).$$

We integrate (B.1) with respect to variables x_-,v_-,A_- to get

$$\begin{aligned}
&\partial_t f^{N:1} + I_1 + I_2 + I_3 = 0,\\
&I_1 = \sum_{i=1}^N \int_{\Omega^{N-1}} \mathrm{div}_{x_i}\left(v_i f^N\right)dx_-dv_-dA_- = v_1\cdot\nabla_{x_i} f^{N:1},\\
&I_2 = \frac{k_1}{N}\sum_{i=1}^N \int_{\Omega^{N-1}} \mathrm{div}_{v_i}\left(\left(\sum_{j=1}^N \varphi_{ij}(v_j-v_i) - \sum_{j=1}^N \varphi_{ij}\dot{A}_i A_i^{-1}(x_j-x_i)\right)f^N\right)dx_-dv_-dA_-\\
&\quad = I_{21} - I_{22}.
\end{aligned}$$

By the symmetry of f^N

$$\begin{aligned}
I_{21} &= \frac{k_1}{N}\sum_{i=1}^N \int_{\Omega^{N-1}} \mathrm{div}_{v_i}\left(\left(\sum_{j=1}^N \varphi_{ij}(v_j-v_i)f^N\right)\right)dx_-dv_-dA_-\\
&= \frac{k_1}{N}\int_{\Omega^{N-1}} \mathrm{div}_{v_1}\left(\left(\sum_{j=1}^N \varphi_{ij}(v_j-v_1)f^N\right)\right)dx_-dv_-dA_-\\
&= \frac{k_1(N-1)}{N}\int_{\Omega^{N-1}} \mathrm{div}_{v_i}\Big(\left(\varphi_{12}(v_2-v_1)f^N\right)\Big)dx_-dv_-dA_-\\
&= \frac{k_1(N-1)}{N}\mathrm{div}_{v_i}\left(\int_\Omega \varphi_{12}(v_2-v_1)f^{N:2}dx_2dv_2dA_2\right),
\end{aligned}$$

where

$$f^{N:2}(t,x_1,v_1,A_1,x_2,v_2,A_2)=\int_{\Omega^{N-2}} f^N dx_3dv_3dA_3\cdots dx_Ndv_NdA_N,$$

and

$$\begin{aligned}I_{22}&=\frac{k_1}{N}\sum_{i=1}^N\int_{\Omega^{N-1}}\mathrm{div}_{v_i}\left(\left(\sum_{j=1}^N\varphi_{ij}\dot{A}_iA_i^{-1}(x_j-x_i)\right)f^N\right)dx_-dv_-dA_-\\&=\frac{k_1}{N}\int_{\Omega^{N-1}}\mathrm{div}_{v_1}\left(\sum_{j=1}^N\varphi_{1j}\dot{A}_1A_1^{-1}(x_j-x_1)f^N\right)dx_-dv_-dA_-\\&=\frac{k_1(N-1)}{N}\int_{\Omega^{N-1}}\mathrm{div}_{v_1}\left(\varphi_{12}\dot{A}_1A_1^{-1}(x_2-x_1)f^N\right)dx_-dv_-dA_-\\&=\frac{k_1(N-1)}{N}\nabla_{v_1}\cdot\left(\int_\Omega\varphi_{12}\dot{A}_1A_1^{-1}(x_2-x_1)f^{N:2}dx_2dv_2dA_2\right).\end{aligned}$$

Because

$$\dot{A}_1A_1^{-1}=\frac{k_2}{N}\sum_{j=1}^N\psi_{1j}\left(A_jA_1^{-1}-A_1A_j^{-1}\right),$$

by the symmetry of f^N, we have

$$\begin{aligned}I_{22}&=\frac{k_1k_2(N-1)}{N^2}\mathrm{div}_{v_1}\left(\int_\Omega\varphi_{12}\sum_{j=1}^N\psi_{1j}\left(A_jA_1^{-1}-A_1A_j^{-1}\right)(x_2-x_1)f^{N:2}dx_2dv_2dA_2\right)\\&=\frac{k_1k_2(N-1)^2}{N^2}\mathrm{div}_{v_1}\left(\int_\Omega\varphi_{12}\psi_{12}\left(A_2A_1^{-1}-A_1A_2^{-1}\right)(x_2-x_1)f^{N:2}dx_2dv_2dA_2\right).\end{aligned}$$

Then we have

$$\begin{aligned}I_2=&\left(k_1-\frac{k_1}{N}\right)\nabla_{v_1}\cdot\left(\int_\Omega\varphi_{12}(v_2-v_1)f^{N:2}dx_2dv_2dA_2\right)\\&-\left(k_1-\frac{k_1}{N}\right)\nabla_{v_1}\cdot\left(\int_\Omega\left(k_2-\frac{k_2}{N}\right)\varphi_{12}\psi_{12}\left(A_2A_1^{-1}-A_1A_2^{-1}\right)(x_2-x_1)f^{N:2}dx_2dv_2dA_2\right).\end{aligned}$$

By divergence theorem on $\mathrm{SO}(d)$, we have

$$\int_{\mathrm{SO}(d)}\mathrm{div}\left(\dot{A}_if^N\right)dA_i=0,\quad \int_{\Omega^{N-1}}\mathrm{div}\left(\dot{A}_if^N\right)dx_-dv_-dA_-=0,\quad i=2,\ldots,N.$$

Hence,

$$\begin{aligned}I_3&=\sum_{i=1}^N\int_{\Omega^{N-1}}\mathrm{div}_{A_i}\left(\dot{A}_if^N\right)dx_-dv_-dA_-\\&=\int_{\Omega^{N-1}}\mathrm{div}_{A_1}\left(\dot{A}_1f^N\right)dx_-dv_-dA_-\end{aligned}$$

$$
\begin{aligned}
&=\frac{k_2}{N}\int_{\Omega^{N-1}}\mathrm{div}_{A_1}\left(\sum_{j=1}^{N}\psi_{1j}\big(A_jA_1^{-1}-A_1A_j^{-1}\big)A_1f^N\right)dx_-dv_-dA_-\\
&=\frac{k_2(N-1)}{N}\int_{\Omega^{N-1}}\mathrm{div}_{A_1}\big(\psi_{12}\big(A_2-A_1A_2^{-1}A_1\big)f^N\big)dx_-dv_-dA_-\\
&=\frac{k_2(N-1)}{N}\int_{\Omega}\mathrm{div}_{A_1}\big(\psi_{12}\big(A_2-A_1A_2^{-1}A_1\big)f^{N/2}\big)dx_2dv_2dA_2.
\end{aligned}
$$

Then we have

$$
\begin{aligned}
0=&\partial_t f^{N:1}+v_1\cdot\nabla_{x_i}f^{N:1}+\left(k_1-\frac{k_1}{N}\right)\nabla_{v_1}\cdot\left(\int_\Omega\varphi_{12}(v_2-v_1)f^{N:2}dx_2dv_2dA_2\right)\\
&-\left(k_1-\frac{k_1}{N}\right)\nabla_{v_1}\cdot\left(\int_\Omega\left(k_2-\frac{k_2}{N}\right)\varphi_{12}\psi_{12}\big(A_2A_1^{-1}-A_1A_2^{-1}\big)(x_2-x_1)f^{N:2}dx_2dv_2dA_2\right)\\
&+\frac{k_2(N-1)}{N}\nabla_{A_1}\cdot\left(\int_\Omega\psi_{12}\big(A_2-A_1A_2^{-1}A_1\big)f^{N:2}dx_2dv_2dA_2\right).
\end{aligned}
$$

Now we take the mean-field limit $N\to\infty$ and obtain the one- and two-particle limiting densities

$$
\begin{aligned}
f^1&=\lim_{N\to\infty}f^{N:1}(t,x_1,v_1,A_1),\\
f^2&=\lim_{N\to\infty}f^{N:2}(t,x_1,v_1,A_1,x_2,v_2,A_2),
\end{aligned}
$$

which satisfy

$$
\begin{aligned}
0=&\partial_t f^1+v_1\cdot\nabla_{x_i}f^1\\
&+k_1\nabla_{v_1}\cdot\left(\int_\Omega\Big[\varphi_{12}(v_2-v_1)-k_2\varphi_{12}\psi_{12}\big(A_2A_1^{-1}-A_1A_2^{-1}\big)(x_2-x_1)\Big]f^2dx_2dv_2dA_2\right)\\
&+k_2\nabla_{A_1}\cdot\left(\int_\Omega\psi_{12}(A_2-A_1A_2^{-1}A_1)f^2dx_2dv_2dA_2\right).
\end{aligned}
$$

We make the molecular chaos assumption that

$$
f^2(t,x_1,v_1,A_1,x_2,v_2,A_2)=f^1(t,x_1,v_1,A_1)f^1(t,x_2,v_2,A_2)
$$

to close the above equation. Then we can obtain one-particle distribution function f satisfies the following equation:

$$
\begin{aligned}
&\partial_t f+v_1\cdot\nabla_{x_i}f+\nabla_v\cdot(L[f]f)+\nabla_A\cdot(Q[f]f)=0,\\
&L[f](t,x,v,A)=k_1\int_\Omega\varphi(|x-y|)\Big[(v_*-v)-k_2\psi(|x-y|)(A_*A^{-1}-AA_*^{-1})(y-x)\Big]\\
&\qquad\qquad\qquad\times f(t,y,v_*,A_*)dydv_*dA_*,\\
&Q[f](t,x,v,A)=k_2\int_\Omega\psi(|x-y|)\big(A_*-AA_*^{-1}A\big)f(t,y,v_*,A_*)dydv_*dA_*.
\end{aligned}
$$

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (Grant No. 12371180), by the Natural Science Foundation of Hunan Province (Grant No. 2022JJ30655) and by the Hunan Provincial Graduate Student Innovation Program (Grant No. CX20230004).

References

[1] G. Albi, N. Bellomo, L. Fermo, S.-Y. Ha, J. Kim, L. Pareschi, D. Poyato, and J. Soler, *Vehicular traffic, crowds, and swarms: From kinetic theory and multiscale methods to applications and research perspectives*, Math. Models Methods Appl. Sci., 29(10):1901–2005, 2019.

[2] V. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, 1978.

[3] J. Cañizo, J. Carrillo, and J. Rosado, *A well-posedness theory in measures for some kinetic models of collective motion*, Math. Models Methods Appl. Sci., 21(3):515–539, 2011.

[4] J. Carrillo, M. Fornasier, J. Rosado, and G. Toscani, *Asymptotic flocking dynamics for the kinetic Cucker-Smale model*, SIAM J. Math. Anal., 42(1):218–236, 2010.

[5] Z. L. Chen and X. X. Yin, *The kinetic Cucker-Smale model: Well-posedness and asymptotic behavior*, SIAM J. Math. Anal., 51(5):3819–3853, 2019.

[6] F. Cucker and S. Smale, *Emergent behavior in flocks*, IEEE Trans. Automat. Contr., 52(5):852–862, 2007.

[7] P. Degond, A. Diez, and A. Frouvelle, *Body-attitude coordination in arbitrary dimension*, arXiv: 2111.05614, 2021.

[8] P. Degond, A. Diez, and M.-Y. Na, *Bulk topological states in a new collective dynamics model*, SIAM J. Appl. Dyn. Syst., 21(2):1455–1494, 2022.

[9] P. Degond, A. Frouvelle, and S. Merino-Aceituno, *A new flocking model through body attitude coordination*, Math. Models Methods Appl. Sci., 27(6):1005–1049, 2017.

[10] J. G. Dong and X. P. Xue, *Synchronization analysis of Kuramoto oscillators*, Commun. Math. Sci., 11(2):465–480, 2013.

[11] F. Dörfler and F. Bullo, *Synchronization in complex networks of phase oscillators: A survey*, Automatica, 50(6):1539–1564, 2014.

[12] R. C. Fetecau, S. Y. Ha, and H. Park, *An intrinsic aggregation model on the special orthogonal group $SO(3)$: Well-posedness and collective behaviours*, J. Nonlinear Sci., 31(5):74, 2021.

[13] R. C. Fetecau, S. Y. Ha, and H. Park, *Emergent behaviors of rotation matrix flocks*, SIAM J. Appl. Dyn. Syst., 21(2):1382–1425, 2022.

[14] F. Golse and S. Y. Ha, *A mean-field limit of the Lohe matrix model and emergent dynamics*, Arch. Ration. Mech. Anal., 234:1445–1491, 2019.

[15] S. Y. Ha, D. Kim, J. Lee, and N. S. Eun, *Emergent dynamics of an orientation flocking model for multi-agent system*, Discrete Contin. Dyn. Syst., 40(4):2037–2060, 2020.

[16] S. Y. Ha, J. Kim, and X. T. Zhang, *Uniform stability of the Cucker-Smale model and its application to the mean-field limit*, Kinet. Relat. Models, 11:1157–1181, 2018.

[17] S. Y. Ha, D. Ko, and S. W. Ryoo, *Emergent dynamics of a generalized Lohe model on some class of Lie groups*, J. Stat. Phys., 168:171–207, 2017.

[18] S. Y. Ha and J. G. Liu, *A simple proof of the Cucker-Smale flocking dynamics and mean-field limit*, Commun. Math. Sci., 7(2):297–325, 2009.

[19] S. Y. Ha and S. W. Ryoo, *Asymptotic phase-locking dynamics and critical coupling strength for the Kuramoto model*, Comm. Math. Phys., 377:811–857, 2020.

[20] S. Y. Ha and E. Tadmor, *From particle to kinetic and hydrodynamic descriptions of flocking*, Kinet. Relat. Models, 1(3):415–435, 2008.

[21] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, Springer, 1984.

[22] J. R. Lawton and R. W. Beard, *Synchronized multiple spacecraft rotations*, Automatica, 38(8): 1359–1364, 2002.

[23] Z. C. Li and X. P. Xue, *Cucker-Smale flocking under rooted leadership with fixed and switching topologies*, SIAM J. Appl. Math., 70(8):3156–3174, 2010.

[24] M. A. Lohe, *Non-Abelian Kuramoto models and synchronization*, J. Phys. A: Math. Theor., 42(39):395101, 2009.

[25] S. Motsch and E. Tadmor, *Heterophilious dynamics enhances consensus*, SIAM Rev., 56(4):577–621, 2014.

[26] Z. Y. Qiao, Y. C. Liu, and X. Wang, *Multi-cluster flocking behavior analysis for a delayed Cucker-Smale model with short-range communication weight*, J. Syst. Sci. Complex., 35:137–158, 2022.

[27] H. Spohn, *Large Scale Dynamics of Interacting Particles*, Springer, 2012.

[28] R. Tron, R. Vidal, and A. Terzis, *Distributed pose averaging in camera networks via consensus on SE(3)*, in: 2008 Second ACM/IEEE International Conference on Distributed Smart Cameras, 1–10, 2008.

[29] T. Vicsek, A. Czirók, E. B. Jacob, I. Cohen, and O. Shochet, *Novel type of phase transition in a system of self-driven particles*, Phys. Rev. Lett., 75(6):1226, 1995.

[30] C. Villani, *Optimal Transport: Old and New*, Springer, 2009.

[31] X. Y. Wang and X. P. Xue, *Formation behaviour of the kinetic Cucker-Smale model with non-compact support*, P. Roy. Soc. Edinb. A, 153(4):1315–1346, 2022.