Data-driven Discovery of Equations of Motion for a Single Inertial Particle in a Fluid Flow

Darin D. Momayezi and Michael F. Schatz
Physics Department, Georgia Institute of Technology.
(Dated: August 31, 2023)

Understanding single particle-fluid interactions is important to many natural and industrial processes, but exact theoretical descriptions are rare due to the difficulties of properly describing the key mechanisms and solving non-linear equations. Theoretical descriptions are based on approximations that are only justified in certain regimes. For geometries and regimes that are not amenable to certain approximations, it is unclear how to proceed. In this study, we present a data-driven approach to discover the equation of motion of a thin circular disk in a quasi-two-dimensional flow.

I. INTRODUCTION

The motion of particles in flows is an integral problem to many natural and industrial processes. Papermaking, pharmaceutics, and the dispersion of microplastics in the environment, to name a few, rely on knowledge of particle motion in flows. For all but the simplest geometries, such as a sphere [1] and some limiting cases [2] [3] [4], however, fully consistent equations of motion of solid bodies in flows are hard to derive. Today, the standard equation of motion for a rigid sphere in a nonuniform fluid is the Maxey-Riley equation

$$\begin{aligned} m_p\frac{d\boldsymbol{v}_p}{dt} &= (m_p-\rho_f V_p)\boldsymbol{g}-\rho_f V_p\frac{D\boldsymbol{v}_f}{Dt} \\ &-\beta\rho_f V_p\frac{d}{dt}\left(\boldsymbol{v}_p-\boldsymbol{v}_f-\frac{1}{10}R^2\nabla^2\boldsymbol{v}_f\right) \\ &-6\pi\mu R\left(\boldsymbol{v}_p-\boldsymbol{v}_f-\frac{1}{6}R^2\nabla^2\boldsymbol{v}_f\right) \\ &-6\pi\mu R^2\int_0^t\frac{\frac{d}{d\tau}(\boldsymbol{v}_p(\tau)-\boldsymbol{v}_f(\tau)-\frac{1}{6}R^2\nabla^2\boldsymbol{v}_f)}{\sqrt{\pi\nu(t-\tau)}}d\tau, \end{aligned} \quad (1)$$

where p and f are the indices for the particle and fluid, respectively, V_p and R are the volume and radius of the particle, ρ_f is the density of the fluid, μ is the viscosity, $\boldsymbol{v}$ is the velocity, and β is an added mass coefficient that depends on the particle geometry ($\beta = 1/2$ for a sphere). The Maxey-Riley equation (1) describes the motion of a sphere moving through a viscous (i.e., the Reynolds number, $Re = v_f L\rho_f/\mu$, is less than 1, where L is the characteristic flow length scale) non-uniform fluid that has negligible effect on the fluid motion far from it; from left to right in equation (1), it accounts for the force balance, the stress gradient of the undisturbed flow, the added mass, Stokes' drag and the unsteady Basset drag.

The Maxey-Riley equation, however, is an approximate equation of motion due to the mathematical difficulty of retaining higher order and non-linear terms in the fluid velocity. Stokes originally solved the problem of the drag force exerted on a spherical particle by assuming creeping flow conditions [1] and avoiding the higher order and nonlinear terms in the Navier-Stokes equation. Stokes' analysis fails, however, in two-dimensional flows and results in Stokes' paradox where there is no stable solution for an unbounded, creeping flow over a circular cylinder. While the Maxey-Riley equation significantly expands on Stokes' original analysis [5], there is no two-dimensional form of the equation that describes the motion of a circular disk in a two-dimensional flow. Additionally, when one considers effects such as anisotropy or the complexities of a quasi-two-dimensional flow realized in an experimental setting, it is unclear how to proceed with the study of particle-fluid interactions. While the idealized two-dimensional setting raises theoretical difficulties, the experimentally realistic quasi-two-dimensional setting also introduces complexities that are difficult to resolve with numerical or pen and paper methods. To get around these challenges, a data-driven approach to discovering governing equations is needed.

In this paper, we propose applying a systematic method called "model discovery" for discovering the equation of motion of a circular disk in a quasi-two-dimensional flow involving a computational procedure based on sparse linear regression. This two-dimensional system is the historical source of Stokes' paradox and reveals the limited validity of Stokes' approximations. The paradox was resolved in approximation by Oseen who took into account the inertial term in the Navier-Stokes equation omitted by Stokes. A detailed discussion of the paradox and its resolution, as well as a detailed description of the Maxey-Riley equation, are given in the Background (Section III). Oseen's solution does not solve the entire Navier-Stokes equation nor does it account for forces known to be generally important such as those in the Maxey-Riley equation. Model discovery, therefore, responds to a need for data-driven approaches to aid in the theoretical study of solid particle-fluid interactions. The model discovery algorithm is called Sparse Physics-Informed Discovery of Empirical Relations (SPIDER) [6], which is initialized with a library of candidate terms and uses sparse regression to arrive at the best continuum description for the given data. An in depth discussion of SPIDER and its application to active nematics is given by Golden [6]. SPIDER combines relevant domain knowledge (e.g. symmetries) and experimental observables (e.g. position and velocity) to compute interpretable relations in the form of partial differential equations. A discussion

of the library of candidate terms, which are manually inputted into the SPIDER algorithm, is also given in the Background (Section III). The SPIDER algorithm returns a set of partial differential equations that is compared to the Maxey-Riley equation of motion. While the Maxey-Riley equation describes the motion of a spherical particle, we posit, due to arguments given by Lamb [7] and Talei [5], that the same forces of different forms should govern the motion of a circular disk in a two-dimensional flow. Therefore, by using the Maxey-Riley equation as a theoretical baseline, we propose the use of SPIDER, a data-driven model discovery algorithm, to discover empirical relations and derive an equation of motion for a circular disk moving in a quasi-two-dimensional flow.

Having introduced the context and importance of the two-dimensional system, section II provides a theoretical baseline for the motion of spherical particles in viscous, non-uniform flows. Section III delves into the experimental procedure and specifications and section IV discusses the basis for model discovery using sparse regression. The results obtained with model discovery are compared to the experimental data and Maxey-Riley equation (1) in section V and we conclude with section VI.

II. BACKGROUND

The criteria for an equation that describes a particle's trajectory in turbulent flow are that it accounts for all relevant forces and satisfies the Navier-Stokes Equations (NSE). The first attempt to do this was carried out by Stokes who derived the drag force exerted on rigid spheres by an axisymmetric steady flow, requiring the assumptions of Stokes' flow. This assumption is only valid in the regime where the Reynolds number tends to zero ($Re \to 0$), indicating that viscous forces dominate over inertial forces. Therefore, the particles must effectively be non-inertial, which makes Stokes' Law a one-way coupling method where particles have mass and experience a drag force, but the fluid is undisturbed by its presence. These considerations justify a linearization of the Navier-Stokes equations by eliminating the advective terms, giving the Stokes' equations

$$-\nabla p + \mu \nabla^2 \boldsymbol{v} + \rho \boldsymbol{g} = 0 \tag{2}$$

$$\nabla \cdot \boldsymbol{v} = 0, \tag{3}$$

where the advective term, $\boldsymbol{v} \cdot \nabla \boldsymbol{v}$, and the time derivative, $\rho \frac{\partial \boldsymbol{v}}{\partial t}$, have been dropped from (2). Solving for the flow in spherical coordinates, and then the pressure and force gives Stokes' drag

$$F_D = 6\pi \mu R \nu, \tag{4}$$

where μ is the dynamic viscosity, R is the radius of the sphere and ν is the flow velocity relative to the sphere. The drag force is corrected by additional Faxén terms, the terms with $\frac{1}{6}R^2\nabla^2 \boldsymbol{v}_f$ in equation (1), that account for the velocity curvature of the flow around the sphere in non-uniform flow. Stokes' method, however, leads to Stokes' paradox where there is no stable solution for an unbounded flow around a circular cylinder in two dimensions (see Appendix B). Oseen corrected Stokes' analysis by making a linear approximation of the advective term, which adds correction terms to Stokes' drag. The Maxey-Riley equation, however, does not consider Oseen's correction.

Beyond the drag force, Stokes' Law does not take into account forces that are known to be generally important. One such force is known as the "added mass" term, which considers the fluid displaced by the particle's movement. As the particle moves through an incompressible fluid, it displaces a certain volume of fluid, resulting in an acceleration of the surrounding fluid. By Newton's second law, this acceleration is associated with a mass and a force opposing the particle's motion. The magnitude of this opposing force depends on the amount of displaced fluid, which has the effect of an "added mass" on the particle. Newton's second law shows that the opposing force is proportional to the mass and acceleration of the surrounding fluid

$$\beta \rho_f V_p \frac{d\boldsymbol{v}_f}{dt}, \tag{5}$$

where the density of the fluid ρ_f multiplied by the volume of the particle V_p gives the amount of displaced fluid, the fluid's acceleration is given by the derivative of its velocity $\boldsymbol{v}_f$, and β is a geometric factor that depends on the particle's shape. The added mass for a sphere is $\frac{m_f}{2}\frac{d\boldsymbol{v}_f}{dt}$, but we have generalized this by introducing the geometric factor β, which is certainly different for a circular disk. The added mass term for a sphere is recovered by the substitution of $\beta = 1/2$ since $\rho_f V_p = m_f$, where m_f is the mass of the displaced fluid.

In addition to Stokes' steady drag, Basset and Boussinesq considered unsteady drag in what is called the "history term"

$$\int_0^t \frac{\frac{d}{d\tau}(\boldsymbol{v}_p(\tau) - \boldsymbol{v}_f(\tau) - \frac{1}{6}R^2\nabla^2 \boldsymbol{v}_f)}{\sqrt{\pi\nu(t-\tau)}} d\tau. \tag{6}$$

The history term accounts for a viscous effect that causes temporal lag in the boundary layer formed around the sphere due to its changing relative velocity [8]; it can be found by solving for the local shear stress of an impulsive flow over a sphere at low Reynolds number [9]. Basset noticed that Stokes did not include the linear term $\frac{\partial \psi}{\partial t}$, where ψ is the Stokes' stream function (recall that we eliminated $\frac{\partial \boldsymbol{v}}{\partial t}$ from (2)), in his analysis and suggested that it be replaced by $U_0 \frac{\partial \psi}{\partial z}$. Basset

[10] then calculated the drag force on the surface of the sphere as

$$6\pi\mu R U_0 \left(1 + \frac{R}{\sqrt{\pi\mu t}}\right). \quad (7)$$

Finally, we present an adapted form of the Maxey-Riley equation to compare to the data-driven model

$$m_p \frac{d\boldsymbol{v}_p}{dt} = (m_p - \rho_f V_p)\boldsymbol{g} - \beta\rho_f V_p \frac{d}{dt}(\boldsymbol{v}_p - \boldsymbol{v}_f)$$
$$-6\pi\mu R\left(\boldsymbol{v}_p - \boldsymbol{v}_f + R\int_0^t \frac{\frac{d}{d\tau}(\boldsymbol{v}_p(\tau) - \boldsymbol{v}_f(\tau))}{\sqrt{\pi\nu(t-\tau)}} d\tau\right). \quad (8)$$

While the Maxey-Riley equation does not describe the motion of a disk in a two-dimensional flow, it summarizes the forces that we expect to be relevant to its motion and the equation of motion for a disk should be similar in form to the Maxey-Riley equation for a sphere.

It is important to note that the Maxey-Riley equation neglects higher order terms in the flow velocity, therefore, it is not a particular solution to the full Navier-Stokes equation. The goal of model discovery is to retain non-linear terms and higher powers of the flow velocity, which fully solve the Navier-Stokes equations. The particular solution found by SPIDER will be used to find the pressure field and forces acting on the disk. A discussion of how the SPIDER library is built with these considerations is given in Section IV.

III. SPARSE PHYSICS-INFORMED DISCOVERY OF EMPIRICAL RELATIONS (SPIDER)

SPIDER is a model discovery algorithm [6] [11] [12] [13] that takes in data pertaining to a set of physical observables and performs sparse regression on a library of candidate terms to arrive at the best continuum description of the data. SPIDER outputs a set of partial differential equations, which makes it an interpretable learning method, contrary to machine learning methods which often lack physical interpretability. The library $\mathcal{L}$ of candidate terms is created by the user who must identify the physical quantities of interest that may play a role in the dynamics of the system. For that reason, SPIDER should not be used as an all encompassing method of discovering physical models, but in conjunction with a theoretical baseline, it can be used to push the boundaries of current models to encorporate greater complexities and consider more systems. Since SPIDER is initialized with a set of tensors that represent physical quantities, such as the flux over boundaries or velocity gradients, it remains a very general way to study the physics of new systems.

In our study of the motion of a circular disk in a quasi-two-dimensional nonuniform flow, there are several complications due to the mathematical difficulties of accounting for higher-order and nonlinear terms and the complexities of a quasi-two-dimensional flow. SPIDER allows us to pursue an equation of motion for the circular disk despite these difficulties by simply recording the physical observables of interest, such as the position of the disk and the velocity field of the flow, and enumerating the terms that should play a role in the dynamics of the disk, such as the squares of the fluid velocity and the rotation of the disk. We have constructed a library of candidate terms by considering the terms that appear in the Maxey-Riley equation, which summarizes the effects explored theoretically to date, and those that ought to play a role in the dynamics of the disk. The library generation process is detailed in the Results section as well as the procedure for refining the model and deriving and equation of motion.

Once a library $\mathcal{L}$ of candidate terms is constructed, SPIDER constructs a feature matrix A based on the weak formulation of differential equations in $\mathcal{L}$. The elements of A correspond to numerical integrals of the library terms over space-time domains Ω_i instead of their values at individual points. This method for using the weak formulation of differential equations (library terms) is more robust to noise. Evaluating the library terms over space-time intervals Ω_i instead of at individual points means that the data and model are being averaged. The larger the space-time interval the more averaging is being done to fit the model to the data. Therefore, a natural time scale should be identified to accurately model the dynamics of the system. In our case, a natural time scale is the time it takes the particle to complete a full rotation, which is about two hundred time steps or about 3.33 seconds.

IV. EXPERIMENT

In this section we discuss the experimental setup of the particles in two-dimensional fluid flow and the data collection and preparation procedure for model discovery. The first section handles the procedure for the particles and the second discusses the flow.

A. Particles

The particles consist of 3D-printed intersecting rods, forming both jacks (3D particles) and crosses (2D particles). Marcus [14] demonstrated, using resistive force theory and arguments by Bretherton [15], that jacks behave equivalently to spheres and crosses behave equivalently to disks. Therefore, we opted to 3D-print jacks and crosses to facilitate the imaging and tracking process. to achieve neutral buoyancy, we added glycerin to the fluid, matching the density of the particles. The jack arms are 1cm in length, so the jacks are 2cm across, and have a density of 1.25g/cm^3.

Since the fluid beneath the particles is dark, the bodies of the particles are colored black, while the tips are

left white and visible for imaging.

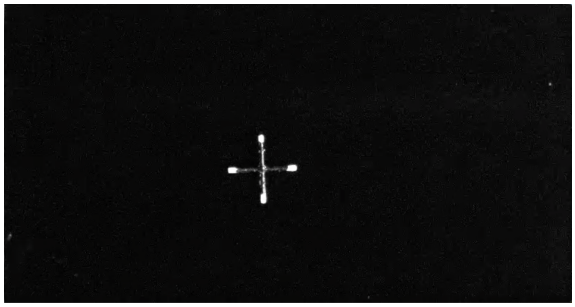

FIG. 1. 2D Cross

B. Flow Setup

In this experiment, we use the flow setup previously reported by Suri et al. in their study of two-layer Kolmogorov-like flow, which we now describe [16]. While no flow is truly two-dimensional, we have implemented a quasi-two-dimensional (Q2D) flow in which the velocity profile along one of three orthogonal directions is greatly diminished. This is achieved by a two-immiscible-layer system of a dielectric *lubricating* layer on the bottom and an electrolyte solution on top in a box. The bottom dielectric layer protects the top layer from the no slip boundary condition at the bottom of the box, which makes the flow approximately two-dimensional. The dielectric fluid is perfluorooctane of viscosity $\mu_d = 1.30\text{mPa} \cdot \text{s}$ and density 1769kg/m^3 and the conducting fluid is a low-viscosity electrolyte consisting of a 0.3M solution of $CuSO_4$ with viscosity $\mu_c = 1.12\text{mPa} \cdot \text{s}$ and density $\rho_c = 1045\text{kg/m}^3$.

A spatially periodic magnetic field is created by an array of 14 NdFeB magnets (Grade N42), each 15.2 cm long, 1.27 cm wide, and 0.32 cm thick, at the bottom of the box. The magnets are placed side-by-side along their width and are magnetized along the vertical direction with a surface field strength of 0.3 T. Experimental measurements using a F. W. Bell Model 6010 Gauss- meter show that the z-component of the magnetic field, B_z, is sinusoidal in the y-direction with a period $T = 2.45\text{cm}$ beyond a height of $z = 0.25\text{cm}$.

V. RESULTS AND DISCUSSION

In this section, we detail the data collection process and the procedure for using SPIDER to derive an equation of motion.

A. Particle Tracking and Fluid Measurement

The particle's center of mass (COM) motion and orientation are the key pieces of information for characterizing its motion. Since the particles are rigid bodies, we only need to track one of their four tips. However, for the purpose of model discovery, we have tracked all four tips. The COM is defined in the usual way

$$\vec{r}_{\text{COM}} = \sum_{i=0}^{N} \frac{\vec{r}_i}{N}, \tag{9}$$

where $\vec{r}_i$ is the coordinate vector of the i-th tip and $N = 4$ for the number of tips. The orientation vector is then defined as the vector extending from the COM to any one of the tips.

The particle and fluid are recorded together with an overhead camera recording at 60 fps to capture the interactions between the particle and fluid. The recording is loaded into TrackPy [17], a particle tracking library in Python, as a sequence of images. These images are intensity filtered to almost entirely remove the flow from the background. Although some dense regions of tracer particles may survive the intensity filtering, they are typically located at the boundaries and are thus spatially filtered. This approach allows us to isolate the particle motion in the images and track the particle tips.

The TrackPy algorithm works by identifying features in each image based on intensity and using a tracking algorithm to link features between images. While TrackPy is originally designed for soft matter systems where particles can temporarily disappear and reappear, the rigid-body nature of the particles precludes the possibility of "features" (particle tips) disappearing between frames. Therefore, we filter out all trajectories that do not persist throughout the entire image sequence, eliminating extraneous trajectories and streamlining the tracking procedure. A simple example of almost purely rotational motion with slight gyration is presented below in figure 2.

Each color represents a separate trajectory and we count four, one for each of the particle's tips. In this example, we expect the coordinates of the orientation vector to lie on a circle and very little COM motion. This is exactly what we observe in figures 3 and 4, respectively, noticing that the length of the axes on figure 4 is very small so the COM displacement is indeed small.

This procedure fully characterizes the motion of the particles with the COM motion and the orientation vector.

For visualization of the flow field, we use white glass tracer particles with a diameter of $70\mu m$ and an overhead camera at 60 fps for imaging. The flow is then measured with Particle Image Velocimetry (PIV) using MatLAB's PIVLab application [18] shown in figure 5. Once the flow recording is analyzed, we calculate the u and v-components of the velocity field, velocity gradient, vorticity field, and velocity divergence with PIVLab. These measurements and calculations are performed over the entire spatial domain. However, since we only anticipate local effects, we are specifically interested in their values at the positions of the particles, so

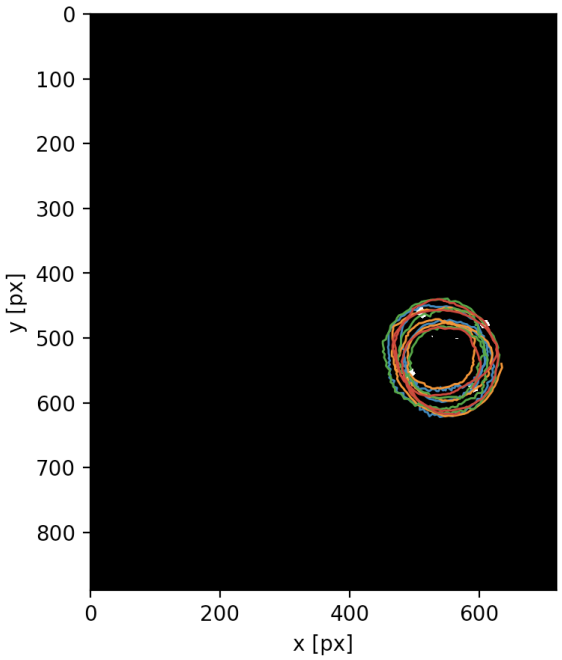

FIG. 2. Particle Trajectory

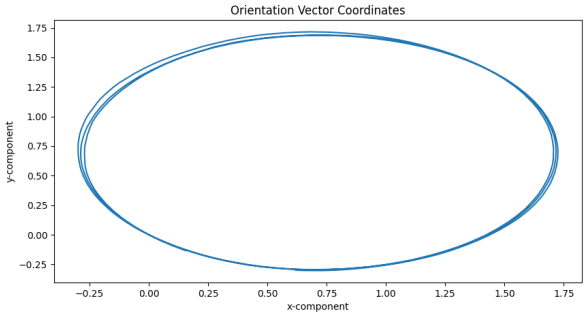

FIG. 3. Orientation Vector Coordinates

we interpolate them onto the positions of the particle's tips using MatLab's interp2 function.

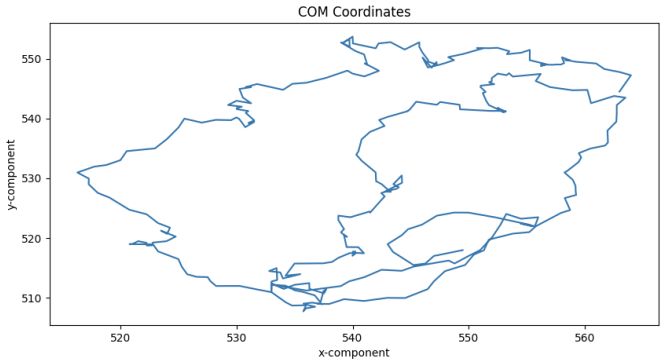

FIG. 4. COM Coordinates

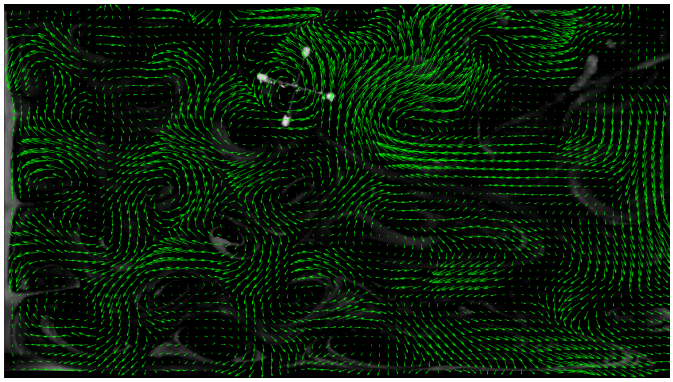

FIG. 5. Velocity Field with PIVLab

B. Model Discovery with SPIDER

In this section, we detail the procedure for model discovery using the algorithm known as Sparse Physics-Informed Discovery of Empirical Relations (SPIDER) [6] [11] [12] [13]. The data pipeline ending in a set of interpretable partial differential equations starts by generating a library of candidate terms from the physical observables measured in the experiment

$$\{u_{ij}, r_{ij}\},$$

where u_{ij} is the flow velocity field and r_{ij} is the center of mass position from which we derive the center of mass velocity (v_i). These observables determine the set of useful physical quantities

$$\{v_i, u_i, \nabla_i, \eta_{ij}, \int dl_i\},$$

where η_{ij} is the anti-symmetric rank-two rotation tensor and $\int dl_i$ denotes the line integral on the circle which passes through the four points around the jack. From this set of quantities, we build tensors that form the basis of the SPIDER library. The equation of motion for the disk must then be of the form

$$\frac{d}{dt}v_i = f_i(v_i, u_i, \nabla_i, \eta_{ij}, \int dl_i), \tag{10}$$

where f_i is a function of all the tensors that form the basis of the library. The tensors of interest up to rank-3 are

$$\tau_1 = v_i \tag{11}$$

$$\tau_2 = \eta_{ij}, v_i v_j, \int dl_i u_j \tag{12}$$

$$\tau_3 = \eta_{ij} v_k, v_i v_j v_k, \int dl_i u_j u_k, \int dl_i \nabla_j u_k, \tag{13}$$

and there are no rank-0 tensors. The SPIDER library is then built with all the possible contractions of these tensors, so our finally library of candidate terms contains tensors of rank-1 and rank-2, labeled as L,

$$L_0 = \tau_0 \cup Tr\tau_2 \tag{14}$$

$$L_1 = \tau_1 \cup Tr\tau_3, \tag{15}$$

where Tr denotes contractions of a given tensor, not just the trace. To demonstrate this process, let's take the rank-3 tensor $Q_{ijk} = \int dl_i \nabla_j u_k$ as an example. The possible contractions of this tensor are

$$Q_i = \int dl_i \nabla_j u_j = \int dl_i \nabla_k u_k =$$
$$[(\int dx \nabla_x u_x + \int dx \nabla_y u_y), (\int dy \nabla_x u_x + \int dy \nabla_y u_y)]$$
$$Q_j = \int dl_i \nabla_j u_i = \int dl_k \nabla_j u_k =$$
$$[(\int dx \nabla_x u_x + \int dy \nabla_x u_y), (\int dx \nabla_y u_x + \int dy \nabla_y u_y)]$$
$$Q_k = \int dl_i \nabla_i u_k = \int dl_j \nabla_j u_k =$$
$$[(\int dx \nabla_x u_x + \int dy \nabla_y u_x), (\int dx \nabla_x u_y + \int dy \nabla_y u_y)],$$

therefore, the rank-3 tensor Q_{ijk} contributes three terms to the SPIDER library. We repeat this process for all of the leaf tensors τ to obtain the full library of candidate terms.

These tensors have been chosen with both physical and mathematical considerations. For example, the equation of motion for the disk ought to depend on its own velocity, so we consider not only the tensor v_j, but all of the other possible tensor up to rank-3 as well: $v_i v_j$, $v_i v_j v_k$, and $\eta_{ij} v_k$. Similarly for the fluid velocity, the tensor $I_{ijk} = \int dl_i u_j u_k$ is included in the library because the theoretical methods to date do not fully take into account higher-order or non-linear terms in the Navier-Stokes equation, which is discussed in the Background. However, this term is more than it appears and we should take a step back to Oseen's correction of Stokes' drag to understand it. Faced with Stokes' paradox, Oseen criticised Stokes' approximations and took into account the advective term of the Navier-Stokes equation, $(\boldsymbol{u} \cdot \nabla)\boldsymbol{u}$, via the linear approximation $U_0 \nabla \boldsymbol{u}$. Oseen's corrections are based on this linearization, so the current theoretical methods still cannot fully resolve this non-linear term. However, by applying Green's theorem, we can see that the tensor $\int dl_i u_j u_k$ actually contains the advective term in its full non-linear form,

$$\iint_R dA(\vec{\boldsymbol{u}} \cdot \vec{\nabla})\vec{\boldsymbol{u}} = \iint_R dA \nabla_i (u_i u_j) =$$
$$\iint_R dA(\underbrace{u_j \nabla_i u_i}_{0} + u_i \nabla_i u_j) = \int_C dl_i u_i u_j,$$

where $\iint_R$ denotes the integral over the surface of the disk, $\int_C$ denotes the integral over the closed loop around the disk, and Green's theorem is applied in the last step. Therefore, the full advective term is accounted for in the j- and k-components of the tensor I_{ijk}

$$I_j = \int dl_i u_i u_k = \int dl_j u_j u_k$$
$$I_k = \int dl_i u_j u_i = \int dl_k u_j u_k.$$

Additionally, the $Q_{ijk} = \int dl_i \nabla_j u_k$ tensor is related to the stresses and contains the drag force exerted on the disk by the surrounding fluid. Stokes' drag is found by integrating the viscous stress tensor,

$$\epsilon_{ij} = -p \cdot \delta_{ij} + \mu \left((\nabla u)_{ij} + (\nabla u)_{ji} \right), \tag{16}$$

over the surface of the sphere to obtain the drag force (see Appendix I). We do not have information about the pressure, so we neglect the first term in ϵ_{ij} (16). The second term evaluates to

$$F_D = \underbrace{\int dl_i \nabla_i u_j}_{Q_k} + \underbrace{\int dl_i \nabla_j u_i}_{Q_j}, \tag{17}$$

but Q_k ought to be zero since we are integrating a derivative. Therefore,

$$F_D = Q_j, \tag{18}$$

which is the drag force exerted by the fluid on the disk. Indeed, the sparse model, which is presented below, includes Q_j as one of the five important terms.

Finally, the tensor $E_i = \eta_{ij} v_j$, where η_{ij} is the rank-2 anti-symmetric rotation tensor, contributed the last two terms in the library. These terms represents the angular

motion of the disk, which is proportional to the COM velocity.

To summarize, the full SPIDER library consists of all the (possible) contractions of the following leaf tensors

$$\begin{aligned} E_i &= \eta_{ij} v_i \\ I_{ijk} &= \int dl_i u_j u_k \\ Q_{ijk} &= \int dl_i \nabla_j u_k, \end{aligned} \tag{19}$$

where E_i accounts for the rotational motion of the disk, I_{ijk} takes into account the products of the the fluid velocity, and Q_{ijk} accounts for the stresses on the disk.

C. Results

Given a library of terms, whose values at different spatiotemporal locations constitute the feature matrix A, SPIDER performs regression to fit a sparse coefficient vector $\boldsymbol{c}$ corresponding to the dynamics (in our case, the COM velocity v_i), where $A\boldsymbol{b} \approx \boldsymbol{c}$. SPIDER then sparsifies the library to arrive at the best continuum model for the dynamics of the system defined by

$$m_p \frac{dv_i}{dt} = \mathcal{L}(E_i, I_{ijk}, Q_{ijk}). \tag{20}$$

The residual (error) of the discovered model measures how well the coefficient matrix has has fit the dynamics of the system and it is defined as

$$\frac{1}{||\boldsymbol{c}||} ||A\boldsymbol{b} - \boldsymbol{c}||. \tag{21}$$

We find that only five terms constitute the best continuum description of our system with a residual of 0.05 (5%), which is shown by the plot of the normalized residual versus the number of library terms V C.

This result can be visually verified by plotting the fit of the model to the real dynamics as a function of library terms.

Figure V C shows the model fit to the COM acceleration where the data are displayed in order of increasing magnitude. This way of presenting the results allows for easier visual confirmation since the COM acceleration has no discernable pattern to the human eye.

The five-term model, which we identify as the best continuum description of the COM acceleration, includes the coefficient vector

$$\boldsymbol{b} = [-0.0010, -0.0011, 13.6277, 5.8182, -13.1558]$$

corresponding to the library terms

$$\mathcal{L}' = [\eta_i, \eta_j, Q_j, I_i, I_j].$$

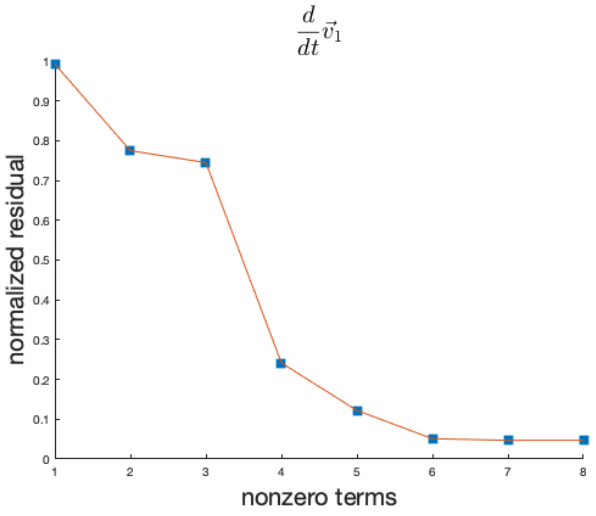

FIG. 6. Normalized Residual

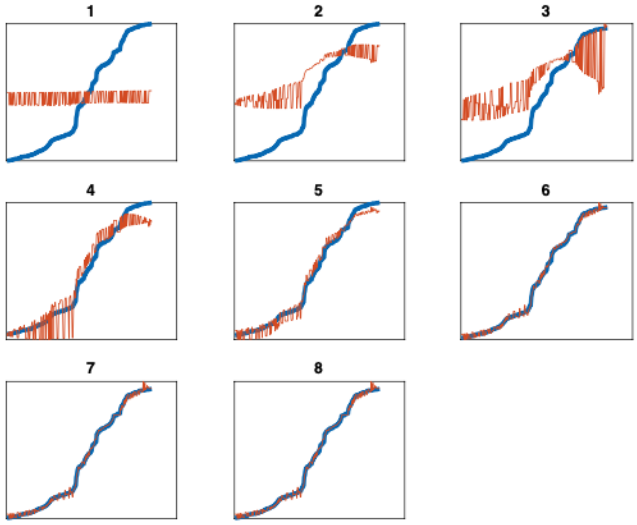

FIG. 7. Model Fit: model (orange), COM acceleration (blue)

D. Discussion

We conclude that SPIDER very successfully derives an equation of motion for the for the circular disk in a quasi-two-dimensional flow, which is evidenced by the low residual (0.05) and the close fit observed in figure V C. All of the leaf tensors are represented in the model, but not necessarily all of their contractions. For example, both components of E_i are represented, unsurprisingly, but Q_j is the only component of Q_{ijk} represented in the model. Recall that we identified Q_j as the drag force in equation (18) in the previous section. We also predicted that Q_k would not contribute, and SPIDER has determined that Q_i is not important either. Both I_i and I_k (identical to I_j) are important. In the previous section, we also identified both I_j and I_k as representing

the advective term $(\boldsymbol{u}\cdot\nabla)\boldsymbol{u}$, which has been the source of much mathematical difficulty since the time of Stokes. The importance of the full nonlinear advective term on the dynamics provides verification for this model as the system is not in Stokes' regime.

We can now present an equation of motion for a circular disk in a quasi-two-dimensional flow based on these findings,

$$m_p\frac{d\boldsymbol{v}}{dt} = \alpha\eta\boldsymbol{v} + \beta(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \kappa\int\nabla\boldsymbol{u}\cdot dl + \\ +\gamma\int\boldsymbol{u}^2 dl, \qquad (22)$$

where the coefficients $[\alpha,\beta,\kappa,\omega]$ correspond to the sparse coefficient vector $\boldsymbol{b}$ found by SPIDER (recall, in tensor form $\eta\boldsymbol{v}$ accounts for two terms, so α corresponds to two coefficients). It is clear that this equation is not in full agreement with the Maxey-Riley equation (1). The discovered model does not take into account the balance between the forces of gravity and buoyancy, the added mass, or the Basset force. To address these, the force balance must be added in by hand, but since the disk floats on the surface of the fluid, it is clearly negligible. The model is not yet able to take into account the added mass and Basset force, however, because no information about the time derivative of the fluid velocity has been provided to it. New library terms with information about the time dependence of the fluid velocity must be constructed in order to account for these forces. The discovered model, however, identifies forces that are not present in the Maxey-Riley equation such as the rotational force or forces due to the products of fluid velocities. These forces were added to the library due to their physical and mathematical validity. It is evident that a rotational force on the disk is very important, and the forces due to the products of fluid velocities may also be significant. The two models do agree on the importance of the advective acceleration and the drag force.

Beyond the different formulations of the two equations, namely due to the different geometry of the two-dimensional system and lack of fluid velocity time derivatives, discrepancies between the Maxey-Riley equation of motion and the discovered model may be due to limitations of the data. The data used in this study describes almost pure rotational motion with only slight COM motion. Therefore, SPIDER is fitting a near-zero COM acceleration and is approximately finding the null space of the feature matrix A. This limits the importance of forces relating to advection or the added mass and increases the importance of the rotational force. Additionally, viscosity affects the importance of certain forces such as the drag force, which is more important in lower Reynolds number regimes. Further study should consider richer data where the circular disk undergoes both COM and rotational motion and different fluid properties.

Finally, SPIDER has determined the numerical value of the coefficients multiplying each term in the equation of motion (22), but further study is required to interpret them.

VI. CONCLUSION

In this study, we have discussed a theoretical baseline for modeling single particle-fluid interactions, namely the Maxey-Riley equation for a rigid sphere in non-uniform flow, an experimental procedure for studying the two-dimensional case of a thin circular disk moving in (quasi-)two-dimensional flow, and a data-driven procedure for discovering the disk's equation of motion and comparing it to the Maxey-Riley equation.

While the Maxey-Riley equation cannot be applied to two dimensions, it summarizes the important forces acting on particles in fluid flows: the gravity-buoyancy force balance, the stress gradient of the undisturbed flow, the added mass, Stokes' drag and the unsteady Basset drag. Having explained each of these forces, we proposed an experimental procedure for studying the motion of a thin circular disk (jack) in a quasi-two-dimensional flow. The data collection procedure entailed recording the experiment with a camera, tracking the ends of the disk with a python package called TrackPy, and measuring the flow field with MatLab's PIVLab application. We then performed model discovery with SPIDER, which we instantiated with a library of candidate terms based on the physical observables measure in the experiment: the fluid velocity field and the COM motion. With a residual of 0.05, or an accuracy of 95%, we found that the best continuum description of the disk's COM motion (22) accounted for the rotational force, the advective acceleration, the drag force, and the products of fluid velocities. These findings agree with the Maxey-Riley equation (1) on the importance of the advective term and the drag force, but it fails to account for terms such as the added mass and Basset history term. This is unsurprising, however, since no information about the time derivative of fluid velocities is provided to SPIDER. More work is needed to include library terms that account for the time dependence of the fluid velocity field.

The low residual of 0.05, however, suggests that terms that are present in the Maxey-Riley equation (1) and absent in the discovered model (22) are not important to the dynamics of the disk. This discrepancy between the theoretical baseline and the discovered model is, in part, due to the fact that they describe different regimes. While the Maxey-Riley equation (1) describes the motion of a rigid sphere in a three-dimensional flow, the discovered model describes the motion of a thin disk in a quasi-two-dimensional flow. The model discovery procedure is data driven, so the discrepancy can also be attributed to limitations of the data, which shows almost entirely rotational motion. Since the data is largely limited to rotational motion and very little COM motion,

the rotational force is very important, while the drag force and advective term are less important. The data also only captures dynamics in a single Reynolds number regime, but data pertaining to more and less viscous regimes should be used to obtain a more generalized model. The latter exploration would shed more light on the effect of the advective term, for example, which should be more important in a high viscosity regime. Future work should focus on obtaining richer dynamics showing both interesting rotational and COM motion, and exploring different fluid and particle Reynolds number regimes to discover a more general model.

Appendix A: Stokes' Law

Stokes' Law is a solution for the frictional force exerted by a steady, low-Reynolds number flow on a sphere given by

$$F_D = 6\pi\mu R\nu, \tag{A1}$$

where μ is the dynamic viscosity, R is the radius of the sphere, and ν is the flow velocity relative to the sphere. The validity of Stokes' Law rests on the assumption that the Reynolds number is small ($Re << 1$), which does not hold far away as Stokes' solution grows logorithmically at infinity instead of returning to the fluid's normal flow. This is certainly unphysical as it suggests the spherical particle affects the fluid infinitely far away. Basset argued that Stokes' formula leads to terminal velocities that are larger than experimentally observed due to his neglect of the advection term. Stokes' approach is also not applicable to two dimensions where there can be no creeping flow around a disk in two dimensions, which is called Stokes' Paradox. Oseen in 1910 calculated correction terms to the drag force F_D on a sphere to rectify the asymptotic behavior of Stokes' solution by taking into account the advective term and solving for the drag force of certain two dimensional geometries. While the theoretically correct approach to calculating the drag force on a thin disk in two dimensional flow would use Oseen's approach, we posit that the two-dimensional drag force should only differ from the three-dimensional case by a constant factor. Therefore, we now present the derivation for the baseline model of the drag force on a thin disk following Stokes' original approach for a sphere.

We would like to calculate the force exerted by a two-dimensional fluid flow of density ρ and dynamic viscosity μ over a thin disk with radius R and vertical velocity V. To simplify the calculations we make two assumptions: (i) the Reynolds number is small

$$Re = \frac{2\rho VR}{\mu} < 1 \tag{A2}$$

and therefore, (ii) the flow around the disk is axisymmetric (i.e. $\frac{\partial}{\partial\phi} = 0$ and $v_\phi = 0$). We define the coordinate whose origin is at the center of the disk, which is fixed while the fluid flows past the disk. No-slip boundary conditions, therefore, require

$$v_r(R,\theta) = 0 \tag{A3}$$

$$v_\theta(R,\theta) = 0. \tag{A4}$$

The flow velocity should asymptote at $r \to \infty$ to $V = V\hat{z}$, which in polar coordinates implies

$$v_r(r \to \infty, \theta) = V\cos\theta \tag{A5}$$

$$v_\theta(r \to \infty, \theta) = -V \sin(\theta). \tag{A6}$$

Stokes' flow is characterized by $\mathcal{L}^2(\psi) = 0$, where ψ is Stokes' stream function and $\mathcal{L}$ is an operator defined by

$$\mathcal{L} = \frac{\partial^2}{\partial r^2} + \frac{\sin\theta}{r^2}\frac{\partial}{\partial\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}.$$

Stokes' stream function is defined by

$$v_r = \frac{1}{r^2\sin\theta}\frac{\partial\psi}{\partial\theta} \tag{A7}$$

$$v_\theta = -\frac{1}{r\sin\theta}\frac{\partial\psi}{\partial r}, \tag{A8}$$

which reduces the boundary conditions at the edge of the disk to

$$\frac{\partial\psi}{\partial\theta}\Big|_{r=R} = 0 \tag{A9}$$

$$\frac{\partial\psi}{\partial r}\Big|_{r=R} = 0. \tag{A10}$$

Setting the components of velocity (A7) and (A8) equal to their respective asymptotic values at $r \to \infty$ we obtain

$$\psi(r \to \infty, \theta) = \frac{1}{2}Vr^2\sin^2\theta, \tag{A11}$$

which implies that the stream function for all values has the form $\psi = f(r)\sin^2\theta$. Therefore, we apply the operator $\mathcal{L}$ and the boundary conditions (A9) and (A10) to determine the exact form of the stream function.

The action of $\mathcal{L}^2$ on ψ simplifies to

$$\left(\frac{d^2}{dr^2} - \frac{2}{r^2}\right)^2 f(r) = 0. \tag{A12}$$

This suggests a power solution $f(r) = Cr^n$, which gives

$$\frac{d^4}{dr^4}Cr^n - \frac{2}{r^2}\frac{d^2}{dr^2}Cr^n - \frac{d^2}{dr^2}2Cr^{n-2} + \frac{4}{r^4}Cr^n =$$

$$= n(n-1)(n-2)(n-3)Cr^{n-4} - 2n(n-1)Cr^{n-4} -$$

$$-2(n-2)(n-3)Cr^{n-4} + 4Cr^{n-4} = 0.$$

In factored form the solution can be easily read off

$$C[n(n-1)-2][(n-2)(n-3)-2] = 0$$

$$\Rightarrow n = -1, 1, 2, 4,$$

which gives

$$f(r) = \frac{A}{r} + Br + Cr^2 + Dr^4. \tag{A13}$$

To determine the coefficients of (A13) we apply the boundary conditions (A9), (A10), and (A11). From (A11) we deduce that $D = 0$ and $C = \frac{1}{2}V$. (A10) and (A11) give a fully determined set of equations

$$\frac{\partial\psi}{\partial\theta}\Big|_{r=R} = \frac{A}{R} + BR + \frac{1}{2}VR^2 = 0 \tag{A14}$$

$$\frac{\partial\psi}{\partial r}\Big|_{r=R} = -\frac{A}{R^2} + B + VR = 0. \tag{A15}$$

Adding (A14) and (A15) gives $B = -\frac{3}{4}VR$ and substituting back into (A15) gives $A = \frac{1}{4}VR$. The stream function ψ that now satisfies all boundary conditions is

$$\psi(r,\theta) = \frac{V(R^3 - 3Rr^2 + 2r^3)}{4r}\sin^2\theta. \tag{A16}$$

The only non-zero component of vorticity is

$$\omega_\phi(r,\theta) = \frac{\mathcal{L}(\phi)}{r\sin\theta} = \frac{3VR\sin\theta}{2r^2}. \tag{A17}$$

The pressure gradient is defined as $\nabla P = -\mu\nabla\times\omega$, which gives

$$\frac{\partial P}{\partial r} = -\frac{3\mu VR\cos\theta}{r^3} \tag{A18}$$

$$\frac{\partial P}{\partial\theta} = -\frac{3\mu VR\sin\theta}{2r^2} \tag{A19}$$

and it is easily seen (or evaluated by integration) that the general pressure distribution neglecting gravity is

$$P(r,\theta) = p_0 + \frac{3\mu VR\cos\theta}{2r^2}. \tag{A20}$$

In the case of Stokes' flow, the viscous stress tensor is

$$\sigma_{ij} = -p \cdot \delta_{ij} + \mu \left((\nabla u)_{ij} + (\nabla u)_{ji} \right). \tag{A21}$$

The tangential components of the stress tensor are

$$f_r(\theta) = \sigma_{rr}(R,\theta) = \left(-p + \underbrace{2\mu \frac{\partial v_r}{\partial r}}_{0} \right)_{r=R} \tag{A22}$$

$$f_\theta(\theta) = \sigma_{r\theta}(R,\theta) = \mu \left(\frac{1}{r}\frac{\partial v_r}{\partial \theta} + \frac{\partial v_\theta}{\partial r} - \frac{v_\theta}{r} \right)_{r=R}, \tag{A23}$$

which give

$$f_r(\theta) = -p_0 - \frac{3\mu V}{2R}\cos\theta \tag{A24}$$

$$f_\theta(\theta) = \frac{3\mu V}{2R}\sin\theta. \tag{A25}$$

This gives the drag force density on the sphere as

$$F_D = -\frac{3\mu V}{2R}\hat{\boldsymbol{z}} \tag{A26}$$

and integrating over the surface of the sphere gives the total drag force,

$$\oint_S F_D \cdot \hat{\boldsymbol{z}} dA = -6\pi\mu R V. \tag{A27}$$

Lamb calculated the drag force acting on a circular disk move edgeways through a viscous fluid by considering its gravitational potential using Dirichilet's formula [7]. Lamb finds the same expression as (A27), where $R = 16c/9\pi$, and $a = 0$ and $b = c$ are the axis of an ellipsoid.

Appendix B: A Note on Stokes' Paradox and Oseen's Equations

In Appendix A we presented the derivation of the drag force experienced by a spherical particle in creeping flow. Now, in two dimensions as was the case in three, the solution is found with a stream function by solving the biharmonic equation

$$\mu\nabla^2\omega = \mu\nabla^4\psi = 0. \tag{B1}$$

In polar coordinates, the stream function can be separated as $\psi = f(r)\sin\theta$ and substituting into (B1) yields

$$\left[\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{1}{r^2} \right]^2 f = 0. \tag{B2}$$

The no slip boundary condition requires that $\psi(R) = \frac{\partial\psi}{\partial r}(R) = 0$ and the condition of an unbounded flow at infinity requires that $\psi_r(r \to \infty) = Ur$. The general solution is then

$$f(r) = Ar^3 + Br\ln r + Cr + Dr^{-1}. \tag{B3}$$

The first two terms can immediately be elimated due to the condition at infinity and the no-slip boundary condition yields

$$CR + DR^{-1} = C - DR^{-2} = 0, \tag{B4}$$

which has the unique solution $C = D = 0$. Therefore, there is no solution for an unbounded flow around a circular disk in two dimensions. This is known as Stokes' paradox. Oseen corrected Stokes' analysis by making a linear approximation of the advective term, $(\boldsymbol{v}\cdot\nabla)\boldsymbol{v} \to U_0\nabla\boldsymbol{v}$, which adds correction terms to Stokes' drag

$$F_D = 6\pi\mu R\nu\left(1 + \frac{3}{8}Re + ...\right). \tag{B5}$$

Appendix C: Reproducibility

All materials used in this study are published to the GitHub repository linked in Supplemental Materials: the COM and fluid velocity field data, the code for particle tracking and the SPIDER algorithm. As long as the SPIDER folder is stored as one file it should run smoothly in Matlab. SPIDER includes functions that are defined in separate files, so it is imperative that everything related to SPIDER remain in the same folder. All of the results of this study should be exactly reproduced by the materials uploaded to the repository. If, however, you would like to examine other data, all import statements are written at the beginning of each script and nowhere else.

Appendix D: Supplemental Materials

The repository containing all of the relevant code for this work including particle tracking and model discovery is hosted on GitHub: repository.

[1] G. G. Stokes, On the effect of the internal friction of fluids on the motion of pendulums, Transactions of the Cambridge Philosophical Society. **9**, 52 (1851).

[2] G. A. Voth and A. Soldati, Anisotropic particles in turbulence, Annual Reviews of Fluid Mechanics **49**, 249 (2017).

[3] G. K. Batchelor, Slender-body theory for particles of arbitrary cross-section in stokes flow, Journal of Fluid Mechanics **44**, 419 (1970), DOI: https://doi.org/10.1017/S002211207000191X.

[4] G. K. Batchelor, The motion of long slender bodies in a viscous fluid. part 1. general theory, Journal of Fluid Mechanics **44**, 791 (1970), DOI: https://doi.org/10.1017/S002211207000215X.

[5] A. Talei and T. J. Garrett, On the maxey–riley equation of motion and its extension to high reynolds numbers, Arxiv (2020), DOI: https://arxiv.org/pdf/2006.16577.pdf.

[6] M. Golden, R. O. Grigoriev, J. Nambisan, and A. Fernandez-Nieves, Physically informed data-driven modeling of active nematics, Science Advances **9** (2023), DOI: https://doi.org/10.1126/sciadv.abq6120.

[7] H. Lamb, Hydrodynamics (Cambridge University Press, 1945).

[8] J. Boussinesq, Sur la résistance qu'oppose un fluide indéfini en repos, sans pesanteur, au mouvement varié d'une sphère solide qu'il mouille sur toute sa surface, quand les vitesses restent bien continues et assez faibles pour que leurs carrés et produits soient négligeables, C. R. Acd. Sci. Paris **100**, pp. 935 (1885).

[9] C. T. Crowe, J. D. Schwarzkopf, M. Sommerfeld, and Y. Tsuji, Multiphase Flows with Droplets and Particles (CRC Press, 2012).

[10] A. B. Basset, Treatise on hydrodynamics, Nature **40** (1889), DOI: https://doi.org/10.1038/040412a0.

[11] D. R. Gurevich, P. A. Reinbold, and R. O. Grigoriev, Robust and optimal sparse regression for nonlinear pde models, Chaos (2019), DOI: https://doi.org/10.1063/1.5120861.

[12] P. A. Reinbold, D. R. Gurevich, and R. O. Grigoriev, Using noisy or incomplete data to discover models of spatiotemporal dynamics, Physical Review E (2020).

[13] P. A. Reinbold, L. M. Kageorge, M. F. Schatz, and R. O. Grigoriev, Robust learning from noisy, incomplete, high-dimensional experimental data via physically constrained symbolic regression, Nature Communications (2021), DOI: https://doi.org/10.1038/s41467-021-23479-0.

[14] M. Do-Quang, G. Amberg, G. Brethouwer, and A. V. Johansson, Simulation of finite-size fibers in turbulent channel flows, Physical Review E **89** (2014), DOI: http://dx.doi.org/10.1088/1367-2630/16/10/102001.

[15] F. P. Bretherton, The motion of rigid particles in a shear flow at low reynolds number, Journal of Fluid Mechanics **14** (1962), DOI: https://doi.org/10.1017/S002211206200124X.

[16] B. Suri, J. Tithof, R. M. Jr., R. O. Grigoriev, and M. F. Schatz, Velocity profile in a two-layer kolmogorov-like flow, Physics of Fluids **26** (2014).

[17] J. C. Crocker and D. G. Grier, Methods of digital video microscopy for colloidal studies, Journal of Colloid and Interface Science **179**, 298 (1996), DOI: http://doi.org/10.1006/jcis.1996.0217.

[18] W. Thielicke and R. Sonntag, Particle image velocimetry for matlab: Accuracy and enhanced algorithms in pivlab, Journal of Open Research Software **9** (2021), DOI: 10.5334/jors.334.