A PROBABILISTIC APPROACH FOR THE MEAN-FIELD LIMIT TO THE CUCKER-SMALE MODEL WITH A SINGULAR COMMUNICATION

SEUNG-YEAL HA
Department of Mathematical Sciences and Research Institute of Mathematics
Seoul National University
Seoul 08826, Korea
and
Korea Institute for Advanced Study
Hoegiro 87, Seoul 02455, Korea

JEONGHO KIM
Department of Mathematical Sciences
Seoul National University
Seoul 08826, Korea

PETER PICKL
Mathematisches Institut
Ludwig-Maximilians-Universität
Theresienstr. 39, Munich 80333, Germany

XIONGTAO ZHANG
Center for Mathematical Sciences
Huazhong University of Science and Technology
1037 Luoyu Road, Wuhan 430074, China

(Communicated by Changjiang Zhu)

ABSTRACT. We present a probabilistic approach for derivation of the kinetic Cucker-Smale (C-S) equation from the particle C-S model with singular communication. For the system we are considering, it is impossible to validate effective description for certain special initial data, thus such a probabilistic approach is the best one can hope for. More precisely, we consider a system in which kinetic trajectories are deviated from a microscopic model and use a suitable probability measure to quantify the randomness in the initial data. We show that the set of “bad initial data” does in fact have small measure and that this small probability decays to zero algebraically, as $N$ tends to infinity. For this, we introduce an appropriate cut-off in the communication weight. We also provide a relation between the order of the singularity and the order of the cut-off such that the machinery for deriving classical mean-field limits introduced in [3] can be applied to our setting.

2010 Mathematics Subject Classification. Primary: 70F10, 35Q70.
Key words and phrases. Flocking, the Cucker-Smale model, mean-field limit, singular interaction.

* Corresponding author.
1. **Introduction.** Collective behaviors of classical and quantum many-body systems are ubiquitous in nature. Typical examples include the flocking of birds, schooling of fishes and herding of sheep, to name a few. The notion “flocking” denotes the phenomenon in which disordered states are organized into an ordered, coherent motion. The mathematical modeling of such flocking dynamics has been initiated by Vicsek and his collaborators in [41] and has boosted further studies on the collective dynamics of many-body systems in control theory and statistical physics communities. After Vicsek’s seminal work on flocking, several mathematical models have been studied. Among those, our main interest lies on the Cucker-Smale model which was motivated by the Vicsek’s model. Denote the position and velocity of the \( i \)-th C-S particle in \( \mathbb{R}^d \) by \( x_i \) and \( v_i \) respectively. Then the C-S dynamics is given by the following second-order ODE system:

\[
\frac{dx_i}{dt} = v_i, \quad t > 0, \quad i = 1, 2, \ldots, N,
\]

\[
\frac{dv_i}{dt} = \frac{\kappa}{N-1} \sum_{j \neq i} \omega(|x_i - x_j|)(v_j - v_i),
\]

where \(|\cdot|\) is the standard \( \ell^2 \)-norm in \( \mathbb{R}^d \), \( \kappa \) and \( \omega \) denote a positive coupling strength and communication weight which is radially symmetric and non-increasing in its argument:

\[
\omega(r) \geq 0 \quad \text{and} \quad (\omega(r_1) - \omega(r_2))(r_1 - r_2) \leq 0, \quad r_1, r_2 \in \mathbb{R}_+ \cup \{0\}. \quad (2)
\]

For regular and bounded communication weights, the C-S model (1)–(2) has been extensively studied in the literature, among the features that have been studied are collision avoidance [1, 14], random effects [2, 15, 24], mean-field limit and measure-valued solutions [4, 7], collective dynamics [8, 9, 10, 16, 17, 18, 19, 21, 22, 25, 26, 27, 40], local flocking [11, 12, 23], bonding force [30, 35], generalized flocking [33], singular and hyperbolic limits [39], kinetic equation [20, 25, 27], application to flight formation [36] and flocking with leaders [32] (see a recent survey article on the C-S model [13] and references therein). In contrast, there are very few manuscripts dealing with the C-S model with singular communication weights, however, flocking dynamics [25], collision avoidance [1, 6], global existence of weak and measure-valued solutions for particle and kinetic C-S models [5, 34, 37, 38] have been studied so far. When the number of particles is significantly large, it is well-known [25] that the corresponding mean-field system (Vlasov-type equation) can describe the effective dynamics of a many-body system with \( N \gg 1 \) for a bounded and Lipschitz communication weight. Recently, there has been much activity [3, 28, 29, 31] around the rigorous justification of the Vlasov-Poisson equation from Newtonian dynamics with singular forces. The main goal of this paper is to extend such results to the case of the C-S model. More precisely, we set \( f = f(x, v, t) \) to be the probability density function of finding particle with position \( x \) and velocity \( v \) for the ensemble of C-S particles. Then, assuming molecular chaos and using the standard BBGKY hierarchy-argument as in [27], the mean-field density function \( f \) is expected to satisfy

\[
\frac{\partial_t f}{f} + \mathbf{v} \cdot \nabla_x f + \nabla_v \cdot (F[f]f) = 0, \quad x, v \in \mathbb{R}^d, \quad t > 0,
\]

\[
F[f](x, v, t) := -(\eta * f)(x, v, t),
\]

where \( \eta(x, v) := \omega(|x|)v \). Throughout the paper, we also call \( \eta \) as a communication weight as well as \( \omega \) itself if there is no confusion.
In this paper, we are interested in the rigorous justification of mean-field limit of (1)–(2) to (3) where communication weight is singular. As discussed in [25], the mean-field limit for the bounded and Lipschitz communication weight has been established using the empirical measure and particle-in-cell (PIC) method. However, the particle-in-cell method cannot be applied to the model with singular communication weight. The reason is that the interaction of particles in the same position is not well defined. Thus the discretization process fails. Another obstacle in applying the traditional mean-field limit technique is that it is hard to control the divergence of the vector field. The divergence of the vector field generated by (3) is

\[ \nabla \cdot \left( v, F[f] \right) = -d \int_{\mathbb{R}^{2d}} \omega(||x - x_*||) f(x_*, v_*, t) \, dx_* \, dv_* \]

which is hard to estimate when \( \omega \) is singular. In fact, this divergence term is also cumbersome when it comes to global well-posedness of kinetic equation (3) when communication weight is singular. Therefore, as far as the authors know, for a singular communication weight, there is no previous work on the rigorous mean-field limit for the Cucker-Smale model. For a rigorous proof, however, one needs to show that the assumption of molecular chaos is in fact valid, at least in a suitable probabilistic sense.

In this paper, we adopt a probabilistic approach to deal with the mean-field limit, which is mainly motivated by the technique in [3]. One might think that verification of such a probabilistic result is a weakness of our method and, with more technical effort, one can find a statement which holds for all initial data which approximate the initial density in a weak sense. However, for the singular communication weight case under consideration, it can be shown that such a general statement is in fact wrong, and validity of the effective description for typical initial data, i.e., a set of initial data with probability close to one, is the best one that we can hope for. For a repulsive Newtonian model, an example of an initial condition which shows deviation from the effective descriptions can be found in the following heuristic example: Consider a Newtonian system with repulsive interactions and an initial state made of a large number (\( \varepsilon^{-1} \) for some small \( \varepsilon \)) of clusters, i.e., each cluster consists of \( \varepsilon N \) particles. Now let us distribute the clusters such that their density converges weakly to the initial density of the effective description and think of the limit \( \varepsilon \to 0 \) and \( N \to \infty \). The potential energy of the clusters will be enormous due to the large number in the cluster and the singularity in the interaction. In fact, the potential energy per particle goes to infinity in this limit. This potential energy will be transferred to kinetic energy in a short amount of time, thus the clusters will explode and most particles will propagate away with a speed tending to infinity in the given limit. Thus the phase space density of the cloud will not converge and thus deviate from the effective description. With similar arguments, one can also find initial data which break away from the effective description under time evolution for attractive systems and/or systems with velocity dependent force. Therefore, as we mentioned above, it is impossible to attain a mean-field limit result for all initial data, when we consider a singular interaction. Instead, one may obtain a singular mean-field limit by restricting the set of initial data. As far as the authors know, there are two approaches for a singular mean-field limit: the deterministic approach and the probabilistic approach. The deterministic approach provides the conditions for initial data under which the mean-field limit is valid, in contrast, the probabilistic approach provides estimates on the upper bound of the size of initial
data that the mean-field limit does not hold. Thus, the probabilistic approach has an advantage of quantifying the set of “bad” initial data over the deterministic one. Note that, for the kinetic C-S equation (3) with regular $\omega$, there have been numerous studies on the kinetic C-S equation. So far, there is no global well-posedness for the kinetic C-S equation (3) with singular interactions. We shall consider a $N$-dependent communication weight $\eta_N$, with $\eta := \sup N \eta_N$ is singular at $x = 0$, in other words, a singular interaction with a $N$-dependent cut-off. The precise definition of $\eta_N$ can be found in Section 2. To fix the idea, we consider an algebraically decaying singular communication weight $\omega$:

$$\omega(r) = \frac{1}{r^\alpha}, \quad \text{for} \quad r > 0, \quad \alpha > 0. \quad (4)$$

Although we concentrate on the specific case where $\omega(r) := r^{\beta}$, we note that our analysis does not strongly depend on the form of communication weight.

Now, let us briefly explain the probabilistic estimate on the discrepancy between the particle and the kinetic flow we shall establish: let $\Psi_{t_0}(Z)$ and $\Phi_{t_0}(Z)$ be the microscopic flow and kinetic flow for the C-S model (1) and the kinetic equation (3) with initial data $Z$ at time $t$ using $N$-dependent cut-off function, respectively (For more precise definition, we refer Definition 2.1). Then, under conditions on the parameters $\alpha, \beta, \gamma$ and $\delta$ appearing in (4), (5) and (6), our main result can be summarized as follows. For any finite time $T \in (0, \infty)$,

$$\mathbb{P}_0 \left\{ \left| \Psi_{t_0}(Z) - \Phi_{t_0}(Z) \right|_{\infty} > \frac{1}{N^\gamma} \right\} \leq \frac{\tilde{C}_T}{N^{\delta - \gamma}}, \quad (5)$$

where $\mathbb{P}_0$ is the probability measure with density function being a $N$-fold product of initial data $f_0$ (see notation below).

The rest of the paper is organized as follows: In Section 2, we formulate our main result on the closeness of the $N$-body description to the effective mean-field description, which is the main result of this paper. In Section 3, we provide basic properties of the particle C-S model and kinetic C-S equation. In Section 4, we present the detailed proof of Theorem 2.2 following the analytical technique in [3]. Finally, Section 5 is devoted to a summary of this paper and discussion on future direction.

**Gallery of notation.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For any random variable $X$ and a measurable subset $A \in \mathcal{F}$, we define $\mathbb{E}^A(X)$ to be the expectation of $X$ restricted on the set $A$:

$$\mathbb{E}^A(X) := \mathbb{E}(X 1_A),$$

where $1_A$ is the characteristic function on the set $A$. Throughout the paper, $|\cdot|$ denotes the Euclidean $l_2$-norm in $\mathbb{R}^k$ and $|\cdot|_{\infty}$ denotes the supremum norm on $\mathbb{R}^k$. We denote the position and velocity vectors as follows:

$$X := (x_1, \ldots, x_N) \in \mathbb{R}^{Nd}, \quad V := (v_1, \ldots, v_N) \in \mathbb{R}^{Nd},$$

$$z_i := (x_i, v_i), \quad Z := (z_1, \ldots, z_N) \in \mathbb{R}^{2Nd}.$$  

Moreover, we write $f_t(\cdot) = f(\cdot, t) : \mathbb{R}^{2Nd} \rightarrow \mathbb{R}_+ \cup \{0\}$ for the one-particle probability density function solving (3) subjected to initial data $f_0$. All probabilities and expectation values at time $t$ are calculated using the product measure $\Pi_{j=1}^N f(z_j, t)$ generated by $f_t$: for any random variable $H : \mathbb{R}^{2Nd} \rightarrow \mathbb{R}$ and any element $A$ of the
Borel $\sigma$-algebra, we set
\[ \mathbb{P}_t(H \in A) := \int_{H^{-1}(A)} \prod_{j=1}^N f(z_j, t) dZ, \quad \mathbb{E}_t(H) := \int_{\mathbb{R}^{2Nd}} H(Z) \prod_{j=1}^N f(z_j, t) dZ. \]
We also use the abbreviated notation $f = f(z, t)$ and $f_* = f(z_*, t)$, when there is no confusion. For any function $f : \mathbb{R}^k \to \mathbb{R}$, we use the notation
\[ \|f\|_{1\vee\infty} := \inf_{f_1 + f_\infty = f} \{ \|f_1\|_1 + \|f_\infty\|_\infty \}. \]
Throughout the paper, $C$ denotes a generic positive constant, which can vary from line to line.

2. **Statement of the main results.** We now state our main result which gives a comparison between microscopic and mesoscopic flows. For this, we will first define the respective flows in the $N$-particle phase space as follows. Since the coupling strength $\kappa$ plays the role of the decay exponent in the flocking estimate, we will set it to be one without loss of generality.

For a given $N \geq 2$, we introduce an "communication weight with cut-off" $\omega_N(r)$ and $\eta_N(x, v)$ defined as follows
\[ \omega_N(r) := \begin{cases} \frac{1}{r^\alpha}, & \text{if } r \geq \frac{1}{N^\beta}, \\ N^\alpha \beta, & \text{if } r < \frac{1}{N^\beta}, \end{cases} \quad \eta_N(x, v) = \omega_N(|x|) v. \quad (6) \]

Given this communication weight with cut-off, we consider the modified particle and kinetic C-S models:
\[
\begin{align*}
\frac{dx_i}{dt} &= v_i, \quad t > 0, \quad i = 1, 2, \ldots, N, \\
\frac{dv_i}{dt} &= \frac{1}{N-1} \sum_{j \neq i} \eta^N(x_i - x_j, v_j - v_i),
\end{align*}
\]
and
\[
\partial_t f^N + v \cdot \nabla_x f^N + \nabla_v \cdot [\eta^N * f^N] f^N = 0. \quad (8)
\]
Let $Z := (X, V) := (x_1, \ldots, x_N, v_1, \ldots, v_N)$ be the solution of the system (7) where the corresponding force $F : \mathbb{R}^{2Nd} \to \mathbb{R}^{Nd}$ is defined by
\[ F(X, V)_i := \frac{1}{N-1} \sum_{j \neq i} \eta^N(x_i - x_j, v_j - v_i), \quad i = 1, \ldots, N. \]

We are now able to define the microscopic flow on the $N$-particle space.

**Definition 2.1.** (1) We introduce the microscopic C-S flow $\Psi_{t,s}^N(Z) = Z_t = (X_t, V_t)$ on the phase space $\mathbb{R}^{2Nd}$ which is generated by the vector field $(V, F(X, V))$:
\[
\begin{align*}
\frac{d}{dt} \Psi_{t,s}^N(Z) &= U(\Psi_{t,s}^N(Z)), \quad t > s, \\
\Psi_{s,s} = Z, \quad t = s,
\end{align*}
\]
where $U(\Psi_{t,s}^N(Z)) := (V_t, F(X_t, V_t))$ is a vector field on the $N$-particle phase space $\mathbb{R}^{Nd} \times \mathbb{R}^{Nd}$. 


(2) For every probability density \( f_0 : \mathbb{R}^{2d} \to \mathbb{R}_+ \cup \{0\} \) and \( z = (x, v) \in \mathbb{R}^{2d} \), we
deﬁne the bi-characteristics \( \phi^N_{t,s}(z) = z_t = (x_t, v_t) \) for (8) as
\[
\begin{aligned}
&\frac{d}{dt} \phi^N_{t,s}(z) = u(\phi^N_{t,s}(z)), \quad t > s, \\
&\phi^N_{s,s}(z) = (x, v), \quad t = s,
\end{aligned}
\]
where \( u \) is given by
\[
u(\phi^N_{t,s}(z)) := \left( v_t, \int_{\mathbb{R}^{2d}} \eta^N(x_t - x_s, v_s - v_t)f^N(z_s, t)dz_s \right)
= \left( v_t, -\left( \eta^N * f^N \right)(x_t, v_t) \right).
\]

Now we can construct the time-dependent probability density \( f^N : \mathbb{R}^{2d} \times (\mathbb{R}_+ \cup \{0\}) \to \mathbb{R}_+ \cup \{0\} \) for finding a particle at the phase space point \( z \) at
time \( t \):
\[
f^N(z, t) = f_0(\phi^N_{0,t}(z)) \exp \left[ dt \int_0^t \left( \int_{\mathbb{R}^{2d}} \omega^N(|x - x_s|)f^N(z_s, s)dz_s \right) ds \right].
\]

(3) This mean-ﬁeld ﬂow can be lifted to the \( N \)-particle phase space straightforwardly and deﬁne the
following effective C-S ﬂow:
\[
\frac{d}{dt} \Phi^N_{t,s}(Z) = \bar{U}_t(\Phi^N_{t,s}(Z)),
\]
where \( \bar{U}_t(\Phi^N_{t,s}(Z)) = (V_t, \bar{F}_t(X_t, V_t)) \) and \( \bar{F}_t(X, V); := -\left( \eta^N * f^N \right)(x, v, t). \)
Equivalently, we may write it as
\[
\Phi^N_{t,s}(Z) = (\phi^N_{t,s}(z_1), \cdots, \phi^N_{t,s}(z_N)).
\]

Note that the existence of the time-dependent probability density function \( f^N(z, t) \) can be
guaranteed since we use cut-off communication weight, which is Lipschitz.

Having deﬁned the (microscopic) C-S ﬂow and the corresponding effective ﬂow, we can present at our main result in the following theorem.

**Theorem 2.2.** For a given \( T \in (0, \infty) \), suppose that the positive parameters \( \alpha, \beta, \gamma \)
and \( \delta \) satisfy the following relations:
\[
\alpha \beta + \delta < \frac{1}{2}, \quad \beta < \gamma < \delta, \quad \alpha < d - 1.
\]
Then, for any solution \( f^N = f^N(x, v, t) \) to (7) satisfying the following condition
\[
\exists \ C_T < \infty \text{ such that } \sup_{0 \leq t \leq T} \| \rho^N(t) \|_{\infty} \leq C_T, \text{ where } \rho^N(x, t) := \int_{\mathbb{R}^d} f^N(x, v, t)dv,
\]
there exists a constant \( \tilde{C}_T < \infty \) such that
\[
P_0 \left( \left\{ Z \in \mathbb{R}^{2Nd} : \sup_{0 \leq t \leq T} \| \Psi^N_{t,0}(Z) - \Phi^N_{t,0}(Z) \|_{\infty} > \frac{1}{N^\gamma} \right\} \right) \leq \frac{\tilde{C}_T}{N^{\delta - \gamma}}.
\]

**Remark 1.** 1. Similar to [31], one can prove that the bound on this probability
is almost exponential in the sense that it is smaller than any polynomial, i.e.,
one can show that for any \( \mu > 0 \) there is a constant \( \tilde{C}_{T, \mu} \) such that
\[
P_0 \left( \left\{ Z \in \mathbb{R}^{2Nd} : \sup_{0 \leq t \leq T} \| \Psi^N_{t,0}(Z) - \Phi^N_{t,0}(Z) \|_{\infty} > \frac{1}{N^\gamma} \right\} \right) \leq \frac{\tilde{C}_{T, \mu}}{N^{\mu(\delta - \gamma)}}.
\]
Since the argument is worked out in [31] in details and messes up the notation, we prove here the weaker result as stated in Theorem 2.2, which is enough to prove mean-field limit.

2. When the communication weight $\omega$ is regular, we can obtain a similar result to Theorem 2.2, without introducing any cut-off. We discuss this at the end of Section 4.

3. Preliminaries. In this section, we provide a theoretical background for the particle C-S model and mean-field kinetic C-S equation.

3.1. The Cucker-Smale model. Let us establish several properties for the C-S model to be used in later sections. For any velocity configuration $V = (v_1, \cdots, v_N)$, we set

$$|V|_{2,\infty} := \max_{1 \leq i \leq N} |v_i|.$$

**Lemma 3.1.** For a given $T \in (0, \infty)$, let $(X,V)$ be a smooth solution to the C-S model (1) with initial data $(X^0, V^0)$ in the time interval $[0,T)$. Then, the following assertions hold.

1. The total momentum is a constant of motion:

$$\sum_{i=1}^{N} v_i(t) = \sum_{i=1}^{N} v_i^0, \quad t \in [0,T).$$

2. The total kinetic energy is non-increasing:

$$\sum_{i=1}^{N} |v_i(t)|^2 + \frac{1}{N-1} \sum_{i \neq j} \int_{0}^{t} \omega(|x_i(s) - x_j(s)|) |v_j(s) - v_i(s)|^2 ds = \sum_{i=1}^{N} |v_i^0|^2, \quad t \in [0,T).$$

3. If $(X,V)$ is a solution to cut-off system (7), then the mixed norm $|V|_{2,\infty}$ is contractive:

$$|V(t)|_{2,\infty} \leq |V(s)|_{2,\infty}, \quad \text{for} \quad s \leq t.$$  

**Proof.** 1. We use the standard index-exchanging trick and symmetry of $\psi$ to get

$$\frac{d}{dt} \sum_{i=1}^{N} v_i(t) = \frac{1}{N-1} \sum_{i \neq j} \omega(|x_i - x_j|) (v_j - v_i) = \frac{1}{N-1} \sum_{i \neq j} \omega(|x_i - x_j|) (v_i - v_j)$$

$$= - \frac{1}{N-1} \sum_{i \neq j} \omega(|x_i - x_j|) (v_j - v_i) = 0.$$

Thus, the total momentum is a constant of motion.

2. We take the inner product of the second line of (1) with $2v_i$, and then sum it over all $i$ to obtain

$$\frac{d}{dt} \sum_{i=1}^{N} |v_i|^2 = \frac{2}{N-1} \sum_{i \neq j} \omega(|x_i - x_j|) v_i \cdot (v_j - v_i)$$

$$= - \frac{2}{N-1} \sum_{i \neq j} \omega(|x_i - x_j|) v_j \cdot (v_j - v_i)$$

$$= - \frac{1}{N-1} \sum_{i \neq j} \omega(|x_i - x_j|) |v_j - v_i|^2.$$  

We integrate the relation above with respect to time to get the desired estimate.
3. Let $M$ be one of the indices of maximal velocity, i.e. for $t \geq 0$,

$$
|v_M(t)| := \max_{1 \leq i \leq N} |v_i(t)|, \quad \text{i.e., } |v_M(t)| \geq |v_j(t)| \quad \text{for all } j \neq M.
$$

Then, it is straightforward to see that

$$
\frac{1}{2} \frac{d|v_M|^2}{dt} = v_M \cdot \frac{dv_M}{dt} = \frac{1}{N-1} \sum_{j \neq M} \omega^N(|x_M - x_j|)(v_j - v_M) \cdot v_M \leq 0,
$$

where we used the following relation in the last inequality:

$$(v_j - v_M) \cdot v_M = v_j \cdot v_M - |v_M|^2 \leq |v_j||v_M| - |v_M|^2 \leq 0.$$

\[\square\]

**Remark 2.**

1. Although the statement of Lemma 3.1 is for communication weight without cut-off, the results also hold for communication weight with cut-off, since the estimate does not depend on the choice of communication weight.
2. In [6], it was shown that for a strong singular communication weight $\alpha \geq 1$ in (4), collisions between C-S particles do not occur for non-collisional initial configurations. Thus, the standard Cauchy-Lipschitz theory can be applied to yield a global classical solution. In the case of a long-range communication weight with $\alpha \in [0, 1]$, the mono-cluster flocking estimate can be obtained using the Lyapunov functional approach in [25] under the a priori assumption of a global classical solution.
3. The reason why we consider the cut-off system (7) for the third assertion of Lemma 3.1 is to guarantee the Lipschitz continuity of $v_M$.

### 3.2. The kinetic Cucker-Smale equation

In this subsection, we study some basic *a priori* estimates for the kinetic C-S model with a singular communication (4).

**Proposition 1.** For a given $T \in (0, \infty)$, let $f = f(x, v, t)$ be a regular classical solution to (3) with initial datum $f_0$ which decays sufficiently fast in $z := (x, v)$ at infinity. Then, the following assertions hold:

(i) $\frac{d}{dt} \int_{\mathbb{R}^d} f(z, t) \, dz = 0$, \quad $\frac{d}{dt} \int_{\mathbb{R}^d} vf(z, t) \, dz = 0$.

(ii) $\frac{d}{dt} \int_{\mathbb{R}^d} |v|^2 f(z, t) \, dz + \int_{\mathbb{R}^d} \omega(|x - x_\ast|)|v - v_\ast|^2 f \, dz \, dz_\ast = 0$.

**Proof.** The proof is based on the same idea as for the particle model in Lemma 3.1. Hence, we omit its details. \[\square\]

When the communication weight is Lipschitz continuous, the well-posedness of the kinetic C-S equation can be obtained straightforwardly [25, 27]. However, there are two existence results available for weak and measure-valued solutions to the kinetic C-S equation (3) - (4) with singular communication. Let us recall the concept of weak solution to (3). Since the definition of measure-valued solution with respect to Theorem 3.4 is quite sophisticated, we refer it to the original work [34].

**Definition 3.2.** [5] Let $L^1_+$ be a family of function $f \in L^1$ with $f \geq 0$ a.e., and let $\mathcal{P}_1$ be the probability measure with finite first moment. For a given $T \in (0, \infty)$, $f$
is a weak solution of (3) subjected to initial data \( f_0 \in L^1_+ \cap L^p \) in the time-interval \([0, T]\) if and only if the following relations hold.

\[(i) \quad f \in L^\infty([0, T]; (L^1_+ \cap L^p)(\mathbb{R}^{2d})) \cap C([0, T]; \mathcal{P}_1(\mathbb{R}^{2d})).\]

\[(ii) \quad \text{For all } \psi \in C_\infty^{\infty}(\mathbb{R}^{2d} \times [0, T]),
\int_{\mathbb{R}^{2d}} f(x, v, T)\psi(x, v, T)dvdx - \int_0^T \int_{\mathbb{R}^{2d}} f(\partial_t \psi + v \cdot \nabla_x \psi + F[f] \cdot \nabla_v \psi)dvdxdt
= \int_{\mathbb{R}^{2d}} f_0(x, v)\psi(x, v, 0)dvdx.\]

**Theorem 3.3.** (Existence of weak solution [5]) Suppose that the communication weight \( \omega \) and the initial datum \( f_0 \) satisfy the following conditions:

\[0 < \alpha < \left( \frac{p - 1}{p} \right) - 1, \quad f_0 \in (L^1_+ \cap L^p)(\mathbb{R}^{2d}) \cap \mathcal{P}_1(\mathbb{R}^{2d}), \quad \text{for some } p.\]

Then, there exists a \( T^* > 0 \) and a unique weak solution \( f \) to (3) on the time interval \([0, T^*] \) in the sense of Definition 3.2.

**Theorem 3.4.** [34] Suppose that the communication weight \( \omega \) and the initial datum \( f_0 \) satisfy the following conditions:

\[0 < \alpha < \frac{1}{2}, \quad f_0 \in \mathcal{M}_+(\mathbb{R}^{2d}), \quad \text{supp} f_0 \text{ is a compact subset of } \mathbb{R}^{2d}.\]

Then, at least one measure-valued solution \( f \) to the kinetic C-S equation (3) in the sense defined in [34] defined in the time-interval \([0, T]\).

**Remark 3.** Theorem 3.3 guarantees the existence of a unique weak solution of kinetic C-S equation (3) in \( L^p \) space, where \( p \) depends on the dimension and singularity of communication weight. However, this theorem guarantees the existence of a weak solution only locally in time. On the other hand, Theorem 3.4 gives us the existence of a measure-valued solution. In this case, we have a global-in-time existence result, but we cannot guarantee uniqueness. In fact, as the authors’ knowledge, the global existence of a unique solution to the kinetic C-S equation is still an open problem. Finally, we mention here that in [34], they provide the mean-field approach to prove the existence of a measure-valued solution. However, they did not mention about propagation of chaos, which compares the microscopic and effective flow and this motivated us to begin this work.

4. **Propagation of molecular chaos.** In this section, we will present a proof for Theorem 2.2 which states that the two different flows typically stay close as formulated in Theorem 2.2.

4.1. **Preparatory estimates.** We provide several estimates needed in proving Theorem 2.2. To estimate the difference between microscopic flow and kinetic flow, we introduce a stochastic process \( J_t \) depending on the initial data and time which measures the variation between microscopic flow and kinetic flow.

**Definition 4.1.** Let \( J : \mathbb{R}^{2Nd} \times \mathbb{R}_+ \to \mathbb{R} \) be the stochastic process given by

\[J_t(Z) = J(Z, t) := \min \left\{ 1, N^7 \sup_{0 \leq s \leq t} \| \Psi_{s,t}^N(Z) - \Phi_{s,t}^N(Z) \|_\infty \right\}.\]
In order to derive Grönwall inequality (12), we will split the time difference of flow and kinetic flow.

Note that we have the following relation:
\[
\mathbb{P}_0 \left( \left\{ Z \in \mathbb{R}^{2Nd} : \sup_{0 \leq t \leq T} \left| \Psi^N_{t,0}(Z) - \Phi^N_{t,0}(Z) \right|_\infty > \frac{1}{N^\gamma} \right\} \right) = \mathbb{P}_0(J_t = 1) \leq 1 \cdot \mathbb{P}_0(J_t = 1) + \mathbb{E}_0^{\{J_t \neq 1\}}(J_t) = \mathbb{E}_0(J_t).
\]

Thus, Theorem 2.2 can be directly obtained, if one can show the following inequality:
\[
\mathbb{E}_0(J_t) \leq CN^{-\delta + \gamma}.
\] (11)

To control the decay rate of \( \mathbb{E}_0(J_t) \), we will derive a Grönwall inequality:
\[
\frac{d}{dt} \mathbb{E}_0(J_t) \leq C\mathbb{E}_0(J_t) + CN^{-\delta + \gamma}, \quad t > 0 \quad \mathbb{E}_0(J_0) = 0.
\] (12)

Thus for any finite time interval \([0, T]\) where \( f^N \) exists, we have
\[
\mathbb{E}_0(J_t) \leq CN^{-\delta + \gamma}e^{CT} \leq CN^{-\delta + \gamma}e^{CT}, \quad t \in [0, T).
\]

In order to derive Grönwall inequality (12), we will split the time difference of \( \mathbb{E}_0(J_t) \) into three parts: First, we consider the initial configuration satisfying the relation:
\[
\sup_{0 \leq s \leq t} \left| \Psi^N_{s,0}(Z) - \Phi^N_{s,0}(Z) \right| > N^{-\gamma}.
\]

Note that on the remaining set, we have
\[
J_t = N^\gamma \sup_{0 \leq s \leq t} \left| \Psi^N_{s,0}(Z) - \Phi^N_{s,0}(Z) \right|_\infty.
\]

To estimate \( \left| \Psi^N_{s,0}(Z) - \Phi^N_{s,0}(Z) \right|_\infty \), we have to estimate the difference between the flows \( U_i \) and \( \tilde{U}_i \), which is at leading order – essentially given by the fluctuation of the forces:
\[
\left| F \left( \Psi^N_{s,0}(Z) \right) - \tilde{F} \left( \Phi^N_{s,0}(Z) \right) \right| \\
\leq \left| F \left( \Psi^N_{s,0}(Z) \right) - F \left( \Phi^N_{s,0}(Z) \right) \right| + \left| F \left( \Phi^N_{s,0}(Z) \right) - \tilde{F} \left( \Phi^N_{s,0}(Z) \right) \right|.
\] (13)

To estimate the first term of the right-hand side of (13), we need the following auxiliary function which controls the derivative of the interaction force \( \eta^N \).

**Definition 4.2.** Let \( \zeta = \zeta(x, v) \) be defined as
\[
\zeta(x, v) := \begin{cases} 
\left( \frac{x^{\alpha+1}}{|x|^{\alpha+1}} + \frac{x^\alpha}{|x|^\alpha} \right) \max \{1, |v|\}, & \text{if } |x| > 3N^{-\beta}, \\
(\alpha N^{(\alpha+1)\beta} + N^{\alpha\beta}) \max \{1, |v|\}, & \text{otherwise}.
\end{cases}
\]

Then, for such \( \zeta \), we define
\[
G(X, V)_i := \frac{1}{N-1} \sum_{j \neq i} \zeta(x_i - x_j, v_j - v_i),
\]
\[
(\tilde{G}_t(Z))_i := \int_{\mathbb{R}^d} \zeta(x_i - x_\ast, v_\ast - v_i) f^N(z_\ast, t) \, dz_\ast.
\]

Note that these two terms will be used to control the second term of the right-hand side of (13) (See Lemma 4.7).
Lemma 4.3. Suppose that the vectors $\delta_x, \delta_v \in \mathbb{R}^d$ satisfy
$$|\delta_x|, |\delta_v| < 2N^{-\beta}.$$ 
Then, we have
$$|\eta^N(x + \delta_x, v + \delta_v) - \eta^N(x, v)|_\infty \leq \zeta(x, v) \max\{|\delta_x|_\infty, |\delta_v|_\infty\}. \quad (14)$$

Proof. The proof consists of two steps.

- **Step A** (Estimate of $|\nabla_{(x,v)} \eta(x,v)|$): For $|x| > N^{-\beta}$, we use (6) to obtain
$$|\partial_x \eta^N| \leq \frac{\alpha |v|}{|x|^{\alpha+1}} \leq \alpha N^{(\alpha+1)\beta} \max\{1, |v|\}, \quad |\partial_v \eta^N| \leq \frac{1}{|x|^{\alpha}} \leq N^{\alpha \beta}.$$ 

On the other hand, for $|x| \leq N^{-\beta}$,
$$|\partial_x \eta^N| = 0, \quad |\partial_v \eta^N| \leq N^{\alpha \beta}.$$ 

Therefore, we have
$$|\nabla_{(x,v)} \eta^N(x, v)| \leq \alpha N^{(\alpha+1)\beta} \max\{1, |v|\} + N^{\alpha \beta}.$$ 

Note that $\omega^N$ is not differentiable at $x = N^{-\beta}$ but piecewise differentiable on all $\mathbb{R}$. Thus, the quantity $|\nabla_{(x,v)} \eta^N(x, v)|$ can be estimated by $|(\delta_x, \delta_v)|$ times the maximal gradient, i.e., the gradient near $x = N^{-\beta}$.

- **Step B** (Verification of (14)): We split this step into two cases.
  
  (i) For $|x| \leq 3N^{-\beta}$, $\zeta(x, v)$ is exactly same as $(\alpha N^{(\alpha+1)\beta} + N^{\alpha \beta}) \max\{1, |v|\}$ which is an upper bound of the gradient of $\eta(x, v)$ and (14) follows.
  
  (ii) For $|x| > 3N^{-\beta}$, we have
$$|\delta_x| < 2N^{-\beta} < \frac{2|x|}{3}, \quad \text{so} \quad |x| - |\delta_x| > \frac{|x|}{3}. \quad (15)$$

Hence, we have
$$|\eta^N(x + \delta_x, v + \delta_v) - \eta^N(x, v)|_\infty$$
$$\leq |\eta^N(x + \delta_x, v + \delta_v) - \eta^N(x + \delta_x, v)|_\infty + |\eta^N(x + \delta_x, v) - \eta^N(x, v)|_\infty$$
$$=: I_{11} + I_{12}. \quad (16)$$

\(\diamond\) (Estimate of $I_{12}$): We use (15) to estimate $I_{12}$:
$$I_{12} \leq \left| \frac{d}{dr} (r^{-\alpha}) \right|_{r=|x| - |\delta_x|} \frac{|v||\delta_x|_\infty}{(|x| - |\delta_x|)^{\alpha+1}} \leq \frac{\alpha |v|}{|x|^{\alpha+1}} |\delta_x|_\infty \leq \frac{3^{\alpha+1} \alpha |v|}{|x|^{\alpha+1}} |\delta_x|_\infty. \quad (17)$$

\(\diamond\) (Estimate of $I_{11}$): Similar to $I_{12}$, we have
$$|I_{11}| \leq \frac{|\delta_v|_\infty}{(|x| - |\delta_x|)^{\alpha}} \leq \frac{3^\alpha |\delta_v|_\infty}{|x|^{\alpha}}. \quad (18)$$

In (16), we combine the estimates (17) and (18) to get
$$|\eta^N(x + \delta_x, v + \delta_v) - \eta^N(x, v)|_\infty \leq \zeta(x, v) \max\{|\delta_x|_\infty, |\delta_v|_\infty\}. \quad (19)$$

Well adapted to the estimates in Lemma 4.3, we introduce the following bad sets consisting of initial configurations:
Definition 4.4. For any fixed time $t$ and any positive constant $\delta$, we define the sets $A, B, C \subseteq \mathbb{R}^{2Nd}$ as follows:

\[
A := \left\{ Z \in \mathbb{R}^{2Nd} : |J_t(Z)| = 1 \right\},
\]

\[
B := \left\{ Z \in \mathbb{R}^{2Nd} : \left| F\left(\Phi_{t,0}^N(Z)\right) - \bar{F}_t\left(\Phi_{t,0}^N(Z)\right)\right|_{\infty} > N^{-\delta} \right\},
\]

\[
C := \left\{ Z \in \mathbb{R}^{2Nd} : \left| G\left(\Phi_{t,0}^N(Z)\right) - \bar{G}_t\left(\Phi_{t,0}^N(Z)\right)\right|_{\infty} > 1 \right\}.
\]

In the following, we will first show that the time derivatives of $J_t$ restricted to these sets are small. Later, we will give a Grönwall estimate under the restriction of being in no one of these sets which finishes the proof. Therefore, let us briefly explain the meaning of these three sets: set $A$ is the set of “bad” situations where the deviation of the effective description from the C-S time evolution is already large. Since the stochastic process $J_t(Z)$ has reached its maximal value in this set, one gets zero as trivial upper bound for the time derivative of $J_t$ restricted to that set. The time derivative of $J_t$ restricted to the sets $B$ and $C$ is controlled using the law of large numbers argument. The probability to be in one of these sets is, as we will show below, decays faster than any polynomial in $N$. Since the supremum of the time derivative of $J_t$ over all possible $Z$ is polynomially bounded in $N$, we get that the time derivative of $J_t$ restricted to these two sets is small.

Assuming that $Z$ is in the complement of the union of $A$, $B$ and $C$, we can estimate (13). Being in the complement of $C$, one gets a good estimate for the second term of R.H.S. of (13), being in the complement of $B$ and using Lemma 4.3 one gets good control of the first term of the right-hand side of (13) (See Lemma 4.7 below).

From now on, we assume that the following conditions on the exponents $\alpha$, $\beta$ and $\gamma$ are satisfied:

\[
\alpha \beta + \delta < \frac{1}{2}, \quad \beta < \gamma < \delta, \quad \alpha < d - 1.
\]

Note that

\[
E_0(J_{t+\Delta t} - J_t) = E_0(J_{t+\Delta t} - J_t \mid A) + E_0(J_{t+\Delta t} - J_t \mid A^c).
\]

On the other hand, we use $A^c = (A^c \cap (B \cup C)) \cup (A^c \cap (B^c \cap C^c))$ to obtain

\[
E_0 \left( \left| U\left(\Phi_{t,0}^N(Z)\right) - \bar{U}_t\left(\Phi_{t,0}^N(Z)\right)\right|_{\infty} \mid \mathcal{A}^c \right) N^\gamma \Delta t
\]

\[
= E_0 \left( \left| U\left(\Psi_{t,0}^N(Z)\right) - \bar{U}_t\left(\Phi_{t,0}^N(Z)\right)\right|_{\infty} \mid \mathcal{A}^c \cap (B \cup C) \right) N^\gamma \Delta t
\]

\[
+ E_0 \left( \left| U\left(\Psi_{t,0}^N(Z)\right) - \bar{U}_t\left(\Phi_{t,0}^N(Z)\right)\right|_{\infty} \mid \mathcal{A}^c \cap B^c \cap C^c \right) N^\gamma \Delta t.
\]

Finally, we combine (20) and (21) to get

\[
E_0(J_{t+\Delta t} - J_t)
\]

\[
= E_0(J_{t+\Delta t} - J_t \mid A)
\]

\[
+ \left[ E_0(J_{t+\Delta t} - J_t \mid A^c) - E_0 \left( \left| U\left(\Psi_{t,0}^N(Z)\right) - \bar{U}_t\left(\Phi_{t,0}^N(Z)\right)\right|_{\infty} \mid \mathcal{A}^c \right) N^\gamma \Delta t \right]
\]

\[
+ E_0 \left( \left| U\left(\Psi_{t,0}^N(Z)\right) - \bar{U}_t\left(\Phi_{t,0}^N(Z)\right)\right|_{\infty} \mid \mathcal{A}^c \cap (B \cup C) \right) N^\gamma \Delta t
\]

\[
+ E_0 \left( \left| U\left(\Psi_{t,0}^N(Z)\right) - \bar{U}_t\left(\Phi_{t,0}^N(Z)\right)\right|_{\infty} \mid \mathcal{A}^c \cap B^c \cap C^c \right) N^\gamma \Delta t
\]

\[
=: I_{21} + I_{22} + I_{23} + I_{24}.
\]
Let us now present the estimates of $I_{2i}$.

**Lemma 4.5.** Given the definitions above, the following estimates hold: There exist a positive constant $P_0$ such that

(i) $I_{21} \leq 0$, $I_{22} = o(\Delta t)$.

(ii) $I_{23} \leq \left( \sup_{Z \in \mathbb{R}^{2N_d}} |F(Z)|_{\infty} + \sup_{Z \in \mathbb{R}^{2N_d}} |\bar{F}(Z)|_{\infty} + N^{-\beta} \right) (\mathbb{P}_0(B) + \mathbb{P}_0(C)) N^\beta \Delta t$. 

(iii) $E_0(J_{t+\Delta t} - J_t) \leq \left( 4P_0N^{\alpha \beta + \gamma} + 1 \right) (\mathbb{P}_0(B) + \mathbb{P}_0(C)) \Delta t$

\[ + E_0 \left( |U \left( \Psi_{t,0}^N(Z) \right) - \bar{U}_t \left( \Phi_{t,0}^N(Z) \right) |_{\infty} \right) \left( (A \cup B \cup C)^c \right) N^\gamma \Delta t \]

\[ + o(\Delta t). \]

**Proof.** We estimate the terms $I_{2i}$ separately in the sequel.

(i) • (Estimate of $I_{21}$): It follows from the defining condition for $J_t$ in Definition 4.1 that

\[ J_t \leq 1, \quad t \geq 0, \]

and since $J_t = 1$ on $A$

\[ J_{t+\Delta t} - J_t = J_{t+\Delta t} - 1 \leq 0 \quad \text{on } A. \]

Thus, we have

\[ E_0(J_{t+\Delta t} - J_t \mid A) \leq 0. \tag{23} \]

• (Estimate of $I_{22}$): It follows from Definition 3.2 that

\[ \Psi_{t+\Delta t,0}^N(Z) = \Psi_{t,0}^N(Z) + U \left( \Psi_{t,0}^N(Z) \right) \Delta t + o(\Delta t), \]

\[ \Phi_{t+\Delta t,0}^N(Z) = \Phi_{t,0}^N(Z) + \bar{U}_t \left( \Phi_{t,0}^N(Z) \right) \Delta t + o(\Delta t). \]

We use triangle inequality to get

\[ \left| \Psi_{t+\Delta t,0}^N(Z) - \Phi_{t+\Delta t,0}^N(Z) \right|_{\infty} \leq \left| \Psi_{t,0}^N(Z) - \Phi_{t,0}^N(Z) \right|_{\infty} + \left| U \left( \Psi_{t,0}^N(Z) \right) - \bar{U}_t \left( \Phi_{t,0}^N(Z) \right) \right|_{\infty} \Delta t + o(\Delta t). \]

Hence for $Z \in A^c$

\[ J_{t+\Delta t}(Z) - J_t(Z) \leq \left| U \left( \Phi_{t,0}^N(Z) \right) \right|_{\infty} N^\gamma \Delta t + o(\Delta t). \]

Now we take the expectation value of both sides to obtain

\[ I_{22} = o(\Delta t). \tag{24} \]

(ii) (Estimate of $I_{23}$): Note that on $A^c$, we have

\[ N^\gamma \sup_{0 \leq s \leq t} \left| \Psi_{s,0}^N(Z) - \Phi_{s,0}^N(Z) \right|_{\infty} < 1. \]

Thus, the velocity difference in $U \left( \Psi_{t,0}^N(Z) \right) - \bar{U}_t \left( \Phi_{t,0}^N(Z) \right)$ is bounded by $N^{-\gamma}$. We use this fact to conclude that

\[ I_{23} \leq \left( \sup_{Z \in \mathbb{R}^{2N_d}} \{|F(Z)|_{\infty}\} + \sup_{Z \in \mathbb{R}^{2N_d}} \{|\bar{F}(Z)|_{\infty}\} + N^{-\gamma} \right) (\mathbb{P}_0(B) + \mathbb{P}_0(C)) N^\gamma \Delta t. \tag{25} \]
(iii) In (22), we combine all estimates (23), (24) and (25) to obtain the desired estimate:

\[
\mathbb{E}_0(J_{t+\Delta t} - J_t) \leq \left( \sup_{Z \in \mathbb{R}^{2Nd}} |F(Z)|_{\infty} + \sup_{Z \in \mathbb{R}^{2Nd}} |\hat{F}(Z)|_{\infty} + N^{-\gamma} \right) \left( \mathbb{P}_0(B) + \mathbb{P}_0(C) \right) N^\gamma \Delta t \tag{26}
\]

\[+ \mathbb{E}_0 \left( \left| U \left( \Psi_{t,0}^N(Z) \right) - \hat{U}_t \left( \Phi_{t,0}^N(Z) \right) \right| \right) \left( A \cup B \cup C \right)^c N^\gamma \Delta t + o(\Delta t). \]

Next, we estimate the terms \(|F(Z)|_{\infty}\) and \(|\hat{F}(Z)|_{\infty}\). It follows from Lemma 3.1 and Remark 2 (1) that the velocity radius of the C-S model is bounded by the initial support: there exists a positive constant \(P_0 := \max_{1 \leq i \leq N} |v_i(0)| < \infty\) such that

\[P_0 \geq \sup_{t \geq 0} \max_{1 \leq i \leq N} |v_i(t)|.\]

Thus, we have

\[|v_j(t) - v_i(t)| \leq |v_j(t)| + |v_i(t)| \leq 2P_0, \quad t \geq 0.\]

We use this contraction principle to get a bound on the \(N\)-particle force and corresponding effective mean-field force:

\[|F(Z)| \leq \frac{1}{N-1} \sum_{j \neq i} \eta^N(x_j - x_i, v_j - v_i) \leq \frac{1}{N-1} \sum_{j \neq i} \omega^N(x_i - x_j)|v_j - v_i|.\]

This implies

\[\sup_{Z \in \mathbb{R}^{2Nd}} |F(Z)|_{\infty} < 2P_0N^{\alpha\beta} \quad \text{and similarly} \quad \sup_{Z \in \mathbb{R}^{2Nd}} |\hat{F}(Z)|_{\infty} < 2P_0N^{\alpha\beta}. \tag{27}\]

Finally, we combine (26) and (27) to derive the estimate (iii).

As explained above, our next step is to estimate the probability of the two atypical sets \(\mathbb{P}_0(B)\) and \(\mathbb{P}_0(C)\). We therefore establish a suitable version of the Law of Large numbers which will be used in a Corollary 1 below to give the desired estimates on \(\mathbb{P}_0(B)\) and \(\mathbb{P}_0(C)\).

**Lemma 4.6.** Let \(\tilde{h} : \mathbb{R}^d \to \mathbb{R}\) be a function satisfying

\[|\tilde{h}(x)| \leq C \min \left\{ N^{-(1-\delta)} \left( \frac{1}{|x|^\alpha} + \frac{1}{|x|^{\alpha+1}} \right), N^{-(1-l)} \right\}, \quad l := \alpha\beta + \delta, \tag{28}\]

and define \(h(x, v) = \tilde{h}(x) \max\{1, |v|\}\) and \(H_t(Z) := \sum_{j \neq i} h(x_i - x_j, v_j - v_i)\). Moreover, let \(D_i \subset \mathbb{R}^{2Nd}\) be given by

\[D_i := \left\{ Z \in \mathbb{R}^{2Nd} : \left| H_t(Z) - (N-1) \int_{\mathbb{R}^{2Nd}} h(x_i - x_s, v_s - v_i) \eta^N(z_s)dz_s \right| > 1 \right\}\]

and their union

\[D := \bigcup_{i=1}^N D_i.\]

Then, there exist a positive constant \(C_\omega\) for all \(\omega > 0\) such that \(\mathbb{P}_t(D) \leq C_\omega N^{-\omega}\).

**Proof.** Basically, we follow the method introduced in [3]. We use the following version of the Markov inequality

\[\mathbb{P}(|X| > a) \leq \frac{\mathbb{E}[X^M]}{a^M},\]

for any random variable \(X\) from \((\Omega, \mathcal{F}, \mathbb{P})\) to \(\mathbb{R}\) with \(M\) even.
Applied to $X = H_1(Z) - (N - 1) \int_{\mathbb{R}^d} h(x_1 - x_*, v_* - v_1) f_t^N(z_*) \, dz_*$ and $a = 1$, we obtain

$$\mathbb{P}_t(D_1) \leq \mathbb{E}_t \left[ \left( H_1(Z) - (N - 1) \int_{\mathbb{R}^d} h(x_1 - x_*, v_* - v_1) f_t^N(z_*) \, dz_* \right)^M \right].$$

Next, let $\mathcal{M}$ be the set of all multi-indices $\alpha = (\alpha_2, \ldots, \alpha_N) \in \mathbb{N}^{(N-1)}$ with $\sum_{j=2}^N \alpha_j = M$. We define for any multi-index $\alpha$ in this set

$$G^\alpha := \prod_{j=2}^N \left( h(x_1 - x_j, v_j - v_1) - \int_{\mathbb{R}^d} h(x_1 - x_*, v_* - v_1) f_t^N(z_*) \, dz_* \right)^{\alpha_j}.$$

With this notation, we have

$$\mathbb{E}_t \left[ \left( H_1(Z) - (N - 1) \int_{\mathbb{R}^d} h(x_1 - x_*, v_* - v_1) f_t^N(z_*) \, dz_* \right)^M \right]$$

$$= \mathbb{E}_t \left( \sum_{\alpha \in \mathcal{M}} \mathbb{E}_t(G^\alpha) = \sum_{|\alpha| \leq M/2} \mathbb{E}_t(G^\alpha),ight.$$ where $|\alpha|$ stands for the total number of indices $j$ with $\alpha_j \neq 0$. The last equality comes from the fact that if $|\alpha| > M/2$, then there exist at least one $j$ such that $\alpha_j = 1$. Integrating over this variable first one gets zero by cancellation with the expectation value, thus $\mathbb{E}_t(G^\alpha) = 0$. This is, of course, the standard argument for the proof of the Law of Large Numbers theorem.

Moreover, note that

$$\mathbb{E}_t(G^\alpha) = \int_{\mathbb{R}^d} \prod_{j=2}^N \int_{\mathbb{R}^d} \left( h(x_1 - x_j, v_j - v_1) \right.$$ 

$$\left. - \int_{\mathbb{R}^d} h(x_1 - x_*, v_* - v_1) f_t^N(z_*) \, dz_* \right)^{\alpha_j} f_t^N(z_j) \, dz_j f_t^N(z_1) \, dz_1.$$

Inspired by this observation, we will focus on estimation of the following value.

$$K :$$

$$= \left| \int_{\mathbb{R}^d} h(x_1 - x_j, v_j - v_1) - \int_{\mathbb{R}^d} h(x_1 - x_*, v_* - v_1) f_t^N(z_*) \, dz_* \right|^n f_t^N(z_j) \, dz_j.$$

Note that under the assumption of finite density

$$\sup_{0 \leq t \leq T} \left\| \rho_t^N \right\|_\infty < C_T,$$

we have

$$\int_{\mathbb{R}^d} h(x_1 - x_*, v_* - v_1) f_t^N(x_*, v_*) \, dx_* \, dv_* \leq (1 + 2P_0) \left\| \hat{h} * \rho_t^N(x_1) \right\|_\infty \leq C \left\| \hat{h} \right\|_{1/\infty},$$
where the last inequality comes from the following estimate. For any \( g_1 \in L^1 \) and \( g_\infty \in L^\infty \) such that \( g = g_1 + g_\infty \),
\[
\|fg\|_\infty \leq \|fg_1\|_\infty + \|fg_\infty\|_\infty \leq \|f\|_\infty \|g_1\|_1 + \|f\|_1 \|g_\infty\|_\infty \leq (\|f\|_\infty + \|f\|_1)(\|g_1\|_1 + \|g_\infty\|_\infty).
\] (29)

We take an infimum over all such pair \((g_1, g_\infty)\) to obtain
\[
\|fg\|\leq (\|f\|_\infty + \|f\|_1)|g|\|1\|_\infty.
\]

Moreover, note that
\[
\begin{aligned}
\|\hat{h}\|_{1\vee\infty} &\leq \int_{|x|<1} |\hat{h}(x)| \, dx + \sup_{|x|\geq 1} |\hat{h}(x)| \lesssim \int_{|x|\leq 1} N^{-\frac{1}{2}} \, dx + N^{-\frac{1}{2}} \leq CN^{-\frac{1}{2}}
\end{aligned}
\]
and we use this estimate to obtain
\[
\left|\int_{\mathbb{R}^d} h(x_1 - x, v_* - v_1)f_t^N(x_*, v_*) \, dz_* \right| \leq CN^{-\frac{1}{2}},
\]
and consequently,
\[
\mathcal{J} := \left| h(x_1 - x_j, v_j - v_1) - \int_{\mathbb{R}^d} h(x_1 - x, v_* - v_1)f_t^N(z_*) \, dz_* \right| \leq CN^{-\frac{1}{2}}.
\]
This implies
\[
K \leq \sup_{(x_j, v_j) \in \mathbb{R}^d} \mathcal{J}^n \leq CN^{-n\frac{1}{2}}.
\]

This estimate again gives
\[
\mathbb{E}_t(G^n) \leq \int_{\mathbb{R}^d} \left( \prod_{\alpha \neq 0} C^{N^{-\alpha_M(1-l)}} \right) f_t^N(z_1) \, dz_1
= \int_{\mathbb{R}^d} (C^{N^{-M(1-l)}}) f_t^N(z_1) \, dz_1 = C^k N^{-M(1-l)},
\]
where \( k = |\alpha| \). Using the standard combinatorics, the number of the multi-indices \( \alpha \) with \( |\alpha| = k \) can be estimated by
\[
\sum_{|\alpha|=k} 1 \leq \binom{N}{k} M^k \leq N^k M^k.
\]

Finally, we collect all estimates to obtain
\[
\begin{aligned}
\mathbb{E}_t \left( H_1(Z) - (N-1) \int_{\mathbb{R}^d} h(x_1 - x_*, v_* - v_1)f_t^N(x_*, v_*) \, dz_* \right)^M \\
\leq \sum_{|\alpha| \leq M/2} \mathbb{E}_t(G^n) \leq \sum_{|\alpha| \leq M/2} C^{N^{-M(1-l)}} \\
\leq \sum_{k \leq M/2} C^k M^k N^{-M(1-l)} \leq C(M) N^{-M(1-l) + \frac{M}{2}} \\
\leq C(M) N^{-M(\frac{1}{2}-l)}.
\end{aligned}
\]

Since \( l < \frac{1}{2} \) by (10), we choose \( M \) sufficiently large to get the desired decay power.

Now, we are ready to estimate \( \mathbb{P}_0(B) \) and \( \mathbb{P}_0(C) \) in Lemma (4.5).
Corollary 1. For any $\omega > 0$, there exist a constant $C_\omega$ such that
$$\mathbb{P}_0(\mathcal{B}) + \mathbb{P}_0(\mathcal{C}) \leq C_\omega N^{-\omega}.$$

Proof. Since $\Phi^N_{t,0}(\mathcal{B})$ and $\Phi^N_{t,0}(\mathcal{C})$ are the push-forward of $\mathcal{B}$ and $\mathcal{C}$ respectively, we have
$$\mathbb{P}_0(\mathcal{B}) = \mathbb{P}_t(\Phi^N_{t,0}(\mathcal{B})),$$
and
$$\mathbb{P}_0(\mathcal{C}) = \mathbb{P}_t(\Phi^N_{t,0}(\mathcal{C})).$$

Moreover, from the hypothesis (10), the functions $\frac{N^i}{N} \eta^N$ and $\frac{1}{N^\gamma} \zeta$ satisfy the bound condition (28). Hence, we can apply Lemma 4.6. Note that
$$Z \in \Phi^N_{t,0}(\mathcal{B}) \iff N^\delta |F(Z) - \bar{F}_t(Z)|_\infty > 1,$$
and
$$Z \in \Phi^N_{t,0}(\mathcal{C}) \iff |G(Z) - \bar{G}_t(Z)|_\infty > 1.$$

Hence, by Lemma 4.6, we have
$$\mathbb{P}_0(\mathcal{B}) = \mathbb{P}_t(\Phi^N_{t,0}(\mathcal{B})) \leq C_\omega N^{-\omega}, \quad \mathbb{P}_0(\mathcal{C}) = \mathbb{P}_t(\Phi^N_{t,0}(\mathcal{C})) \leq C_\omega N^{-\omega}.$$

Next, we estimate the difference $|U(\Psi^N_{t,0}(Z)) - \bar{U}_t(\Phi^N_{t,0}(Z))|_\infty$ for $Z$ which are neither in $\mathcal{A}$, $\mathcal{B}$ or $\mathcal{C}$.

Lemma 4.7. For $Z \in (\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}^c)$, the following estimate holds:
$$|U(\Psi^N_{t,0}(Z)) - \bar{U}_t(\Phi^N_{t,0}(Z))|_\infty \leq CJ_1(Z)N^{-\gamma} + N^{-\delta}.$$

Proof. For notational simplicity, we set
$$Z_1 := (X_1, V_1) := \Psi^N_{t,0}(Z), \quad Z_2 := (X_2, V_2) := \Phi^N_{t,0}(Z).$$

Then we use the definitions of $U$ and $\bar{U}$ and triangle inequality to obtain
$$|U(Z_1) - \bar{U}_t(Z_2)|_\infty \leq |F(Z_1) - \bar{F}_t(Z_2)|_\infty + |Z_1 - Z_2|_\infty \leq |F(Z_1) - F(Z_2)|_\infty + |F(Z_2) - \bar{F}_t(Z_2)|_\infty + |Z_1 - Z_2|_\infty.$$

Since $Z \in \mathcal{A}^c \cap \mathcal{B}^c$ and $\beta < \gamma$, for $N \gg 1$, we have
$$|Z_1 - Z_2| \leq d|Z_1 - Z_2|_\infty < dN^{-\gamma} < N^{-\beta} \quad \text{and} \quad |F(Z_2) - \bar{F}_t(Z_2)|_\infty \leq N^{-\delta}.$$

On the other hand
$$|(F(Z_1) - F(Z_2))_j|_\infty = \frac{1}{N-1} \left| \sum_{k \neq j} \eta^N(x_{1j} - x_{1k} + v_{1k} - v_{1j}) - \eta^N(x_{2j} - x_{2k} + v_{2k} - v_{2j}) \right|_\infty \quad (30)$$

The last difference can be estimated with help of (19):
$$|\eta^N(x_{1j} - x_{1k} + v_{1k} - v_{1j}) - \eta^N(x_{2j} - x_{2k} + v_{2k} - v_{2j})|_\infty$$
$$\leq \zeta(x_{2k} - x_{2j} + v_{2k} - v_{2j}) \max \{ |x_{2k} - x_{2j} + x_{1j} - x_{1k}|_\infty, |v_{2k} - v_{2j} - v_{1k} + v_{1j}|_\infty \}$$
$$\leq C\zeta(x_{2k} - x_{2j}, v_{2k} - v_{2j})|Z_1 - Z_2|_\infty.$$

(31)
We use (30) and (31) to get
\[ |(F(Z_1) - F(Z_2))_j| \leq \frac{C|Z_1 - Z_2|}{N - 1} \sum_{k \neq j} \zeta(x_{2k} - x_{2j}, v_{2k} - v_{2j}) \]
\[ = C|Z_1 - Z_2| \infty (G(Z_2))_j. \]

On the other hand, since \( Z \notin C \),
\[ |G(Z_2)|_\infty = |G(\Phi^N_{t,0}(Z))|_\infty \leq \tilde{G}_t(\Phi^N_{t,0}(Z)) \infty + 1 = |\tilde{G}_t(Z_2)|_\infty + 1. \]

However, with estimate (29) again and the definition of \( \zeta \), we have
\[ |\tilde{G}_t(Z_2)|_\infty \leq C \left( 1 + \int_{|x|<1} \frac{1}{|x|^\alpha+1} d x \right) (\|\rho^N(t)\|_\infty + \|\rho^N(t)\|_1) \]
\[ \leq C \left( 1 + \int_{|x|<1} \frac{1}{|x|^\alpha+1} d x \right). \]

Since \( \alpha < d - 1 \), we have \( |\tilde{G}_t(Z_2)|_\infty < C \) which does not depend on \( N \) and from this estimation, we conclude
\[ |(F(Z_1) - F(Z_2))_j| \leq C|Z_1 - Z_2|_\infty. \]

Finally, since \( Z \notin A \), we obtain
\[ |F(Z_1) - F(Z_2)|_\infty + |Z_1 - Z_2|_\infty \leq C|Z_1 - Z_2|_\infty = CJ_t(Z)N^{-\gamma}. \]

\[ \square \]

4.2. Proof of Theorem 2.2. We are now ready to provide the proof of Theorem 2.2. It follows from Lemma 4.5 (iii), combined with Corollary 1 and Lemma 4.7 that
\[ E_0(J_{t+\Delta t} - J_t) \]
\[ \leq (4P_0N^{\alpha+\beta+\gamma} + 1)(P_0(B) + P_0(C))\Delta t \]
\[ + E_0 \left( |U(\Phi^N_{t,0}(Z)) - \tilde{U}_t(\Phi^N_{t,0}(Z))|_\infty \right) (A \cup B \cup C) \infty N^\gamma \Delta t + o(\Delta t) \]
\[ \leq C_\omega (4P_0N^{\alpha+\beta+\gamma} + 1) N^{-\alpha} \Delta t + C (E_0(J_t(Z)) + N^{-\delta+\gamma}) \Delta t + o(\Delta t) \]
\[ \leq CN^{-1} \Delta t + CE_0(J_t(Z))\Delta t + N^{-\delta+\gamma} \Delta t + o(\Delta t). \]

Now we divide the above relation by \( \Delta t \) and then let \( \Delta t \) tend to zero to get
\[ \frac{d}{dt} E_0(J_t) \leq CE_0(J_t(Z)) + \left( CN^{-1} + N^{-\delta+\gamma} \right) \]
\[ \leq CE_0(J_t(Z)) + CN^{-\delta+\gamma}, \quad t > 0. \]

where we use \( \delta - \gamma < 1 \) since \( 0 < \gamma < \delta < \frac{1}{2} \). Then, Grönwall’s lemma yields
\[ E_0(J_t) \leq CN^{-\delta+\gamma}e^{CT}. \]

Here, we use the fact that \( J_0(Z) = 0 \). This completes the proof of Theorem 2.2.

Before we end this section, we comment on an analogous result for regular communication weights \( \omega \) which has not been known before. As we mentioned before, previous mean-field results for the kinetic C-S equation in [25] use the particle-in-cell method after the derivation of the kinetic equation formally under the molecular chaos assumption, i.e., the approach in [25, 27] avoids the rigorous justification of
the assumption of molecular chaos. Here, we briefly mention the result of propagation of molecular chaos for regular communication weight as a direct corollary for the singular case employed in previous subsection. For a bounded and Lipschitz continuous $\omega$, there is no need to introduce a cut-off in the arguments of Section 4.1. Although the following argument can be applied to general communication weight $\omega$, we will concentrate on the Cucker-Smale communication weight $\omega$ and $\eta$ which are given by

$$\omega(r) = \frac{1}{(1 + r^2)^{\alpha/2}}, \quad \eta = \omega(|x|)v$$

for simplicity. Then we can bound the derivative of $\eta$ by

$$\left| \partial_x \eta \right| \leq \frac{\alpha}{2} \frac{|x||v|}{(1 + |x|^2)^{\alpha/2}}, \quad \left| \partial_v \eta \right| \leq \frac{1}{(1 + |x|^2)^{\alpha/2}}.$$  

Therefore, we can take $\zeta(x,v)$ as follows:

$$\zeta(x,v) = \frac{\alpha}{2} \frac{|x||v|}{(1 + |x|^2)^{\alpha/2}} + \frac{1}{(1 + |x|^2)^{\alpha/2}},$$

and for the bound of $h$ in (28) we get the estimate

$$|h(x,v)| \leq C N^{1-\delta} \max \{1, |v|\} \left(1 + |x|^2\right)^{\alpha/2}.$$  

For this form of $h$, we can choose $\beta = 0$ since we don’t need to take a cut-off and therefore, we only need the following conditions:

$$\gamma < \delta < \frac{1}{2}$$  

(32)

which is less restrictive compared to the singular case (10). This gives a result for regular communication weights, which is analogous to Theorem 2.2.

**Theorem 4.8.** Suppose that the communication weight $\omega$ and parameters satisfy

$$\omega(r) = \frac{1}{(1 + r^2)^{\alpha/2}}, \quad 0 < T < \infty, \quad \gamma \text{ and } \delta \text{ satisfy condition (32)}.$$  

Then, for any solution $f = f(x,v,t)$ to (3) satisfying the following condition

$$\exists C_T < \infty \quad \text{such that} \quad \sup_{0 \leq t < T} \|\rho(t)\|_{\infty} \leq C_T, \quad \text{where } \rho(x,t) := \int_{\mathbb{R}^d} f(x,v,t) dv,$$

there exists a constant $\tilde{C}_T < \infty$ such that

$$\mathbb{P}_0 \left( \left\{ Z \in \mathbb{R}^{2Nd} : \sup_{0 \leq t \leq T} \left| \Psi_{t,0}^\infty(Z) - \Phi_{t,0}^\infty(Z) \right|_{\infty} > \frac{1}{N^{\gamma}} \right\} \right) \leq \frac{\tilde{C}_T}{N^{\delta - \gamma}},$$

where $\Psi_{t,0}^\infty(Z)$ and $\Phi_{t,0}^\infty(Z)$ are flows defined in Definition 2.1 generated by communication weight without cutoff.

**Remark 4.** The assumption $\sup_{0 \leq t \leq T} \|\rho(t)\|_{\infty} \leq C_T$ in Theorem 4.8 can be justified if the initial data has compact support. Since the communication weight is now regular, it is well-known, for example as in [27], that $f(t)$ is contained in $L^\infty(\mathbb{R}^{2d})$. However, the estimate in Lemma 3.1 (3) implies that the velocity support of $f(t)$ can be uniformly bounded by that of the initial data. Hence,

$$\rho(x,t) = \int_{\mathbb{R}^d} f(x,v,t) dv \leq C \|f\|_{\infty} < C_T.$$


4.3. **Comparison between the mean-field equations with and without cut-off.** In this part, we directly compare the solutions of mean field equations with and without cut-off. To develop further estimates, we make the further assumption on singularity

$$0 < \alpha < \frac{d}{2}.$$  

Now, let $f^N$ be the solution of equation with cut-off and suppose that the solution $f \in L^\infty((0,T); L^2(\mathbb{R}^d))$ to equation without cut-off exists up to some finite time $T$. Note that this existence assumption can be verified if we assume stronger assumptions $\alpha < \frac{d}{2} - 1$, according to Theorem 3.3. Finally we assume the uniform-in-$N$ boundedness of $L^2$-norm of velocity gradient of $f^N$:

$$\|\nabla_v f^N\|_{L^2} < C.$$  

Under these existence and boundedness assumption, we have

$$\partial_t (f - f^N) + v \cdot \nabla_x (f - f^N) + \nabla_v \cdot [F[f] - F[N] f^N] = 0,$$

where

$$F^N[f] := \int_{\mathbb{R}^d} \omega^N(|x - x_*)|v_*) f(x_*, v_*, t) dz_*.$$

Now, we let $g := f - f^N$. Then, we can rewrite equation as

$$\partial_t g + v \cdot \nabla_x g + \nabla_v \cdot [F[f] g + F[g] f^N + (F - F^N) [f^N] f^N] = 0.$$  

Therefore, we investigate $L^2$ norm of $g$ as

$$\frac{1}{2} \frac{d}{dt} \|g\|_{L^2}^2 \leq C \|g\|_{L^2} \|\nabla_v \cdot [F[f] - F[N] f^N]\|_{L^2} \|\nabla_v \cdot [F[f] - F[N] f^N]\|_{L^2}$$

$$+ C \|g\|_{L^2} \|\nabla_v f^N\|_{L^2} \|F[g]\|_{L^\infty} + C \|g\|_{L^2} \|\nabla_v f^N\|_{L^2} \|F - F^N\|_{L^\infty} \|f^N\|_{L^\infty}$$

$$+ \|g\|_{L^2} \|f^N\|_{L^2} \|\nabla_v \cdot [(F - F^N) [f^N]]\|_{L^\infty}.$$  

However, since we have the low-singularity condition $\alpha < \frac{d}{2}$, together with compact support, we can estimate several terms above as

$$\|\nabla_v \cdot [F[f]]\|_{L^\infty} \leq C \|\omega\|_{L^\infty} \|f\|_{L^2},$$

$$\|\nabla_v \cdot [F[f]]\|_{L^\infty} \leq C \|\omega\|_{L^2} \|g\|_{L^2}, \quad \|F[g]\|_{L^\infty} \leq C \|\omega\|_{L^2} \|g\|_{L^2},$$

$$\|\nabla_v \cdot [(F - F^N) [f^N]]\|_{L^\infty} \leq C \|\omega - \omega^N\|_{L^2} \|f^N\|_{L^2},$$

and

$$\|\omega - \omega^N\|_{L^2}^2 = \int_{|x| < N^{-\beta}} \left( \frac{1}{|x|^{\alpha}} - N^{\alpha \beta} \right)^2 dx \leq \int_{|x| < N^{-\beta}} \frac{1}{|x|^{2\alpha}} dx$$

$$= C \int_0^{N^{-\beta}} r^{-2\alpha + d - 1} dr = O \left( \frac{1}{N^\beta (d - 2\alpha)} \right).$$

Therefore, we have

$$\frac{d}{dt} \|g\|_{L^2} \leq C \|g\|_{L^2} + O \left( \frac{1}{N^\beta (d - 2\alpha)}/2 \right), \quad \|g(\cdot, 0)\|_{L^2} = 0.$$  

This implies

$$\|g\|_{L^2} \leq e^{CT} O \left( \frac{1}{N^\beta (d - 2\alpha)/2} \right). \quad (33)$$
Therefore, the $L^2$-norm of difference between solutions of equation with cut-off and equation without cut-off decays algebraically as $N$ tends to infinity.

**Remark 5.** We briefly discuss about the exponent in (33). We will focus on the two parameter, namely $\alpha$ and $\beta$. First, recall that the $\alpha$ denotes singularity of communication weight at the origin. Therefore (33) implies that the weaker singularity is, the faster the difference decays. Moreover, parameter $\beta$ measures the location of cut-off. If $\beta$ is large, then it means cut-off will be located very close to the origin, making cut-off communication weight much similar to original cut-off. Hence, if $\beta$ is large, heuristically it means that with cut-off equation and original one are similar and their solutions as well.

5. **Conclusion.** In this paper, we have addressed a rigorous justification of the persistence of molecular chaos for the particle C-S model with a singular communication weight. We have shown the convergence of singular C-S dynamics to the solution of the mean-field kinetic equation. We used the systematic machinery developed by the third author and his collaborators in a series of works [3, 31]. We identified a bad set of initial conditions for which the particle flow and kinetic flow deviate and provide a probability for the size of this bad set. In a large $N$-particle limit, this probability tends to zero algebraically in $N$ in any finite time interval. There are two immediate remaining issues left for a future work. First, we have employed a truncation in the singular coupling function and used a regularized approximate system instead of the original singular one. Thus a natural question is whether we can treat the original singularity without introducing a cut-off by exploitation of flocking estimates. Second, as noticed by the first author and his collaborators in [23], for a regular collision kernel one can use the flocking estimate for the particle system to derive a mean-field limit which is valid uniformly in time. Whether the same analysis can be applied for the singular case or not will be an interesting question which needs further study.

**Acknowledgments.** The work of S.-Y. Ha is supported by the National Research Foundation of Korea(NRF-2017R1A2B2001864) and the work of J. Kim is supported by the German Research Foundation (DFG) under the program IRTG 2235. The work of X. Zhang is supported by the National Natural Science Foundation of China (No. 11801194).

**REFERENCES**

[1] S. Ahn, H. Choi, S.-Y. Ha and H. Lee, On collision-avoiding initial configurations to Cucker-Smale type flocking models, Commun. Math. Sci., 10 (2012), 625–643.
[2] S. Ahn and S.-Y. Ha, Stochastic flocking dynamics of the Cucker-Smale model with multiplicative white noises, J. Math. Phys., 51 (2010), 103301, 17pp.
[3] N. Boers and P. Pickl, On mean field limits for dynamical systems, J. Stat. Phys., 164 (2016), 1–16.
[4] F. Bolley, J. A. Cañizo and J. A. Carrillo, Stochastic mean-field limit: Non-Lipschitz forces and swarming, Math. Models Methods Appl. Sci., 21 (2011), 2179–2210.
[5] J. A. Carrillo, Y.-P. Choi and M. Hauray, Local well-posedness of the generalized Cucker-Smale model with singular kernels, MMCS, Mathematical Modeling of Complex Systems, ESAIM Proc. Surveys, EDP Sci., Les Ulis, 47 (2014), 17–35.
[6] J. A. Carrillo, Y.-P. Choi, P. B. Mucha and J. Peszek, Sharp conditions to avoid collisions in singular Cucker-Smale interactions, Nonlinear Anal. Real World Appl., 37 (2017), 317–328.
[7] J. A. Cañizo, J. A. Carrillo and J. Rosado, A well-posedness theory in measures for some kinetic models of collective motion, Math. Models Methods Appl. Sci., 21 (2011), 515–539.
[8] J. A. Carrillo, M. R. D’Orsogna and Y. Panferov, Double milling in self-propelled swarms from kinetic theory, *Kinet. Relat. Models*, 2 (2009), 363–378.

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

[10] J. A. Carrillo, A. Klar, S. Martin and S. Tiwari, Self-propelled interacting particle systems with roosting force, *Math. Models Methods Appl. Sci.*, 20 (2010), 1533–1552.

[11] J. Cho, S.-Y. Ha, F. Huang, C. Jin and D. Ko, Emergence of bi-cluster flocking for the Cucker-Smale model, *Math. Models Methods Appl. Sci.*, 28 (2016), 1191–1218.

[12] J. Cho, S.-Y. Ha, F. Huang, C. Jin and D. Ko, Emergence of bi-cluster flocking for agent-based models with unit speed constraint, *Anal. Appl.*, 14 (2016), 39–73.

[13] Y.-P. Choi, S.-Y. Ha and Z. Li., Emergent dynamics of the Cucker-Smale flocking model and its variants, *Active Particles, Advances in Theory, Models, and Applications*, Model. Simul. Sci. Eng. Technol., Birkhäuser/Springer, Cham, 1 (2017) 299–331.

[14] F. Cucker and J.-G. Dong, Avoiding collisions in flocks, *IEEE Trans. Automat. Control*, 55 (2010), 1238–1243.

[15] F. Cucker and E. Mordecki, Flocking in noisy environments, *J. Math. Pures Appl.*, 89 (2008), 278–296.

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

[17] P. Degond and S. Motsch, Macroscopic limit of self-driven particles with orientation interaction, *C.R. Math. Acad. Sci. Paris*, 345 (2007), 555–560.

[18] P. Degond and S. Motsch, Large scale dynamics of the persistent turning walker model of fish behavior, *J. Stat. Phys.*, 131 (2008), 989–1021.

[19] P. Degond and S. Motsch, Continuum limit of self-driven particles with orientation interaction, *Math. Models Methods Appl. Sci.*, 18 (2008), 1193–1215.

[20] R. Duan, M. Fornasier and G. Toscani, A kinetic flocking model with diffusion, *Comm. Math. Phys.*, 300 (2010), 95–145.

[21] M. Fornasier, J. Haskovec and G. Toscani, Fluid dynamic description of flocking via Povzner-Boltzmann equation, *Phys. D*, 240 (2011), 21–31.

[22] S.-Y. Ha, T. Ha and J. Kim, Asymptotic dynamics for the Cucker-Smale-type model with the Rayleigh friction, *J. Phys. A*, 43 (2010), 315201, 19pp.

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

[24] S.-Y. Ha and K. Lee and D. Levy, Emergence of time-asymptotic flocking in a stochastic Cucker-Smale system, *Commun. Math. Sci.*, 7 (2009), 453–469.

[25] S.-Y. Ha and J.-G. Liu, A simple proof of Cucker-Smale flocking dynamics and mean field limit, *Commun. Math. Sci.*, 7 (2009), 325–327.

[26] S.-Y. Ha and M. Slemrod, Flocking dynamics of singularly perturbed oscillator chain and the Cucker-Smale system, *J. Dynam. Differential Equations*, 22 (2010), 325–330.

[27] M. Hauray and P.-E. Jabin, N-particles Approximation of the Vlasov Equations with Singular Potential, *Arch. Ration. Mech. Anal.*, 183 (2007), 489–524.

[28] M. K.-H. Kiessling, The microscopic foundations of Vlasov theory for jellium-like Newtonian N-body systems, *J. Stat. Phys.*, 155 (2014), 1299–1328.

[29] J. Kim and J. Peszek, Cucker-Smale model with a bonding force and singular interaction kernel, preprint.

[30] D. Lazarovici and P. Pickl, A mean field limit for the Vlasov-Poisson system, *Arch. Ration. Mech. Anal.*, 225 (2017), 1201–1231.

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

[32] S. Motsch and E. Tadmor, A new model for self-organized dynamics and its flocking behavior, *J. Stat. Phys.*, 144 (2011), 923–947.

[33] P. B. Mucha and J. Peszek, The Cucker-Smale equation: singular communication weight, measure solutions and weak-atomic uniqueness, *Arch. Ration. Mech. Anal.*, 227 (2018), 273–308.

[34] J. Park, H. Kim and S.-Y. Ha, Cucker-Smale flocking with inter-particle bonding forces, *IEEE Trans. Automat. Control*, 55 (2010), 2017–2023.
[36] L. Perea, P. Elosegui and G. Gómez, Extension of the Cucker-Smale control law to space flight formations, J. Guid. Control Dynam., 32 (2009), 526–536.

[37] J. Peszek, Existence of piecewise weak solutions of a discrete Cucker-Smale’s flocking model with a singular communication weight, J. Differential Equations, 257 (2014), 2900–2925.

[38] J. Peszek, Discrete Cucker-Smale flocking model with a weakly singular weight, SIAM J. Math. Anal., 47 (2015), 3671–3686.

[39] D. Poyato and J. Soler, Euler-type equations and commutators in singular and hyperbolic limits of kinetic Cucker-Smale models, Math. Models Methods Appl. Sci., 27 (2017), 1089–1152.

[40] J. Shen, Cucker-Smale flocking under hierarchical leadership, SIAM J. Appl. Math., 68 (2007), 694–719.

[41] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen and O. Shochet, Novel type of phase transition in a system of self-driven particles, Phys. Rev. Lett., 75 (1995), 1226–1229.

Received August 2018; revised February 2019.

E-mail address: syha@snu.ac.kr
E-mail address: jhkim206@snu.ac.kr
E-mail address: pickl@math.lmu.de
E-mail address: xtzhang@hust.edu.cn