.

Characterizing Non-Equilibrium Behavior in a Simulated Magnetic Colloidal Fluid

Shiv Seshan
(Dated: May 3rd 2024)

When driven by an external magnetic field, a two-dimensional colloid of microscopic magnets behaves like a chiral fluid. On account of its exotic chiral behavior, such a fluid remains poorly characterized by physical theory. We apply a novel computational technique from active learning–a curiosity-driven search (curiosity)–to probe the qualitative behaviors that a simulated chiral fluid can exhibit. Based on the results of this search, we also propose order parameters to quantify these qualitative beehaviors. Our *in silica* findings suggest that curiosity holds promise for the characterization of behavior in driven magnetic colloidal fluids.

I. INTRODUCTION

A. Motivation

1. Magnetic Colloidal Fluids

Symmetries under time-reversal and parity are ubiquitous in the natural world. As such, fluids which are engineered to break these symmetries can exhibit exotic physical properties. Recently, a fluid was designed to violate these symmetries on account of the organized rotational motion of its particles [7]. This fluid exhibits a poorly-understood, exotic form of hydrodynamics–chiral hydrodynamics–which lacks a clear analogue in the largely achiral natural world.

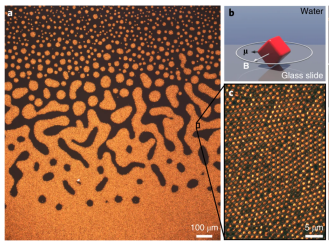

FIG. 1. a): An optical image of a colloid of magnets in bulk (scalebar 100 μm); b) A schematic of an individual microscopic magnet with magnetic dipole moment μ under the influence of a constantly-rotating external field; c) A 10-μm-scale image of the individual magnets [7].

Although easy to imagine, controlling the microscopic rotation of fluid particles in an organized fashion is no easy task. One way to induce synchronized rotational behavior in a fluid is to construct a two-dimensional array of microscopic magnets and subject them to a spatially uniform magnetic field of the form:

$$\mathbf{B}(t) = B\cos(B_\theta(t))\hat{x} + B\sin(B_\theta(t))\hat{y} \tag{1}$$

for a time-invariant magnitude B and a periodic function $B_\theta(t)$ specifying the polar angle of the B-field. Taking $B_\theta(t)$ to be the constantly-rotating function ωt, [7] were able to successfully use such an experimental setup to observe chiral hydrodynamic phenomena, as evinced by the exotic patterns in the optical images in parts a) and c) of FIG 1.

2. Overview of the Curiosity-Driven Search

The constantly-rotating $B_\theta(t) = \omega t$ is just one of potentially many other field protocols $\{B_\theta(t)\}$ which can induce novel behavior in the magnetic colloidal fluid. In the absence of an underlying theory of chiral hydrodynamics, it is difficult to determine which other $\{B_\theta(t)\}$ can produce previously unrealized hydrodynamical behavior. Randomly testing various $\{B_\theta(t)\}$ in the hopes of stumbling across a new behavior is impractical; the periodic functions in $\{B_\theta(t)\}$ have the same dimensionality as the space of continuous functions on the unit interval–infinity.

To resolve this issue of finding interesting $\{B_\theta(t)\}$, we consider the curiosity-driven search (curiosity). A very abstract technique which has proven useful in fields as disparate as video game studies [5] and developmental psychology [4], the curiosity=driven search can be adapted into an efficient procedure for finding and characterizing the behaviors of non-equilibrium many-body systems like the magnetic colloidal fluid. Indeed, in a recent paper, curiosity was employed to comprehensively map out the allowed behaviors in the theoretical Kuramoto model of coupled oscillators [3].

The basic premise of curiosity is to reformulate a search for dynamical phenomena over a high-dimensional parameter space as a more tractable search over a lower-dimensional latent space of behaviors. Over each iteration it is run, curiosity works from some sample of points in parameter space to construct a model of behavior space which facilitates the sampling of parameter space points for the subsequent iteration of the search. Through running many iterations of curiosity on an initial seed of parameter points in the magnetic colloidal fluid, we can converge onto a precise characterization of the system's dynamical behaviors.

B. Aims of Research

We are interested in applying curiosity to a *simulated* magnetic colloidal fluid subjected to an external magnetic field. In so doing, we hope to

1. Document the important classes of behavior the simulated system is capable of exhibiting

2. Suggest "order parameters" to characterize the behaviors of the system

To achieve our first goal, we plan on running a curiosity-driven search over many iterations on the driven colloidal fluid. With each successive iteration, a new set of points in parameter space is sampled, and behaviors qualitatively distinct from anything observed previously may result from these new parameters. If the search is conducted over sufficiently many iterations, we can exhaust the qualitative range of distinct behaviors the system can exhibit; beyond a certain threshold, the new parameters in the curiosity search only yield dynamical behaviors which resemble those which have been previously observed.

Our second goal adds quantitative heft to the qualitative description achieved in our first goal. As a curiosity search progresses, we construct increasingly accurate latent spaces which host the observed behaviors of the system. The axes of these behavior spaces are aligned with the directions in which the behaviors exhibit the greatest variation, and therefore measure a set of quantities which distill the system's most salient behavioral features. In the limit of many iterations of curiosity, these quantities will converge to "order parameters" for the non-equilibrium magnetic colloidal system [3].

II. METHODS

A. Simulation-work with HOOMD Molecular Dynamics

HOOMD-Blue is a high-performance molecular dynamics software designed to model the evolution of many-particle systems with respect to a specified force-field [1]. Following the lead of [2] and [7]–who employed HOOMD to discover novel physical phenomena in non-equilibrium systems–we used HOOMD's capabilities in our curiosity search.

In our two-dimensional HOOMD simulations of the magnetic colloidal fluid, the simulated particles numbered in the hundreds, and were initially distributed in an energy-minimizing circular lattice configuration. The simulated force $\mathbf{F}$ imparted on a particle p with position $\mathbf{x}$ was determined by the following Langevin equation:

$$\mathbf{F} = -\gamma\dot{\mathbf{x}} + \mathbf{F}_{\text{WCA}} + \mathbf{F}_{\text{iso-dd}} + \mathbf{F}_{\text{rot-vis}} + \mathbf{F}_{\text{B}} \qquad (2)$$

The first term of the above expression for $\mathbf{F}$ is $-\gamma\dot{\mathbf{x}}$, and reflects the losses in p's momentum due to dissipative forces. γ is sufficiently large that p's motion is overdamped. Meanwhile, the second term $\mathbf{F}_{\text{WCA}}$ captures the steric interactions between each non-p particle and p, and is specified by the spherically symmetric Weeks-Chandler-Anderson potential. In terms of the Lenard-Jones potential $U_{\text{LJ}}(r) = 4\epsilon((\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6})$ describing the potential at the location of p due to another particle a distance r away [6], the WCA potential U_{WCA} is truncated such that the potential vanishes at some finite cutoff distance r_{cutoff} instead of $r = \infty$. More precisely, U_{WCA} is defined piecewise as

$$\begin{cases} U_{\text{LJ}}(r) - U_{\text{LJ}}(1.5\sigma) & x < 1.5\sigma \\ 0 & x \geq 1.5\sigma \end{cases}$$

The particles in the simulated fluid are magnetic, and the third term $\mathbf{F}_{\text{iso-dd}}$ accounts for the longitudinal pairwise interaction between p's magnetic dipole and those of other particles. Unlike the steric repulsion described by the $\mathbf{F}_{\text{WCA}}$ term, $\mathbf{F}_{\text{iso-dd}}$ is an interaction which does not vanish at some threshold distance; it instead decays as the reciprocal of the cube of the distance separating the particles.

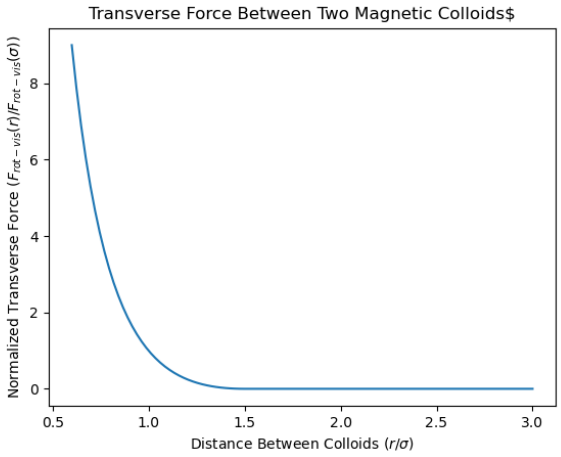

FIG. 2. The magnitude of the transverse force between two colloids with fixed angular velocities, normalized to the magnitude of the force when the colloids are separated by a distance of $r = \sigma$ (for the same σ which appears in the WCA potential).

Contrasting with the longitudinal $\mathbf{F}_{\text{WCA}}$ and $\mathbf{F}_{\text{iso-dd}}$ terms, $\mathbf{F}_{\text{rot-vis}}$ is a *transverse* term which accounts for the rotational, viscous interactions between the particles. Since the colloidal particles have non-vanishing magnetic moments, the external magnetic field $\mathbf{B}(t)$ exerts a torque which imparts a time-varying angular velocity $\overrightarrow{\omega}_i(t)$ to the ith of these particles. For the ith and jth particles, the transverse force $\mathbf{F}^{ij}_{\text{rot-vis}}$ between them admits

the following piecewise expression in terms of their angular velocities $\vec{\omega}_i$ and $\vec{\omega}_j$ (consult FIG 4 for a plot of the magnitude of $\mathbf{F}_{\text{rot-vis}}$):

$$\mathbf{F}^{ij}_{\text{rot-vis}} \propto \begin{cases} (\frac{1}{r} - \frac{1}{1.5\sigma})^2(\vec{\omega}_i + \vec{\omega}_j) \times \hat{\mathbf{r}} & r < r_{\text{cutoff}} \\ 0 & r \geq r_{\text{cutoff}} \end{cases} \quad (3)$$

Finally, $\mathbf{F}_{\text{B}}$ is a term which reflects the effects of thermal noise. This term is dominated by the other terms since the simulations are run at a temperature T such that $k_B T$ is much smaller than the natural energy scale of the system.

Numerically integrating the equation of motion of the simulated particles in (2) from the initial configuration, we can evolve the positions and the velocities of the particles in time. With this setup, we can associate dynamical information to the field protocol $B_\theta(t)$ via the $\mathbf{F}^{\text{rot-vis}}$ term; the external field $B_\theta(t)$ determines the angular velocities of the simulated particles, which in turn determine the force field $\mathbf{F}$ from the expression in (3).

B. Applying Curiosity to the Magnetic Colloidal Setup

1. *Synopsis of The Curiosity Algorithm*

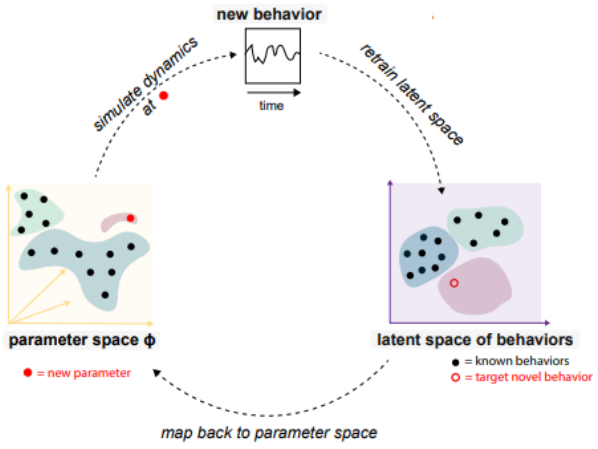

FIG. 3. A schematic of the curiosity-driven search procedure used in [3], showing the reformulation of a high-dimensional search over parameter space in terms of a much easier lower-dimensional search over a constructed latent space

The algorithm which powers the curiosity-driven search is defined by a repeated series of mappings between a parameter space and an evolving space of latent behaviors (Figure 3). In applying curiosity to the simulated magnetic colloidal fluid, we dictated the form and the evolution of these mappings by defining a parameter space, encoding the parameters' dynamical output in a representation, and constructing a low-dimensional latent space from this representation by means of dimensionality-reduction.

2. *The Parameter Space*

In our setup, the field protocols $\{B_\theta(t)\}$ are the tunable parameter for controlling the dynamics of the colloidal fluid. These field protocols are periodic over a time interval T, so admit the following Fourier expansion:

$$B_\theta(t) = a_0 + \sum_{n=1}^{\infty}(a_n \cos(\frac{2\pi n\theta}{T}) + b_n \sin(\frac{2\pi n\theta}{T})) \quad (4)$$

for real coefficients $a_0, a_1, a_2, a_3, b_1, b_2, b_3$.

Owing to its infinite-dimensionality, this parameter space is difficult to work with *in silico*. Instead, for our work the discrete world of simulations, we approximate B_θ by its seven lowest-frequency terms, such that:

$$B_\theta(t) \approx a_0 + \sum_{n=1}^{3}(a_n \cos(\frac{2\pi n\theta}{T}) + b_n \sin(\frac{2\pi n\theta}{T})) \quad (5)$$

In order to sample this space for the curiosity algorithm, we must demand that the seven expansion coefficients are contained in some finite interval. In restricting each of these seven coefficients to the interval $[-1, 1]$, we still preserve a wide diversity of waveforms among the $\{B_\theta(t)\}$ (see FIG 4).

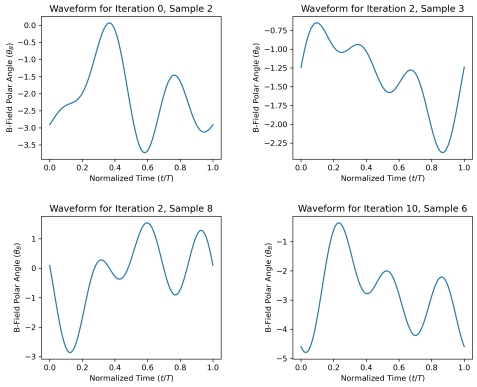

FIG. 4. Examples of $B_\theta(t)$-waveforms of varied shapes sampled during a run of our curiosity algorithm

3. *Producing Representations from Dynamical Data*

After sampling a field protocol $B_\theta(t)$ in our seven-dimensional parameter space $[-1, 1]^7$, we construct the

relevant force field **F** described in (2). Applying HOOMD's molecular dynamics software, this force field yields a trajectory which describes how the colloidal fluid evolves under the specified field protocol (see FIG 5).

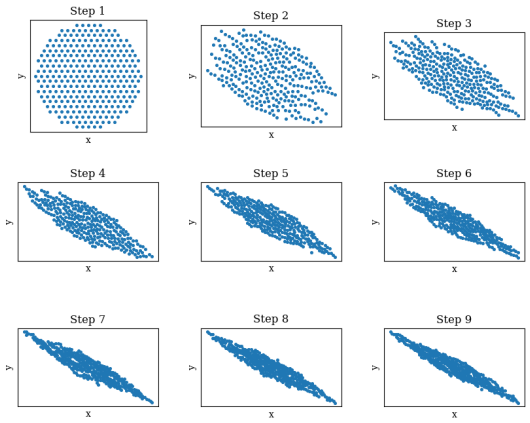

FIG. 5. A time-series depicting the evolution of the shape of an initially circular lattice of colloids for a randomly generated $B_\theta(t)$ in our seven-dimensional parameter space. Successive terms in this time-series are separated by a duration of $5T$, i.e, 5 passes through the $B_\theta(t)$ waveform during which **B** exerts a torque on the colloid's magnetic dipoles.

For our curiosity-driven search, we simulated 301 magnetic colloidal particles initially in a circular configuration (see Step 1 of FIG 5). These simulations had a duration of $50T$. In other words, we simulated the effect of the 50 passes of the driving $B_\theta(t)$ waveform on the fluid.

Our simulations produced a wealth of dynamical data, including the positions and the orientations of the colloids throughout the course of the simulation. However, much of this data was either redundant or irrelevant to the qualitative behaviors we documented. Indeed, the most dramatic of these behavioral features–the evolution in colloid's morphology–could be gleaned from just the final positions of the colloid's particles at simulation time $t = 50T$.

Even when restricting our attention to the final positions of the colloids, our data possessed features which are irrelevant to the fluid's morphology. For instance, the order in which the colloids' are labeled when compiling their final positions does not reflect the colloid's shape, but would impact the list of final positions associated to the colloid. As such, the raw final positions of the colloid need to be further refined to more faithfully capture the colloid's morphology.

To do this, we construct a *representation* of the colloids' final positions, and demand the following three conditions for this representation:

1. The representation is invariant to changes in the labelling of the simulated particles (permutation invariance)

2. The representation is invariant to changes in the position of the origin of a Cartesian coordinate system given to the simulated particles (translational invariance)

3. The representation is invariant to changes in the position of the $\theta = 0$ ray of a polar coordinate system given to the simulated particles (rotational invariance)

The above conditions ensure that the representation captures the physical properties of the simulated system; any representation which somehow depended either on the means in which the system's identical particles were labeled or on the choice of coordinates for the system would simply be unphysical.

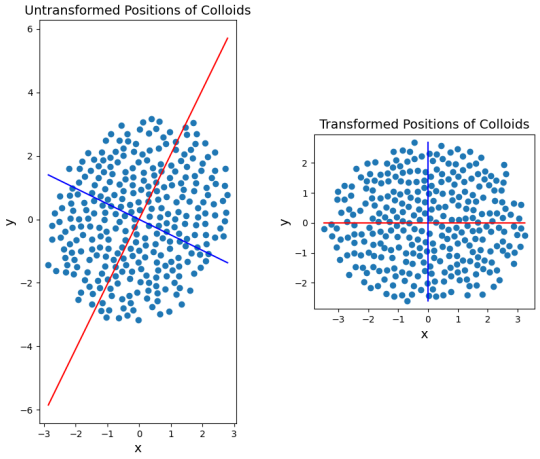

FIG. 6. The final configuration of the fluid's particles for one of the simulated trajectories our curiosity search. The left panel depicts the fluid's particles in untransformed coordinates, whereas the right panel depicts the resulting positions after the coordinate transformation has been performed. In both panels, the red line denotes the *major* axis and the blue line denotes the *minor* axis.

To construct this representation, we first determine the axes along which the particles' positions exhibit the greatest variation and least variation. This is done by taking vectorized lists of the mean-centered x and the y positions of the particles in the colloid, and computing the eigenvectors of the covariance matrix between these two vectors. From the construction of this covariance matrix, the eigenvector with the greater eigenvalue is aligned in the planar direction in which the colloids' positions exhibit the greatest variation, while the eigenvector with the smaller eigenvalue points in the direction of least positional variation. Since the final positions of the colloids approximately constitute a well-defined shape in the plane, we refer to the former eigenvector as lying on

the shape's *major axis*, while the latter lies on the shape's *minor axis*.

Having identified these axes, we then transform the colloids' positions to a coordinate system where the major axis coincides with the x-axis, the minor axis coincides with the y-axis, and the intersection between the major and minor axis at the colloids' center of mass lies at the origin (see FIG 6). This coordinate transformation is an isometry that is a composition of a mean-centering translation with an axis-aligning rotation.

In this new coordinate system, we then considered the distribution of the 301 transformed x coordinates and the distribution of 301 transformed y coordinates of the colloids. We binned the colloids into 40 groups on the basis of the z-scores of their x coordinates and y-coordinates in the respective coordinate distributions.

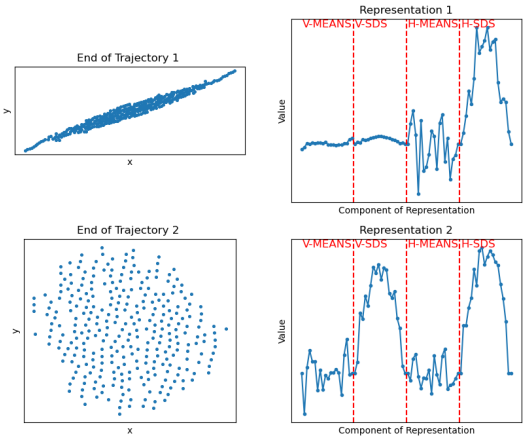

FIG. 7. The final configurations of two trajectories encountered in our curiosity search, and their associated representations. The representations, depicted on the right, are 80-dimensional vectors which capture many features pertaining to the final shapes the colloids assume. For instance, the near rotational symmetry of the second trajectory is reflected in the commensurate sizes of the V-SDS and H-SDS peaks in its representation, while the discrepant size of these peaks in the first representation reflects the absence of rotational symmetry in the first trajectory.

More precisely, we grouped the colloids whose x-positions had a z-score between -2.5 and 2.5 into 20 bins which each spanned 0.25 standard deviations of the x-coordinate distribution. We then extracted two pieces of information pertaining to the distribution of y-coordinates for the colloids each of these 20 x-cordinate bins: (1) their mean and (2) their standard deviation. Intuitively, these 20 means (termed "V-MEANS" in FIG 7) and 20 standard deviations (termed "V-SDS" in FIG 7) measure how far above and how dispersed the colloids generally are with respect at varying points along this axis. Proceeding similarly for the 20 y-coordinate bins, we could gather analogous set of 20 means (termed "H-MEANS" in FIG 7) and 20 standard deviations (termed "H-SDS" in FIG 7) for the minor axis. All together, this leaves a total of 80 data points to be used in a representation.

We note that the above procedure for constructing a representation respects all three symmetries we imposed on the representation. The inputs to the representation are determined from a coordinate system in which the center of mass of the colloids lies at the origin, so the representation is impervious to a spatial translation of the particles. Moreover, this coordinate system was constructed such that the major axis always coincides with the x-axis, so a uniform rotation of the particles would have no impact on the representation. Finally, the representation is permutation invariant since no references for the labels of the particles were made in constructing it.

C. Constructing the Latent Behavior Space

Our curiosity algorithm has a seed of $n = 10$ randomly-generated points in the seven-dimensional parameter space $\{(a_0, a_1, a_2, a_3, b_1, b_2, b_3) \in [-1, 1]^7\}$ as input. As described in Section II.B, these n points determine equations of motion which can be integrated to produce n trajectories from which n representations can be obtained.

However, the eighty-dimensional representation space is much too large to be an appropriate setting for understanding the behaviors of the driven colloidal fluid; such a space contains many large dimensional subspaces which have minimal relevance to the observed behaviors of the fluid. In order to create a space which more precisely reflects the behaviors of the fluid, we need to collapse representation space onto a lower-dimensional space on which the fluid exhibits a wide variety of behavior. To this end, we apply principal component analysis (PCA) to the n points of the initial seed in the fluid's representation space to find seven linearly independent vectors along which the n points display the greatest variation. These seven vectors constitute a basis for our first version of latent space.

In addition to being an important structure in classifying a system's behaviors, the latent space is an important auxiliary to the parameter space; by construction, the latent space is tailored for describing the behaviors of the colloidal system, and therefore can be used to define a sampling protocol for parameter space which favors behavior-rich regions.

To wit, we consider the smallest axial hyperrectangle which bounds the n sampled points in latent space, dilate this hyperrectangle about its center in all seven dimensions by a factor of 1.1, and uniformly sample n latent space points in this dilated hyperectangle (see upper right panel of FIG 8). These n sampled latent points almost certainly will not correspond precisely to points in the parameter space, but will each have a nearest neighbor among the n latent behaviors produced by the initial seed of parameter space points.

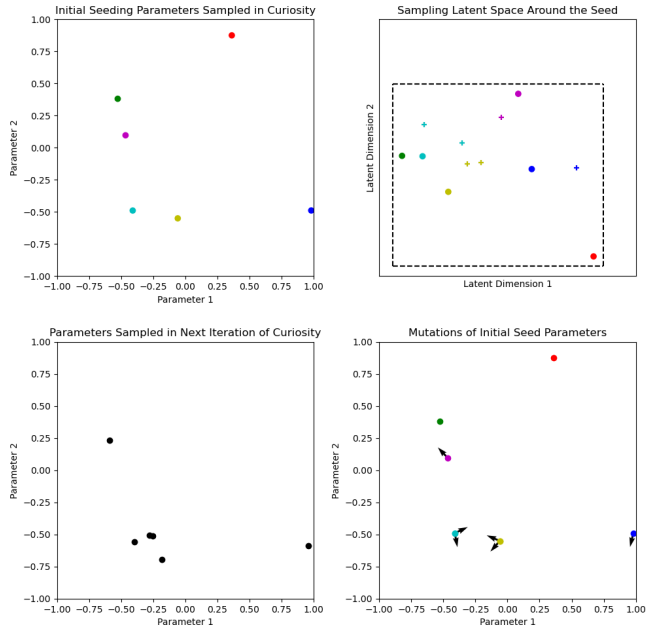

FIG. 8. A series of plots describing how the latent space facilitates sampling in a simpler *two-dimensional* parameter and latent space. An initial seed of $n = 6$ parameter space points are sampled from a parameter space $[-1, 1]^2$ in the top left panel. After performing PCA to the representations associated with these points, the n color-coded parameter points inhabit a latent behavior space depicted in the top right panel. A dilated 2-hyperrectangle, i.e, a rectangle, depicted by the black dashed lines is uniformly sampled n times to identify the nearest neighbors. The perturbations to the parameter space points associated with the nearest neighbors are depicted in the bottom right panel, and the resulting collection of parameter space points to be used in the next iteration of curiosity is depicted in the bottom left panel.

For each of these n nearest neighbor points, we find the associated seed point in parameter space and vary each of the seven components of this seed point by a real number uniformly drawn from the interval $[-0.2, 0.2]$, an interval whose radius constitutes 10% of the length of one side of the parameter space (bottom right panel of FIG 8). If the resulting seed point lies outside of the parameter space hypercube $[-1, 1]^7$, we repeat the procedure of perturbing the seed point until the perturbed point is contained in our parameter space.

The n resulting perturbations to the parameter space points will then serve as the sampled parameter points for the second iteration of curiosity. By looking to latent space to determine which points of parameter space are to be perturbed, we bias the selected parameters towards behavior-rich regions of parameter space, increasing the likelihood of finding novel behavior.

Subsequent iterations of the curiosity-driven search are almost identical to the first, with one slight twist. In the kth iteration of curiosity for $k \geq 2$, only half of the seeding parameter space points in the $k + 1$st iteration are generated by perturbing the seeding points of the kth iteration. The other $\frac{n}{2}$ points in the seed for the $k+1$st iteration come from randomly selecting–and leaving unperturbed–parameter space points from *all* of the k previous iterations of curiosity. In so doing, we continually inject diversity into the sampled parameter space points, substantially reducing the likelihood that the behaviors of these sampled points cluster around just one behavior-dense region of parameter space. We anticipate the formation of multiple clusters of behavior, and the above procedure strikes an appropriate balance between charting the points in parameter space which correspond to any particular behavioral cluster and exploring for other potential clusters.

As the curiosity search progressed, the mapping from representation space to latent space was performed through a PCA of *all* previously collected representations. Since representation space accumulated points with more varied qualitative behaviors by the later iterations of the curiosity search, we anticipate that the latent spaces constructed from PCA during these later iterations will more precisely reflect the system's salient behavioral features.

III. RESULTS

A. Formation of Behavioral Clusters

We ran our curiosity-driven search for the colloidal fluid system over 18 iterations. Since 10 new parameter points were sampled during each iteration, this curiosity-driven search left us with a total of 180 trajectories whose behaviors could be analyzed.

Each iteration of curiosity produced a new latent space spanned by seven principal components of an eighty-dimensional representation space. Examining the coordinates of the latent space points with respect to just the first two of these principal components, we observed the formation of three distinct clusters in latent space (FIG 9).

Although the number of latent space clusters was hard-coded to be three on the basis of our qualitative observations, the assignment of the individual points to these clusters was performed through a non-observationally-dependent procedure. We employed Ward's minimum variance method [8], a well-known greedy algorithm in algomerative clustering, to identify the three clusters. The clusters suggested by Ward's method are apparently robust, since they emerged relatively early on in our curiosity-driven search, and persisted thereafter.

B. Analysis of the Behavioral Clusters

Given the anticipated association between latent space and the behaviors of the driven colloidal fluid, it is natural to wonder if the three clusters identified in the

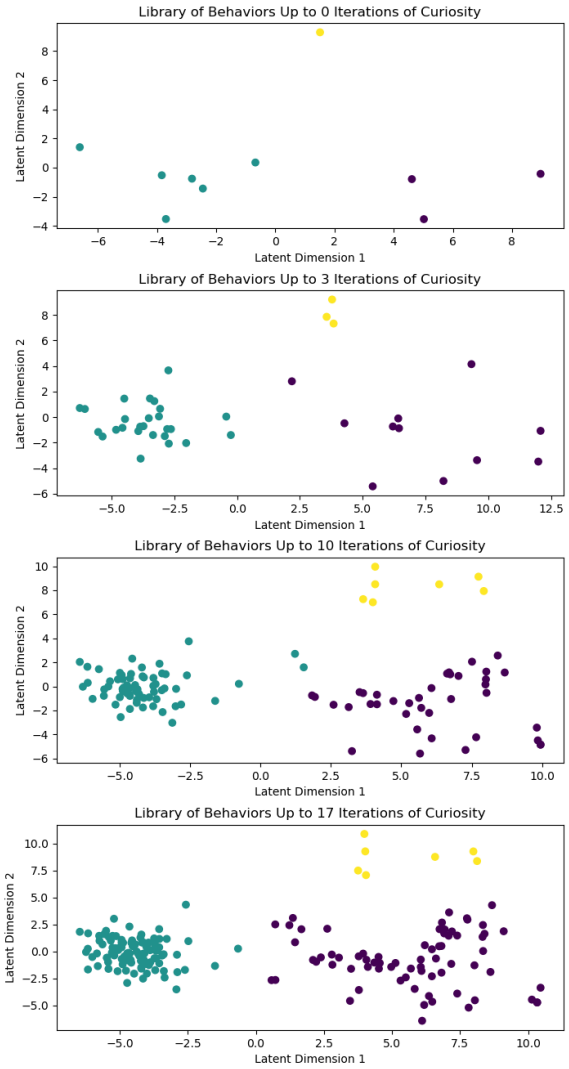

FIG. 9. Projections of latent space onto its two principal components at various points during the curiosity algorithm. As the algorithm proceeds and more latent space points are produced, the emergence of three distinct clusters of points (colored in green, purple, and yellow), is clearly evident

curiosity-driven search correspond to distinct behavioral modes of the system. Indeed, an examination of the fluid trajectories belonging to each of the three clusters revealed characteristic fluid morphologies associated with each cluster.

1. *Morphology 1: "Blobs"*

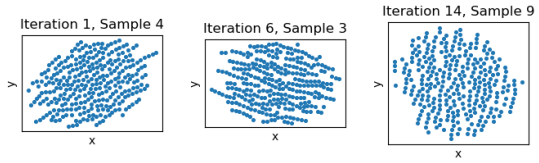

FIG. 10. Configurations encountered in our curiosity search associated with the green cluster of FIG 9

The trajectories for the points in the green cluster of FIG 9 all terminate in a roughly circular configuration. Three examples of these configurations are presented in FIG 10. Many of these blobs do have discernible major axes to which strings of particles are aligned, but these axes are nowhere near as pronounced as what will be encountered in the string morphology. It appears that the green cluster reflects the qualitative description of a trajectory as lacking "pointedness".

2. *Morphology 2: "Strings"*

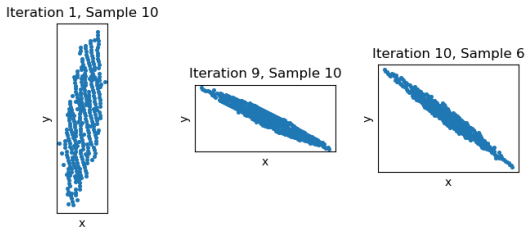

FIG. 11. Configurations encountered in our curiosity search associated with the purple cluster of FIG 9

Now, the final configurations for the points in the purple cluster of FIG 9 can be united under the qualitative description of being "elongated". FIG 11 displays three examples of such stretched configurations of varying orientations in the plane.

3. *Morphology 3: "S-curves"*

Finally, the yellow cluster of FIG 9 corresponds to the final configurations of trajectories which qualitatively appear to be "S"-shaped. Indeed, FIG 12 depicts a few example configurations which resemble the letter "S" with varying aspect ratios and orientations. Much like the strings, all of the "S"-curves configurations observed were also elongated.

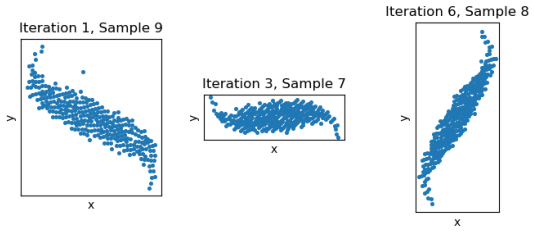

FIG. 12. Configurations encountered in our curiosity search associated with the yellow cluster of FIG 9

C. Order Parameters for the Driven Magnetic Colloidal Fluid

1. *Emergence of Order Parameters*

Over the course of our curiosity search, we arrived at models of latent space which increasingly came to reflect the behaviors of the driven colloidal fluid. When taken as an abstract mathematical structure, the latent space facilitates comparisons *between* trajectories on the grounds of their behavioral similarity. However, in order to understand these behaviors in *absolute*, physical terms, we must understand how the axes of the latent space were constructed from the representation space. To do this, we consider the PCA projections which define the mappings from representation space to latent space for each iteration of curiosity.

In our curiosity-driven search, the first and the second principal components of the PCA projection converged to particular representation space vectors (the differences between the components of these two vectors between the bottom two plots of FIG 13 is barely discernible). Consequently, the principal two axes of the latent stabilized over the course of the curiosity-driven search, just as would be expected if these latent space axes came to measure precise physical features about the colloidal system.

On these grounds, we had good reason to believe that the two principal components converged towards order parameters for our system of interest, prompting us to examine the representation space components of the principal components to determine these order parameters' physical significance.

2. *First Dimension of Latent Space*

The first dimension of latent space lies in a direction determined by the first principal component of the PCA which acts on representation space. The final version of this first principal component has representation space components whose values are remarkably stable in each of the four 20-dimensional subspaces of representation space, i.e, the V-MEANS, V-SDS, H-MEANS, and H-SDS subspaces; namely, these components hover around

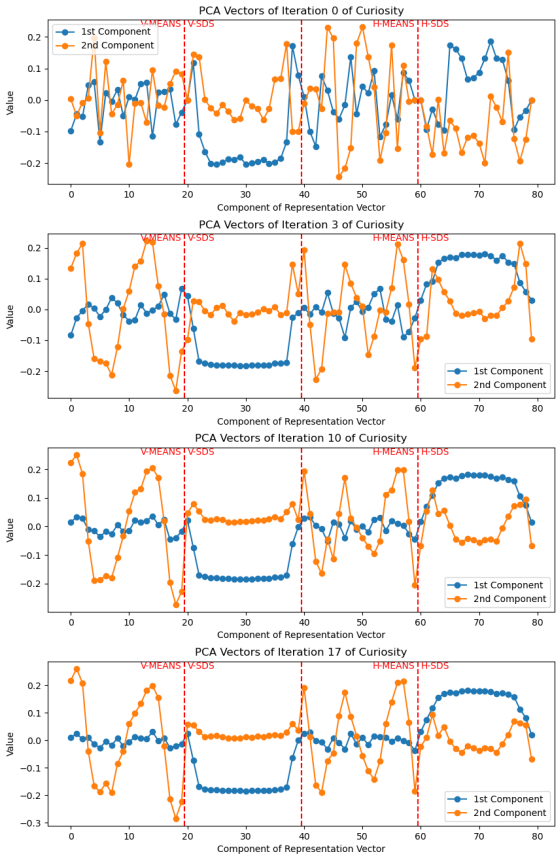

FIG. 13. Plots displaying the entries of the first (blue) and second principal (orange) components of the PCA performed on representation space during selected iterations of our curiosity search.

zero for both the V-MEANS and H-MEANS subspaces, and generally hover around a negative value and a positive value for the V-SDS and the H-SDS subspaces, respectively (see the blue line in the bottom panel of FIG 13).

As mentioned in Section B.III, the entries of V-MEANS and H-MEANS reflect the deviations of a colloidal configuration away from its major and minor axes, respectively. Thus, the fact that the first principal component assigns relatively little weight to the representation space vector in the V-MEANS and H-MEANS subspaces implies that it does not capture information pertaining to the asymmetry of the fluid about its major or minor axes. Instead, by negatively weighting V-SDS componnts and positively weighting the H-SDS components, this first principal component measures the stretching of the fluid along the direction of its major

axis; a small dispersion of the fluid away from its major axis and a large dispersion of the fluid away from its orthogonal minor axis both correspond to a stretched fluid shape. **Therefore, the first principal component defines an order parameter which quantifies the elongation of the system's end shape.**

Indeed, such an interpretation of the first principal component is borne out by the relative elongations of the characteristic morphologies associated to the three clusters of Figure 9; the clusters which lie further along the direction specified by the first principal component on latent dimension 1–the purple and green clusters–have string and S-curve morphologies which are more elongated than the green cluster's blob morphology.

3. *Second Dimension of Latent Space*

Unlike the final first principal component of our curiosity-driven search, the final second prinicpal component hovers around zero for most of the components in the V-SDS and the H-SDS subspaces, only slightly peaking to positive values near the edge of these subspaces (see bottom panel of FIG 13). This suggests that the the second component is largely insensitive to stretching and compressing transformations of the colloidal fluid about its major and minor axes; only stretching effects among the peripheral particles of the fluid can be captured by this second component.

Rather than reflecting the elongation of the bulk colloidal fluid in the vein of the first principal component, the second principal component appears to be picking up on a very specific asymmetry of the colloidal fluid about its major and minor axes. As seen in the bottom panel of FIG 13, the entries of the second component exhibit well-defined oscillatory patterns in both the V-SDS and the H-SDS regions of representation space, suggesting that this second component is sensitive to colloidal fluids which meander around their major and minor axes in some specific fashion.

To be precise, a fluid shape which exhibited a sinusoidal pattern of deviations above and below its major axis for 1.5 periods across the length of its major axis would have a high second principal component due to the shape of this component in the V-MEANS region. Such a fluid would resemble an "S", suggesting that the second principal component corresponds to an order parameter which measures the "S"-like character of the fluid shape.

Moreover, a fluid shape which exhibited a sinusoidal pattern of deviations above and below its minor axis for 2.5 periods across the smaller length of its minor axis would also have a high second principal component due to the shape of this component in the H-MEANS region. Such a fluid would in all likelihood have a more nuanced shape than the letter "S", but would evince many morphological characteristics which resemble the letter "S". For instance, the tails of such a shape would be on opposite ends of the minor axis, and the fluid would have substantial portions of its bulk which lie on both sides of its minor axis.

Based on the oscillatory shape of the second principal component in the V-MEANS and H-MEANS region and the relative flatness of this component in the V-SDS and H-SDS regions, **the second principal component defines an order parameter which reflects the qualitative degree to which the shape reflects an "S".** This is corroborated by a comparison of the morphologies of the yellow and purple clusters of FIG 9. The yellow cluster is situated further up on latent dimension 2 than the purple cluster, and features S-strings with more S-like morphologies–but roughly equal elongations–to the strings of the purple cluster.

4. *Higher Principal Components of Latent Space*

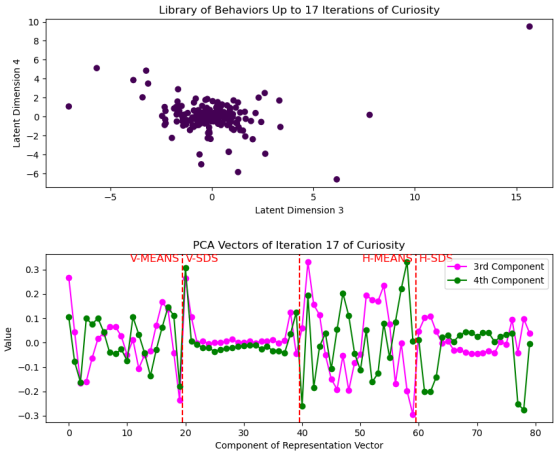

FIG. 14. The top panel depicts the distribution of the sampled latent space points along its third and fourth dimensions, i.e, the third and fourth principal components of the PCA projection from representation space, for the final iteration of our curiosity-driven search. The bottom panel plots the entries of the third and fourth principal components of the final PCA considered during the final iteration.

Contrasting to our first two latent space dimensions, which we could associate to order parameters, the higher latent space dimensions did not appear to have much relevance to the physical behaviors of the colloidal fluid. The projections of the latent space points down into the subspace spanned by these five higher dimensions was notably devoid of clusters which could correspond to distinct behaviors of the colloidal fluid (see the top panel of FIG 14 for an example of this lack of clustering).

An analysis of the representation space entries of the higher principal components lent further credence to the notion that these principal components were physically meaningless. For instance, in the bottom panel of FIG 14,

we see that the entries for the third and fourth principal components wildly varied between successive entries in the V-MEANS and H-MEANS regions, suggesting that these principal components cannot be extracting any significant information from these two regions. Even though these components exhibited a relative stability around 0 in the V-SDS and the H-SDS regions, this more reflects the insensitivity of these components towards elongation behavior in the fluid than any *positive* ability to capture any fluid behavior; by construction, these components are orthogonal to the first principal component, and must be patterned to be oblivious to any physical data the first component measures.

IV. DISCUSSION

Without any *a priori* knowledge about the range of behaviors in the driven magnetic colloidal fluid, we were able to use the curiosity algorithm to document three distinct classes of behaviors: the blob, string, and s-curve morphologies. Moreover, we were able to suggest two order parameters, an "elongation" parameter and an "S-like" parameter, to classify the non-equilibrium behaviors we observed. Our validation of the capabilities of curiosity for simulated colloidal fluid extends [3]'s conclusion about the applicability of the search to theoretical models.

One of the greatest shortcomings of our curiosity-driven search is its inability to describe the *mechanisms* through which the found behaviors are produced. Curiosity simply exhibits the behaviors which can arise in a system of interest, and further analysis is needed to understand *how* specific points in parameter space facilitate the emergence of specific types of behaviors. As such, one of our next goals of the project is to map out which regions of parameter space correspond to each of the three morphologies we observed. By examining these parameter space regions, we should obtain a better sense of the physical processes which drive the evolution of particular shapes in the colloidal fluid. We might also get a better understanding of why certain shapes do *not* appear in the colloidal fluid, such as the lack of a morphology with an intermediate elongation between the blobs and the strings.

Of all of the mappings between parameter space, representation space, and latent space which constitute our curiosity algorithm, the one which admits the greatest possibility for improvement is the projection from representation space to latent space. Currently, we are using a relatively crude PCA technique to project down onto latent space, but [3] has employed a neural network to perform this projection to great success. With a more sophisticated means of projecting onto latent space, we may be able to tease out more order parameters from the axes of this latent space; using PCA, we were able to identify two latent space axes with order parameters, but it is unlikely that these order parameters are the only two which govern the dynamics of the driven colloidal fluid.

With our notions of a parameter space among the periodic field protocols $\{B_\theta(t)\}$ and behaviors in the final shape of the simulated colloidal fluid, we can be fairly confident that curiosity has documented all of the major classes of distinct behaviors which arise in the simulated parameter space; from just a few iterations into our 18-iteration search, no new classes of behavior emerged, suggesting that the major behavioral classes were encountered very early on in the search. However, in the absence of experimental results, these simulated behaviors are nothing more than speculative hypothesis for what could occur in the physical world.

Fortunately, since we chose a simulated system which admits an experimental realization [7], we hope to soon connect the results of our curiosity search to experimental findings. Through experimentation, we hope to resolve the following two questions

1. Can the simulated behaviors we found through curiosity, i.e, the blob, string, and S-shape morphologies, be documented in an experimental setting?

2. Do these simulated behaviors comprehensively represent the range of behaviors which is encountered in experiment?

If we can answer the above two questions in the affirmative, it would be a major validation of curiosity-driven search techniques. While curiosity has been established as a successful procedure for characterizing the behaviors of theoretical physical models, the literature has little to say about the procedure's success in probing the behaviors of real-world physical systems.

[1] J.A Andersom, M. E. Irrgang, and S. C. Glotzer. Hoomdblue: A python package for high-performance molecular dynamics and hard particle monte carlo simulations. *Computational Materials Science 173*, 173, 2020.

[2] Ephraim S. Bililign, Florencio Balboa Usabiaga, Ganan Yehuda A., Alexis Poncet, Vishal Soni, Sofia Magkiriadou, Michael J. Shelley, Denis Bartolo, and William T. M. Irvine. Motile dislocations knead odd crystals into whorls. *Nature Physics*, 18(2):212–218, 2022.

[3] Martin J. Falk, Finnegan D. Roach, William Gilpin, and Arvind Murugan. Curiosity-driven search for novel non-equilibrium behaviors, 2023.

[4] Pierre-Yves Oudeyer, Frédéric Kaplan, Alexei A. Efros, and Verena V. Hafner. Intrinsic motivation systems for autonomous mental development. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, 2007.

[5] Deepak Pathak, Pulkit Agrawal, Alexei A. Efros, and Trevor Darrell. Curiosity-driven exploration by self-

supervised prediction. *arXiv preprint arXiv:1705.05363*, 2017.

[6] Peter Schwerdtfeger and David J. Wales. 100 years of the lennard-jones potential. *Journal of Chemical Theory and Computation*, 0(0):null, 0. PMID: 38669689.

[7] Vishal Soni, Ephraim S. Bililign, Sofia Magkiriadou, Stefano Sacanna, Denis Bartolo, Michael J. Shelley, and William T. M. Irvine. The odd free surface flows of a colloidal chiral fluid. *Nature Physics*, 15(2):1188–1194, 2019.

[8] Jr. Ward, Joe H. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58:236–244, 1963.