Impact of anticipation in dynamical systems

P. Gerlee,[1,][*] K. Tunstrøm,[2] T. Lundh,[1] and B. Wennberg[1]

[1]*Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 412 96 Göteborg, Sweden*
[2]*Department of Physics, Chalmers University of Technology, 412 96 Göteborg, Sweden*

Many animals, including humans, have predictive capabilities and, presumably, base their behavioral decisions—at least partially—upon an anticipated state of their environment. We explore a minimal version of this idea in the context of particles that interact according to a pairwise potential. Anticipation enters the picture by calculating the interparticle forces from linear extrapolations of the particle positions some time τ in the future. Simulations show that for intermediate values of τ, compared to a transient time scale defined by the potential and the initial conditions, the particles form rotating clusters in which the particles are arranged in a hexagonal pattern. Analysis of the system shows that anticipation induces energy dissipation and we show that the kinetic energy asymptotically decays as $1/t$. Furthermore, we show that the angular momentum is not necessarily conserved for $\tau > 0$, and that asymmetries in the initial condition therefore can cause rotational movement. These results suggest that anticipation could play an important role in collective behavior, since it may induce pattern formation and stabilizes the dynamics of the system.

I. INTRODUCTION

The interactions of individuals within a collective are often explained in terms of simple rules of interaction between animals, e.g., attraction, repulsion, and alignment [1,2]. Typically, swarming behavior is modeled as a collection of particles that represent the organisms in question. Taking inspiration from physics, the interactions between individuals are often assumed to result from pairwise forces that depend only on the distance between the individuals. Typically the force is defined via a potential that contains a repulsive and attractive region such that at some intermediate distance the interaction energy is minimized. In terms of animals this would correspond to some preferred distance between neighboring individuals [3]. In addition, most physics-inspired models rely on self-propulsion to drive pattern formation [4] and often include a noise term as well [5].

An underlying assumption in most current collective motion models is that individuals react and update their velocity according to the current state of other individuals. Contrary to this, we know that many animals including humans have the ability to anticipate movement and act on predicted states. Humans anticipate the movement of visual cues, such as other people in a moving crowd, by extrapolating their positions in time [6]. This process occurs within the retina itself and is, in fact, necessary if we are to respond to rapid visual cues, since the delay induced by phototransduction is of the order of 100 ms [7]. In addition to performing extrapolation the retina also detects when its prediction fails and signals this downstream to the visual cortex [8]. Anticipation is not restricted to humans but has also been detected among insects [9], amphibians [10], and fish [11]. Therefore anticipation most likely plays an important role in the behavior of animals that engage in flocking and swarming.

A recent study of human interactions in crowds quantified the effect of anticipation by showing that the strength of a physical interaction depends not on the distance, but on the time to collision [12]. This suggests that humans act in an anticipatory fashion and use current positions and velocities to extrapolate possible future collisions and update their current velocity to avoid such collisions. In terms of the above-mentioned interaction potential one might then assume that individuals interact according to anticipated future positions and adjust their current velocities according to the predicted state of the system. This idea has been investigated by Morin *et al.* [13], who considered the continuous-time Viscek model, which assumes self-propelled particles that interact with neighboring particles within a certain interaction radius. The classical Viscek model was modified so that the difference in angle between particle i and particle j, $\theta_i - \theta_j$, was altered to $\theta_i - (\theta_j + \alpha\sigma_j)$, where σ_j is the sign of the angular velocity of particle j and α is some positive parameter. Depending on the value of α and the magnitude of the noise, that model can exhibit isotropic behavior, spinning and flocking, which shows that anticipation indeed can have important consequences. The effect of anticipation has also been investigated in a lattice-based model inspired by the swarming of soldier crabs [14]. They showed that mutual anticipation leads to dense collective motion with a high degree of polarization and that the turning response depends on the distance between two individuals rather than the relative heading, which is in agreement with empirical data.

In this paper we also explore the idea of anticipation, but in contrast to Morin *et al.* [13], who considered alignment interaction, we investigate the impact of anticitipation in a model with pairwise forces. We consider a system of interacting particles where the interaction forces are calculated not from current positions, but from positions extrapolated some time τ in the future. For the sake of simplicity we assume that the future positions are given by a linear extrapolation from the current velocities, although more elaborate means of extrapolation are conceivable (see Sec. IV). In order to get a thorough understanding of the dynamics that anticipation induces we disregard self-propulsion, alignment, and other processes commonly found in models of collective behavior

---
[*]gerlee@chalmers.se

and focus on the effects in a simple model where individuals are represented as interacting particles.

II. A MODEL OF ANTICIPATION IN DYNAMICAL SYSTEMS

To introduce the concept of anticipation let us begin with a simple example. Consider two particles of mass m connected by a spring with rest length 0 and spring constant $k/2$, where $k > 0$. The distance x between the two particles obeys the equation $m\ddot{x} = -kx$, which is simply that of a harmonic oscillator. Now let us assume that the particles anticipate the movement of one another and, hence, that the forces acting on the particles are given not by their instantaneous positions, but by their anticipated positions some time τ in the future, i.e., we calculate the forces based on predicted positions $x_i^p(t + \tau) = x_i(t) + \tau v_i(t)$ for particles $i = 1$ and 2. In this case the equation of motion of the interparticle distance x is given by $m\ddot{x} = -kx - k\tau\dot{x}$. This is the equation for a damped harmonic oscillator, and hence we conclude that in this simple system anticipation has a damping effect on the dynamics. In contrast to the undamped system there is dissipation of energy and for any initial condition the system will reach a stationary state with $x = 0$ as $t \rightarrow \infty$.

We now extend the idea of anticipation to N identical particles that interact via some potential $U(r)$ and hence obey the equations of motion

$$
\begin{aligned}
\dot{\vec{x}}_i(t) &= \vec{v}_i(t), \\
m\dot{\vec{v}}_i(t) &= -\sum_{j \neq i} \vec{\nabla} U\big(\vec{x}_i^p(t + \tau) - \vec{x}_j^p(t + \tau)\big),
\end{aligned}
\tag{1}
$$

where $m = 1$ is the mass and $\vec{x}_i, \vec{v}_i \in \mathbb{R}^2$, i.e., the system is defined in the plane. Again, we assume that the individuals make use of a linear prediction of future positions such that $\vec{x}_i^p(t + \tau) = \vec{x}_i(t) + \tau\vec{v}_i(t)$. We note that for $\tau = 0$ the system reduces to a standard system of interacting particles.

We use a generalized Morse potential,

$$
U(x) = C_r e^{-|x|/l_r} - C_a e^{-|x|/l_a},
\tag{2}
$$

to define the interaction between pairs of particles, where C_r and C_a represent the amplitudes of the repulsive and attractive component of the potential, and l_r and l_a denote their respective ranges. In the simulations presented here we use $C_r = 15$, $C_a = 2.5$ and $l_a = 0.1$, $l_r = 0.05$. We note that the force between particle i and particle j is in the direction of the vector,

$$
r_{i,j}^\tau = \vec{x}_i(t) - \vec{x}_j(t) + \tau(\vec{v}_i(t) - \vec{v}_j(t)).
\tag{3}
$$

This implies that for $\tau > 0$ the direction of the force is generally not parallel to the vector $\vec{x}_i(t) - \vec{x}_j(t)$, pointing from particle i to particle j, but is influenced by the velocity of the two particles.

In the following we consider particles moving in two dimensions with open boundaries. We typically initialize the system with the particles at random positions within a disk and zero initial velocities. All numerical solutions have been carried out using a Runge-Kutta method of order 8 from the GNU Scientific Library [15].

III. RESULTS

Before investigating the impact of anticipation on the dynamics we need to set the anticipation time τ into relation with other time and length scales present in the system. The model has one natural spatial scale, given by the minimum of the Morse potential,

$$
\bar{r}_p = \mathrm{argmin}\, U(r),
\tag{4}
$$

which, with the parameters used here, is given by

$$
\bar{r}_p = \frac{l_a l_r}{l_a - l_r} \ln \frac{C_r/C_a}{l_r/l_a} \approx 0.25.
\tag{5}
$$

The anticipation time, τ, gives one natural time scale, but because the system is defined in the plane without boundaries it is difficult to define time scales or spatial scales valid asymptotically for long time intervals, for comparison with $\bar{r}_p$ and τ. However, it is possible to define a relevant transient time scale expressed in terms of the initial data. We let

$$
\bar{f}_T = \mathbb{E}[|\nabla U(|\vec{x}_j - \vec{x}_k|)|]
\tag{6}
$$

be the expected interparticle force, where the expectation is taken over a randomly chosen pair of initial positions of the particles. From this we define a transient time scale

$$
\bar{t}_T = \sqrt{2\bar{r}_p/\bar{f}_T},
\tag{7}
$$

which is the time needed for a particle with initial velocity 0 to move a distance $\bar{r}_p$ when subjected to a constant force of strength $\bar{f}_T$.

By solving the equations of motion, (1), numerically with initial positions uniform random within a disk and initial velocities equal to 0, we observe for $\tau/\bar{t}_T \approx 1$, independent of the particle number, the rapid formation of milling structures, where the particles organize into rotating patterns. For smaller $\tau/\bar{t}_T \ll 0.1$ the particles behave much like the classical system, where nearby particles tend to form clusters, whereas others disperse. For $\tau/\bar{t}_T \gg 10$ particles also disperse, but do so in smaller clusters of aligned particles. The following subsections are devoted to understanding this behavior for both small and large particle numbers.

A. Small particle numbers

To get a better understanding of the effect of anticipation we start by looking at the dynamics of small systems with $N \in [2,11]$ and set $\tau = 1$. The particles are placed at random positions in the unit circle and the initial velocities are 0 for all particles. This implies that the initial density and therefore the transient time scale will vary with N, but for small N we have that τ and $\bar{t}_T$ are of the same order of magnitude.

For $N = 2$ the system behaves much like the two-particle system mentioned in the introduction. The particles attract one another and rapidly converge to a distance given by the equilibrium distance of the pairwise potential. This suggests that energy is being dissipated, a fact we return to later.

For $N = 3$, the particles start rotating in the same orbit around their common center of gravity. A similar behavior is seen for larger systems, but the number of unique orbits and their radii depend on the number of particles N (see Fig. 1). We note that the configurations closely resemble
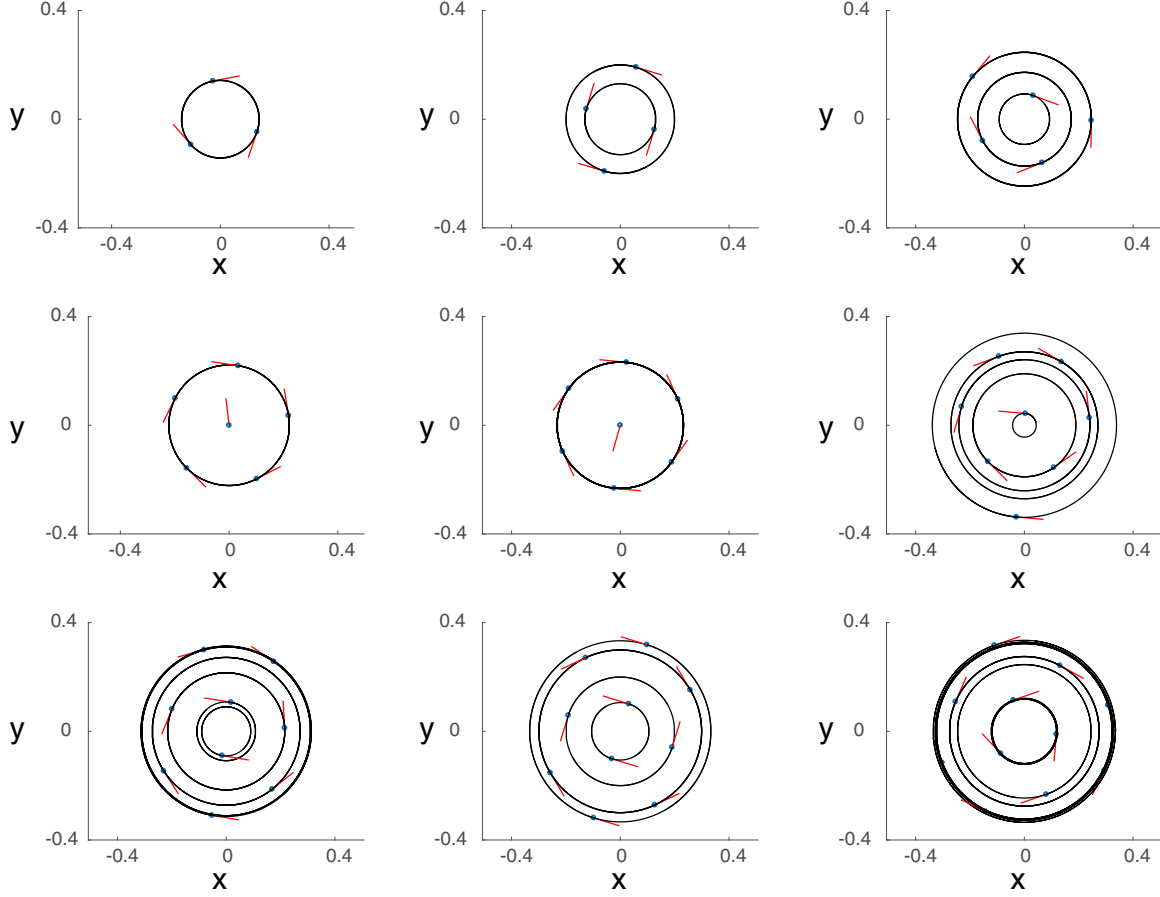
FIG. 1. Rotational configurations observed for $N \in [3,11]$. Initial position were chosen at random in the unit circle and initial velocities were 0. Particles are shown as circles and red lines correspond to the unit velocity vector of each particle. Lines show the trajectories of the particles, where the first 10 time steps have been discarded.

a partial hexagonal lattice built from equilateral triangles. These patterns are rigid rotations of a locally crystalline configuration, much like the dynamics reported in region VI in [4]. The main difference is that in our case there is no self-propulsion, and the particles are only influenced by a pairwise potential with anticipation.

1. Analysis of milling patterns

The reason behind the observed rotation for $N > 2$ can be understood by considering the forces acting on each particle (see Fig. 2). For simplicity we consider the case $N = 3$ and formulate the following simplified model, which only describes the case where the particles are symmetrically distributed around the origin.

We assume that the three particles are symmetrically distributed, i.e., with angle $\frac{2\pi}{3}$ between them; lie on a circle of radius $r(t)$; and move at the same speed $|\vec{v}(t)|$ in the clockwise direction. The particles move at velocity $\vec{v}(t)$, which has both a tangential and a radial component. The nonzero velocity gives rise to anticipated positions that lie on a circle with radius $R(t)$. A third radius is also of importance, R_0, which is the radius at which the interparticle distance equals the equilibrium distance of the potential. For the case $N = 3$, we

have that the equilibrium (or resting) radius is

$$R_0 = \frac{l_a l_r \ln\left(\frac{C_r l_a}{C_a l_r}\right)}{\sqrt{3}(l_a - l_r)}.$$

A symmetric configuration on R_0 implies that the force from the potential $\vec{F}(R_0) = 0$, where $\vec{F}(R)$ is the net force acting on each particle when at a distance R from the center. If $R(t) > R_0$, the net force will be radially inwardly directed from the point $\vec{R}(t)$. Now this force, $\vec{F}(R(t))$ is applied not at point $\vec{R}(t)$, which is the anticipated position on the outer circle but, instead, at the current positions at $\vec{r}(t)$. By decomposing this force into its radial and tangential components we get a triangle which is similar to the right-angled triangle that has $\vec{R}(t)$ as the hypotenuse (see Fig. 2). The net force can be decomposed into a radial component, F_r, which is generating a central force, and a tangential force F_θ, which is retarding the rotation. Due to the fact that the anticipated position is outside R_0, the particles will experience attractive forces, although they in fact are located in the repulsive region of the potential. However, note that this is only true if $|\vec{v}(t)|\tau$ is large enough that the anticipated position lies outside of R_0.

To obtain a quantitative understanding of the three-particle system we formulated an ODE system for the special case of N particles located on a circle (as in Fig. 2). We describe the dynamics of a single particle, but since we assume a symmetric
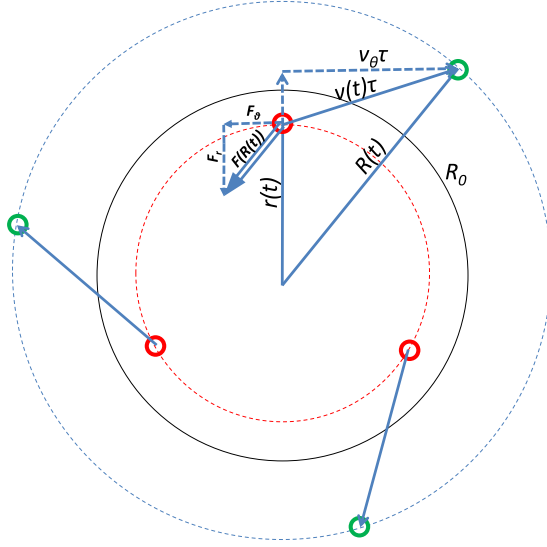
FIG. 2. Schematic of a symmetric three-particle system that is rotating on a circle with a radius $r(t)$ which is smaller than the equilibrium radius R_0. Small red circles show the current positions and green circles denote the anticipated positions obtained by adding the vector $\vec{v}(t)\tau$ to the current position, i.e., $\vec{R}(t) = \vec{r}(t) + \vec{v}(t)\tau$. Since the anticipated position is located outside the equilibrium circle, there will be an attracting force applied at the actual position $\vec{r}(t)$. By decomposing this force, we get a central acceleration and a tangential retardation that eventually makes the speed of the particles go to 0.

configuration the positions and velocities of the other particles can be obtained by shifting the solution appropriately. The equations of motion are given by

$$\dot{\vec{r}}(t) = \vec{v}(t),$$

$$\dot{\vec{v}}(t) = -F(R(t))\frac{\vec{R}(t)}{R(t)},$$
(8)

where $\vec{R}(t) = \vec{r}(t) + \tau\vec{v}(t)$ and $R(t) = |\vec{R}(t)|$. The force is given by

$$F(R(t)) = 2\sum_{k=1}^{\frac{N-1}{2}}\left(\frac{C_a}{l_a}e^{-\alpha_k/l_a} - \frac{C_r}{l_r}e^{-\alpha_k/l_r}\right)\cos\left(\frac{\pi - k\frac{2\pi}{N}}{2}\right),$$
(9)

where

$$\alpha_k = \sqrt{2}R(t)\sqrt{1 - \cos\left(k\frac{2\pi}{N}\right)}$$
(10)

is the interparticle distance between the focal particle and particle k.

We solve this system numerically for the the case $N = 3$ (see Fig. 3), and observe two distinct phases: (i) an accelerating phase, where the particles move onto a circle with a radius slightly smaller than the equilibrium distance of the potential, and (ii) a milling phase, in which the particles in unison slowly spiral outwards towards the equilibrium circle with radius R_0.

From equation system (8) and Fig. 2, we can obtain an analytic expression of how the speed decays in the milling phase in the following way. The second equation in system (8)
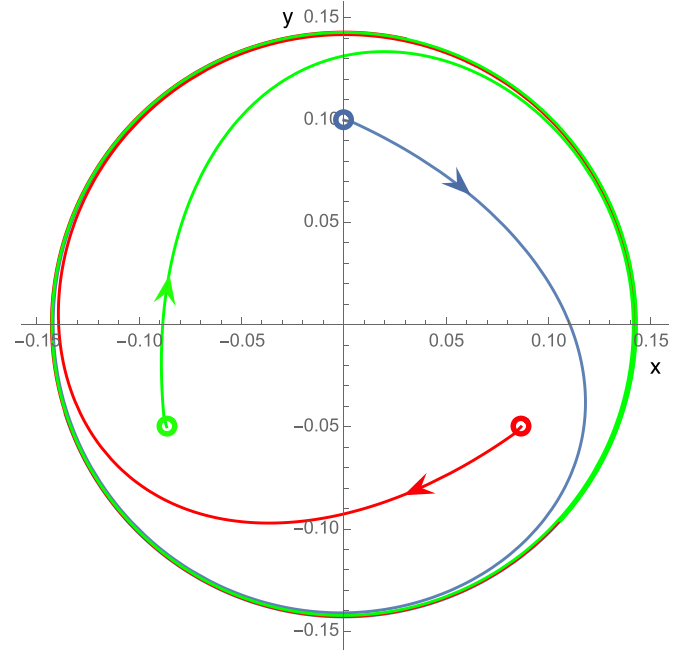

FIG. 3. Solution of the ODE system, (8), with the focal particle (blue circle) at initial position $\vec{r}(0) = (0, 0.1)$ with $\vec{v}(0) = (0, 0.35)$. Particles spiral out towards the equilibrium radius R_0.

gives that

$$\dot{v}_\theta = -F_\theta,$$

where v_θ is the tangential velocity. By the similarity of the two right-angled triangles in Fig. 2 we have that

$$F_\theta = \frac{F(R(t))}{R(t)}v_\theta\tau.$$

Note that both $F(R(t))$ and v_θ tend to 0 as $t \to \infty$. We are interested in studying how quickly that happens and, therefore, assume that the system has reached the milling phase. This means that the particles are orbiting very close to the equilibrium circle and that $F(R(t)) \approx F_r$ and $R(t) \approx R_0$. By making these assumptions, we can simplify the differential equation in the following way:

$$\dot{v}_\theta = -F_\theta = -\frac{F(R(t))}{R(t)}v_\theta\tau \approx -\frac{F_r}{R_0}v_\theta\tau = -\frac{v_\theta^2}{R_0^2}v_\theta\tau,$$

where the last equality holds because the particles are following a circular movement of radius R_0, for which we have $F_r = v_\theta^2/R_0$. Thus, we end up with the asymptotic separable differential equation

$$\dot{v}_\theta = -\frac{v_\theta^3\tau}{R_0^2}$$

with solution

$$v_\theta(t) = \frac{1}{\sqrt{\frac{2\tau t}{R_0^2} + \frac{1}{v_\theta(0)^2}}}.$$
(11)

From this expression it is clear that $|v(t)| \sim t^{-1/2}$ as $t \to \infty$, and from this we conclude that the kinetic energy should go as $K(t) = m|v(t)|^2/2 \sim t^{-1}$. This result is in
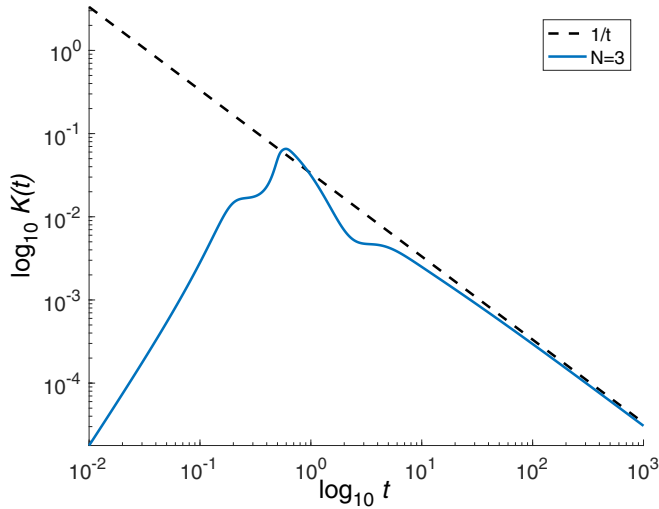
FIG. 4. Kinetic energy $K(t)$ as a function of time for $N = 3$. The dashed line corresponds to $1/t$. After an initial acceleration the kinetic energy decays approximately as $1/t$ in agreement with (11).

agreement with the long-term kinetic energy calculated for a three-particle system with random initial conditions (see Fig. 4). This indeed shows that the milling phase is transient, but also that the rate of energy dissipation is low and the kinetic energy scales as $1/t$.

B. Large particle numbers

We now consider systems with $N \gg 10$ particles and look for large-scale patterns in the dynamics. Figure 5 contains snapshots from a simulation with $N = 100$ particles which show that the particles first aggregate into local clusters ($t \approx 12$) that eventually coalesce and form a single milling structure ($t \approx 90$). Similar dynamics are seen with $N = 1000$ particles initialized at the same density (see Fig. 6), with the difference that not all clusters merge within the duration of the simulation. Although not visible due to the large system size, the particles are organized in a hexagonal lattice.

In the case of $N = 1000$ particles initialized at a 10-fold higher density (see Fig. 7) we observe an initial repulsive phase
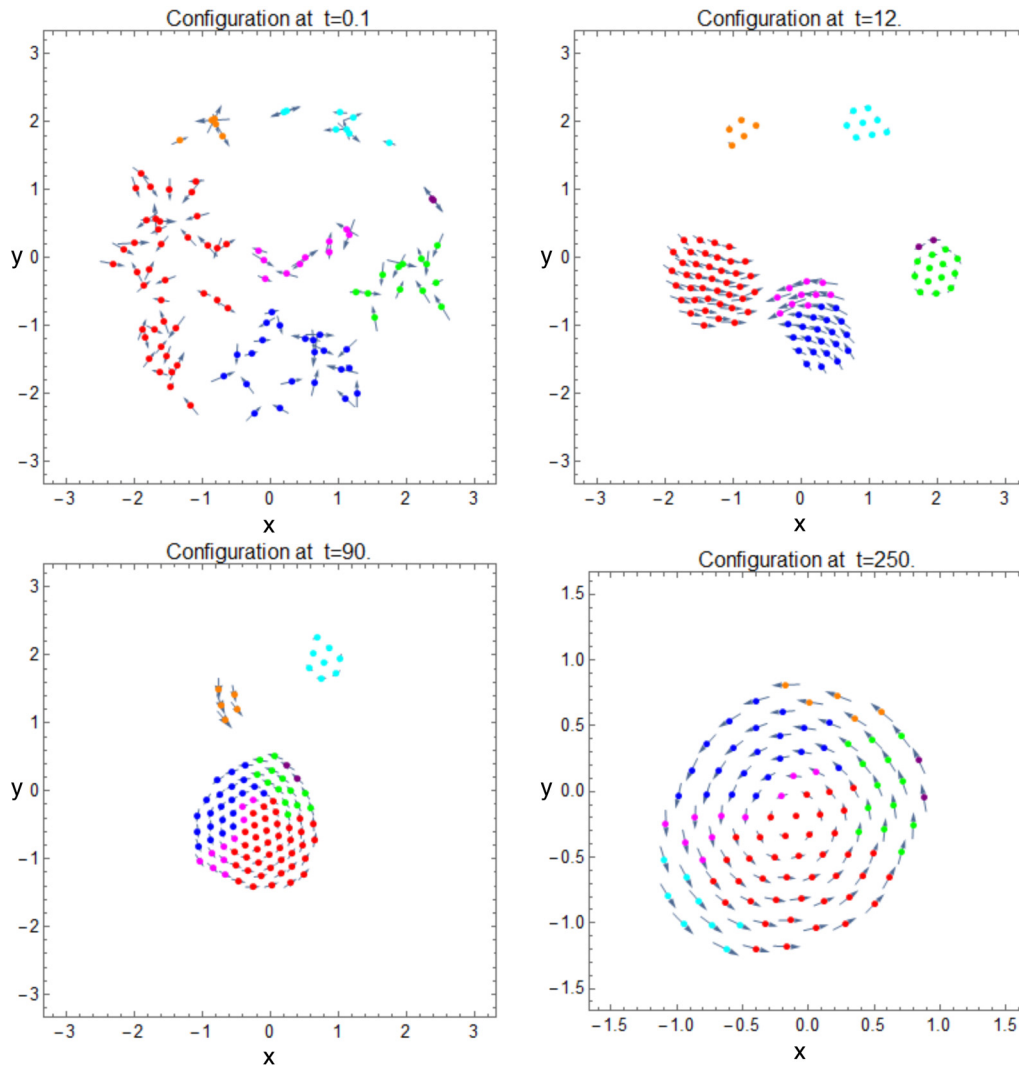

FIG. 5. Positions of 100 particles at different times, starting from a random initial configuration with zero velocities. The initial distribution was an asymmetric inside a disk of diameter 2.5. The anticipation time $\tau = 1$, and the transient time scale is approximately 1.8. Particles are coloured according to the cluster they belong to at $t = 12$.
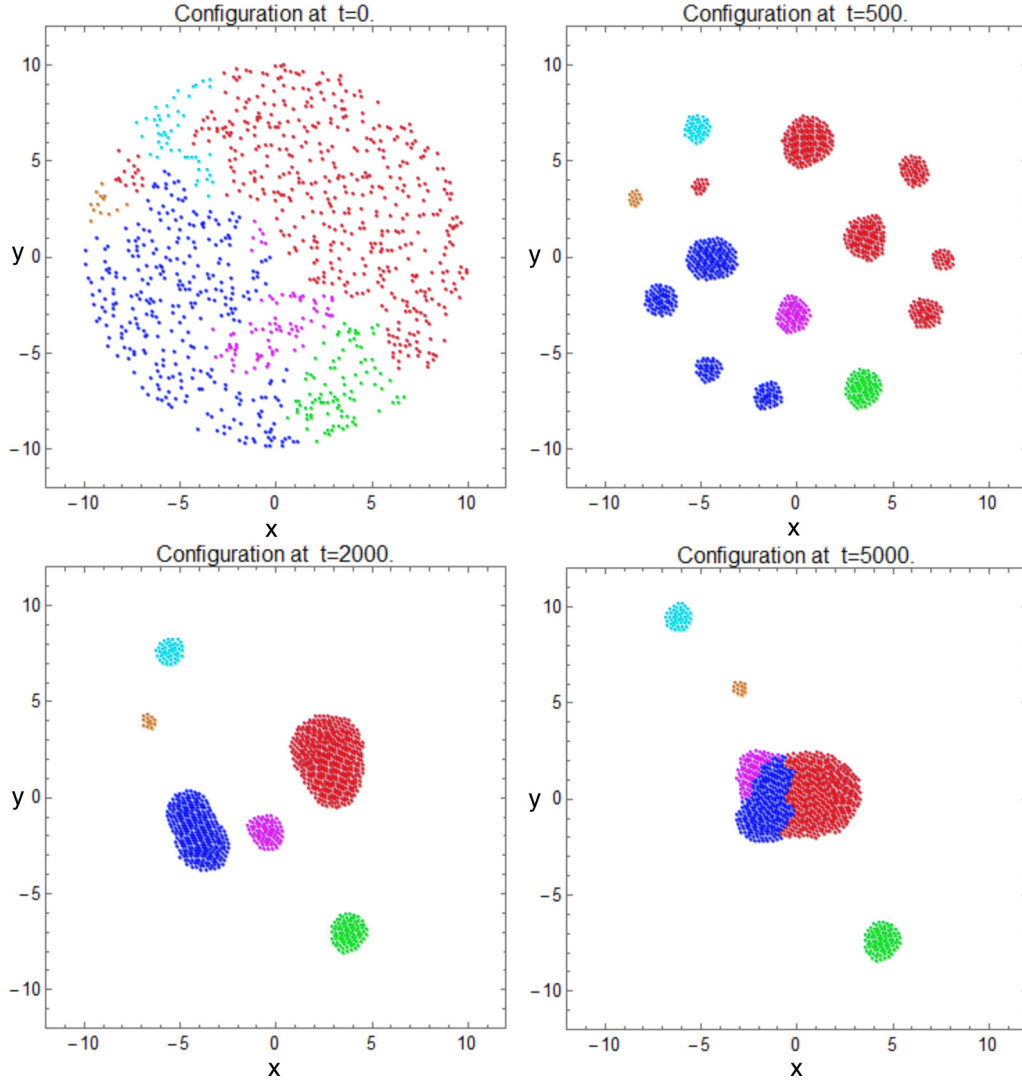
FIG. 6. Positions of 1000 particles at different times, starting from a random initial configuration with zero velocities. The initial distribution was uniform in a disk with radius 10.0. The anticipation time $\tau = 0.25$, and the transient time scale equals approximately 9. Particles are colored according to the cluster they belong to at $t = 2000$.

(at $t \approx 4.5$), which leads to an expansion of the cluster (at $t \approx 15$). However, the cluster remains coherent and contracts into an approximately hexagonal lattice, which rotates around its center of gravity. In conclusion, we observe similar dynamics independent of the system size. The system, which is initially disordered, self-organizes into an approximately hexagonal lattice that rotates as a rigid body around its center of gravity.

It appears as if the inclusion of anticipation allows the system to dissipate energy and reach a state which minimizes the potential energy. In order to investigate this we compared the configuration in Fig. 5 at $t = 250$ with a configuration of stationary particles that minimizes the potential energy (calculated using Mathematica's function `Minimize`). The comparison is shown in Fig. 8 and reveals almost-perfect agreement between the milling configuration obtained with anticipation (black circles) and the stationary configuration that minimizes the potential energy, which consists of a partial hexagonal lattice.

We now move on to the question of energy dissipation for large N, which seems to play a crucial role in the dynamics of the system.

C. Energy dissipation

The example with two spring-coupled particles and the special case of a symmetric configuration of three particles show that anticipation may result in energy dissipation. The following calculation shows that the anticipated energy, $E(t)$, computed as the sum of the kinetic energy and the anticipated potential energy,

$$E(t) = \sum_{i=1}^{N} |\vec{v}_i|^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} U(\vec{x}_i - \vec{x}_j + \tau(\vec{v}_i - \vec{v}_j)), \quad (12)$$

is strictly decreasing along the orbits when not all velocities are constant. Here we consider a more general anticipating
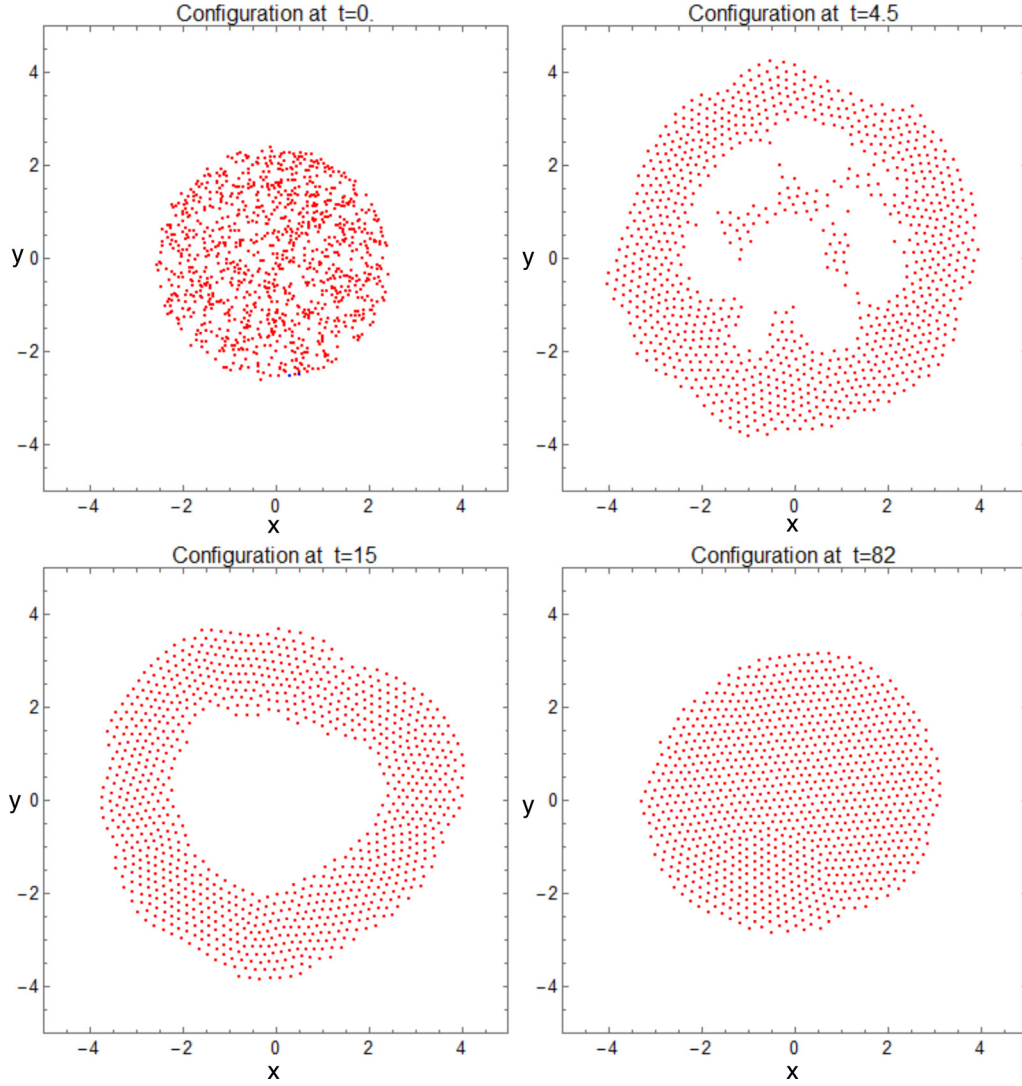
FIG. 7. Positions of 1000 particles at different times, starting from a random initial configuration with zero velocities. The initial distribution was uniform in a disk with radius 2.5. The anticipation time $\tau = 0.05$, and the transient time scale equals approximately 2.3.

dynamical system, of which the particle system studied above is but a special case:

$$\dot{\vec{x}} = \vec{v}, \quad \dot{\vec{v}} = -\nabla U(\vec{x} + \tau\vec{v}), \qquad (13)$$

where $\vec{x}, \vec{v} \in \mathbb{R}^n$, and $U : \mathbb{R}^n \rightarrow \mathbb{R}$ is a smooth and bounded potential. Then

$$\frac{d}{dt}(E(t)) = \frac{d}{dt}\left(\frac{|\vec{v}|^2}{2} + U(\vec{x} + \tau\vec{v})\right) \qquad (14)$$

$$= \vec{v} \cdot \dot{\vec{v}} + \nabla U(\vec{x} + \tau\vec{v}) \cdot (\dot{\vec{x}} + \tau\dot{\vec{v}})$$

$$= \vec{v} \cdot \underbrace{(\dot{\vec{v}} + \nabla U(\vec{x} + \tau\vec{v}))}_{=0} - \tau|\dot{\vec{v}}|^2. \qquad (15)$$

Hence the sum of the kinetic energy and the anticipated potential energy is strictly decreasing until all velocities are constant. This behavior is shown in Fig. 9(a), which shows how the energy evolves for the simulation presented in Fig. 5 with $N = 100$ particles. The anticipated energy is strictly decreasing and the jumps correspond to collision and merger of clusters. Note that the classical energy (calculated with

$\tau = 0$) is not strictly decreasing but increases transiently during collisions.

1. The rate of dissipation

For the three-particle system we could show that the energy decays as $1/t$. We now analyze the rate of energy dissipation for a general potential $U(x)$ and arbitrary system size N.

With a potential given by a positive definite quadratic form, $U(x) = \vec{x}^T A \vec{x}$, the system is again a damped oscillator, whose solutions converge exponentially to 0, but for a semidefinite form, say $U = \frac{1}{2}\lambda_2 x_2^2 + \cdots + \frac{1}{2}\lambda_n x_n^2$, the system never comes to rest, because the potential is constant along the degenerate x_1 direction. The anticipation does not change this, and seen from the point of view of simulations, the same will be observed if the quadratic form has an almost-degenerate subspace, in the sense that some of the eigenvalues of A are several magnitudes smaller than the the dominant eigenvalues.

However, if the potential U is constant, or nearly constant, on a curved manifold, e.g., a curve, the situation is different. Three particles on a circle is an example of this more
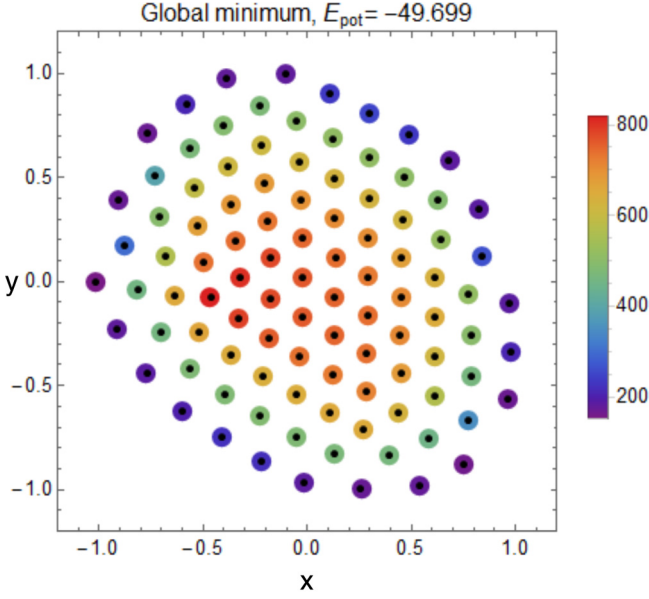
FIG. 8. Black circles show the final point configuration; colored circles, a local energy minimum. The colors represent the stiffness of each particle position, calculated from the largest eigenvalue of the linearized system as a function of only one particle, given that the others are fixed.
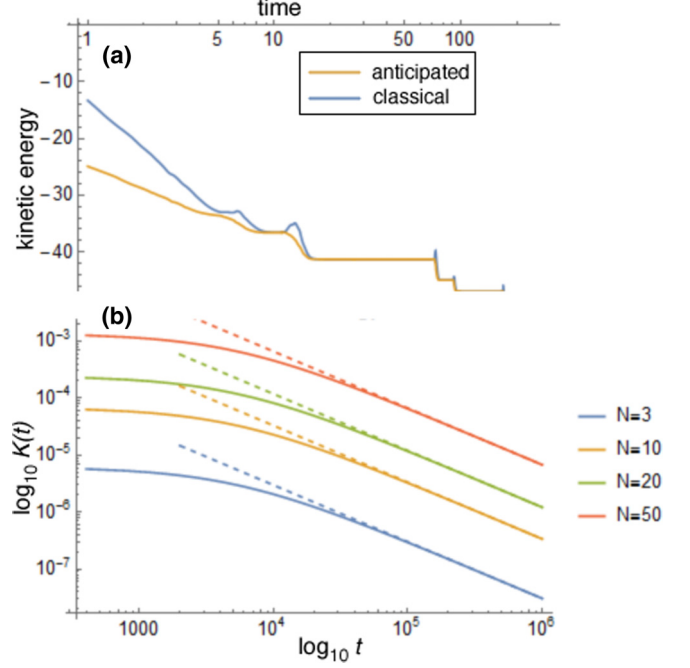

FIG. 9. (a) Evolution of the energy for the simulation represented in Fig. 5. The jumps in the two curves correspond to the merger of clusters. (b) Evolution of the kinetic energy for $N = 3$, 10, 20, and 50. The initial condition was chosen with positions according to a global minimum of the potential and velocities corresponding to a rigid rotation of the point configuration. The slope of the dashed lines is -1.

general situation. Here the potential attains its minimum on a one-dimensional curved submanifold, the circle, with $\vec{x}(t)$ very close to the circle and with $\vec{v}(t)$ almost tangential to the manifold. A Taylor expansion of the potential U then gives

$$\frac{d}{dt}\vec{v}_{\parallel}(t) = \mathcal{O}(|\vec{v}_{\parallel}|^2),$$

at least if $\tau|\vec{v}_{\parallel}|$ is small compared to the curvature of the manifold. This calculation can be carried out in a more general setting, for example, when the minimum of U is attained along a curve in $\mathbb{R}^n$: Assume that $\Gamma = \{\vec{\phi}(x) = (\phi_1(s),\ldots,\phi_n(s)) \mid s \in I \subset \mathbb{R}\}$ is a segment of a curve parametrized by s and assume that $U(x)$ is a quadratic function of $\vec{x} - \vec{x}^*$, where $\vec{x}^*$ is the point on Γ closest to $\vec{x}$. Though not completely general, this would cover many generic situations. We may choose a coordinate system so that

$$\phi_k(0) = 0 \quad \text{for } k = 1,\ldots,n,$$
$$\phi_1(s) = s,$$
$$\phi_k'(0) = 0 \quad \text{for } k = 2,\ldots,n,$$
$$\phi_k''(0) = \mu_k \quad \text{for } k = 1,\ldots,n,$$

so that the curve is tangent to the x_1 axis of the coordinate system and it may be parametrized with x_1. The point $\vec{x}^* = (s,\phi_2(s),\ldots,\phi_n(s))$ is found by minimizing the function $\sum(x_k - \phi_k(s))^2$, and at a minimum, this must satisfy

$$s - x_1 + \sum_{k=2}^{n}(\phi_k(s) - x_k)\phi_k'(s) = 0, \qquad (16)$$

which determines s as a function of $\vec{x}$, at least near the origin. Hence

$$U(x_1,\ldots,x_n) = \frac{1}{2}(\vec{x} - \vec{\phi}(s))^T A(\vec{x} - \vec{\phi}(s))$$
$$= \frac{1}{2}v_1(x_1 - s)^2 + \frac{1}{2}\sum_{k=2}^{n}v_k(\phi_k(s) - x_k)^2,$$

where, for simplicity, we have assumed that A is diagonal with diagonal elements v_k. Restricting Eq. (16) to points on the x_1 axis, we find that

$$x_1(s) = s + \sum_{k=2}^{n}\frac{1}{2}\mu_k s^3 + \mathcal{O}(s^4) \quad \text{or}$$
$$s(x_1) = x_1 - \sum_{k=2}^{n}\frac{1}{2}\mu_k x_1^3 + \mathcal{O}(x_1^4),$$

so that

$$U(x_1,0,\ldots,0) = \mathcal{O}(s^6) + \frac{1}{2}\sum_{k=2}^{n}v_k\left(\frac{1}{2}\mu_k s^2 + \mathcal{O}(s^3)\right)^2$$
$$= \frac{1}{2}\sum_{k=2}^{n}\sum_{j=2}^{n}\mu_j\mu_k v_k x_1^4 + \mathcal{O}(x_1^5).$$

Therefore, with $\vec{v} = (v_1,0,\ldots,0)$ and $\vec{x} = \vec{0}$,

$$\frac{d}{dt}\vec{v} = -\nabla U(\vec{x} + \tau\vec{v}) = -\nabla U(\tau v_1,0,\ldots,0),$$

and, finally,

$$\frac{d}{dt}\frac{|\vec{v}|^2}{2} = -\vec{v} \cdot \nabla U(\tau v_1, 0, \ldots, 0) \qquad (17)$$

$$= -\gamma v_1 \tau^3 v_1^3 + \mathcal{O}(v_1^5) \qquad (18)$$

$$= -\gamma \tau^3 v_1^4 + \mathcal{O}(v_1^5), \qquad (19)$$

where $\gamma = 2\sum_{k=2}^{n}\sum_{j=2}^{n}\mu_j\mu_k \nu_k$, and therefore $\frac{d}{dt}|\vec{v}|^2 \sim -|\vec{v}|^4$, which implies that

$$|\vec{v}(t)|^2 \sim \frac{1}{t}$$

when t is large, and the exact expression depends on the constants μ_k, which determine the curvature of the curve. Of course this is not a full proof but, rather, motivation for the observed behavior. A complete proof would require an additional calculation showing that the component of $\vec{F}$ normal to Γ is sufficiently strong to confine $\vec{x}(t)$ to remain sufficiently close to Γ, and $\vec{v}(t)$ to rest almost tangent to Γ, that the estimates above are still valid.

For the N-particle systems with pair interactions that are the main subject of this paper, the potential energy is invariant under rotations and translations, and hence the potential minimum is attained on a three-dimensional curved submanifold of the $2n$-dimensional configuration space. In simulations with a large number of particles it is difficult to observe the $1/t$ behavior, however, because the dynamics is complex and one may need to wait a very long time before this asymptotic behavior can be seen. To circumvent this problem we have computed numerical solutions with initial data such that the positions $\vec{x}_i(0)$ correspond to a minimum for the potential energy, and with initial velocities corresponding to a rigidly rotating body, and find excellent agreement with the predicted $1/t$ decay of energy [see Fig. 9(b)].

Situations where the minimum manifold has a higher dimension appear in simulations with a large number of particles, where we often observe the system breaking up into smaller clusters that essentially do not interact with the other clusters and, hence, represent a case of a many-dimensional subspace of almost-constant potential energy, defined by translations and rotations of each cluster.

D. Linear and angular momenta

So far we have shown that anticipation induces dissipation and generally leads to a $1/t$ decay in kinetic energy. This leads to a gradual reduction in the rotation of the cluster, but we have not explained how this rotation comes about in the first place. To do this we need to analyze the time evolution of the linear and angular momenta, $\vec{M}(t)$ and $L(t)$, which we define as

$$\vec{M}(t) = \sum_{i=1}^{N}\vec{v}_i \quad \text{and} \quad L(t) = \sum_{i=1}^{N}\vec{v}_i \times \vec{x}_i, \qquad (20)$$

where $\vec{v}_i \times \vec{x}_i$ denotes the cross product of $\vec{v}$ and $\vec{x}$, which is a scalar in this two-dimensional setting. First,

$$\frac{d}{dt}\vec{M}(t) = \sum_{i=1}^{N}\frac{d}{dt}\vec{v}_i = \sum_{\substack{i,j \\ i \neq j}} F_{i,j}$$

$$= \frac{1}{2}\sum_{\substack{i,j \\ i \neq j}}(F_{i,j} + F_{j,i}) = 0. \qquad (21)$$

Upon a change of variables we may assume that $M(t) = \vec{0}$ and, hence, that the center of mass $\frac{1}{N}\sum_i \vec{x}_i$ is constant and may, therefore, be set equal to 0. The angular momentum satisfies

$$\frac{d}{dt}L(t) = \sum_i (\dot{\vec{v}}_i \times \vec{x}_i + \vec{v}_i \times \dot{\vec{x}}_i)$$

$$= \sum_{\substack{i,j \\ i \neq j}} F_{i,j} \times \vec{x}_i = \frac{1}{2}\sum_{\substack{i,j \\ i \neq j}} F_{i,j} \times (\vec{x}_i - \vec{x}_j). \qquad (22)$$

Note that the force $F_{i,j}$ may be written

$$F_{i,j} = -\frac{U'(|r_{i,j}^\tau|)}{|r_{i,j}^\tau|}r_{i,j}^\tau \equiv -V(|r_{i,j}^\tau|)r_{i,j}^\tau, \qquad (23)$$

where $r_{i,j}^\tau = \vec{x}_i - \vec{x}_j + \tau(\vec{v}_i - \vec{v}_j)$, and we define $V(r) = U'(r)/r$. Therefore

$$\frac{d}{dt}L(t) = -\frac{1}{2}\tau \sum_{\substack{i,j \\ i \neq j}} V(|r_{i,j}^\tau|)(\vec{v}_i - \vec{v}_j) \times (\vec{x}_i - \vec{x}_j). \qquad (24)$$

So with $\tau = 0$, the angular momentum is constant, as it should be for a classical system. All simulations presented here start with all initial velocities equal to 0, and therefore also $\dot{L}(0) = 0$, but the second derivative need not be. Computing $\ddot{L}(t)$ gives

$$\frac{d^2}{dt^2}L(t) = \sum_{\substack{i,j \\ i \neq j}} \Bigg(V'(|r_{i,j}^\tau|)\frac{r_{i,j}^\tau \cdot \dot{r}_{i,j}^\tau}{|r_{i,j}^\tau|}(\vec{v}_i - \vec{v}_j) \times (\vec{x}_i - \vec{x}_j)$$

$$+ V(|r_{i,j}^\tau|)(\dot{\vec{v}}_i - \dot{\vec{v}}_j) \times (\vec{x}_i - \vec{x}_j) \Bigg). \qquad (25)$$

At $t = 0$, when the velocities are 0, the first terms in the sum disappear, and using the symmetries in the sum this yields

$$\frac{d^2 L}{dt^2}(0) \qquad (26)$$

$$= \frac{\tau}{2}\sum_{\substack{i,j,k \\ i \neq j \neq k}} V(|\vec{x}_i - \vec{x}_j|)V(|\vec{x}_i - \vec{x}_k|)$$

$$\times (\vec{x}_i - \vec{x}_j) \times (\vec{x}_i - \vec{x}_k), \qquad (27)$$

which typically will be nonzero for a random initial configuration. Therefore, contrary to conservative systems, asymmetries in the initial configuration of a particle system initially at rest may result in a nonzero angular momentum, which is then observed as a milling structure.

The time evolution of the angular momentum of a typical simulation is exhibited in Fig. 10(a), which shows $L(t)$ for the simulation presented in Fig. 5. The jumps in the curve correspond to the merger of clusters, and we note that collisions typically increase the absolute value of the angular momentum, while it decreases between collisions due to the $1/t$ dissipation
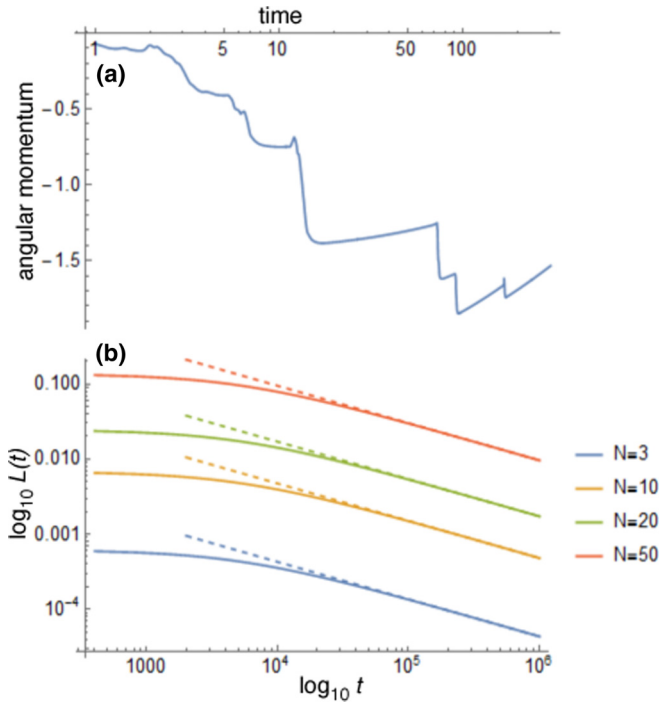
FIG. 10. (a) Evolution of the angular momentum for the simulation represented in Fig. 5. The jumps in the curve correspond to the merger of clusters. (b) Evolution of the angular momentum for $N = 3, 10, 20$, and 50. The initial condition was chosen with positions according to a global minimum of the potential and velocities corresponding to a rigid rotation of the point configuration. The slope of the dashed lines is $-1/2$.

of kinetic energy discussed above. Since $L(t)$ is proportional to the speed $v(t)$, which scales as $1/\sqrt{t}$, we expect the angular momentum to asymptotically scale as $1/\sqrt{t}$, which is precisely what is observed in simulations [see Fig. 10(b)].

IV. DISCUSSION

In this paper we have shown that internal prediction or anticipation may lead to milling behavior in a system of particles that interact via a pairwise potential. The behavior of the system depends on the relation between the anticipation time τ and the transient time scale $\bar{t}_T$, (7), such that when τ is of the same order of magnitude as $\bar{t}_T$ we observe the rapid formation of rotating clusters that merge upon collision. These clusters have a crystalline structure made up of a hexagonal arrangement of particles and rotate around the center of gravity as a rigid body. We have shown analytically that rotation

emerges due to an increase in the angular momentum, which only occurs for $\tau > 0$. Further, we have shown that anticipation leads to dissipative behavior where the kinetic energy typically decreases as $1/t$.

These results show that motion anticipation—the ability to extrapolate the future location of their peers—which is known to be present in humans and other animals, has a large impact on the dynamics of interacting particle models that are often used for describing collective behavior. It acts to stabilize the system, which leads to an ordered crystalline structure and, also, induces milling, which is a ubiquitous feature of flocking—even in the absence of self-propulsion.

Using models that incorporate pairwise forces and anticipation it would be possible to analyze existing data on, e.g., pedestrians, to infer the anticipation time τ that is used by humans. In contrast to the work by Karamouzas *et al.* [12], in which an interaction potential that depends on the "time to collision" is inferred, our approach would allow for an estimate of the prediction horizon used by humans. This could then be compared with the anticipation times obtained from other species and other systems and could give novel insight into how anticipation influences collective behavior at different spatial and temporal scales.

In relation to experimental data we note that the fixed value of τ most likely is a simplification and that most animals (including humans) make use of a dynamical value of τ which depends on the distance to the object whose trajectory one is trying to predict. For longer distances a larger value of τ is beneficial, while for shorter distances one runs the risk of overshooting unless τ is chosen sufficiently small.

In the future it would be interesting to investigate the dynamics of this type of adaptive anticipation and, also, the impact of anticipation in more realistic models that contain self-propulsion. Another possible extension of the current model is to take into account the fact that an individual might have knowledge about the predictions made by other individuals. That knowledge could be captured in an internalized model that would give individual predictions that go beyond the simple linear extrapolation used in the current model.

[1] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, J. Theor. Biol. **218**, 1 (2002).

[2] K. Tunstrøm, Y. Katz, C. C. Ioannou, C. Huepe, M. J. Lutz, and I. D. Couzin, PLoS Comput. Biol. **9**, e1002915 (2013).

[3] Y. Katz, K. Tunstrøm, C. C. Ioannou, C. Huepe, and I. D. Couzin, Proc. Natl. Acad. Sci. USA **108**, 18720 (2011).

[4] M. R. D' Orsogna, Y. L. Chuang, A. L. Bertozzi, and L. S. Chayes, Phys. Rev. Lett. **96**, 104302 (2006).

[5] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Phys. Rev. Lett. **75**, 1226 (1995).

[6] R. Nijhawan, Nature **370**, 256 (1994).

[7] M. J. Berry, I. H. Brivanlou, T. A. Jordan, and M. Meister, Nature **398**, 334 (1999).

[8] G. Schwartz, S. Taylor, C. Fisher, R. Harris, and M. J. Berry II, Neuron **55**, 958 (2007).

[9] T. S. Collett and M. F. Land, J. Comp. Physiol. **125**, 191 (1978).

[10] B. G. Borghuis and A. Leonardo, J. Neurosci. **35**, 15430 (2015).

[11] S. Rossel, J. Corlija, and S. Schuster, J. Exp. Biol. **205**, 3321 (2002).

[12] I. Karamouzas, B. Skinner, and S. J. Guy, Phys. Rev. Lett. **113**, 238701 (2014).

[13] A. Morin, J.-B. Caussin, C. Eloy, and D. Bartolo, Phys. Rev. E **91**, 012134 (2015).

[14] H. Murakami, T. Niizato, and Y.-P. Gunji, Sci. Rep. **7**, 46447 (2017).

[15] B. Gough, *GNU Scientific Library Reference Manual* (Network Theory Ltd., 2009); http://www.network-theory.co.uk/.