Arch. Rational Mech. Anal. 240 (2021) 203–241
Digital Object Identifier (DOI) https://doi.org/10.1007/s00205-021-01609-8

Anticipation Breeds Alignment

Ruiwen Shu & Eitan Tadmor

Communicated by A. Figalli

Abstract

We study the large-time behavior of systems driven by radial potentials, which react to *anticipated* positions, $\mathbf{x}^\tau(t) = \mathbf{x}(t) + \tau\mathbf{v}(t)$, with anticipation increment $\tau > 0$. As a special case, such systems yield the celebrated Cucker–Smale model for alignment, coupled with pairwise interactions. Viewed from this perspective, such anticipation-driven systems are expected to emerge into *flocking* due to alignment of velocities, and *spatial concentration* due to confining potentials. We treat both the discrete dynamics and large crowd hydrodynamics, proving the decisive role of anticipation in driving such systems with attractive potentials into velocity alignment and spatial concentration. We also study the concentration effect near equilibrium for anticipated-based dynamics of pair of agents governed by attractive–repulsive potentials.

Contents

1. Introduction and Statement of Main Results 203
2. A Priori L^∞ Bounds for Confining Potentials 212
3. Anticipation with Convex Potentials and Positive Kernels 217
4. Local Versus Global Weighted Means 222
5. Anticipation Dynamics with Attractive Potentials 228
6. Anticipation Dynamics with Repulsive–Attractive Potential 230
7. Anticipation Dynamics: Hydrodynamic Formulation 235
References 239

1. Introduction and Statement of Main Results

Consider the dynamical system

$$\left\{\begin{array}{l} \dot{\mathbf{x}}_i(t) = \mathbf{v}_i(t) \\ \dot{\mathbf{v}}_i(t) = -\nabla_i \mathcal{H}_N(\mathbf{x}_1^\tau, \ldots, \mathbf{x}_N^\tau), \quad \mathbf{x}_i^\tau = \mathbf{x}_i^\tau(t) := \mathbf{x}_i(t) + \tau\mathbf{v}_i(t), \end{array}\right. \qquad i = 1, \ldots, N.$$

When $\tau = 0$, this is the classical N-particle dynamics for positions and velocities, $(\mathbf{x}_i(t), \mathbf{v}_i(t)) \in (\mathbb{R}^d, \mathbb{R}^d)$, governed by the general Hamiltonian $\mathcal{H}_N(\cdots)$. If we fix a small time step $\tau > 0$, then the system is not driven instantaneously but reacts to the positions $\mathbf{x}^\tau(t) = \mathbf{x}(t) + \tau\mathbf{v}(t)$, *anticipated* at time $t + \tau$, where τ is the anticipation time increment. Anticipation is an important feature in social dynamics of N-agent and N-player systems, [20,22,30]. A key feature in the the large time behavior of such anticipated dynamics is the dissipation of the *anticipated energy*

$$\mathcal{E}(t) = \frac{1}{2N}\sum_i |\mathbf{v}_i|^2 + \frac{1}{N}\mathcal{H}_N(\mathbf{x}_1^\tau, \ldots, \mathbf{x}_N^\tau),$$

at a rate given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) = \frac{1}{N}\sum_i \mathbf{v}_i \cdot \dot{\mathbf{v}}_i + \frac{1}{N}\sum_i \nabla_i\mathcal{H}_N(\mathbf{x}_1^\tau, \ldots, \mathbf{x}_N^\tau)\cdot(\mathbf{v}_i + \tau\dot{\mathbf{v}}_i) = -\frac{\tau}{N}\sum_i |\dot{\mathbf{v}}_i|^2, \quad \tau > 0.$$

We refer to the quantity on the right, $\frac{\tau}{N}\sum_i |\dot{\mathbf{v}}_i|^2$, as the *enstrophy* of the system.

1.1. Pairwise Interactions

In this work we study the anticipation dynamics of pairwise interactions

$$\left\{\begin{aligned} &\dot{\mathbf{x}}_i(t) = \mathbf{v}_i(t) \\ &\dot{\mathbf{v}}_i(t) = -\frac{1}{N}\sum_{j=1}^N \nabla U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|), \quad \mathbf{x}_i^\tau = \mathbf{x}_i(t) + \tau\mathbf{v}_i(t), \end{aligned}\right. \qquad i = 1, \ldots, N, \tag{AT}$$

governed by a radial interaction potential $U(r)$, $r = |\mathbf{x}|$. This corresponds to the Hamiltonian $\mathcal{H}_N(\mathbf{x}_1^\tau, \ldots, \mathbf{x}_N^\tau) = \dfrac{1}{2N}\sum_{j,k} U(|\mathbf{x}_j^\tau - \mathbf{x}_k^\tau|)$, where the conservative N-body problem ($\tau = 0$) is now replaced by N-agent dynamics with anticipated energy dissipation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) = -\frac{\tau}{N}\sum_i |\dot{\mathbf{v}}_i|^2, \qquad \mathcal{E}(t) := \frac{1}{2N}\sum_i |\mathbf{v}_i|^2 + \frac{1}{2N^2}\sum_{i,j} U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|), \quad \tau > 0. \tag{1.1}$$

To gain a better insight into (AT), we consider the general system

$$\left\{\begin{aligned} &\dot{\mathbf{x}}_i = \mathbf{v}_i \\ &\dot{\mathbf{v}}_i = \frac{\tau}{N}\sum_{j=1}^N \Phi_{ij}(\mathbf{v}_j - \mathbf{v}_i) - \frac{1}{N}\sum_{j=1}^N \nabla U(|\mathbf{x}_i - \mathbf{x}_j|), \end{aligned}\right. \qquad i = 1, \ldots, N. \tag{ΦU}$$

The anticipation (AT) is recovered as a special case of (ΦU) in terms of the 'intermediate' Hessians, $\Phi_{ij} = \overline{D^2U_{ij}}$, but since these intermediate Hessians are not readily available,[1] we will consider (ΦU) for a general class of *communication matrices*, $\Phi_{ij} \in \boldsymbol{\Phi}$, which respect the symmetry property satisfied by the $\overline{D^2U}_{ij}$'s,

$$\boldsymbol{\Phi} := \left\{\Phi(\cdot,\cdot) \in \text{Sym}_{d\times d} \mid \Phi_{ij} := \Phi\big((\mathbf{x}_i,\mathbf{v}_i),(\mathbf{x}_j,\mathbf{v}_j)\big) = \Phi_{ji}\right\}.$$

The so-called (ΦU) system provides a unified framework for anticipation dynamics by coupling general symmetric communication matrices, $\{\Phi \in \boldsymbol{\Phi}\}$, together with pairwise interactions induced by the potential U. In the particular case of $U = 0$, the general system (ΦU) yields the celebrated Cucker–Smale (CS) model [13,14], $\dot{\mathbf{v}}_i = \frac{\tau}{N}\sum_j \Phi_{ij}(\mathbf{v}_j - \mathbf{v}_i)$, a prototypical model for alignment dynamics in which $\max_{i,j} |\mathbf{v}_i(t) - \mathbf{v}_j(t)| \stackrel{t\to\infty}{\longrightarrow} 0$. There is, however, one distinct difference: while the CS model is governed by a *scalar* kernel involving geometric distances, $\Phi_{ij} = \phi(|\mathbf{x}_i - \mathbf{x}_j|)\mathbb{I}_{d\times d}$, here (ΦU) allows for a larger class of communication protocols based on *matrix* kernels, e.g., $\Phi_{ij} = \Phi(\mathbf{x}_i,\mathbf{x}_j)$, with a possible dependence on topological distances, [37]. The flocking behavior of such *matrix-based* CS models, proved in Section 3.1, is considerably more intricate than in the scalar case, due to the lack of a maximum principle.

The main purpose of this paper is to study the decisive role of anticipation in driving the emergent behavior of (AT) and (ΦU), proving, under appropriate assumptions, *flocking* and *spatial concentration*, see (1.9),(1.18) below,

$$|\mathbf{x}_i(t) - \big(\overline{\mathbf{x}}_0 + t\overline{\mathbf{v}}_0\big)| + |\mathbf{v}_i(t) - \overline{\mathbf{v}}_0| \stackrel{t\to\infty}{\longrightarrow} 0.$$

Viewed from the perspective of Cucker–Smale alignment dynamics, the large time flocking behavior of (AT),(ΦU) is expected due to alignment of velocities. Moreover, our study [38] shows that confinement due to external forcing, $-\nabla V(|\mathbf{x}_i|)$, leads to spatial concentration, and it is known, e.g., [38, p. 351], that the dynamics with external forcing coincides with pairwise interactions, $-\frac{1}{N}\sum_j \nabla U(|\mathbf{x}_i - \mathbf{x}_j|)$, in the special case of quadratic potential $V(r) \sim \frac{1}{2}r^2$. From this perspective, here we prove spatial concentration for a (much) larger class of attractive potentials U.

We begin in Section 3 with the general system (ΦU). The basic bookkeeping associated with (ΦU) quantifies its decay rate of the (instantaneous) energy

$$E(t) := \frac{1}{2N}\sum_i |\mathbf{v}_i|^2 + \frac{1}{2N^2}\sum_{i,j} U(|\mathbf{x}_i - \mathbf{x}_j|),$$

which is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}E(t) = -\frac{\tau}{2N^2}\sum_{i,j}(\mathbf{v}_j - \mathbf{v}_i)^\top \Phi_{ij}(\mathbf{v}_j - \mathbf{v}_i). \tag{1.2}$$

---

[1] Expanding (AT) in τ we obtain (ΦU) with matrices $\Phi_{ij} = \overline{D^2U}_{ij} := \int_0^1 D^2U(|(\mathbf{x}_i - \mathbf{x}_j) + \tau s(\mathbf{v}_i - \mathbf{v}_j)|)\,\mathrm{d}s$, depending on states $(\mathbf{x}_i,\mathbf{v}_i)$ and $(\mathbf{x}_j,\mathbf{v}_j)$. Their (pq) entries are given by $(\overline{D^2U}_{ij})_{pq} = (D^2U)_{pq}(|\mathbf{x}_i(t;\tau_{ij}^{pq}) - \mathbf{x}_j(t;\tau_{ij}^{pq})|)$, evaluated in anticipated positions, $\mathbf{x}(t;\tau_{ij}^{pq}) = \mathbf{x} + \tau_{ij}^{pq}\mathbf{v}$, at some intermediate times, $\tau_{ij}^{pq} \in [0,\tau]$.

This will be contrasted with the dissipation of anticipated energy (1.1) in Section 5 below. To explore the enstrophy on the right of (1.2) we need to further elaborate on properties of the communication matrices, $\Phi_{ij} = \Phi(\cdot, \cdot)$, and their relations to the potential U.

We start by rewriting the Hessian of the radial potential in the form

$$D^2U(|\mathbf{x}_i - \mathbf{x}_j|) = \frac{U'(r_{ij})}{r_{ij}}(\mathbb{I} - \widehat{\mathbf{x}}_{ij}\widehat{\mathbf{x}}_{ij}^\top) + U''(r_{ij})\widehat{\mathbf{x}}_{ij}\widehat{\mathbf{x}}_{ij}^\top, \quad r_{ij} := |\mathbf{x}_i - \mathbf{x}_j|, \quad \widehat{\mathbf{x}}_{ij} := \frac{\mathbf{x}_i - \mathbf{x}_j}{r_{ij}}, \tag{1.3}$$

and observe that the symmetric matrix $D^2U(|\mathbf{x}_i - \mathbf{x}_j|)$ has a single eigenvalue $U''(r_{ij})$ in the radial direction, $\mathbf{x}_i - \mathbf{x}_j$, and $d-1$ multiple of the eigenvalues $\frac{U'(r_{ij})}{r_{ij}}$ in tangential directions $(\mathbf{x}_i - \mathbf{x}_j)^\perp$. We study the dynamics induced by potentials U which are at least C^2. As a result, $U'(0) = 0$, and we may assume $U(0) = 0$ by adding a constant to it. We specify two main classes of potentials we will be working with: *convex* potentials, $U''(r) \gtrsim \langle r\rangle^{2-\beta}$, studied in Section 3, and the larger class of *attractive* potentials, $\frac{U'(r)}{r} \gtrsim \langle r\rangle^{2-\beta}$, studied in Section 5.[2] In either case, we postulate that the potential is *bounded*, in the sense that

$$\text{there exists a constant } A > 0 \text{ such that } |U''(r)| \leqslant A. \tag{1.4}$$

It follows that $|U'(r)| \leqslant \displaystyle\int_0^r |U''(s)|\,\mathrm{d}s \leqslant \int_0^r A\,\mathrm{d}s = Ar$, and hence that the communication matrices are bounded:

$$-A\mathbb{I}_{d\times d} < D^2U(\cdot) \leqslant A\mathbb{I}_{d\times d}.$$

In particular, this rules out the important class of singular kernels (in both first- and second-order dynamics), e.g., [8,9,17,21,26,29,32,34–36], which is left for a future study. Finally, we mention the larger class of *confining* potentials, (2.1), which includes the repulsive–attractive potentials studied in Section 6.

1.2. Anticipation Dynamics with Convex Potentials

Recall that the flocking behavior of CS model , see (3.8) below, is guaranteed for *scalar* communication kernels, $\Phi(r) = \phi(r)\mathbb{I}$, which satisfy a so-called *fat tail condition*, [23], [31, Proposition 2.9],

$$\int \phi(r)\,\mathrm{d}r = \infty,$$

or — expressed in terms of its decay rate, $\phi(r) \sim \langle r\rangle^{-\beta}$ for $0 \leqslant \beta \leqslant 1$. Since the anticipation model (AT) can be viewed as a special case of (ΦU), it is natural to quantify the convexity of U and positivity of Φ in terms of their 'fat tail' decay.

---

[2] Throughout this paper, we use the notation $\langle r\rangle^s := (1+r^2)^{s/2}$ for scalar r and $\langle \mathbf{z}\rangle = \langle |\mathbf{z}|\rangle$ for vectors $\mathbf{z}$.

Assumption 1.1. (*Convex potentials*) There exist constants $0 < a < A$ and β such that

$$a\langle r\rangle^{-\beta} \leqslant U''(r) \leqslant A, \qquad 0 \leqslant \beta \leqslant 1. \tag{1.5}$$

Typically, the upper and lower bounds associated with U and its derivatives dictate its profile near the origin, $r \ll 1$ and respectively, near infinity, $r \gg 1$. It particular, the lower bound in (1.5) implies $U'(r) = \int_0^r U''(s)\,\mathrm{d}s \geqslant \int_0^r a\langle s\rangle^{-\beta}\,\mathrm{d}s \geqslant a\langle r\rangle^{-\beta} r$, and hence D^2U in (1.3) satisfies the fat tail condition $D^2U(|\mathbf{x}|) \geqslant a\langle \mathbf{x}\rangle^{-\beta}$ with $0 \leqslant \beta \leqslant 1$.

Assumption 1.2. (*Positive kernels*) There exist constants $0 < \phi_- < \phi_+$ and γ such that

$$\phi_-(\langle \mathbf{x}_i - \mathbf{x}_j\rangle + \langle \mathbf{v}_i - \mathbf{v}_j\rangle)^{-\gamma} \leqslant \Phi_{ij} \leqslant \phi_+, \qquad 0 \leqslant \gamma < 1. \tag{1.6}$$

Observe that (ΦU) conserves momentum

$$\begin{cases} \dot{\overline{\mathbf{x}}} = \overline{\mathbf{v}}, & \overline{\mathbf{x}} := \dfrac{1}{N}\displaystyle\sum_i \mathbf{x}_i, \\ \dot{\overline{\mathbf{v}}} = 0, & \overline{\mathbf{v}} := \dfrac{1}{N}\displaystyle\sum_i \mathbf{v}_i. \end{cases} \tag{1.7}$$

It follows that the mean velocity $\overline{\mathbf{v}}$ is constant in time, $\overline{\mathbf{v}}(t) = \overline{\mathbf{v}}_0$, and hence $\overline{\mathbf{x}}(t) = \overline{\mathbf{x}}_0 + t\overline{\mathbf{v}}_0$. Our first main result is expressed in terms of the *energy fluctuations*

$$\delta E(t) := \frac{1}{2N}\sum_i |\mathbf{v}_i - \overline{\mathbf{v}}|^2 + \frac{1}{2N^2}\sum_{i,j} U(|\mathbf{x}_i - \mathbf{x}_j|).$$

Theorem 1. *(*Anticipation dynamics *(ΦU)* — velocity alignment and spatial concentration*) Consider the anticipation dynamics (ΦU). Assume a bounded convex potential U with fat-tail decay of order β,* (1.5)*, and a symmetric kernel matrix Φ with a fat-tail decay of order γ,* (1.6)*. If the decay parameters lie in the restricted range $3\beta + 2\max\{\beta,\gamma\} < 4$, then there is sub-exponential decay of the energy fluctuations*

$$\delta E(t) \leqslant Ce^{-t^{1-\eta}}, \qquad \eta = \frac{2\max\{\beta,\gamma\}}{4-3\beta} < 1. \tag{1.8}$$

We conclude that for large time, the dynamics concentrates in space with global velocity alignment at sub-exponential rate,

$$|\mathbf{x}_i(t) - \overline{\mathbf{x}}(t)| \to 0, \quad |\mathbf{v}_i(t) - \overline{\mathbf{v}}_0| \to 0, \qquad \overline{\mathbf{x}}(t) = \overline{\mathbf{x}}_0 + t\overline{\mathbf{v}}_0. \tag{1.9}$$

The proof of Theorem 1 proceeds in two steps:
(i) A uniform bound, outlined in Lemma 2.1 below, on maximal spread of positions $|\mathbf{x}_i(t)|$,

$$\max_i |\mathbf{x}_i(t)| \leqslant C_\infty\langle t\rangle^{\frac{2}{4-3\beta}}, \quad \max_i |\mathbf{v}_i(t)| \leqslant C_\infty\langle t\rangle^{\frac{2-\beta}{4-3\beta}}, \qquad 0 \leqslant \beta \leqslant 1. \tag{1.10}$$

(ii) Observe that in view of (1.7), $\frac{\mathrm{d}}{\mathrm{d}t}\delta E(t) = \frac{\mathrm{d}}{\mathrm{d}t}E(t)$. The energy dissipation (1.2) combined with the bounds (1.6),(1.10) imply the decay of energy fluctuations

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta E(t) = \frac{\mathrm{d}}{\mathrm{d}t}E(t) \lesssim -\frac{\tau}{2N}\langle t\rangle^{-\frac{2\gamma}{4-3\beta}}\sum_i |\mathbf{v}_i - \overline{\mathbf{v}}|^2.$$

To close the last bound we need a hypocoercivity argument carried out in Section 3, which leads to the sub-exponential decay (1.8). The conclusion of sub-exponential flocking
$(\mathbf{x}_i - \mathbf{x}_j, \mathbf{v}_i - \mathbf{v}_j) \stackrel{t\to\infty}{\longrightarrow} 0$ follows, and naturally, $(\mathbf{x}_i, \mathbf{v}_i) - (\overline{\mathbf{x}}(t), \overline{\mathbf{v}}_0) \to 0$ since this is the only minimizer of $\delta E(t)$.

Since the anticipation dynamics (AT) can be viewed as a special case of (ΦU) system with intermediate Hessians $\Phi_{ij} = \overline{D^2U}_{ij}$ (outlined in footnote 1), Theorem 1 applies with $\gamma = \beta$.

Corollary 1.1. (Anticipation dynamics (AT) with convex potentials) *Consider the anticipated dynamics* (AT) *with bounded convex potential satisfying*

$$a\langle r\rangle^{-\beta} \leqslant U''(r) \leqslant A, \qquad 0 \leqslant \beta < \frac{4}{5}.$$

Then there is sub-exponential decay of the energy fluctuations

$$\delta E(t) \leqslant Ce^{-t^{1-\eta}}, \quad \eta = \frac{2\beta}{4-3\beta}. \tag{1.11}$$

The large time flocking behavior follows: the dynamics concentrates in space with global velocity alignment at sub-exponential rate,

$$|\mathbf{x}_i(t) - \overline{\mathbf{x}}(t)| \to 0, \quad |\mathbf{v}_i(t) - \overline{\mathbf{v}}_0| \to 0, \qquad \overline{\mathbf{x}}(t) := \overline{\mathbf{x}}_0 + t\overline{\mathbf{v}}_0. \tag{1.12}$$

Remark 1.1. (*Optimal result with improved fat-tail condition*) Suppose we strengthen assumption 1.1 with a more precise behavior of $U''(r) \sim \langle r\rangle^{-\beta}$, thus replacing (1.5) with the requirement that there exist constants $0 < a < A$ and β such that

$$a\langle r\rangle^{-\beta} \leqslant U''(r), \quad U'(r) \leqslant A\langle r\rangle^{1-\beta}, \qquad 0 \leqslant \beta \leqslant 1. \tag{1.13}$$

Then the anticipation dynamics (ΦU) with a fat-tail kernel matrix Φ of order γ, (1.6), satisfies the sub-exponential decay

$$\delta E(t) \leqslant Ce^{-t^{1-\eta}}, \qquad \eta = \min\left\{1, \frac{2}{4-3\beta}\right\}\cdot\max\{\beta,\gamma\} < 1. \tag{1.14}$$

This improved decay follows from the corresponding improvement of the uniform bound in Lemma 2.1 below which reads $\max_i |\mathbf{x}_i(t)| \lesssim \langle t\rangle$. In the particular case of $\beta = \gamma$, we recover an improved corollary 1.1 for anticipated dynamics (AT), where the anticipated energy (see (1.16) below), satisfies an optimal decay of order

$$\delta\mathcal{E}(t) \leqslant Ce^{-t^{1-\eta}}, \qquad \eta = \min\left\{\frac{2\beta}{4-3\beta}, \beta\right\} < 1.$$

1.3. Anticipation Dynamics with Purely Attractive Potential

We now turn our attention to the main anticipation model (AT). We already know the flocking behavior of (AT) for convex potentials, from the general considerations of the (ΦU) system, summarized in corollary 1.1. In fact, the corresponding communication matrix of (AT) prescribed in (1.3), D^2U, has a special structure of rank-one modification of the scalar kernel $\dfrac{U'(r)}{r}$. This enables us to treat the flocking behavior of (AT) for a larger class of purely attractive potentials.

Assumption 1.3. (*Purely attractive potentials*) There exist constants $0 < a < A$ and β such that

$$a\langle r\rangle^{-\beta} \leqslant \frac{U'(r)}{r}, \quad |U''(r)| \leqslant A, \qquad 0 \leqslant \beta \leqslant 1. \tag{1.15}$$

Our result is expressed in terms of fluctuations of the *anticipated energy*

$$\delta\mathcal{E}(t) := \frac{1}{2N}\sum_i |\mathbf{v}_i - \overline{\mathbf{v}}|^2 + \frac{1}{2N^2}\sum_{i,j} U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|). \tag{1.16}$$

Theorem 2. *(*Anticipation dynamics (AT) with attractive potential*) Consider the anticipation dynamics* (AT)*, and assume a bounded, purely attractive potential with a fat tail decay of order β,* (1.15)*. If the decay parameter $\beta < \frac{1}{3}$, then there is sub-exponential decay of the anticipated energy fluctuations*

$$\delta\mathcal{E}(t) \leqslant Ce^{-t^{1-\eta}}, \qquad \eta = \frac{2\beta}{1-\beta} < 1. \tag{1.17}$$

It follows that for large time, the anticipation dynamics concentrates in space with global velocity alignment at sub-exponential rate,

$$|\mathbf{x}_i(t) - \overline{\mathbf{x}}(t)| \to 0, \quad |\mathbf{v}_i(t) - \overline{\mathbf{v}}_0| \to 0, \qquad \overline{\mathbf{x}}(t) = \overline{\mathbf{x}}_0 + t\overline{\mathbf{v}}_0. \tag{1.18}$$

Remark 1.2. This result is surprising if one interprets (AT) in its equivalent matrix formulation (ΦU), since attractive potentials do *not* necessarily induce communication matrix $\Phi = D^2U$ which is positive definite. In particular, the corresponding 'regular' (instantaneous) energy $E(t)$ referred to in corollary 1.1 is not necessarily decreasing; only the *anticipated* energy is.

The proof of Theorem 2, carried out in Section 5, involves two main ingredients. **(i)**. First, we derive an a priori uniform bound on the maximal spread of *anticipated* positions $|\mathbf{x}_i^\tau(t)|$,

$$\max_i |\mathbf{x}_i^\tau(t)| \leqslant C_\infty \langle t\rangle^{\frac{1}{2-2\beta}}, \qquad 0 \leqslant \beta < 1. \tag{1.19}$$

(ii). A second main ingredient for the proof of Theorem 2 is based on the energy dissipation (1.1). The key step here is to relate the enstrophy in (1.1),

$$\frac{\tau}{N}\sum_i |\dot{\mathbf{v}}_i|^2 = \frac{\tau}{N}\sum_i \left|\frac{1}{N}\sum_j c_{ij}(\mathbf{x}_i^\tau - \mathbf{x}_j^\tau)\right|^2, \qquad c_{ij} = \frac{U'(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|)}{|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|}, \tag{1.20}$$

to the fluctuations of the (expected) *positions*. This is done by the following proposition, interesting for its own sake, which deals with the local vs. global means of arbitrary $\mathbf{z}_j \in \mathbb{R}^d$:

Lemma 1.1. (Local and global means are comparable) *Fix* $0 < \lambda \leqslant \Lambda$ *and weights* c_{ij}

$$0 < \lambda \leqslant c_{ij} \leqslant \Lambda.$$

Then, there exists a constant $C = C(\lambda, \Lambda) \leqslant 32\frac{\Lambda^2}{\lambda^4}$ *such that, for arbitrary* $\mathbf{z}_j \in \mathbb{R}^d$*, it holds that*

$$\frac{1}{N^2}\sum_{i,j}\left|\mathbf{z}_i - \mathbf{z}_j\right|^2 \leqslant \frac{C(\lambda,\Lambda)}{N}\sum_i\left|\frac{1}{N}\sum_j c_{ij}(\mathbf{z}_i - \mathbf{z}_j)\right|^2, \qquad C(\Lambda,\lambda) \leqslant 32\frac{\Lambda^2}{\lambda^4}. \tag{1.21}$$

Remark 1.3. (*Why a lemma on means?*) The sum on the left of (1.21) quantifies the fluctuations relative to the average $\overline{\mathbf{z}} := \frac{1}{N}\sum_j \mathbf{z}_j$,

$$\frac{1}{N^2}\sum_{i,j}\left|\mathbf{z}_i - \mathbf{z}_j\right|^2 = \frac{2}{N}\sum_i |\mathbf{z}_i - \overline{\mathbf{z}}|^2 .$$

Hence, (1.21) implies (and in fact is equivalent, up to scaling, to the statement about the local means induced by weights θ_{ij})

$$\frac{\lambda}{N} \leqslant \theta_{ij} \leqslant \frac{\Lambda}{N}, \qquad \sum_j \theta_{ij} = 1.$$

If $\overline{\mathbf{z}}_i(\theta) := \sum_j \theta_{ij}\mathbf{z}_j$ are the local means, then (1.21) with $c_{ij} = N\theta_{ij}$ implies

$$\frac{1}{N}\sum_i |\mathbf{z}_i - \overline{\mathbf{z}}|^2 \leqslant 16\frac{\Lambda^2}{\lambda^4}\frac{1}{N}\sum_i |\mathbf{z}_i - \overline{\mathbf{z}}_i(\theta)|^2. \tag{1.22}$$

Thus, the deviation from the local means is comparable to the deviation from the global mean.

Applying (1.21) to (1.20) with the given bounds (1.15),(1.19), yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) &= -\frac{\tau}{N}\sum_i |\dot{\mathbf{v}}_i|^2 \\ &\lesssim -\frac{\tau}{N^2}\frac{A^2}{a^4}\Big(\max_{i,j}\langle \mathbf{x}_i^\tau - \mathbf{x}_j^\tau\rangle\Big)^{-4\beta}\sum_{i,j}|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|^2 \lesssim -\frac{\tau}{2N^2}\langle t\rangle^{-\frac{2\beta}{1-\beta}}\sum_{i,j}|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|^2.\end{aligned} \tag{1.23}$$

Observe that in this case, the enstrophy of the anticipated energy is bounded by the fluctuations of the anticipated positions (compared with velocity fluctuations in the 'regular' energy decay (1.2)). We close the last bound by hypocoercivity argument carried out in Section 5 which leads to the sub-exponential decay (1.17).

1.4. Anticipation Dynamics with Attractive–Repulsive Potential

For attractive–repulsive potentials, the large time behavior of (AT) is significantly more complicated, for the following two reasons:

- The topography of the total potential energy $\frac{1}{2N^2}\sum_{i,j} U(|\mathbf{x}_i - \mathbf{x}_j|)$ which includes multiple local minima with different geometric configurations could be very complicated, see e.g., [1,8,10,12,18,27,28,33] and the references therein.
- It is numerically observed in [20] that the decay of $E(t)$ is of order $\mathcal{O}(t^{-1})$. Therefore, one cannot expect for sub-exponential energy dissipation rate, $\dot{E}(t) \lesssim -\langle t\rangle^{-\eta} E(t)$, or that its hypocoercivity counterpart will hold.

Here we focus on the second difficulty, and give a first rigorous result in this direction.

Theorem 3. (Anticipation with repulsive–attractive potential) *Consider the 2D anticipated dynamics* (AT) *of* $N = 2$ *agents subject to repulsive–attractive potential which has a local minimum at* $r = r_0 > 0$ *where* $U''(r_0) = a > 0$. *Then there exists a constant* $\epsilon > 0$, *such that if the initial data is close enough to equilibrium,*

$$\big||\mathbf{x}_1(0) - \mathbf{x}_2(0)| - r_0\big|^2 + |\mathbf{v}_1(0) - \mathbf{v}_2(0)|^2 \leqslant \epsilon, \tag{1.24}$$

then the solution to (AT) *satisfies the following algebraic decay:*

$$\big||\mathbf{x}_1(t)-\mathbf{x}_2(t)|-r_0\big| \leqslant C\langle t\rangle^{-1}\ln^{1/2}\langle 1+t\rangle, \quad |\mathbf{v}_1(t)-\mathbf{v}_2(t)| \leqslant C\langle t\rangle^{-1/2}. \tag{1.25}$$

The proof, based on nonlinear hypocoercivity argument for the anticipated energy is carried out in Section 6.

Remark 1.4. The detailed description of the dynamics outlined in the proof, reveals that the radial component of the velocity, $v_r \lesssim \langle t\rangle^{-1}\ln^{1/2}\langle 1+t\rangle$, decays faster than its tangential part, $v_\theta \lesssim \langle t\rangle^{-1/2}$. Therefore, although the dynamics of (6.1) can be complicated at the beginning, it will finally settles as a circulation around the equilibrium, provided the initial data is close enough to equilibrium.

1.5. Anticipation Hydrodynamics

The large crowd (hydro-)dynamics associated with (AT) is described by density and momentum $(\rho, \rho\mathbf{u})$ governed by

$$\begin{cases} \rho_t + \nabla_{\mathbf{x}}\cdot(\rho\mathbf{u}) = 0 \\ (\rho\mathbf{u})_t + \nabla_{\mathbf{x}}\cdot(\rho\mathbf{u}\otimes\mathbf{u}) = -\displaystyle\int \nabla U(|\mathbf{x}^\tau - \mathbf{y}^\tau|)\,\mathrm{d}\rho(\mathbf{y}), \quad \mathbf{x}^\tau := \mathbf{x} + \tau\mathbf{u}(t,\mathbf{x}). \end{cases} \tag{1.26}$$

[3] The large-time flocking behavior of (1.26) is studied in terms of Lemma 4.1—a continuum version of the discrete lemma of means proved in Section 4. That is, we obtain a sub-linear time bound on the spread of supp $\rho(t,\cdot)$, which in turn is used

[3] Under a simplifying assumption of a mono-kinetic closure.

to control the enstrophy of the anticipated energy. In Section 7 we outline the proof of our last main result, which states that *if* (1.26) admits a global smooth solution then such smooth solution must flock, in agreement with the general paradigm for Cucker–Smale dynamics discussed in [25,39].

Theorem 4. (Anticipation hydrodynamics: smooth solutions must flock) *Let* $(\rho, \mathbf{u})$ *be a smooth solution of the anticipation hydrodynamics* (1.26) *with an attractive potential subject to a fat tail decay,* (1.15)*, of order* $\beta < \frac{1}{3}$*. Then there is sub-exponential decay of the anticipated energy fluctuations*

$$\int\int \left(\frac{1}{2m_0}|\mathbf{u}(\mathbf{x}) - \overline{\mathbf{u}}_0|^2 + U(|\mathbf{x}^\tau - \mathbf{y}^\tau|)\right) \mathrm{d}\rho(\mathbf{x})\,\mathrm{d}\rho(\mathbf{y}) \leqslant Ce^{-t^{1-\eta}}, \quad \eta = \frac{2\beta}{1-\beta} < 1. \tag{1.27}$$

It follows that there is large time flocking, with sub-exponential alignment

$$|\mathbf{u}(t,\mathbf{x}) - \overline{\mathbf{u}}_0|^2\,\mathrm{d}\rho(\mathbf{x}) \stackrel{t\to\infty}{\longrightarrow} 0, \qquad \overline{\mathbf{u}}_0 = \frac{1}{m_0}\int (\rho\mathbf{u})_0(\mathbf{x})\,\mathrm{d}\mathbf{x}, \quad m_0 = \int \rho_0(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

In proposition 7.1 we verify the existence of global smooth solution (and hence flocking) of the 1D system, (1.26), provided the threshold condition, $u_0'(x) \geqslant -C(\tau, m_0, a)$ holds, for a proper negative constant depending on τ, m_0 and the minimal convexity $a = \min U'' > 0$.

2. A Priori L^∞ Bounds for Confining Potentials

In this section we prove the uniform bounds asserted in (1.10) and (1.19), corresponding to the anticipation dynamics in (ΦU) and, respectively, (AT). Due to the momentum conservation (1.7), we may assume without loss of generality that in both cases $\overline{\mathbf{x}}(t) = \overline{\mathbf{v}}(t) \equiv 0$. This will always be assumed in the rest of this paper.

We recall the dynamics of (ΦU) assumes that U lies in the class of convex potentials, (1.5), and the dynamics of (AT) assumes a larger class of attractive potentials, (1.15). In fact, here we prove uniform bounds under a more general setup of *confining potentials*.

Assumption 2.1. (*Confining potentials*) There exist constants $a > 0$, $L \geqslant 0$ and β such that

$$U(r) \geqslant a\left(\langle r\rangle^{2-\beta} - L\right), \qquad 0 \leqslant \beta \leqslant 1. \tag{2.1}$$

Observe that attractive potentials (1.15) satisfy (2.1) with $L = 1$,

$$U(r) = \int_0^r U'(s)\,\mathrm{d}s \geqslant \int_0^r a\langle s\rangle^{-\beta} s\,\mathrm{d}s = \frac{a}{2-\beta}\left(\langle r\rangle^{2-\beta} - 1\right). \tag{2.2}$$

Thus, we have the increasing hierarchy of convex, attractive and confining potentials. The class of confining potentials is much larger, however, and it includes repulsive–attractive potentials (discussed in Section 6 below).

Remark 2.1. General confining potentials need not be positive. But taking into account that purely-attractive (and likewise — convex) potentials are positive, then we can improve (2.2),

$$r^2 \leqslant C \max\{r^\beta, 1\} U(r) \text{ for some } c > 0. \tag{2.3}$$

Indeed, for $r < 1$, (1.15) implies $U'(r) \gtrsim r$ and since $U(0) = 0$ then $U(r) \gtrsim r^2$. For $r \geqslant 1$ we have $r^{2-\beta} \leqslant C\big(\langle r\rangle^{2-\beta} - 1\big)$ with large enough C_1 $\big($e.g., $C_1 > \frac{\sqrt{2}}{\sqrt{2}-1}\big)$, and (2.2) implies (2.3) with $C = \frac{2-\beta}{a} C_1$.

Lemma 2.1. (Uniform bound on positions for (ΦU) system) *Consider the anticipation dynamics (ΦU) with bounded positive communication matrix* $0 \leqslant \Phi \leqslant \phi_+ \mathbb{I}_{d\times d}$*, and bounded confining potential* (1.4),(2.1)*. Then the solution* $\{(\mathbf{x}_i(t), \mathbf{v}_i(t))\}$ *satisfies the a priori estimate*

$$\max_i |\mathbf{x}_i(t)| \leqslant C_\infty \langle t\rangle^{\frac{2}{4-3\beta}}, \quad \max_i |\mathbf{v}_i(t)| \leqslant C_\infty \langle t\rangle^{\frac{2-\beta}{4-3\beta}}, \qquad 0 \leqslant \beta \leqslant 1. \tag{2.4}$$

Remark 2.2. Note that we require a positive communication matrix Φ but otherwise, we do not insist on any fat tail condition (1.6).

Proof. Our proof is based on the technique introduced in [38, §2.2], in which we prove uniform bounds in terms of the *particle energy*

$$E_i(t) := \frac{1}{2}|\mathbf{v}_i|^2 + \frac{1}{N}\sum_j U(|\mathbf{x}_i - \mathbf{x}_j|). \tag{2.5}$$

[4] We start by relating the local energy to the position of particle i: using (2.1) followed by Jensen inequality for the convex mapping[5] $\mathbf{x} \mapsto \langle \mathbf{x}\rangle^{2-\beta}$, we find that

$$\frac{E_i(t)}{a} \geqslant \frac{1}{N}\sum_j \big(\langle \mathbf{x}_i - \mathbf{x}_j\rangle^{2-\beta} - L\big) \geqslant \big\langle \frac{1}{N}\sum_j (\mathbf{x}_i - \mathbf{x}_j)\big\rangle^{2-\beta} - L = \langle \mathbf{x}_i\rangle^{2-\beta} - L.$$

It follows that the maximal spread of positions, $\max_i |\mathbf{x}_i(t)|$ does not exceed

$$X(t) \leqslant \left(\frac{E_\infty(t)}{a} + L\right)^{\frac{1}{2-\beta}}, \qquad X(t) := \max_i |\mathbf{x}_i(t)|, \quad E_\infty(t) := \max_i E_i(t). \tag{2.6}$$

[4] In fact E_i is not a proper particle energy, since $\sum_i E_i \neq NE$ (the pairwise potential is counted twice). However, it is the ratio of the kinetic energy and potential energy in (2.5) which is essential, as one would like to eliminate all the positive terms with indices i in (2.5), in order to avoid exponential growth of E_i.

[5] $(\langle r\rangle^{2-\beta})'' = -\beta(2-\beta)r^2\langle r\rangle^{-2-\beta} + (2-\beta)\langle r\rangle^{-\beta} = (2-\beta)\big((1-\beta)r^2+1\big)\langle r\rangle^{-2-\beta} > 0$ for $\beta \leqslant 1$.

Next we bound the energy dissipation rate of each particle. By (1.6) the communication matrices Φ_{ij} are non-negative and bounded,[6] $0 \leqslant \Phi_{ij} \leqslant \phi_+ \mathbb{I}_{d\times d}$, and since $\sum_j \mathbf{v}_j = 0$,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} E_i(t) &= \mathbf{v}_i \cdot \left(-\frac{1}{N}\sum_j \nabla U(|\mathbf{x}_i - \mathbf{x}_j|) + \frac{1}{N}\sum_j \Phi_{ij}(\mathbf{v}_j - \mathbf{v}_i) \right) \\
&\quad + \frac{1}{N}\sum_j \nabla U(|\mathbf{x}_i - \mathbf{x}_j|)\cdot(\mathbf{v}_i - \mathbf{v}_j) \\
&= \frac{1}{N}\sum_j \Phi_{ij}(\mathbf{v}_j - \mathbf{v}_i)\cdot \mathbf{v}_i - \frac{1}{N}\sum_j \nabla U(|\mathbf{x}_i - \mathbf{x}_j|)\cdot \mathbf{v}_j \\
&= -\frac{1}{2N}\sum_j \Phi_{ij}\mathbf{v}_i\cdot\mathbf{v}_i - \frac{1}{2N}\sum_j \Phi_{ij}(\mathbf{v}_j - \mathbf{v}_i)\cdot(\mathbf{v}_j - \mathbf{v}_i) \\
&\quad + \frac{1}{2N}\sum_j \Phi_{ij}\mathbf{v}_j\cdot\mathbf{v}_j - \frac{1}{N}\sum_j \Big(\nabla U(|\mathbf{x}_i - \mathbf{x}_j|) - \nabla U(|\mathbf{x}_i|)\Big)\cdot \mathbf{v}_j \\
&\leqslant \phi_+ E(0) + \sqrt{2E(0)}\left(\frac{1}{N}\sum_j \left|\nabla U(|\mathbf{x}_i - \mathbf{x}_j|) - \nabla U(|\mathbf{x}_i|)\right|^2 \right)^{1/2}.
\end{aligned}
\tag{2.7}
$$

To bound the sum on the right, we use the fact that D^2U is bounded, (1.4), followed by (2.6), to find that

$$
\begin{aligned}
&\frac{1}{N}\sum_j |\nabla U(|\mathbf{x}_i - \mathbf{x}_j|) - \nabla U(|\mathbf{x}_i|)|^2 \\
&\qquad \leqslant \sup_{\mathbf{x}} |D^2U(|\mathbf{x}|)|^2 \frac{1}{N}\sum_j |\mathbf{x}_j|^2 \leqslant \frac{A^2}{N}\sum_j |\mathbf{x}_j|^2 \\
&\qquad = \frac{A^2}{2N^2}\sum_{i,j} |\mathbf{x}_i - \mathbf{x}_j|^2 \leqslant 2^\beta A^2 \max_i |\mathbf{x}_i|^\beta \times \frac{1}{N^2}\sum_{i,j}|\mathbf{x}_i - \mathbf{x}_j|^{2-\beta} \\
&\qquad \leqslant 2^\beta A^2 X^\beta \frac{1}{2N^2}\sum_{i,j}|\mathbf{x}_i - \mathbf{x}_j|^{2-\beta} \leqslant 2^\beta A^2 X^\beta \frac{1}{2N^2}\sum_{i,j}\left(\frac{U(|\mathbf{x}_i - \mathbf{x}_j|)}{a} + L \right) \\
&\qquad \leqslant 2^\beta A^2 X^\beta \left(\frac{E(0)}{a} + \frac{L}{2} \right).
\end{aligned}
\tag{2.8}
$$

[6] Observe that we do not use the fat tail decay (1.6).

Therefore

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}E_i(t) &\leqslant \phi_+ E(0) + \sqrt{2E(0)}\left(2^\beta A^2 X^\beta\Big(\frac{E(0)}{a}+\frac{L}{2}\Big)\right)^{1/2} \\ &\leqslant \phi_+ E(0) + \sqrt{2E(0)}\left(2^\beta A^2\Big(\frac{E_\infty}{a}+L\Big)^{\frac{\beta}{2-\beta}}\Big(\frac{E(0)}{a}+\frac{L}{2}\Big)\right)^{1/2}\end{aligned} \tag{2.9}$$

and taking maximum among all i's we have[7]

$$\frac{\mathrm{d}}{\mathrm{d}t}E_\infty(t) \leqslant \phi_+ E(0) + \sqrt{2E(0)}\left(2^\beta A^2\Big(\frac{E_\infty(t)}{a}+L\Big)^{\frac{\beta}{2-\beta}}\Big(\frac{E(0)}{a}+\frac{L}{2}\Big)\right)^{1/2}. \tag{2.10}$$

Set $f(t) := E_\infty(t) + aL$, then the last inequality tells us $f' \leqslant C_1 + C_2 f^\alpha$ with $\alpha := \frac{\beta}{(4-2\beta)}$, and since by assumption $\alpha < 1/2$, then

$$f \lesssim \langle t\rangle^{\frac{1}{1-\alpha}} = \langle t\rangle^{\frac{2(2-\beta)}{4-3\beta}},$$

which implies the uniform bound on velocities in (2.4),

$$\max_i |\mathbf{v}_i(t)| \leqslant 2\sqrt{E_\infty(t)+aL} \lesssim \langle t\rangle^{\frac{2-\beta}{4-3\beta}}.$$

The uniform bound on positions, $\max_i |\mathbf{x}_i(t)|$, follows in view of (2.6). $\square$

Lemma 2.1 applies, in particular, to the anticipation dynamics (AT) with convex potential, so that D^2U is positive definite. Next, we prove uniform bounds for more general confining U's.

Lemma 2.2. (Uniform bound on anticipated positions) *Consider the anticipation dynamics* (AT) *with bounded confining potential* (1.4),(2.1). *Then the solution of the anticipation dynamics* (AT) *satisfies the a priori estimate*

$$\max_i |\mathbf{x}_i^\tau(t)| \leqslant C_\infty \langle t\rangle^{\frac{1}{2-2\beta}}, \qquad 0 \leqslant \beta < 1. \tag{2.11}$$

Remark 2.3. The a priori bound (2.11) is weaker than Lemma 2.1 and may not be optimal for β close to 1. We do not pursue an improved bound since it does not provide an increased range of β's for which Theorem 2 holds.

Proof of Lemma 2.2. The key quantity for proving the priori bound (2.11) is the 'anticipated particle energy' in (AT),

$$\mathcal{E}_i(t) := \frac{1}{2}|\mathbf{v}_i|^2 + \frac{1}{N}\sum_j U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|). \tag{2.12}$$

[7] To be pedantic at this point, the time derivative on the left (2.10) exists for almost all t's by Rademacher theorem, where it coincides with the maximal time derivatives on the left of $(2.9)_i$.

Similar to the previous proof, the confining property of U implies that the diameter of *anticipated* positions, $\max_i |\mathbf{x}_i^\tau(t)|$, does not exceed

$$\mathcal{X}(t) \leqslant \left(\frac{\mathcal{E}_\infty(t)}{a} + L\right)^{\frac{1}{2-\beta}}, \qquad \mathcal{X}(t) := \max_i |\mathbf{x}_i^\tau(t)|, \quad \mathcal{E}_\infty(t) := \max_i \mathcal{E}_i(t). \tag{2.13}$$

Next we bound the energy dissipation rate of each particle: since $\sum_j (\mathbf{v}_j + \tau\dot{\mathbf{v}}_j) = 0$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_i(t) &= \mathbf{v}_i \cdot \dot{\mathbf{v}}_i + \frac{1}{N}\sum_j \nabla U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|)\cdot(\mathbf{v}_i + \tau\dot{\mathbf{v}}_i - \mathbf{v}_j - \tau\dot{\mathbf{v}}_j)\\ &= -\tau|\dot{\mathbf{v}}_i|^2 - \frac{1}{N}\sum_j \nabla U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|)\cdot(\mathbf{v}_j + \tau\dot{\mathbf{v}}_j)\\ &= -\tau|\dot{\mathbf{v}}_i|^2 - \frac{1}{N}\sum_j \big(\nabla U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|) - \nabla U(|\mathbf{x}_i^\tau|)\big)\cdot(\mathbf{v}_j + \tau\dot{\mathbf{v}}_j).\end{aligned} \tag{2.14}$$

As before, the boundedness of D^2U followed by (2.13) to find imply

$$\frac{1}{N}\sum_j \big|\nabla U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|) - \nabla U(|\mathbf{x}_i^\tau|)\big|^2 \leqslant 2^\beta A^2 \mathcal{X}^\beta \left(\frac{\mathcal{E}(0)}{a} + \frac{L}{2}\right). \tag{2.15}$$

Inserting (2.15) into the RHS of (2.14) and adding the energy-enstrophy balance (1.1) we find[8]

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{E}(t) + \mathcal{E}_i(t))\\ &\quad\leqslant -\frac{\tau}{N}\sum_j |\dot{\mathbf{v}}_j|^2 - \tau|\dot{\mathbf{v}}_i|^2 + \frac{c}{N}\sum_j |\mathbf{v}_j|^2 + \frac{c\tau^2}{N}\sum_j |\dot{\mathbf{v}}_j|^2\\ &\qquad + \frac{1}{4cN}\sum_j \big|\nabla U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|) - \nabla U(|\mathbf{x}_i^\tau|)\big|^2\\ &\quad\leqslant -\frac{\tau(1-c\tau)}{N}\sum_j |\dot{\mathbf{v}}_j|^2 - \tau|\dot{\mathbf{v}}_i|^2 + 2c\big(\mathcal{E}(0) + aL\big) + \frac{1}{4c}2^\beta A^2 \mathcal{X}^\beta \left(\frac{\mathcal{E}(0)}{a} + \frac{L}{2}\right)\\ &\quad\leqslant \frac{2}{\tau}\big(\mathcal{E}(0) + aL\big) + \frac{\tau}{4}2^\beta A^2 \mathcal{X}^\beta \left(\frac{\mathcal{E}(0)}{a} + \frac{L}{2}\right), \qquad \text{(taking } c = 1/\tau).\end{aligned}$$

By taking maximum among all i,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{E}(t) + \mathcal{E}_\infty(t)) &\leqslant \frac{2}{\tau}\big(\mathcal{E}(0) + aL\big) + \frac{\tau}{4}2^\beta A^2 \mathcal{X}^\beta \left(\frac{\mathcal{E}(0)}{a} + \frac{L}{2}\right)\\ &\leqslant \frac{2}{\tau}\big(\mathcal{E}(0) + aL\big) + \frac{\tau}{4}2^\beta A^2 \left(\frac{\mathcal{E}_\infty(t)}{a} + L\right)^{\frac{\beta}{2-\beta}} \left(\frac{\mathcal{E}(0)}{a} + \frac{L}{2}\right)\end{aligned}$$

[8] Note that a confining potential need not be positive yet $U \geqslant -aL$ and hence $1/2N\sum_j |\mathbf{v}_j|^2 \leqslant \mathcal{E}(0) + aL$.

The last inequality tells us that $f(t) := \mathcal{E}(t)+\mathcal{E}_\infty(t)+aL$ satisfies $f' \leqslant C_1+C_2 f^\alpha$ with $\alpha := \frac{\beta}{2-\beta}$. Since by assumption $\alpha < 1$, then

$$f \lesssim \langle t\rangle^{\frac{1}{1-\alpha}} = \langle t\rangle^{\frac{2-\beta}{2-2\beta}},$$

and the uniform bound (2.11) follows in view of (2.13). $\square$

3. Anticipation with Convex Potentials and Positive Kernels

Equipped with the uniform bound (2.11), we turn to prove Theorem 1 by hypocoercivity argument. In [38] we use hypocoercivity to prove the flocking with *quadratic potentials*. Here, we make a judicious use of the assumed fat tail conditions, (1.6),(1.5), to extend these arguments for general convex potentials.

Proof of Theorem 1. We introduce the *modified energy*, $\widehat{E}(t)$, by adding a multiple of the cross term $1/N\sum_i \mathbf{x}_i\cdot\mathbf{v}_i$,

$$\widehat{E}(t) := E(t) + \frac{\epsilon(t)}{N}\sum_i \mathbf{x}_i(t)\cdot\mathbf{v}_i(t).$$

We claim that with a proper choice $\epsilon(t)$, the modified energy is positive definite. Indeed, the convex (hence attractive) potential satisfies the pointwise bound (2.3), and together with the uniform bound (1.10) they imply

$$\begin{aligned}|\epsilon(t)\frac{1}{N}\sum_i \mathbf{x}_i\cdot\mathbf{v}_i| &\leqslant \frac{1}{4N}\sum_i|\mathbf{v}_i|^2 + \frac{\epsilon^2(t)}{N}\sum_i|\mathbf{x}_i|^2 = \frac{1}{4N}\sum_i|\mathbf{v}_i|^2 + \frac{\epsilon^2(t)}{2N^2}\sum_{i,j}|\mathbf{x}_i-\mathbf{x}_j|^2\\
&\leqslant \frac{1}{4N}\sum_i|\mathbf{v}_i|^2 + \epsilon^2(t)C\max\left\{(2X(t))^\beta,1\right\}\frac{1}{2N^2}\sum_{i,j}U(|\mathbf{x}_i-\mathbf{x}_j|)\\
&\leqslant \frac{1}{4N}\sum_i|\mathbf{v}_i|^2 + \epsilon^2(t)C\big((2C_\infty)^\beta+1\big)\langle t\rangle^{\frac{2\beta}{4-3\beta}}\frac{1}{2N^2}\sum_{i,j}U(|\mathbf{x}_i-\mathbf{x}_j|).\end{aligned}$$

Therefore it suffices to choose

$$\epsilon(t) = \epsilon_0\langle t\rangle^{-\alpha}, \qquad \alpha > \frac{\beta}{4-3\beta}, \tag{3.1}$$

with small enough $\epsilon_0 > 0$ and *any* $\alpha > \frac{\beta}{4-3\beta}$ which is to be determined later, to guarantee $|\epsilon(t)/N\sum_i \mathbf{x}_i\cdot\mathbf{v}_i| \leqslant E(t)/2$, hence the positivity of $\widehat{E}(t) \geqslant E(t)/2 > 0$.

Next, we turn to verify the coercivity of $\widehat{E}(t)$. First notice that Lemma 2.1 implies the following L^∞ bound on $\mathbf{x}_i^\tau$:

$$|\mathbf{x}_i^\tau| \leqslant |\mathbf{x}_i| + \tau|\mathbf{v}_i| \leqslant (1+\tau)C_\infty\langle t\rangle^{\frac{2}{4-3\beta}}$$

This together with the assumed fat-tails of Φ and D^2U, imply their lower-bounds: by (1.6) $\Phi_{ij}(t)$ are bounded from below by

$$\Phi_{ij}(t) \geqslant \phi_-(t) := c\langle t\rangle^{-\frac{2\gamma}{4-3\beta}}, \tag{3.2}$$

and integrating (1.5) $U''(r) \geqslant a\langle r\rangle^{-\beta}$ twice, implies U has the lower bound (2.2)

$$U(|\mathbf{x}_i-\mathbf{x}_j|) \geqslant c|\mathbf{x}_i-\mathbf{x}_j|^2\langle|\mathbf{x}_i-\mathbf{x}_j|\rangle^{-\beta} \geqslant |\mathbf{x}_i-\mathbf{x}_j|^2\psi_-(t), \quad \psi_-(t) := c\langle t\rangle^{-\frac{2\beta}{4-3\beta}}. \tag{3.3}$$

Now, we turn to conduct hypocoercivity argument based on the energy estimate (1.2). To this end, we append to $E(t)$, a proper multiple of the cross term $\sum \mathbf{x}_i \cdot \mathbf{v}_i$, consult e.g., [16,38]. Using the symmetry of Φ_{ij}, the time derivative of this cross term is given by

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{N}\sum_i \mathbf{x}_i\cdot\mathbf{v}_i \\
&\quad= \frac{1}{N}\sum_i|\mathbf{v}_i|^2 + \frac{1}{N}\sum_i \mathbf{x}_i\cdot\left(-\frac{1}{N}\sum_j \nabla U(|\mathbf{x}_i-\mathbf{x}_j|) + \frac{1}{N}\sum_j \Phi_{ij}(\mathbf{v}_j-\mathbf{v}_i)\right) \\
&\quad= \frac{1}{2N^2}\sum_{i,j}|\mathbf{v}_i-\mathbf{v}_j|^2 - \frac{1}{2N^2}\sum_{i,j}(\mathbf{x}_i-\mathbf{x}_j)\cdot\nabla U(|\mathbf{x}_i-\mathbf{x}_j|) \\
&\qquad+ \frac{1}{2N^2}\sum_{i,j}\Phi_{ij}(\mathbf{v}_j-\mathbf{v}_i)\cdot(\mathbf{x}_i-\mathbf{x}_j).
\end{aligned} \tag{3.4}$$

We prepare three bounds. Noticing that since U is convex $U'(r)$ is increasing, hence $U(r) = \displaystyle\int_0^r U'(s)\,\mathrm{d}s \leqslant rU'(r)$ implies

$$\begin{aligned}
\frac{1}{2N^2}\sum_{i,j}(\mathbf{x}_i-\mathbf{x}_j)\cdot\nabla U(|\mathbf{x}_i-\mathbf{x}_j|) &= \frac{1}{2N^2}\sum_{i,j}U'(|\mathbf{x}_i-\mathbf{x}_j|)|\mathbf{x}_i-\mathbf{x}_j| \\
&\geqslant \frac{1}{2N^2}\sum_{i,j}U(|\mathbf{x}_i-\mathbf{x}_j|).
\end{aligned} \tag{3.5a}$$

Using the weighted Cauchy-Schwarz twice — weighted by the positive definite $0 < \Phi_{ij} \leqslant \phi_+$, and then by the yet-to-be determined $\kappa(t) > 0$,

$$\begin{aligned}
&\left|\frac{1}{2N^2}\sum_{i,j}\Phi_{ij}(\mathbf{v}_j-\mathbf{v}_i)\cdot(\mathbf{x}_i-\mathbf{x}_j)\right| \\
&\quad\leqslant \frac{\kappa}{4N^2}\sum_{i,j}\Phi_{ij}(\mathbf{v}_i-\mathbf{v}_j)\cdot(\mathbf{v}_i-\mathbf{v}_j) + \frac{1}{4\kappa N^2}\sum_{i,j}\Phi_{ij}(\mathbf{x}_i-\mathbf{x}_j)\cdot(\mathbf{x}_i-\mathbf{x}_j) \\
&\quad\leqslant \frac{\kappa(t)}{4N^2}\sum_{i,j}\Phi_{ij}(\mathbf{v}_i-\mathbf{v}_j)\cdot(\mathbf{v}_i-\mathbf{v}_j) + \frac{\phi_+}{4\kappa(t)N^2}\sum_{i,j}|\mathbf{x}_i-\mathbf{x}_j|^2
\end{aligned} \tag{3.5b}$$

Recall that with the choice of $\epsilon(t)=\epsilon_0\langle t\rangle^{-\alpha}$ in (3.1), we have $|\dot\epsilon(t)|\leqslant\alpha\dfrac{\epsilon(t)}{\langle t\rangle}$. We have the final bound

$$\begin{aligned}\left|\frac{\dot\epsilon(t)}{N}\sum_i \mathbf{x}_i\cdot\mathbf{v}_i\right| &\leqslant |\dot\epsilon(t)|\,\frac{1}{2\delta(t)N^2}\sum_{i,j}|\mathbf{x}_i|^2+|\dot\epsilon(t)|\,\frac{\delta(t)}{2N^2}\sum_{i,j}|\mathbf{v}_i|^2\\ &\leqslant \frac{\alpha}{2\delta(t)\langle t\rangle}\frac{\epsilon(t)}{2N^2}\sum_{i,j}|\mathbf{x}_i-\mathbf{x}_j|^2+\frac{\alpha\delta(t)}{2\langle t\rangle}\frac{\epsilon(t)}{2N^2}\sum_{i,j}|\mathbf{v}_i-\mathbf{v}_j|^2\end{aligned} \tag{3.5c}$$

Adding (3.4) to the energy decay (1.2) we find that the dissipation rate of the modified energy $\widehat{E}(t):=E(t)+{}^{\epsilon(t)}/_{N}\sum_i\mathbf{x}_i(t)\cdot\mathbf{v}_i(t))$ does not exceed, in view of (3.5a)—(3.5c),

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\widehat{E}(t) &\leqslant \left(-\tau+\frac{\kappa(t)}{2}\epsilon(t)\right)\frac{1}{2N^2}\sum_i\Phi_{ij}(\mathbf{v}_i-\mathbf{v}_j)\cdot(\mathbf{v}_i-\mathbf{v}_j)\\ &\quad+\left(\frac{\phi_+}{2\kappa(t)}\epsilon(t)+\frac{\alpha}{2\delta(t)\langle t\rangle}\epsilon(t)\right)\frac{1}{2N^2}\sum_{i,j}|\mathbf{x}_i-\mathbf{x}_j|^2\\ &\quad+\left(\epsilon(t)+\frac{\alpha\delta(t)}{2\langle t\rangle}\epsilon(t)\right)\frac{1}{2N^2}\sum_{i,j}|\mathbf{v}_i-\mathbf{v}_j|^2\\ &\quad-\epsilon(t)\frac{1}{2N^2}\sum_{i,j}U(|\mathbf{x}_i-\mathbf{x}_j|)\\ &=: I+II+III+IV.\end{aligned} \tag{3.6}$$

To complete the (hypo-)coercivity argument, we guarantee the terms on the right of (3.6) are negative. To this end, set $\kappa(t)=\tau/\epsilon(t)$ so the first pre-factor is less than $-{}^{\tau}/_{2}$ and hence

$$I\leqslant-\frac{\tau}{2}\phi_-(t)\frac{1}{2N^2}\sum_{i,j}|\mathbf{v}_i-\mathbf{v}_j|^2=-\frac{\tau}{2}\phi_-(t)\frac{1}{N}\sum_i|\mathbf{v}_i|^2,\qquad \kappa(t)=\frac{\tau}{\epsilon(t)}.$$

Next, we set $\delta(t)=\dfrac{\delta_0}{\epsilon(t)\langle t\rangle}$ so that the second pre-factor $\leqslant\left(\dfrac{\phi_+}{\tau}+\dfrac{\alpha}{2\delta_0}\right)\epsilon^2(t)$, hence the second term does not exceed, in view of (3.3)

$$II\leqslant\left(\frac{\phi_+}{\tau}+\frac{\alpha}{2\delta_0}\right)\frac{\epsilon^2(t)}{\psi_-(t)}\frac{1}{2N^2}\sum_{i,j}U(|\mathbf{x}_i-\mathbf{x}_j|),\qquad \delta(t)=\frac{\delta_0}{\epsilon(t)\langle t\rangle}.$$

With these choices of κ and δ, the third term does not exceed

$$III\leqslant\left(\epsilon(t)+\frac{\alpha\delta_0}{2\langle t\rangle^2}\right)\frac{1}{2N^2}\sum_{i,j}|\mathbf{v}_i-\mathbf{v}_j|^2=\left(\epsilon(t)+\frac{\alpha\delta_0}{2\langle t\rangle^2}\right)\frac{1}{N}\sum_i|\mathbf{v}_i|^2$$

We conclude that

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{E}(t) \leqslant \left(-\frac{\tau}{2}\phi_-(t) + \epsilon(t) + \frac{\alpha\delta_0}{2\langle t\rangle^2}\right)\frac{1}{N}\sum_i |\mathbf{v}_i|^2 \\ + \left(-\epsilon(t) + \left(\frac{\phi_+}{\tau} + \frac{\alpha}{2\delta_0}\right)\frac{\epsilon^2(t)}{\psi_-(t)}\right)\frac{1}{2N^2}\sum_{i,j} U(|\mathbf{x}_i - \mathbf{x}_j|). \tag{3.7}$$

Now set $\alpha \geqslant \dfrac{2\gamma}{4-3\beta}$ so that $\phi_-(t)$ decays no faster than $\epsilon(t)$; moreover, $\phi_-(t)$ decays no faster than $\langle t\rangle^{-2}$ since $6\beta + 2\gamma \leqslant 8$, and hence, with small enough $\epsilon_0, \delta_0 > 0$, the first pre-factor on the right of (3.7) does not exceed $-\tau\phi_-(t)/4$. Next, let $\alpha \geqslant \dfrac{2\beta}{4-3\beta}$ so that $\epsilon(t)/\psi_-(t)$ is bounded: hence, with small enough $\epsilon_0 \ll \delta_0$, the second pre-factor on the right of (3.7) does not exceed $-\epsilon(t)/2$. We conclude that

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{E}(t) \lesssim -\frac{\phi_-(t)}{N}\sum_i |\mathbf{v}_i|^2 - \frac{\epsilon(t)}{2N^2}\sum_{i,j} U(|\mathbf{x}_i - \mathbf{x}_j|) \lesssim -\langle t\rangle^{-\eta}\widehat{E}(t), \quad \eta = \frac{2\max\{\beta,\gamma\}}{4-3\beta}.$$

This implies the sub-exponential decay of $\widehat{E}$, and thus that of the comparable E. □

3.1. Flocking of Matrix-Based Cucker–Smale Dynamics

The Cucker–Smale model [13,14]

$$\begin{cases} \dot{\mathbf{x}}_i = \mathbf{v}_i \\ \dot{\mathbf{v}}_i = \dfrac{\tau}{N}\displaystyle\sum_{j=1}^{N} \Phi_{ij}(\mathbf{v}_j - \mathbf{v}_i), \end{cases} \tag{3.8}$$

is a special case of (ΦU) with no external potential $U = 0$, which formally corresponds to $\beta = 0$, in which case Theorem 1 would yield flocking for $\gamma < 1/2$. Here we justify these formalities and prove the velocity alignment of (3.8) (no spatial concentration effect, however), under a slightly larger threshold.

Proposition 3.1. *(*Alignment of (3.8) model with positive kernels*) Consider the Cucker–Smale dynamics* (3.8) *with symmetric matrix kernel* Φ *satisfying (compared to* (1.6) *with* $\mathbf{v} \equiv 0$*) for some constants* $0 < \phi_- < \phi_+$ *and* γ*,*

$$\phi_-\langle \mathbf{x}_i - \mathbf{x}_j\rangle^{-\gamma} \leqslant \Phi(\mathbf{x}_i, \mathbf{x}_j) \leqslant \phi_+, \qquad 0 \leqslant \gamma < 2/3. \tag{3.9}$$

Then there is sub-exponential decay of the energy fluctuations

$$\delta E(t) \leqslant Ce^{-t^{1-\eta}}, \qquad \eta = \frac{3\gamma}{2}, \qquad \delta E(t) := \frac{1}{2N}\sum_i |\mathbf{v}_i - \overline{\mathbf{v}}|^2. \tag{3.10}$$

It follows that there is a flock formation around the mean $\overline{\mathbf{x}}(t)$ *with large time velocity alignment at sub-exponential rate:*

$$\mathbf{v}_i(t) \to \overline{\mathbf{v}}_0, \quad \mathbf{x}_i(t) - \overline{\mathbf{x}}(t) \to \mathbf{x}_i^\infty, \qquad \overline{\mathbf{x}}(t) := \overline{\mathbf{x}}_0 + t\overline{\mathbf{v}}_0, \tag{3.11}$$

for some constants $\mathbf{x}_i^\infty$.

The proof is similar but follows a slightly different strategy from that of Theorem 1: we start by a priori estimate for the particle energy E_i, and then proceed to controlling the position $\langle \mathbf{x}_i \rangle$, which in turn gives enough energy dissipation.

Proof. Define the particle energy

$$E_i(t) := \frac{1}{2}|\mathbf{v}_i|^2, \quad E_\infty(t) = \max_i E_i(t). \tag{3.12}$$

Observe that this satisfies

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}E_i(t) =& \frac{1}{N}\sum_{j=1}^N \Phi_{ij}(\mathbf{v}_j - \mathbf{v}_i)\cdot \mathbf{v}_i \\ =& -\frac{1}{2N}\sum_{j=1}^N \Phi_{ij}(\mathbf{v}_j - \mathbf{v}_i)\cdot(\mathbf{v}_j - \mathbf{v}_i) - \frac{1}{2N}\sum_{j=1}^N \Phi_{ij}\mathbf{v}_i\cdot\mathbf{v}_i + \frac{1}{2N}\sum_{j=1}^N \Phi_{ij}\mathbf{v}_j\cdot\mathbf{v}_j \\ \leqslant& -\phi_-(t)\frac{|\mathbf{v}_i|^2}{2} + \phi_+ E(t),\end{aligned} \tag{3.13}$$

where $\phi_-(t)$, is a time-dependent lower-bound of the symmetric $\Phi(\mathbf{x}_i(t), \mathbf{x}_j(t))$ which can be taken, in view of (3.9),

$$\Phi(\mathbf{x}_i(t), \mathbf{x}_j(t)) \geqslant \phi_-(t), \qquad \phi_-(t) := \phi_- \cdot \langle 2X(t)\rangle^{-\gamma}, \qquad X(t) = \max_i |\mathbf{x}_i(t)|. \tag{3.14}$$

Taking i as the particle with the largest E_i, then

$$\frac{\mathrm{d}}{\mathrm{d}t}E_\infty(t) \leqslant -\phi_-(t)E_\infty(t) + \phi_+ E(t) \leqslant -\phi_-(t)E_\infty(t) + \phi_+ E(0). \tag{3.15}$$

This implies

$$E_\infty(t) \leqslant E_\infty(0) + \phi_+ E(0)t. \tag{3.16}$$

Next, we notice that

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) \leqslant \max_i |\mathbf{v}_i| \leqslant \sqrt{2E_\infty(t)} \leqslant \sqrt{2(E_\infty(0) + \phi_+ E(0)t)}. \tag{3.17}$$

This yields $X(t) \leqslant C\langle t\rangle^{3/2}$, and in view of the fat tail (1.6), we end with the lower bound

$$\phi_-(t) \geqslant c\langle t\rangle^{-\frac{3\gamma}{2}} \ \text{ with } \ c = (2C)^{-\gamma}.$$

Therefore the energy dissipation (1.2) gives

$$\frac{\mathrm{d}}{\mathrm{d}t}E(t) \leqslant -\phi_-(t)E(t) \leqslant -c\langle t\rangle^{-\frac{3\gamma}{2}}E(t), \tag{3.18}$$

which implies the sub-exponential decay (3.10), $E(t) \leqslant E(0)e^{-c\langle t\rangle^{1-\eta}}$ with $\eta = \frac{3\gamma}{2} < 1$.

Equipped with this sub-exponential decay of $E(t)$, we revisit (3.15): this time it implies

$$
\begin{aligned}
E_\infty(t) &\leqslant e^{-\Phi_-(t)}E_\infty(0)+\phi_+\int_0^t e^{-\Phi_-(t-s)}E(s)\,\mathrm{d}s\\
&\leqslant C\langle t\rangle e^{-c\langle t\rangle^{1-\eta}}, \qquad \Phi_-(t):=\int_0^t \phi_-(s)\,\mathrm{d}s \geqslant c\langle t\rangle^{1-\eta}
\end{aligned}
\tag{3.19}
$$

This shows the sub-exponential decay of the kinetic energy of each agent, $E_\infty(t)$, independently of N, $|\mathbf{v}_i(t)-\overline{\mathbf{v}}_0|\to 0$. As a result, $\mathbf{x}_i(t)=\mathbf{x}_i(0)+\displaystyle\int_0^t \mathbf{v}_i(s)\,\mathrm{d}s$, converges as $t\to\infty$ since the last integral converges absolutely in view of $|\mathbf{v}_i(t)|\leqslant\sqrt{2E_\infty(t)}$. $\square$

4. Local Versus Global Weighted Means

In this section we prove Lemma 1.1 about discrete means, which in turn will be used in proving the hypocoercivity of the discrete anticipation dynamics (AT). We also treat the corresponding continuum lemma of means in Lemma 4.1, which is used in the hypocoercivity of the hydrodynamic anticipation model (1.26).

We begin with the proof of the Lemma of means 1.1:

Proof of Lemma 1.1. We first treat the scalar setup, in which case we may assume, without loss of generality that the z_i's are rearranged in a decreasing order, $z_1\geqslant z_2\geqslant\cdots\geqslant z_N$, and have a zero mean $\sum_j z_j=0$, and we need to bound the fluctuations on the left of (1.21) which amount to $\frac{1}{N^2}\sum_{i,j}|z_i-z_j|^2=\frac{2}{N}\sum_i|z_i|^2$. Let i_0 be the smallest index i such that

$$
\frac{1}{N}\sum_{j=1}^{i-1}z_j \geqslant \frac{\lambda}{2(\Lambda-\lambda)}z_i. \tag{4.1}
$$

Noticing that if i_+ is the maximal index of the positive entries, $z_{i\leqslant i_+}\geqslant 0$, then (4.1) clearly holds for $i>i_+$ (where LHS $>0>$ RHS), hence $i_0\leqslant i_+$, and since LHS is increasing (for $i\leqslant i_+$) and RHS is decreasing, see Fig. 1 below, (4.1) holds for all $i\geqslant i_0$,

$$
\frac{1}{N}\sum_{j=1}^{i-1}z_j \geqslant \frac{\lambda}{2(\Lambda-\lambda)}z_i, \quad i\geqslant i_0. \tag{4.2}
$$

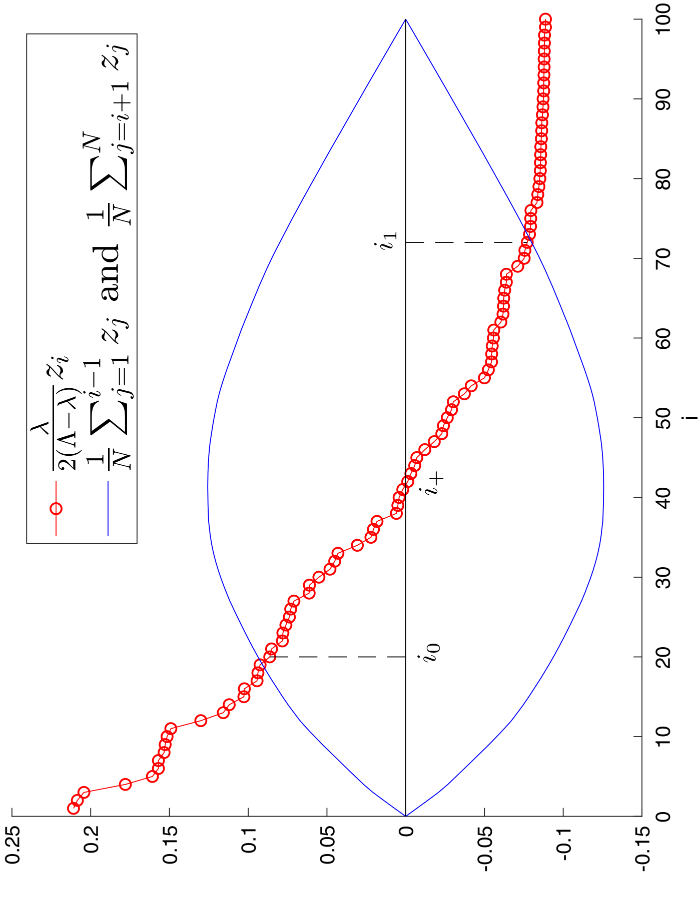

Fig. 1. Comparison of means

For $i < i_0$ we have $z_i \geqslant 0$, hence

$$
\begin{aligned}
\frac{1}{N}\sum_j c_{ij}(z_i - z_j) &= \frac{1}{N}\sum_{j=1}^{i} c_{ij}(z_i - z_j) + \frac{1}{N}\sum_{j=i+1}^{N} c_{ij}(z_i - z_j) \\
&\geqslant \frac{\Lambda}{N}\sum_{j=1}^{i}(z_i - z_j) + \frac{\lambda}{N}\sum_{j=i+1}^{N}(z_i - z_j) \\
&= -\frac{\Lambda}{N}\sum_{j=1}^{i} z_j + \frac{\lambda}{N}\sum_{j=1}^{i} z_j + \frac{\Lambda}{N}\sum_{j=1}^{i} z_i + \frac{\lambda}{N}\sum_{j=i+1}^{N} z_i \\
&\geqslant -\frac{\Lambda-\lambda}{N}\sum_{j=1}^{i-1} z_j + \lambda z_i, \qquad i < i_0,
\end{aligned}
\tag{4.3}
$$

and therefore, by the minimality of i_0 in (4.2)

$$
\frac{1}{N}\sum_j c_{ij}(z_i - z_j) \geqslant -\frac{\Lambda-\lambda}{2(\Lambda-\lambda)}\lambda z_i + \lambda z_i = \frac{\lambda}{2} z_i \geqslant 0, \qquad i < i_0.
$$

It follows that

$$
\frac{1}{N}\sum_{i=1}^{i_0-1}\left|\frac{1}{N}\sum_j c_{ij}(z_i - z_j)\right|^2 \geqslant \frac{\lambda^2}{4}\frac{1}{N}\sum_{i=1}^{i_0-1} z_i^2. \tag{4.4}
$$

Else, for $i \geqslant i_0$, (4.2) implies

$$
z_i \leqslant z_{i_0} \leqslant \frac{2(\Lambda-\lambda)}{\lambda}\frac{1}{N}\sum_{j=1}^{i_0-1} z_j, \quad i \geqslant i_0.
$$

It follows that for all positive entries, $0 \leqslant z_i \leqslant z_{i_0}$,

$$
\frac{1}{N}\sum_{i=i_0}^{i_+} z_i^2 \leqslant z_{i_0}^2 \leqslant \frac{4(\Lambda-\lambda)^2}{\lambda^2}\frac{1}{N^2}\left(\sum_{j=1}^{i_0-1} z_j\right)^2 \leqslant \frac{4(\Lambda-\lambda)^2}{\lambda^2}\frac{1}{N}\sum_{j=1}^{i_0-1} z_j^2 \tag{4.5}
$$

Therefore, by (4.5),(4.4),

$$
\begin{aligned}
\frac{1}{N}\sum_{z_i\geqslant 0} z_i^2 &= \frac{1}{N}\sum_{i=1}^{i_0-1} z_i^2 + \frac{1}{N}\sum_{i=i_0}^{i_+} z_i^2 \\
&\leqslant \left(1 + \frac{4(\Lambda-\lambda)^2}{\lambda^2}\right)\frac{1}{N}\sum_{j=1}^{i_0-1} z_j^2 \\
&\leqslant \frac{4}{\lambda^2}\left(1 + 4\Big(\frac{\Lambda}{\lambda} - 1\Big)^2\right)\frac{1}{N}\sum_i\left|\frac{1}{N}\sum_j c_{ij}(z_i - z_j)\right|^2.
\end{aligned}
\tag{4.6}
$$

Now apply (4.6) to z_i replaced by $-z_i$, to find the same upper-bound on the negative entries

$$\frac{1}{N}\sum_{z_i\leqslant 0} z_i^2 \leqslant \frac{4}{\lambda^2}\left(1+4\Big(\frac{\Lambda}{\lambda}-1\Big)^2\right)\frac{1}{N}\sum_i\Big|\frac{1}{N}\sum_j c_{ij}(z_i-z_j)\Big|^2. \tag{4.7}$$

The scalar result follows from (4.6),(4.7). For the d-dimensional case, notice given that $\sum_i \mathbf{z}_i = 0$

$$\sum_{i,j}|\mathbf{z}_i-\mathbf{z}_j|^2 = 2\sum_i|\mathbf{z}_i|^2 = \sum_{k=1}^{d}\sum_i |z_i^k|^2,$$

$$\sum_i\Big|\frac{1}{N}\sum_j c_{ij}(\mathbf{z}_i-\mathbf{z}_j)\Big|^2 = \sum_{k=1}^{d}\sum_i\Big|\frac{1}{N}\sum_j c_{ij}(z_i^k-z_j^k)\Big|^2$$

where superscript stands for component. Therefore the conclusion follows by applying the scalar result to the components of $\mathbf{z}_i = \{z_i^k\}_k$ for each fixed k, ending with the same constant $C(\Lambda,\lambda)$ which is independent of d. □

Next, we extend the result from the discrete framework to the continuum.

Lemma 4.1. (Local and global means are comparable) *Let* $(\Omega,\mathcal{F},\mu)$ *be a probability measure, and* $\mathbf{X}:\Omega\to\mathbb{R}^d$ *be a random variable with finite second moment,* $\displaystyle\int |\mathbf{X}(\omega')|^2\,\mathrm{d}\mu(\omega') < \infty$. *Then, for any measurable* $c = c(\omega,\omega') : \Omega\times\Omega \mapsto \mathbb{R}$ *satisfying*

$$0<\lambda\leqslant c(\omega,\omega')\leqslant\Lambda,$$

there holds

$$\iint |\mathbf{X}(\omega)-\mathbf{X}(\omega')|^2\,\mathrm{d}\mu(\omega)\,\mathrm{d}\mu(\omega') \leqslant 32\frac{\Lambda^2}{\lambda^4}\int\Big|\int c(\omega,\omega')\big(\mathbf{X}(\omega)-\mathbf{X}(\omega')\big)\,\mathrm{d}\mu(\omega')\Big|^2\,\mathrm{d}\mu(\omega).$$

Observe that the quantity on the left can be equally expressed as the amount of fluctuations relative to the mean $\overline{\mathbf{X}}$

$$\iint |\mathbf{X}(\omega)-\mathbf{X}(\omega')|^2\,\mathrm{d}\mu(\omega)\,\mathrm{d}\mu(\omega') = 2\int |\mathbf{X}(\omega)-\overline{\mathbf{X}}|^2\,\mathrm{d}\mu(\omega), \qquad \overline{\mathbf{X}} := \int \mathbf{X}(\omega')\,\mathrm{d}\mu(\omega'),$$

and in particular, Dirac measure $\mathrm{d}\mu = \displaystyle\frac{1}{N}\sum_j \delta(\mathbf{x}-\mathbf{z}_j)$ recovers the discrete case of Lemma 1.1.

Proof. We first prove the 1D case, for which the map $\mathbf{X}$, denoted by X, may assume a zero mean, $\overline{X} = 0$, without loss of generality. Take ω with $X(\omega) := x \geqslant 0$, then

$$\begin{aligned}
&\int c(\omega,\omega')(x - X(\omega'))\,\mathrm{d}\mu(\omega') \\
&\qquad = -\int_{\omega':X(\omega')>x} c(\omega,\omega')(X(\omega') - x)\,\mathrm{d}\mu(\omega') + \int_{\omega':X(\omega')\leqslant x} c(\omega,\omega')(x - X(\omega'))\,\mathrm{d}\mu(\omega') \\
&\qquad \geqslant -\Lambda\int_{\omega':X(\omega')>x} (X(\omega') - x)\,\mathrm{d}\mu(\omega') + \lambda\int_{\omega':X(\omega')\leqslant x} (x - X(\omega'))\,\mathrm{d}\mu(\omega') \\
&\qquad = -(\Lambda-\lambda)\int_{\omega':X(\omega')>x} (X(\omega') - x)\,\mathrm{d}\mu(\omega') + \lambda x \\
&\qquad \geqslant -(\Lambda-\lambda)\int_{\omega':X(\omega')>x} X(\omega')\,\mathrm{d}\mu(\omega') + \lambda x
\end{aligned}$$

Let

$$x_0 := \sup\{x : Y(x) \geqslant 0\}, \qquad Y(x) = \int_{\omega':X(\omega')>x} X(\omega')\,\mathrm{d}\mu(\omega') - \frac{\lambda}{2(\Lambda-\lambda)}x. \tag{4.8}$$

Since $\lim_{x\to\infty} Y(x) = -\infty$ and $Y(0) \geqslant 0$, x_0 is finite and non-negative. It is clear that $Y(x)$ is decreasing and right-continuous. Therefore $Y(x) \geqslant 0$ for any $x < x_0$, and $Y(x) \leqslant 0$ for any $x \geqslant x_0$.

If $x \geqslant x_0$, then

$$\int c(\omega,\omega')(x - X(\omega'))\,\mathrm{d}\mu(\omega') \geqslant -(\Lambda-\lambda)\frac{\lambda}{2(\Lambda-\lambda)}x + \lambda x = \frac{\lambda}{2}x \tag{4.9}$$

Thus taking square and integrating in ω with $x = X(\omega) \geqslant x_0 \geqslant 0$ gives

$$\int_{\omega:X(\omega)\geqslant x_0}\left(\int c(\omega,\omega')(X(\omega) - X(\omega'))\,\mathrm{d}\mu(\omega')\right)^2 \mathrm{d}\mu(\omega) \geqslant \frac{\lambda^2}{4}\int_{\omega:X(\omega)\geqslant x_0} X^2(\omega)\,\mathrm{d}\mu(\omega). \tag{4.10}$$

Then we claim that the above integral on $\{\omega : X(\omega) \geqslant x_0\}$ is enough to get the conclusion. Notice that for any $\epsilon > 0$, one has $Y(x_0 - \epsilon) \geqslant 0$, i.e.,

$$x_0 - \epsilon \leqslant \frac{2(\Lambda-\lambda)}{\lambda}\int_{\omega:X(\omega)>x_0-\epsilon} X(\omega')\,\mathrm{d}\mu(\omega') \tag{4.11}$$

and therefore

$$\begin{aligned}
(x_0 - \epsilon)^2 &\leqslant \frac{4(\Lambda-\lambda)^2}{\lambda^2}\left(\int_{\omega:X(\omega)>x_0-\epsilon} X(\omega')\,\mathrm{d}\mu(\omega')\right)^2 \\
&\leqslant \frac{4(\Lambda-\lambda)^2}{\lambda^2}\left(\int_{\omega:X(\omega)>x_0-\epsilon} X(\omega')^2\,\mathrm{d}\mu(\omega')\right)\left(\int_{\omega:X(\omega)>x_0-\epsilon} \mathrm{d}\mu(\omega')\right) \\
&\leqslant \frac{4(\Lambda-\lambda)^2}{\lambda^2}\int_{\omega:X(\omega)>x_0-\epsilon} X(\omega')^2\,\mathrm{d}\mu(\omega')
\end{aligned}$$

Taking $\epsilon \to 0$, noticing that the RHS integral domain $\{\omega : X(\omega) > x_0 - \epsilon\}$ converges to $\{\omega : X(\omega) \geqslant x_0\}$, we get

$$x_0^2 \leqslant \frac{4(\Lambda - \lambda)^2}{\lambda^2} \int_{\omega:X(\omega)\geqslant x_0} X(\omega')^2 \,\mathrm{d}\mu(\omega') \tag{4.12}$$

Thus, using (4.12) and (4.10) we find

$$\begin{aligned}
&\int_{\omega':X(\omega')\geqslant 0} X(\omega')^2 \,\mathrm{d}\mu(\omega') \\
&= \int_{\omega':0\leqslant X(\omega')<x_0} X(\omega')^2 \,\mathrm{d}\mu(\omega') + \int_{\omega':X(\omega')\geqslant x_0} X(\omega')^2 \,\mathrm{d}\mu(\omega') \\
&\leqslant \left(\frac{4(\Lambda - \lambda)^2}{\lambda^2} + 1\right) \int_{\omega':X(\omega')\geqslant x_0} X(\omega')^2 \,\mathrm{d}\mu(\omega') \\
&\leqslant \left(\frac{4(\Lambda - \lambda)^2}{\lambda^2} + 1\right) \frac{4}{\lambda^2} \int_{[x_0,\infty)} \left(\int c(\omega,\omega')(X(\omega) - X(\omega'))\,\mathrm{d}\mu(\omega')\right)^2 \mathrm{d}\mu(\omega).
\end{aligned}$$

Apply the last bound with $X(\cdot)$ replaced by $-X(\cdot)$ to find that the $\int_{\omega':X(\omega')\leqslant 0} X(\omega')^2 \,\mathrm{d}\mu(\omega')$ satisfies the same bound on the right, which completes the scalar part of the proof. For the d-dimensional case with $\mathbf{X} = (X_1, \ldots, X_d)$, notice that

$$\int |\mathbf{X}(\omega')|^2 \,\mathrm{d}\mu(\omega') = \sum_{k=1}^{d} \int |X_k(\omega')|^2 \,\mathrm{d}\mu(\omega'),$$

and similarly,

$$\begin{aligned}
&\int \left|\int c(\omega,\omega')(\mathbf{X}(\omega) - \mathbf{X}(\omega'))\,\mathrm{d}\mu(\omega')\right|^2 \mathrm{d}\mu(\omega) \\
&= \sum_{k=1}^{d} \int \left|\int c(\omega,\omega')(X_k(\omega) - X_k(\omega'))\,\mathrm{d}\mu(\omega')\right|^2 \mathrm{d}\mu(\omega)
\end{aligned}$$

Applying the 1D result to the random variable X_k gives

$$\int |X_k(\omega')|^2 \,\mathrm{d}\mu(\omega') \leqslant C(\lambda, \Lambda) \int \left|\int c(\omega,\omega')(X_k(\omega) - X_k(\omega'))\,\mathrm{d}\mu(\omega')\right|^2 \mathrm{d}\mu(\omega). \tag{4.13$_k$}$$

Summing $(4.13)_k$ recovers the desired result with the constant $C(\Lambda, \lambda)$ independent of d. □

5. Anticipation Dynamics with Attractive Potentials

In this section we prove the flocking behavior of (AT) asserted in Theorem 2. Here, we treat the larger class of attractive potentials, thus extending the case of convex potentials of Theorem 1. The starting point is the anticipated energy balance (1.1)

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) = -\frac{\tau}{N}\sum_i |\dot{\mathbf{v}}_i|^2.$$

Remark 5.1. We note in passing that the first-order system

$$\dot{\mathbf{x}}_i = -\frac{1}{N}\sum_{j=1}^{N}\nabla U(|\mathbf{x}_i - \mathbf{x}_j|)$$

satisfies an energy estimate, reminiscent of the energy-enstrophy balance in the anticipation dynamics (AT),

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2N^2}\sum_{i,j} U(|\mathbf{x}_i - \mathbf{x}_j|) = -\frac{1}{N}\sum_i |\dot{\mathbf{x}}_i|^2.$$

Proof of Theorem 2. We aim to conduct a hypocoercivity argument to complement the anticipated energy estimate (1.1). To this end, we use the 'anticipated' cross term

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Big(-\frac{1}{N}\sum_i \mathbf{x}_i^\tau\cdot\mathbf{v}_i\Big) &= \frac{1}{N}\sum_i \big(-(\mathbf{v}_i+\tau\dot{\mathbf{v}}_i)\cdot\mathbf{v}_i - \mathbf{x}_i^\tau\cdot\dot{\mathbf{v}}_i\big)\\
&\leqslant \frac{1}{N}\sum_i\Big(-|\mathbf{v}_i|^2 + \tau\big(\frac{\tau}{2}|\dot{\mathbf{v}}_i|^2 + \frac{1}{2\tau}|\mathbf{v}_i|^2\big) + \frac{1}{2}\big(|\mathbf{x}_i^\tau|^2 + |\dot{\mathbf{v}}_i|^2\big)\Big)\\
&\leqslant -\frac{1}{2}\frac{1}{N}\sum_i |\mathbf{v}_i|^2 + \frac{1}{2}\frac{1}{N}\sum_i |\mathbf{x}_i^\tau|^2 + \frac{\tau^2+1}{2}\frac{1}{N}\sum_i |\dot{\mathbf{v}}_i|^2.
\end{aligned} \tag{5.1}$$

Consider the modified anticipated energy $\widehat{\mathcal{E}}(t) := \mathcal{E}(t) - \epsilon(t)\frac{1}{N}\sum_i \mathbf{x}_i^\tau\cdot\mathbf{v}_i$, where $\epsilon(t) > 0$ is small, decreasing, and is yet to be chosen. We first need to guarantee that this modified energy is positive definite, and in fact — comparable to $\mathcal{E}(t)$,

$$\Big|\epsilon(t)\frac{1}{N}\sum_i \mathbf{x}_i^\tau\cdot\mathbf{v}_i\Big| \leqslant \frac{\mathcal{E}(t)}{2} = \frac{1}{4N}\sum_i |\mathbf{v}_i|^2 + \frac{1}{4N^2}\sum_{i,j} U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|). \tag{5.2}$$

Indeed, notice that

$$\begin{aligned}
|\epsilon(t)\frac{1}{N}\sum_i \mathbf{x}_i^\tau\cdot\mathbf{v}_i| &\leqslant \frac{1}{4N}\sum_i|\mathbf{v}_i|^2 + \epsilon^2(t)\frac{1}{N}\sum_i|\mathbf{x}_i^\tau|^2\\
&\leqslant \frac{1}{4N}\sum_i|\mathbf{v}_i|^2 + \epsilon(t)^2 C\max\big\{(2\mathcal{X})^\beta, 1\big\}\frac{1}{2N^2}\sum_{i,j}U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|)\\
&\leqslant \frac{1}{4N}\sum_i|\mathbf{v}_i|^2 + \epsilon^2(t)C\big((2C_\infty)^\beta + 1\big)\langle t\rangle^{\frac{\beta}{2-2\beta}}\frac{1}{2N^2}\sum_{i,j}U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|).
\end{aligned}$$

The second inequality is obtained similarly to (2.15) and using (2.3), and the third inequality uses Lemma 2.2. Therefore it suffices to choose

$$\epsilon(t) = \epsilon_0(10+t)^{-\alpha}, \quad \alpha \geqslant \frac{\beta}{4-4\beta} \tag{5.3}$$

with small enough ϵ_0 to guarantee (5.2).

We now turn to verify the (hypo-)coercivity of $\widehat{\mathcal{E}}(t)$,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathcal{E}(t) - \epsilon(t)\frac{1}{N}\sum_i \mathbf{x}_i^\tau \cdot \mathbf{v}_i\Big) \\
&\quad\leqslant -\frac{\tau}{N}\sum_i |\dot{\mathbf{v}}_i|^2 - \frac{\epsilon(t)}{2}\frac{1}{N}\sum_i |\mathbf{v}_i|^2 \\
&\qquad + \epsilon(t)\left(\frac{1}{2}\frac{1}{N}\sum_i |\mathbf{x}_i^\tau|^2 + \frac{\tau^2+1}{2}\frac{1}{N}\sum_i |\dot{\mathbf{v}}_i|^2\right) + |\dot{\epsilon}(t)|\frac{1}{N}\sum_i |\mathbf{x}_i^\tau \cdot \mathbf{v}_i| \\
&\quad\leqslant -\left(\tau - \epsilon(t)\frac{\tau^2+1}{2}\right)\frac{1}{N}\sum_i |\dot{\mathbf{v}}_i|^2 \\
&\qquad - \frac{\epsilon(t) - |\dot{\epsilon}(t)|}{2}\frac{1}{N}\sum_i |\mathbf{v}_i|^2 + \frac{\epsilon(t) + |\dot{\epsilon}(t)|}{2}\frac{1}{N}\sum_i |\mathbf{x}_i^\tau|^2
\end{aligned} \tag{5.4}$$

The first pre–factor on the right of (5.4) is less than $-\frac{\tau}{2}$ for small enough ϵ_0. The second pre-factor is negative since

$$|\dot{\epsilon}(t)| = \alpha\epsilon_0(10+t)^{-\alpha-1} \leqslant \frac{\alpha}{10}\epsilon(t).$$

It remains to control the last term on the right of (5.4). To this end we recall that U is assumed attractive, $U'(r)/r \geqslant \langle r\rangle^{-\beta}$, hence, by Lemma 2.2,

$$A \geqslant \frac{U'(r_{ij}^\tau)}{r_{ij}^\tau} \geqslant a\langle r_{ij}^\tau\rangle^{-\beta} \geqslant c\langle t\rangle^{-\frac{\beta}{2-2\beta}}, \qquad r_{ij}^\tau = |\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|.$$

We now invoke Lemma 1.1, which implies

$$\frac{1}{N}\sum_i |\dot{\mathbf{v}}_i|^2 = \frac{1}{N}\sum_i \left|\frac{1}{N}\sum_j \frac{U'(r_{ij}^\tau)}{r_{ij}^\tau}(\mathbf{x}_i^\tau - \mathbf{x}_j^\tau)\right|^2 \geqslant c\langle t\rangle^{-\eta}\frac{1}{N^2}\sum_{i,j}|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|^2, \quad \eta = \frac{2\beta}{1-\beta} \tag{5.5}$$

Therefore, the last term on the right of (5.4) does not exceed $\lesssim \epsilon(t)\langle t\rangle^\eta \frac{1}{N}\sum_i |\dot{\mathbf{v}}_i|^2$ and choosing $\epsilon(t)$ as in (5.3) with $\alpha = \eta$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{E}}(t) \leqslant -\frac{\tau}{4}\frac{1}{N}\sum_i |\dot{\mathbf{v}}_i|^2 - \frac{\epsilon_0(10+t)^{-\eta}}{4}\frac{1}{N}\sum_i |\mathbf{v}_i|^2. \tag{5.6}$$

[9] As before, since U is bounded, it has at most quadratic growth,

$$\frac{1}{2N^2}\sum_{i,j} U(|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|) \leqslant A\frac{1}{2N^2}\sum_{i,j}|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|^2 \leqslant C\langle t\rangle^{\frac{2\beta}{1-\beta}}\frac{1}{N}\sum_i |\dot{\mathbf{v}}_i|^2 = C\langle t\rangle^{\eta}\frac{1}{N}\sum_i |\dot{\mathbf{v}}_i|^2,$$

and we conclude the sub-exponential decay

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{E}}(t) \leqslant -c\langle t\rangle^{-\eta}\widehat{\mathcal{E}}(t) \rightsquigarrow \widehat{\mathcal{E}}(t) \leqslant Ce^{-t^{1-\eta}}, \tag{5.7}$$

which implies the same decay rate of $\mathcal{E}(t)$. $\square$

6. Anticipation Dynamics with Repulsive–Attractive Potential

In this section we prove Theorem 3. The assumption $\sum_i \mathbf{x}_i = \sum_i \mathbf{v}_i = 0$ amounts to saying that $\mathbf{x} := \mathbf{x}_1 = -\mathbf{x}_2$, $\mathbf{v} := \mathbf{v}_1 = -\mathbf{v}_2$. Replacing $U(|\mathbf{x}|)$ by $U(2|\mathbf{x}|)$ and r_0 by $r_0/2$, (AT) becomes

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{v} \\ \dot{\mathbf{v}} = -\nabla U(|\mathbf{x}^\tau|) \end{cases} \tag{6.1}$$

where $U(r)$ has a local minimum at $r = r_0 > 0$ with $U''(r_0) = a > 0$.

We use polar coordinates

$$\begin{cases} x_1^\tau = r\cos\theta \\ x_2^\tau = r\sin\theta \end{cases} \tag{6.2}$$

and

$$\begin{cases} v_r = v_1\cos\theta + v_2\sin\theta \\ v_\theta = -v_1\sin\theta + v_2\cos\theta \end{cases}. \tag{6.3}$$

Then (6.1) becomes

$$\begin{cases} \dot{r} = v_r - \tau U'(r) \\ \dot{\theta} = \dfrac{v_\theta}{r} \\ \dot{v}_r = -U'(r) + \dfrac{v_\theta^2}{r} \\ \dot{v}_\theta = \dfrac{-v_r v_\theta}{r} \end{cases}. \tag{6.4}$$

We will focus on perturbative solutions near $r = r_0$, $v_r = v_\theta = 0$. Write $r := r_0 + \delta_r$, and there hold the approximations

$$U(r) \approx \frac{a}{2}\delta_r^2, \quad U'(r) \approx a\delta_r, \quad U''(r) \approx a \tag{6.5}$$

[9] We may assume without loss of generality, that the two time invariant moments vanish, $\sum \mathbf{x}_i = \sum \mathbf{v}_i = 0$, and hence $\frac{1}{N}\sum_i |\mathbf{x}_i^\tau|^2 = \frac{1}{2N^2}\sum_{i,j}|\mathbf{x}_i^\tau - \mathbf{x}_j^\tau|^2$.

Observe that our assumed initial configuration in (1.24) implies, and in fact is equivalent to the assumption of smallness on the *anticipated* energy, $\mathcal{E}(0) \leqslant 2(1+\tau)\epsilon$. Theorem 3 is a consequence of the following proposition on the polar system (6.4):

Proposition 6.1. *(*polar coordinates*) There exists a constant* $\epsilon > 0$*, such that if the initial data is small enough,*

$$\mathcal{E}_0 := \left(U(r) + \frac{1}{2}v_r^2 + \frac{1}{2}v_\theta^2\right)_{|t=0} \leqslant \epsilon, \tag{6.6}$$

then the solution to (6.4) *decays to zero at the following algebraic rates:*

$$\delta_r \leqslant C\langle t\rangle^{-1}\ln^{1/2}\langle t\rangle, \quad v_r \leqslant C\langle t\rangle^{-1}\ln^{1/2}\langle t\rangle, \quad v_\theta \leqslant C\langle t\rangle^{-1/2}. \tag{6.7}$$

Proof. Fix $0 < \zeta \leqslant \min\{\frac{r_0}{2}, 1\}$ as a small number such that

$$\frac{a}{2} \leqslant U''(r) \leqslant 2a, \quad \forall \delta_r^2 \leqslant \zeta, \tag{6.8}$$

and as a result,

$$\frac{a}{2}|\delta_r| \leqslant |U'(r)| \leqslant 2a|\delta_r|, \quad \frac{a}{4}\delta_r^2 \leqslant U(r) \leqslant a\delta_r^2, \quad \forall \delta_r^2 \leqslant \zeta. \tag{6.9}$$

We start from the energy estimate for the anticipated energy $\mathcal{E}(t) := U(r) + \frac{1}{2}v_r^2 + \frac{1}{2}v_\theta^2$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) = U'(r)\cdot(v_r - \tau U'(r)) + v_r \cdot \left(-U'(r) + \frac{v_\theta^2}{r}\right) + v_\theta \cdot \frac{-v_r v_\theta}{r} = -\tau U'(r)^2$$

Therefore, for any positive $\epsilon \leqslant \frac{a}{4}\zeta$ to be chosen later, if $\mathcal{E}_0 \leqslant \epsilon$, then

$$\delta_r^2 \leqslant \frac{4}{a}U(r) \leqslant \frac{4}{a}\epsilon < \zeta, \qquad v_r^2 \leqslant \epsilon \tag{6.10}$$

hold for all time which in turn implies that (6.9) holds. Next we consider the cross terms

$$\frac{\mathrm{d}}{\mathrm{d}t}(-v_r v_\theta^2) = -\left(-U'(r) + \frac{v_\theta^2}{r}\right)v_\theta^2 - 2v_r v_\theta \frac{-v_r v_\theta}{r} = -\frac{v_\theta^4}{r} + U'(r)v_\theta^2 + 2\frac{v_r^2 v_\theta^2}{r}, \tag{6.11}$$

and

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(-U'(r)v_r) &= -U''(r)v_r \cdot (v_r - \tau U'(r)) - U'(r)\left(-U'(r) + \frac{v_\theta^2}{r}\right) \\ &= -U''(r)v_r^2 + \tau U''(r)v_r U'(r) + U'(r)^2 - U'(r)\frac{v_\theta^2}{r} \end{aligned} \tag{6.12}$$

We now introduce the modified energy

$$\widehat{\mathcal{E}}(t) := U(r) + \frac{1}{2}v_r^2 + \frac{1}{2}v_\theta^2 - cv_r v_\theta^2 - cU'(r)v_r,$$

depending on a small $c > 0$ which is yet to be determined. A straightforward calculation, based on (6.9) shows its decay rate does not exceed

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{E}}(t) &= -\tau U'(r)^2 - \frac{c}{r}v_\theta^4 - cU''(r)v_r^2 \\
&\quad + c\left(U'(r)v_\theta^2 + 2\frac{v_r^2 v_\theta^2}{r}\right) + c\left(\tau U''(r)v_r U'(r) + U'(r)^2 - U'(r)\frac{v_\theta^2}{r}\right) \\
&\leqslant -\tau\frac{a^2}{4}\delta_r^2 - \frac{c}{2r_0}v_\theta^4 - c\frac{a}{2}v_r^2 \\
&\quad + c\left(2a|\delta_r|v_\theta^2 + 4\frac{v_r^2 v_\theta^2}{r_0}\right) + c\left(4\tau a^2|v_r\delta_r| + 4a^2\delta_r^2 + 4a|\delta_r|\frac{v_\theta^2}{r_0}\right),
\end{aligned}$$

and by Cauchy-Schwarz

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{E}}(t) &\leqslant -\tau\frac{a^2}{4}\delta_r^2 - \frac{c}{2r_0}v_\theta^4 - c\frac{a}{2}v_r^2 \\
&\quad + c\left(\frac{1}{\kappa}a\delta_r^2 + \kappa a v_\theta^4 + \kappa\frac{2}{r_0}v_\theta^4 + \frac{1}{\kappa}\frac{2}{r_0}v_r^4\right) \\
&\quad + c\left(2\kappa\tau a^2 v_r^2 + \frac{1}{\kappa}2\tau a^2\delta_r^2 + 4a^2\delta_r^2 + \frac{1}{\kappa}\frac{2a}{r_0}\delta_r^2 + \kappa\frac{2a}{r_0}v_\theta^4\right) \\
&= -\left(\tau\frac{a^2}{4} - \frac{c}{\kappa}\big(a + 2\tau a^2 + 4\kappa a^2 + \frac{2a}{r_0}\big)\right)\delta_r^2 \\
&\quad - c\left(\frac{1}{2r_0} - \kappa\big(a + \frac{2}{r_0} + \frac{2a}{r_0}\big)\right)v_\theta^4 - c\left(\frac{a}{2} - \frac{1}{\kappa}\frac{2}{r_0}v_r^2 - 2\kappa\tau a^2\right)v_r^2,
\end{aligned} \tag{6.13}$$

with $\kappa \in (0, 1)$ which is yet to be determined. We want to guarantee that the three pre-factors on the right are positive. To this end, we first fix the ratio

$$\frac{c}{\kappa} = \frac{\tau\frac{a^2}{8}}{a + 2\tau a^2 + 4a^2 + \frac{2a}{r_0}} \tag{6.14$_1$}$$

so that the first pre-factor is lower-bounded by $\tau\frac{a^2}{8}$. Then we choose

$$\kappa \leqslant \min\left\{1, \frac{\frac{1}{4r_0}}{a + \frac{2}{r_0} + \frac{2a}{r_0}}, \frac{\frac{a}{4}}{2\tau a^2}\right\} \tag{6.14$_2$}$$

so that the second pre-factor, the coefficient of v_θ^4, becomes larger than $\dfrac{c}{4r_0}$. Finally, the third pre-factor is also positive because (i) a small enough κ was chosen in (6.14)$_2$, and (ii) a key aspect in which v_r can be made small enough to compensate for small κ, so that the negative contribution of $-\frac{1}{\kappa}\frac{2}{r_0}v_r^2$ can be absorbed into the rest: indeed, if

$$v_r^2 \leqslant \frac{\frac{a}{8}}{\frac{1}{\kappa}\frac{2}{r_0}} = \frac{ar_0}{16}\kappa \tag{6.14$_3$}$$

then

$$c\left(\frac{a}{2} - \frac{1}{\kappa}\frac{2}{r_0}v_r^2 - 2\kappa\tau a^2\right) \leqslant c\left(\frac{a}{2} - \frac{1}{\kappa}\frac{2}{r_0}\frac{\frac{a}{8}}{\frac{1}{\kappa}\frac{2}{r_0}} - \frac{a}{4}\right) = \frac{ca}{8}.$$

Therefore, (6.13) implies the decay rate

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{E}}(t) \leqslant -\eta_1(\delta_r^2 + v_\theta^4 + v_r^2), \qquad \eta_1 = \min\left\{\tau\frac{a^2}{8}, \frac{c}{4r_0}, \frac{ca}{8}\right\}, \tag{6.15}$$

provided (6.9) and (6.14)$_1$–(6.14)$_3$ are satisfied.

Moreover, we claim that $\widehat{\mathcal{E}}$ is comparable to the original anticipated energy $\mathcal{E}$. Indeed, if in addition

$$c\sqrt{\frac{ar_0}{16}\kappa} \leqslant \frac{1}{4} \tag{6.16$_1$}$$

holds, then in view of (6.14)$_3$, $c|v_r v_\theta^2| \leqslant \frac{1}{4}v_\theta^2$, and if

$$c \leqslant \min\left\{\frac{1}{8}, \frac{1}{4a}\right\}, \tag{6.16$_2$}$$

holds, then in view of (6.9),

$$c|U'(r)v_r| \leqslant ca(\delta_r^2 + v_r^2) \leqslant ca\Big(\frac{4}{a}U(r) + v_r^2\Big) \leqslant \frac{1}{2}\Big(U(r) + \frac{1}{2}v_r^2\Big).$$

It follows that

$$\frac{1}{2}\mathcal{E}(t) \leqslant \widehat{\mathcal{E}}(t) \leqslant 2\mathcal{E}(t), \tag{6.17}$$

provided ((6.16)$_1$)–((6.16)$_2$) are satisfied. These last two conditions are clearly met for small enough κ: recall that the ratio c/κ was fixed in ((6.14)$_1$) then

$$\kappa \leqslant \frac{a + 2\tau a^2 + 4a^2 + \frac{2a}{r_0}}{\tau\frac{a^2}{8}} \min\left\{\sqrt{ar_0}, \frac{1}{8}, \frac{1}{4a}\right\} \tag{6.18$_1$}$$

suffices to guarantee ((6.16)$_1$)–((6.16)$_2$). Thus, we finally choose small enough κ to satisfy both ((6.14)$_2$),((6.18)$_1$), and small enough $\epsilon < \min\left\{\frac{a}{4}\zeta, \frac{ar_0}{16}\kappa\right\}$ so that (6.10) and ((6.14)$_3$) hold. By now we proved (6.15) and (6.17). Finally, notice that for small enough δ_r, v_r we have

$$\delta_r^2 + v_\theta^4 + v_r^2 \geqslant \delta_r^4 + v_\theta^4 + v_r^4 \geqslant \frac{1}{3}(\delta_r^2 + v_\theta^2 + v_r^2)^2 \geqslant \frac{1}{3}\min\left\{\frac{1}{a^2}, 1\right\}\mathcal{E}^2(t).$$

We conclude, in view of (6.15) and (6.17),

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{E}}(t) \leqslant -\eta\widehat{\mathcal{E}}^2(t) \rightsquigarrow \widehat{\mathcal{E}}(t) \leqslant \frac{1}{\eta t + 1/\widehat{\mathcal{E}}(0)}, \quad \eta = \frac{\eta_1}{12}\min\left\{\frac{1}{a^2}, 1\right\}.$$

It follows that $|v_\theta| \leqslant C\langle t\rangle^{-1/2}$.

To get a better decay rate for δ_r and v_r, we use yet another modified energy functional,

$$\breve{\mathcal{E}}(t) := U(r) + \frac{1}{2}v_r^2 - c_1 U'(r)v_r,$$

for which we find

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\breve{\mathcal{E}}(t) =& -\tau U'(r)^2 - c_1 U''(r)v_r^2 + \frac{v_r v_\theta^2}{r} + c_1\left(\tau U''(r)v_r U'(r) + U'(r)^2 - U'(r)\frac{v_\theta^2}{r}\right)\\ \leqslant& -\tau\frac{a^2}{4}\delta_r^2 - c_1\frac{a}{2}v_r^2 + \frac{2v_r v_\theta^2}{r_0} + c_1\left(4\tau a^2|v_r\delta_r| + 4a^2\delta_r^2 + 4a|\delta_r|\frac{v_\theta^2}{r_0}\right)\\ \leqslant& -\left(\tau\frac{a^2}{4} - \frac{c_1}{\kappa_1}\left(2\tau a^2 + 4\kappa_1 a^2 + \frac{2a}{r_0}\right)\right)\delta_r^2 - c_1\left(\frac{a}{2} - \kappa_1\frac{1}{r_0} - \kappa_1\cdot 2\tau a^2\right)v_r^2\\ &+ \left(\frac{1}{c_1\kappa_1 r_0} + c_1\kappa_1\frac{2a}{r_0}\right)v_\theta^4.\end{aligned}$$

By similar choices of c_1 and κ_1, one can guarantee that $\breve{\mathcal{E}}(t)$ is equivalent to $\delta_r^2 + v_r^2$, and the coefficients of δ_r^2 and v_r^2 are positive. Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\breve{\mathcal{E}}(t) \leqslant -\eta_2\breve{\mathcal{E}}(t) + Cv_\theta^4 \leqslant -\eta_2\breve{\mathcal{E}}(t) + C\langle t\rangle^{-2} \tag{6.19}$$

This gives

$$\breve{\mathcal{E}}(t) = e^{-\eta_2 t}\breve{\mathcal{E}}(0) + C\int_0^t e^{-\eta_2(t-s)}(1+s)^{-2}\,\mathrm{d}s \tag{6.20}$$

We estimate the last integral for large enough t,

$$\begin{aligned}\int_0^t e^{-\eta_2(t-s)}(1+s)^{-2}\,\mathrm{d}s &\leqslant \left(\int_0^{t-\frac{1}{\eta_2}\ln\langle t\rangle} + \int_{t-\frac{1}{\mu}\ln\langle t\rangle}^t\right)e^{-\eta_2(t-s)}(1+s)^{-2}\,\mathrm{d}s\\ &\leqslant \langle t\rangle^{-1}\int_0^t(1+s)^{-2}\,\mathrm{d}s + \left(1+(t-\frac{1}{\eta_2}\ln\langle t\rangle)\right)^{-2}\frac{1}{\eta_2}\ln\langle t\rangle\\ &\leqslant \langle t\rangle^{-2} + \frac{2}{\eta_2}\langle t\rangle^{-2}\ln\langle t\rangle.\end{aligned} \tag{6.21}$$

This shows that $\breve{\mathcal{E}}(t) \leqslant C\langle t\rangle^{-2}\ln\langle 1+t\rangle$, and therefore $|v_r| + |\delta_r| \leqslant C\langle t\rangle^{-1}\ln^{1/2}\langle 1+t\rangle$. $\square$

Finally, we conclude by noting that the last bound on δ_r tells us that

$$\big||\mathbf{x}_1^\tau(t)-\mathbf{x}_2^\tau(t)|-r_0\big| \leqslant C\langle t\rangle^{-1}\ln^{1/2}\langle 1+t\rangle, \quad |\mathbf{v}_1(t)-\mathbf{v}_2(t)| \leqslant C\langle t\rangle^{-1/2}.$$

Observe that this bound on relative *anticipated* positions is in fact equivalent to the claimed statement of the current positions, $\big||\mathbf{x}_1(t)-\mathbf{x}_2(t)|-r_0\big| \lesssim \langle t\rangle^{-1}\ln^{1/2}\langle 1+t\rangle$, which concludes the proof of Theorem 3.

Remark 6.1. Numerical examples [20, sec. 1] show that the rate $v_\theta = \mathcal{O}(t^{-1/2})$ is optimal. Therefore,

$$\theta(t) = \theta_0 + \int_0^t \frac{1}{r(s)} v_\theta(s)\,\mathrm{d}s = \mathcal{O}(\sqrt{t}),$$

which means that θ needs not converge to any point, even for near equilibrium initial data. Thus, although we trace the dynamics of δ_r, v_r, v_θ using essentially perturbative arguments, the dynamics of (6.1) is not.

Remark 6.2. The optimal algebraic rate of v_θ in Proposition 6.1 can also be derived via centre manifold reduction. To this end, rewriting the system (6.4) in terms of perturbation variables δ_r, v_r and v_θ,

$$\begin{bmatrix} \dot{\delta}_r \\ \dot{v}_r \\ \dot{v}_\theta \end{bmatrix} = \begin{bmatrix} \tau a & 1 & 0 \\ -a & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_r \\ v_r \\ v_\theta \end{bmatrix} + \begin{bmatrix} -\tau\left(U'(r_0+\delta_r) - a\delta_r\right) \\ -\left(U'(r_0+\delta_r)-a\delta_r\right) + \frac{v_\theta^2}{r_0+\delta_r} \\ -\frac{v_r v_\theta}{r_0+\delta_r} \end{bmatrix}.$$

Since the 2×2 leading minor is stable, the centre manifold near the equilibrium, W^c, can be parametrized by v_θ,

$$W^c = \left\{ (\delta_r, v_r, v_\theta) \,\middle|\, \delta_r = \frac{v_\theta^2}{r_0} + \mathcal{O}(v_\theta^3), \quad v_r = \frac{v_\theta^3}{r_0} + \mathcal{O}(v_\theta^3) \right\}.$$

It follows that v_θ is governed by $\dot{v}_\theta = -\dfrac{v_r v_\theta}{r} = -\tau\dfrac{v_\theta^3}{r_0} + \mathcal{O}(v_\theta^4)$, which implies its algebraic decay of order $\langle t\rangle^{-1/2}$.

7. Anticipation Dynamics: Hydrodynamic Formulation

The large crowd dynamics associated with (AT) is captured by the macroscopic density $\rho(t,\mathbf{x}) : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{R}_+$ and momentum $\rho\mathbf{u}(t,\mathbf{x}) : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{R}^d$, which are governed by the hydrodynamic description (1.26)

$$\begin{cases} \rho_t + \nabla_{\mathbf{x}}\cdot(\rho\mathbf{u}) = 0 \\ (\rho\mathbf{u})_t + \nabla_{\mathbf{x}}\cdot(\rho\mathbf{u}\otimes\mathbf{u}) = -\displaystyle\int \nabla U(|\mathbf{x}^\tau - \mathbf{y}^\tau|)\rho(t,\mathbf{x})\,\mathrm{d}\rho(t,\mathbf{y}), \quad \mathbf{x}^\tau := \mathbf{x} + \tau\mathbf{u}(t,\mathbf{x}). \end{cases}$$

The flux on the left involves additional second-order moment fluctuations, $\mathcal{P}$, which can be dictated by proper closure relations, e.g., [8,11,19,21,24]. As in [24], we will focus on the mono-kinetic case, in which case $\mathcal{P} = 0$.

To study the large time behavior we appeal, as in the discrete case, to the basic balance between energy and enstrophy: here we consider the *anticipated energy*

$$\mathcal{E}(t) := \int \frac{1}{2}|\mathbf{u}(t,\mathbf{x})|^2\rho(t,\mathbf{x})\,\mathrm{d}\mathbf{x} + \frac{1}{2}\int\int U(|\mathbf{x}^\tau(t)-\mathbf{y}^\tau(t)|)\,\mathrm{d}\rho(t,\mathbf{x})\,\mathrm{d}\rho(t,\mathbf{y}). \tag{7.1}$$

Away from vacuum, the velocity field $\mathbf{u}(\mathbf{x}) = \mathbf{u}(t,\mathbf{x})$ satisfies the transport equation

$$\mathbf{u}_t(t,\mathbf{x}) + \mathbf{u}\cdot\nabla_{\mathbf{x}}\mathbf{u}(t,\mathbf{x}) = \mathbf{A}(\rho,\mathbf{u})(t,\mathbf{x}), \tag{7.2a}$$

where $\mathbf{A}(\rho,\mathbf{u})$ denotes the anticipated interaction term

$$\mathbf{A}(\rho,\mathbf{u})(t,\mathbf{x}) := -\int \nabla U(|\mathbf{x}^\tau(t)-\mathbf{y}^\tau(t)|)\,\mathrm{d}\rho(t,\mathbf{y}), \qquad \mathbf{x}^\tau(t) = \mathbf{x}+\tau\mathbf{u}(t,\mathbf{x}). \tag{7.2b}$$

We compute (suppressing the time dependence)

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) =& \int \mathbf{u}(\mathbf{x})\cdot(-\mathbf{u}\cdot\nabla\mathbf{u}+\mathbf{A}(\mathbf{x}))\,\mathrm{d}\rho(\mathbf{x}) + \int\frac{1}{2}|\mathbf{u}(\mathbf{x})|^2(-\nabla\cdot(\rho\mathbf{u}))\,\mathrm{d}\mathbf{x}\\
&+\frac{\tau}{2}\int\int\nabla U(|\mathbf{x}^\tau-\mathbf{y}^\tau|)\cdot(-\mathbf{u}(\mathbf{x})\cdot\nabla\mathbf{u}(\mathbf{x})+\mathbf{A}(\mathbf{x})+\mathbf{u}(\mathbf{y})\cdot\nabla\mathbf{u}(\mathbf{y})\\
&-\mathbf{A}(\mathbf{y}))\,\mathrm{d}\rho(\mathbf{x})\,\mathrm{d}\rho(\mathbf{y})\\
&+\frac{1}{2}\int\int U(|\mathbf{x}^\tau-\mathbf{y}^\tau|)(-\nabla\cdot(\rho\mathbf{u})(\mathbf{y})))\,\mathrm{d}\rho(\mathbf{x})\,\mathrm{d}\mathbf{y}\\
&+\frac{1}{2}\int\int U(|\mathbf{x}^\tau-\mathbf{y}^\tau|)(-\nabla\cdot(\rho\mathbf{u})(\mathbf{x}))\,\mathrm{d}\mathbf{x}\,\mathrm{d}\rho(\mathbf{y})\\
=&\int\mathbf{u}(\mathbf{x})\cdot\mathbf{A}(\mathbf{x})\,\mathrm{d}\rho(\mathbf{x})+\frac{\tau}{2}\int\int\nabla U(|\mathbf{x}^\tau-\mathbf{y}^\tau|)\cdot(\mathbf{A}(\mathbf{x})-\mathbf{A}(\mathbf{y}))\,\mathrm{d}\rho(\mathbf{y})\,\mathrm{d}\rho(\mathbf{x})\\
&+\frac{1}{2}\int\int\nabla U(|\mathbf{x}^\tau-\mathbf{y}^\tau|)\cdot(-\mathbf{u}(\mathbf{y})+\mathbf{u}(\mathbf{x}))\,\mathrm{d}\rho(\mathbf{y})\,\mathrm{d}\rho(\mathbf{x})\\
=&\tau\int\int\nabla U(|\mathbf{x}^\tau-\mathbf{y}^\tau|)\cdot\mathbf{A}(\mathbf{x})\,\mathrm{d}\rho(\mathbf{y})\,\mathrm{d}\rho(\mathbf{x})\\
=&-\tau\int\left|\int\nabla U(|\mathbf{x}^\tau-\mathbf{y}^\tau|)\,\mathrm{d}\rho(\mathbf{y})\right|^2\mathrm{d}\rho(\mathbf{x}).
\end{aligned}$$

This is the continuum analogue of the discrete enstrophy statement (1.1), which becomes apparent when it is expressed in terms of the *material derivative*,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) = -\tau\int\left|\int\nabla U(|\mathbf{x}^\tau-\mathbf{y}^\tau|)\,\mathrm{d}\rho(\mathbf{y})\right|^2\mathrm{d}\rho(\mathbf{x}) = -\tau\int\left|\frac{\mathrm{D}}{\mathrm{D}t}\mathbf{u}(t,\mathbf{x})\right|^2\mathrm{d}\rho(t,\mathbf{x}). \tag{7.3}$$

7.1. Smooth Solutions must Flock

We consider the anticipation hydrodynamics (1.26) with attractive potentials, (1.15)

$$a\langle r\rangle^{-\beta}\leqslant\frac{U'(r)}{r},\quad |U''(r)|\leqslant A,\qquad 0<a<A.$$

The study of its large time 'flocking' behavior proceeds precisely along the lines of our discrete proof of Theorem 2. Here are the three main ingredients in the proof of Theorem 4.

Step (i) We begin where we left with the anticipated energy balance (7.3), which we express as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) &= -\int \Big| \int \nabla U(|\mathbf{x}^\tau - \mathbf{y}^\tau|)\,\mathrm{d}\rho(\mathbf{y})\Big|^2 \mathrm{d}\rho(\mathbf{x}) \\ &= -\int \Big| \int \frac{U'(|\mathbf{x}^\tau - \mathbf{y}^\tau|)}{|\mathbf{x}^\tau - \mathbf{y}^\tau|}(\mathbf{x}^\tau - \mathbf{y}^\tau)\,\mathrm{d}\rho(\mathbf{y})\Big|^2 \mathrm{d}\rho(\mathbf{x}).\end{aligned}$$

We now appeal to the special case of Lemma 4.1 with $\Omega = \mathbb{R}^d$ (with variable $\mathbf{x}$), with probability measure[10] $\mathrm{d}\mu = \rho(\mathbf{x})\,\mathrm{d}\mathbf{x}$, $\mathbf{X}(\mathbf{x}) = \mathbf{x}^\tau$, $\mathbf{X}(\mathbf{y}) = \mathbf{y}^\tau$ and $c(\mathbf{x},\mathbf{y}) = \dfrac{U'(|\mathbf{x}^\tau - \mathbf{y}^\tau|)}{|\mathbf{x}^\tau - \mathbf{y}^\tau|}$, in which case we have

$$\iint |\mathbf{x}^\tau - \mathbf{y}^\tau|^2\,\mathrm{d}\rho(\mathbf{y})\,\mathrm{d}\rho(\mathbf{x}) \leqslant 32\frac{\Lambda^2}{\lambda^4}\int \Big| \int \frac{U'(|\mathbf{x}^\tau - \mathbf{y}^\tau|)}{|\mathbf{x}^\tau - \mathbf{y}^\tau|}(\mathbf{x}^\tau - \mathbf{y}^\tau)\,\mathrm{d}\rho(\mathbf{y})\Big|^2 \mathrm{d}\rho(\mathbf{x}) \tag{7.4}$$

where $\Lambda = A$ and λ are the upper- and respectively, lower-bounds of $\dfrac{U'(|\mathbf{x}^\tau - \mathbf{y}^\tau|)}{|\mathbf{x}^\tau - \mathbf{y}^\tau|}$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) &= -\tau\int \Big| \int \frac{U'(|\mathbf{x}^\tau - \mathbf{y}^\tau|)}{|\mathbf{x}^\tau - \mathbf{y}^\tau|}(\mathbf{x}^\tau - \mathbf{y}^\tau)\,\mathrm{d}\rho(\mathbf{y})\Big|^2 \mathrm{d}\rho(\mathbf{x}) \\ &\lesssim -\Big(\min \frac{U'(|\mathbf{x}^\tau - \mathbf{y}^\tau|)}{|\mathbf{x}^\tau - \mathbf{y}^\tau|}\Big)^4 \iint |\mathbf{x}^\tau - \mathbf{y}^\tau|^2\,\mathrm{d}\rho(\mathbf{y})\,\mathrm{d}\rho(\mathbf{x}).\end{aligned} \tag{7.5}$$

Step (ii). A bound on the spread of the anticipated positions supported on non-vacuous states

$$\max_{\mathbf{x}^\tau \in \mathrm{supp}\,\rho(t,\cdot)} |\mathbf{x}^\tau| \leqslant c\langle t\rangle^\eta. \tag{7.6}$$

Arguing along the lines of Lemma 2.2 one finds that (7.6) holds with $\eta = \dfrac{1}{2(1-\beta)}$, hence

$$a\langle t\rangle^{-\frac{\beta}{2(1-\beta)}} \lesssim \frac{U'(|\mathbf{x}^\tau - \mathbf{y}^\tau|)}{|\mathbf{x}^\tau - \mathbf{y}^\tau|} \leqslant A, \qquad \mathbf{x}^\tau, \mathbf{y}^\tau \in \mathrm{supp}\,\rho(t,\cdot),$$

and (7.5) implies

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) &= -\tau\int \Big| \int \frac{U'(|\mathbf{x}^\tau - \mathbf{y}^\tau|)}{|\mathbf{x}^\tau - \mathbf{y}^\tau|}(\mathbf{x}^\tau - \mathbf{y}^\tau)\,\mathrm{d}\rho(\mathbf{y})\Big|^2 \mathrm{d}\rho(\mathbf{x}) \\ &\lesssim -\tau\langle t\rangle^{-\frac{2\beta}{1-\beta}} \iint |\mathbf{x}^\tau - \mathbf{y}^\tau|^2\,\mathrm{d}\rho(\mathbf{x})\,\mathrm{d}\rho(\mathbf{y}).\end{aligned} \tag{7.7}$$

[10] Without loss of generality we use the normalization $\int \rho_0(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int \rho(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$.

We are now exactly at the point we had with the discrete anticipation dynamics, in which the decay of anticipated energy is controlled by the fluctuations of anticipated position, (1.23).

Step (iii). To close the decay rate (7.7) one invokes hypocoercivity argument on the modified energy,

$$\widehat{\mathcal{E}}(t) := \mathcal{E}(t) - \epsilon(t) \int \mathbf{x}^\tau \cdot \mathbf{u}(\mathbf{x})\,\mathrm{d}\rho(\mathbf{x}).$$

Arguing along the lines of Section 5, one can find a suitable $\epsilon(t) > 0$ which leads to the sub-exponential decay of $\widehat{\mathcal{E}}(t)$ and hence of the comparable $\mathcal{E}(t)$, thus completing the proof of Theorem 4.

7.2. Existence of Smooth Solutions—the 1D Case

We study the existence of smooth solutions of the 1D anticipated hydrodynamic system

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0 \\ \partial_t u + u\partial_x u = -\displaystyle\int U'(|x^\tau - y^\tau|) sgn(x^\tau - y^\tau)\rho(y)\,\mathrm{d}y, \qquad x^\tau = x + \tau u(t,x), \end{cases} \tag{7.8}$$

subject to uniformly convex potential $U''(\cdot) \geqslant a > 0$. Let $\mathtt{d} := \partial_x u$. Then

$$\partial_t \mathtt{d} + u\partial_x \mathtt{d} + \mathtt{d}^2 = -(1 + \tau\mathtt{d}) \int U''(|x^\tau - y^\tau|)\rho(y)\,\mathrm{d}y \tag{7.9}$$

or

$$\mathtt{d}' = -\mathtt{d}^2 - c(1 + \tau\mathtt{d}), \qquad ' := \partial_t + u\partial_x, \tag{7.10}$$

where by uniform convexity $c = c(t,x) := \int U''(|x^\tau - y^\tau|)\rho(y)\,\mathrm{d}y \in [m_0 a, m_0 A]$. The discriminant of RHS, given by $(\tau c)^2 - 4c = c(\tau^2 c - 4)$ is non-negative, provided $\tau^2 m_0 a \geqslant 4$. In this case, the smaller root of (7.10) is given by

$$\frac{1}{2}(-\tau c - \sqrt{c(\tau^2 c - 4)}) \leqslant -\frac{1}{2}(\tau m_0 a + \sqrt{m_0 a(\tau^2 m_0 a - 4)}), \tag{7.11}$$

and the region to its right is an invariant of the dynamics (7.9). We conclude the following:

Proposition 7.1. *(*Existence of global smooth solution*) Consider the 1D anticipation hydrodynamic system* (7.8) *with uniformly convex potential* $0 < a \leqslant U'' \leqslant A$. *It admits a global smooth solution for sub-critical initial data,* (ρ_0, u_0)*, satisfying*

$$\min_x u_0'(x) \geqslant -\frac{1}{2}(\tau m_0 a + \sqrt{m_0 a(\tau^2 m_0 a - 4)}), \qquad \tau \geqslant \frac{2}{\sqrt{m_0 a}}.$$

Acknowledgements. Research was supported by NSF Grants DMS16-13911, RNMS11-07444 (KI-Net) and ONR Grant N00014-1812465. ET thanks the hospitality of the Institut of Mittag-Leffler during fall 2018 visit which initiated this work, and of the Laboratoire Jacques-Louis Lions in Sorbonne University during spring 2019, with support through ERC Grant 740623 under the EU Horizon 2020, while concluding this work. We thank the anonymous reviewers for careful reading and insightful comments which help improving an earlier draft of the paper.

References

1. Balagué , D., Carrillo , T.J.A., Laurent , R.G.: Dimensionality of local minimizers of the interaction energy. *Arch. Rat. Mech. Anal.* **209**, 1055–1088, 2013
2. Balagué , D., Carrillo , J., Yao , Y.: Confinement for repulsive–attractive kernels. *DCDS - B* **19**(5), 1227–1248, 2014
3. Bernoff , A.J., Topaz , C.M.: A primer of swarm equilibria SIAM. *J. Appl. Dyn. Syst.* **10**, 212–250, 2011
4. Bertozzi , A.L., Carrillo , J.A., Laurent , T.: Blowup in multidimensional aggregation equations with mildly singular interaction kernels. *Nonlinearity* **22**, 683–710, 2009
5. Bertozzi , A.L., Kolokolnikov , T., Sun , H., Uminsky , D., von Brecht , J.: Ring patterns and their bifurcations in a nonlocal model of biological swarms. *Commun. Math. Sci.* **13**, 955–985, 2015
6. Bertozzi , A.L., Laurent , T., Léger , F.: Aggregation and spreading via the Newtonian potential: the dynamics of patch solutions. *Math. Models Methods Appl. Sci.* **22**(supp01), 1140005, 2012
7. Carrillo, J.A., Choi, Y.-P., Hauray, M.: The derivation of swarming models: mean-field limit and Wasserstein distances. In: Collective Dynamics from Bacteria to Crowds: An Excursion Through Modeling, Analysis and Simulation, Series: CISM Inter. Centre for Mech. Sci. Springer, vol. 533, pp. 1–45 (2014)
8. Carrillo, J.A., Choi, Y.-P., Perez, S.: A review on attractive–repulsive hydrodynamics for consensus in collective behavior Active Particles, Volume 1. Modeling and Simulation in Science, Engineering and Technology Bellomo, N., Degond, P., Tadmor, E. (eds.). Birkhäuser (2017)
9. Carrillo , J.A., Choi , Y.-P., Tadmor , E., Tan , C.: Critical thresholds in 1D Euler equations with non-local forces. *Math. Models Methods Appl. Sci.* **26**(1), 185–206, 2016
10. Carrillo , J.A., D'Orsogna , M.R., Panferov , V.: Double milling in self-propelled swarms from kinetic theory. *Kinet. Relat. Models* **2**, 363–378, 2009
11. Carrillo , J., Fornasier , M., Rosado , J., Toscani , G.: Asymptotic flocking dynamics for the kinetic Cucker–Smale model. *SIAM J. Math. Anal.* **42**(218), 218–236, 2010
12. Carrillo , J.A., Huang , Y., Martin , S.: Nonlinear stability of flock solutions in second-order swarming models. *Nonlinear Anal. Real World Appl.* **17**, 332–343, 2014
13. Cucker , F., Smale , S.: Emergent behavior in flocks. *IEEE Trans. Autom. Control* **52**(5), 852–862, 2007
14. Cucker , F., Smale , S.: On the mathematics of emergence. *Jpn. J. Math.* **2**(1), 197–227, 2007
15. Danchin , R., Mucha , P.B., Peszek , J., Wróblewski , B.: Regular solutions to the fractional Euler alignment system in the Besov spaces framework. *Math. Models Methods Appl. Sci.* **29**(1), 89–119, 2019

16. DIETERT, H., SHVYDKOY, R.: On Cucker–Smale dynamical systems with degenerate communication. *Anal. Appl.* (2020)
17. DO , T., KISELEV , A., RYZHIK , L., TAN , C.: Global regularity for the fractional Euler alignment system. *Arch. Ration. Mech. Anal.* **228**(1), 1–37, 2018
18. D'ORSOGNA, M.R., CHUANG, Y.L., BERTOZZI, A.L., CHAYES, L.: Self-propelled particles with soft-core interactions: patterns, stability, and collapse. *Phys. Rev. Lett.* **96**, 104302, 2006
19. FIGALLI, A., KANG, M.-J.: A rigorous derivation from the kinetic Cucker–Smale model to the pressureless Euler system with nonlocal alignment. *Anal. PDE* **1293**, 843–866, 2019
20. GERLEE , P., TUNSTRM , K., LUNDH , T., WENNBERG , B.: Impact of anticipation in dynamical systems. *Phys. Rev. E* **96**, 062413, 2017
21. GOLSE , F.: On the dynamics of large particle systems in the mean field limit. In macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity. *Lect. Notes Appl. Math. Mech.* **3**, 1–144, 2016
22. GUÉANT, O., LASRY, J.-M.: Pierre–Louis lions mean field games and applications. Paris-Princeton Lectures on Mathematical Finance, pp. 205-266 (2010)
23. HA , S.-Y., LIU , J.-G.: A simple proof of the Cucker–Smale flocking dynamics and mean-field limit. *Commun. Math. Sci.* **7**(2), 297–325, 2009
24. HA , S.-Y., TADMOR , E.: From particle to kinetic and hydrodynamic descriptions of flocking. *Kinet. Relat. Models* **1**(3), 415–435, 2008
25. HE , S., TADMOR , E.: Global regularity of two-dimensional flocking hydrodynamics. *Comptes rendus - Mathematique Ser.* **I**(355), 795–805, 2017
26. JABIN , P.E.: A review of the mean field limits for Vlasov equations. *KRM* **7**, 661–711, 2014
27. KOLOKONIKOV , T., SUN , H., UMINSKY , D., BERTOZZI , A.: Stability of ring patterns arising from 2d particle interactions. *Phys. Rev. E* **84**, 015203, 2011
28. LEVINE , H., RAPPEL , W.-J., COHEN , I.: Self-organization in systems of self-propelled particles. *Phys. Rev. E* **63**, 017101, 2000
29. MINAKOWSKI, P., MUCHA, P.B., PESZEK, J., ZATORSKA, E.: Singular Cucker–Smale dynamics. In: Bellomo, N., Degond, P., Tadmor, E., (eds.) Active Particles—Volume 2—Theory, Models, Applications. Birkhäuser-Springer, Boston, USA (2019)
30. MORIN , A., CAUSSIN , J.-B., ELOY , C., BARTOLO , D.: Collective motion with anticipation: flocking, spinning, and swarming. *Phys. Rev. E* **91**, 012134, 2015
31. MOTSCH , S., TADMOR , E.: Heterophilious dynamics enhances consensus. *SIAM Rev.* **56**(4), 577–621, 2014
32. POYATO , D., SOLER , J.: Euler-type equations and commutators in singular and hyperbolic limits of kinetic Cucker–Smale models. *Math. Models Methods Appl. Sci.* **27**(6), 1089–1152, 2017
33. SERFATY, S.: Coulomb gases and Ginzburg–Landau vortices. Zurich Lectures in Advanced Mathematics, 70, Eur. Math. Soc. (2015)
34. SERFATY, S.: Mean field limit for Coulomb flows. arXiv:1803.08345
35. SHVYDKOY, R., TADMOR, E.: Eulerian dynamics with a commutator forcing. *Trans. Math. Appl.* **1**(1), tnx001 (2017)
36. SHVYDKOY, R., TADMOR, E.: Eulerian dynamics with a commutator forcing III: fractional diffusion of order $0 \leqslant \alpha \leqslant 1$. *Physica D* 376–377, 131–137 (2018)
37. SHVYDKOY, R., TADMOR, E.: Topologically-based fractional diffusion and emergent dynamics with short-range interactions. ArXiv:1806:01371v3
38. SHU , R., TADMOR , E.: Flocking hydrodynamics with external potentials. *Arch. Rat. Mech. Anal.* **238**, 347–381, 2020
39. TADMOR, E., TAN, C.: Critical thresholds in flocking hydrodynamics with non-local alignment. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **372**(2028), 20130401 (2014)

Ruiwen Shu & Eitan Tadmor
Department of Mathematics,
University of Maryland,
College Park
20742 MD
USA.
e-mail: rshu@cscamm.umd.edu

and

Eitan Tadmor
Center for Scientific Computation and Mathematical Modelling (CSCAMM),
University of Maryland,
College Park
20742 MD
USA.

and

Institute for Physical Sciences and Technology (IPST),
University of Maryland,
College Park
20742 MD
USA.
e-mail: tadmor@umd.edu

(*Received September 11, 2019 / Accepted January 23, 2021*)
Published online February 12, 2021