Soft Matter

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Magneto-capillary particle dynamics at curved interfaces: inference and criticism of dynamical models†

Dimitri Livitz, Kiran Dhatt-Gauthier, and Kyle J. M. Bishop*

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Time-varying fields drive the motion of magnetic particles adsorbed on liquid drops due to interfacial constraints that couple magnetic torques to capillary forces. Such magneto-capillary particle dynamics and the associated fluid flows are potentially useful for propelling drop motion, mixing drop contents, and enhancing mass transfer between phases. The design of such functions benefits from the development and validation of predictive models. Here, we apply methods of Bayesian data analysis to identify and validate a dynamical model that accurately predicts the field-driven motion of a magnetic particle adsorbed at the interface of a spherical droplet. Building on previous work, we consider candidate models that describe particle tilting at the interface, field-dependent contributions to the magnetic moment, gravitational forces, and their combinations. The analysis of each candidate is informed by particle tracking data for a magnetic Janus sphere moving in a precessing field at different frequencies and angles. We infer the uncertain parameters of each model, criticize their ability to describe and predict experimental data, and select the most probable candidate, which accounts for gravitational forces and the tilting of the Janus sphere at the interface. We show how this favored model can predict complex particle trajectories with micron-level accuracy across the range of driving fields considered. We discuss how knowledge of this "best" model can be used to design experiments that inform accurate parameter estimates or achieve desired particle trajectories.

1 Introduction

Magnetic particles adsorbed at liquid interfaces can be actuated by external fields to drive particle motion and fluid flow that provide a basis for dynamic biphasic materials[1–3] such as emulsions,[4,5] foams,[6,7] and bijels.[8] In this context, we distinguish different types of magneto-capillary phenomena based on the nature of the driving field, the role of interparticle interactions, and the curvature of the interface. For an individual particle at a planar interface, field gradients are required to produce magnetic forces that drive particle motion. In the absence of such gradients, time-varying fields can propel particle assemblies mediated by magnetic and capillary interactions.[9–13] Even at low Reynolds numbers, such fields can produce non-reciprocal motions among three or more particles that propel their motion across the interface.[11–13] At *curved* interfaces, spatially uniform fields move individual particles by coupling magnetic and capillary torques to produce forces that scale linearly with the interfacial curvature.[14–16] Such magneto-capillary forces at curved interfaces are potentially useful in addressing magnetic particles adsorbed onto micron-scale droplets in emulsions. Their field-induced motion and the associated fluid flow can serve to accelerate mass transfer[16] at the level of individual droplets and perhaps also propel droplet motion through viscous environments.

The design of magneto-capillary particle dynamics to achieve behaviors such as droplet mixing or propulsion benefits from the development and validation of predictive models. Such models allow for rapid screening of candidate designs—for example, different driving fields, particle shapes, and magnetic properties—to identify and understand the conditions that produce desired functions. Existing models of magneto-capillary particle motion on curved interfaces are based on lattice Boltzmann simulations of the Navier–Stokes equations[14,17] or numerical integration of approximate dynamical equations.[15,16] The former allows for direct simulation of multiple particles of different shapes and/or surface chemistries but requires significant computational resources, which can slow the design process. Approximate models of particle dynamics make simplifying assumptions that significantly accelerate model evaluation but require comparison with experimental data to validate predictions and infer unknown parameters. Previously, we showed that models based on rigid capillary constraints can explain the complex trajectories of magnetic Janus spheres adsorbed at spherical interfaces subject to precess-

Department of Chemical Engineering, 500 W 120 St, New York, NY USA; E-mail: kyle.bishop@columbia.edu

† Electronic Supplementary Information (ESI) available: derivation of model dynamics and details on numerical methods; summary of prior distributions used for parameter estimation; further details on model selection; calculation of the information content of an experiment; implementation details of numerical and statistical methods. See DOI: 10.1039/cXsm00000x/

ing fields.[16] However, the quantitative accuracy of such models remains uncertain thereby limiting their utility in the design of useful functions.

Given one or more candidate models, Bayesian data analysis[18–20] provides a principled approach for inferring their parameters from experimental observations, assessing their descriptive and predictive capabilities, and selecting the most plausible among them. These methods of parameter estimation, model criticism, and model selection require probabilistic descriptions that provide the likelihood of observed data given the values of model parameters. The uncertain parameters are described by prior probability distributions that serve to constrain their values based on relevant information. Using Bayes theorem, the likelihood and the prior are combined to produce a posterior distribution for model parameters, which summarizes their likely values and the associated uncertainty given the data. Importantly, this process of Bayesian parameter estimation (or model training) assumes the validity of the candidate model, which must be criticized on the basis of its predictive accuracy and internal consistency to avoid drawing false conclusions.[21,22] Ideally, a fitted model should produce simulated data which is indistinguishable from the observed data used to inform parameter estimation; discrepancies between simulations and observations indicate limitations in the descriptive capacity of the model. Additionally, a useful model should produce accurate predictions of new observations under unfamiliar conditions not used in training. When multiple candidate models survive such criticism, Bayesian model selection can help to identify the more probable candidate given the available data. This best model can then be used to guide the design of future experiments for the purpose of achieving functional behaviors (e.g., drop mixing) or enhancing knowledge of the modeled system (e.g., reducing parameter uncertainty).

Here, we apply methods of Bayesian data analysis to identify and validate a dynamical model that accurately predicts the field-driven motion of a magnetic particle adsorbed at the interface of a spherical droplet. In addition to previous models of magneto-capillary dynamics,[15,16] we consider candidate models that describe particle tilting at the interface, field-dependent contributions to the magnetic moment, gravitational forces, and their combinations. The analysis of each candidate model is informed by particle tracking data for a magnetic Janus sphere moving in a precessing field at different frequencies and angles. For each candidate, we divide the experiments into batches and estimate the most probable parameters and their uncertainties. The fitted models are criticized to assess (i) their capacity for describing the data, (ii) the consistency of parameter estimates across batches, and (iii) the accuracy of model predictions. For the few models that survive criticism, we use Bayesian model selection to identify the most probable candidate, which accounts for the effects of gravity and tilting of the Janus sphere at the interface. We show how this favored model can predict complex particle trajectories with micron-level accuracy across the range of driving fields considered. Additional experiments on tilted particles reveal star-shaped trajectories, which were previously unobserved but are well explained by the model. We discuss how knowledge of this "best" model can be used to design experiments that inform accurate parameter estimates or achieve desired particle trajectories.

2 Methods

2.1 Experimental Data

Our analysis is based on data from a previous experimental study on the field-driven dynamics of magnetic Janus particles adsorbed at the interface of a water drop in decane (Fig. 1a).[16] The particles are 4 μm polystyrene spheres coated on one hemisphere with a magnetic Ni layer and functionalized with a hydrophilic surface treatment to direct their orientation at the interface.[15,23] Previous analysis of their field-driven motion suggests that these particles have a permanent magnetic moment, $m \approx 3\times10^{-14}$ A m^2, directed parallel to the Janus equator and to the water-decane interface (Fig. 1a).[15] Prior to each experiment, a single adsorbed particle sediments under gravity to the lower pole of a spherical drop where it is imaged by an optical microscope. Application of a precessing magnetic field with frequency ω and angle φ drives the particle to move across the interface and approach one of two possible attractors—termed, pole-orbiting and zig-zag, respectively.[16] The particle moves either along circular trajectories around the lower pole of the drop (pole-orbiting; Fig. 1c) or along sawtooth trajectories around the drop equator (zig-zag; Fig. 1c).

We analyze particle tracking data for a series of experiments conducted at different frequencies ω and angles φ of the precessing field for a single particle-drop pair (Fig. 1). The goal of the analysis is to develop and validate a quantitative model for the magneto-capillary dynamics of particles moving on curved interfaces in time-varying magnetic fields. Ideally, the model should be capable of describing the transient particle trajectories for all experimental designs (i.e., values of ω and φ) to within a specified precision using a minimal number of fitting parameters. As we will show, the model proposed previously[16] to describe the particle dynamics is inadequate in reproducing quantitatively the tracking data although it agrees qualitatively with the experimental results. In the following section, we describe several extensions of this base model that account for previously neglected effects due to gravity, particle tilting at the interface, field-dependent contributions to the magnetic moment, and combinations thereof.

2.2 Candidate Models

Base model: Magnetocapillary dynamics

The base model accounts for the magnetic, capillary, and hydrodynamic forces/torques acting on a rigid spherical particle adsorbed at the interface of a spherical drop (see also ESI, Section 1.1†).[15,16] In a spatially uniform field $\mathbf{B}(t)$, the magnetic torque $\mathbf{m}\times\mathbf{B}$ acts to align the particle's magnetic moment parallel to the field, and there is no magnetic force on the particle. The moment $\mathbf{m}$ is approximated by a constant vector fixed to the particle directed parallel to the water-oil interface. Capillary forces and torques act to constrain particle motion, allowing two degrees of particle translation on the interface and one degree of particle rotation normal to the interface.[16] For a spherical drop of radius

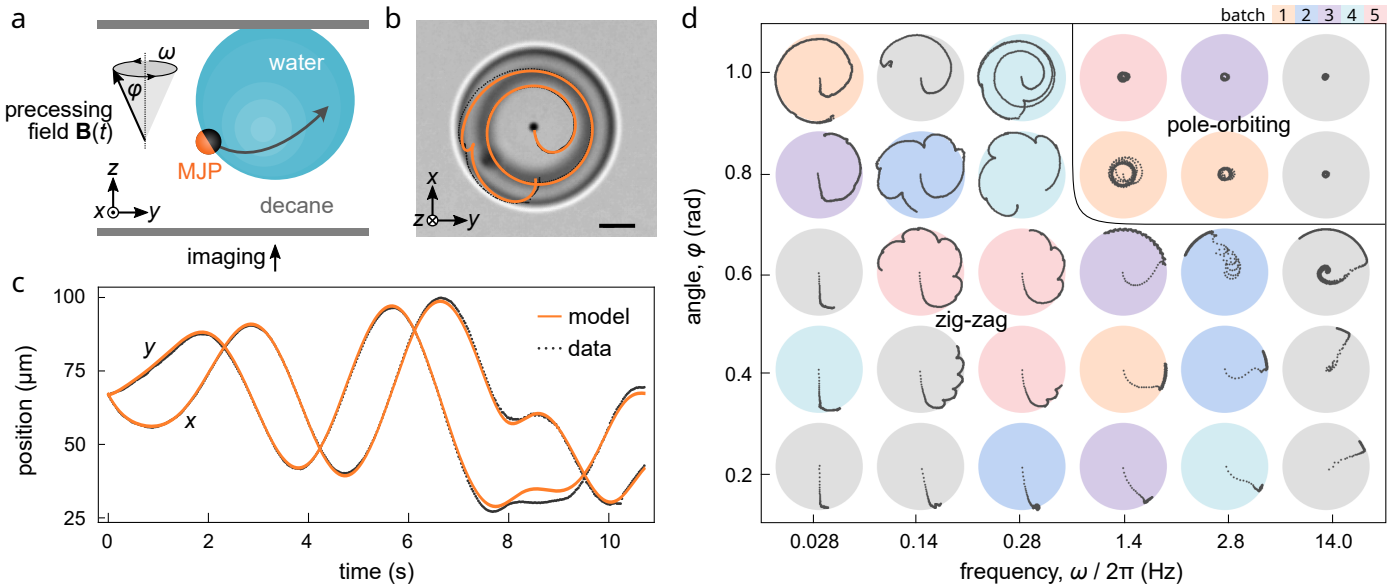

Fig. 1 (a) Schematic illustration of the experimental setup: a magnetic Janus particle is adsorbed at the interface of a water drop in decane and moves under the influence of a precessing field $\mathbf{B}(t)$ with angular frequency ω and precession angle φ. (b) Experimental image with measured trajectory of the particle (gray markers) overlaid with the prediction of the best fit model (orange curve). Scale bar is 20 μm. (c) Particle position x and y as a function of time t for a particular experiment (gray markers) as compared to predictions of the best fit model (orange curves). (d) Experiments are conducted at different frequencies ω and angles φ for a single particle-drop pair. Experiments at frequencies below 14 Hz are randomly assigned to batches 1-5 indicated by the drop color. Measured trajectories (gray markers) are classified as "pole-orbiting" or "zig-zag" based on their asymptotic behaviors.[16]

R, these constraints are summarized by the kinematic condition $\mathbf{U} = R(\mathbf{\Omega} \times \mathbf{n})$, which relates the particle's linear velocity $\mathbf{U}$ to its angular velocity $\mathbf{\Omega}$ and the unit normal vector $\mathbf{n}$ directed out from the drop. At low Reynolds numbers ($\mathrm{Re} = \rho a U/\eta \approx 7\times10^{-6} \ll 1$), inertial effects are negligible, and the particle velocity is linearly related to the applied force/torque by the hydrodynamic resistance tensor. We approximate the components of this tensor by that of a solid sphere of radius a in an unbounded fluid of viscosity η.[16] With these preliminaries, the angular velocity of the particle is given by

$$\mathbf{\Omega} = \left(\frac{\boldsymbol{\delta} - \mathbf{nn}}{8\pi\eta a^3 + 6\pi\eta a R^2} + \frac{\mathbf{nn}}{8\pi\eta a^3} \right) \cdot (\mathbf{m} \times \mathbf{B}) \qquad (1)$$

where $\boldsymbol{\delta}$ is the identity matrix. Due to the kinematic constraint, the dynamical state of the system can be specified by the 3D orientation of the particle-drop complex as parameterized by Euler angles or the unit quaternion.[24] Equation (1) describes the evolution of this orientation—and the associated particle position—due to the time-varying field $\mathbf{B}(t)$. The model is integrated numerically using the DifferentialEquations.jl library in Julia.[25]

The base model is fully specified by ten parameters: the particle radius a, the drop radius R, the average fluid viscosity η, the magnetic moment m, the magnitude B, frequency ω, and angle φ of the precessing field, as well as the initial position and orientation of the particle at the interface. Of these, eight are independent since the dynamics depends on the expression mB/η rather than the component parameters. Some of the parameter values are known a priori: $a = 2$ μm, $\eta = 8.7\times10^{-4}$ Pa s, $B = 5$ mT. Others are specified by the experiment design—namely, the precession angle φ and the angular frequency ω. Still other parameters are uncertain but remain constant from one experiment to the next: the drop radius R and the magnetic moment m. The remaining parameters are uncertain and vary from one experiment to the next—namely, the initial position and orientation of the particle on the drop. As detailed below, the uncertain parameters are fitted to the data using Bayesian inference.

To understand the limits of the base model, it is instructive to consider more closely the assumptions on which it relies. The accuracy of the rigid capillary constraint implies that capillary effects are much stronger than magnetic effects as described by the dimensionless parameter $\gamma a^2/mB \approx 10^3 \gg 1$, where $\gamma = 51$ mN/m is the interfacial tension. The approximation used for the resistance tensor assumes (i) that the Janus particle is adsorbed symmetrically at the interface, (ii) that the viscosity contrast between the two liquids is small (ca. 6% for water-decane), and (iii) that the particle is much smaller than the drop such that curvature effects are negligible ($a/R \approx 0.02 \ll 1$).[26] While these assumptions appear to be reasonably accurate, it is possible that the Janus particle adsorbs at an angle thereby tilting the direction of the magnetic moment relative to the interface. Moreover, the base model neglects additional forces due to gravity which may be non-negligible as suggested by the dimensionless parameter $MgR/mB \approx 0.05$, where M is the buoyant mass of the particle, and g is the acceleration due to gravity. Finally, the particle's magnetic moment may not be constant due to additional contributions induced by the external field.

In analyzing the base model and its extensions, we also account for imperfections in the experimental measurements. In particular, the axis of the precessing field is not exactly parallel to the imaging direction, thereby distorting particle trajectories pro-

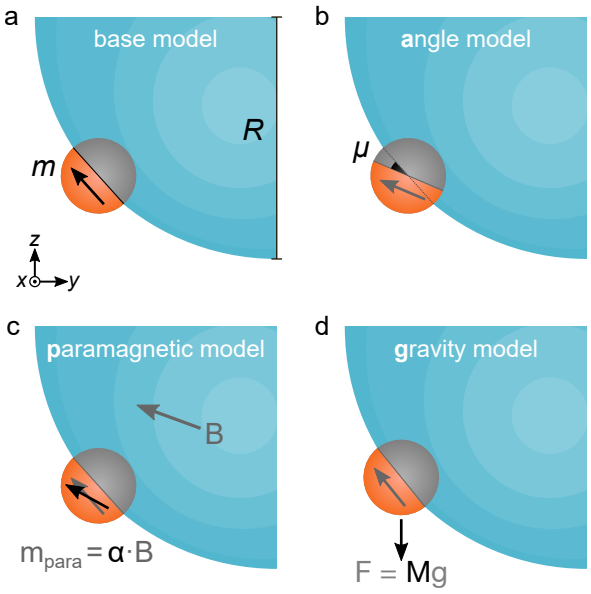

Fig. 2 Candidate models. (a) The *base model* assumes an permanent magnetic moment m of unknown magnitude oriented parallel to the liquid interface. (b) The *angle model* allows the permanent moment to tilt by an angle μ relative to the interface due to asymmetric particle adsorption. (c) The *paramagnetic model* instead considers field-induced contributions to the magnetic moment as characterized by the polarizability difference $\Delta\alpha = \alpha_\perp - \alpha_\parallel$ between directions pependicular and parallel to the Janus axis. (d) The *gravity model* includes the gravitational force on the particle which depends on its buoyant mass M. Combinations of the extended models (b-d) are also considered.

jected onto the imaging plane. For example, pole orbiting trajectories appear displaced from the drop center and distorted from their ideal circular shape. We introduce two angular parameters common to all experiments that describe the tilt of the field axis with respect to the imaging direction. By contrast, the gravity direction is assumed parallel to the imaging direction for models that include gravity. Finally, because imaging is not synchronized with the driving field, the initial phase of the precessing field in each experiment is unknown and treated as a fitting parameter.

Angle model: Effect of particle tilting

In the simplest extension of the base model, we consider that the Janus particle adsorbs onto the water-decane interface at an angle thereby tilting its magnetic moment out of the plane of the interface. The particle dynamics is still described by equation (1); however, the magnetic moment $\mathbf{m}$ now has components parallel and perpendicular the interface. We parameterize the magnetic moment $\mathbf{m}$ in terms of its magnitude m and the tilt angle μ such that $\mathbf{m}\cdot\mathbf{n} = m\sin\mu$.

Paramagnetic model: Effect of field-induced moment

In addition to the permanent magnetic moment considered in the base model, we consider that the moment has an additional field-induced contribution of the form $\mathbf{m} = \mathbf{m}_\mathrm{p} + \boldsymbol{\alpha}\cdot\mathbf{B}$, where $\mathbf{m}_\mathrm{p}$ is the permanent moment, and $\boldsymbol{\alpha}$ is the magnetic polarizability tensor. As in the base model, the permanent magnetic moment is directed parallel to the Janus equator and to the water-decane interface. We assume that the polarizability tensor shares the symmetry of the Janus particle with components $\alpha_\parallel$ and $\alpha_\perp$ parallel and perpendicular to the Janus director. Notably, only the difference between these components $\alpha_\perp - \alpha_\parallel$ is needed to compute the additional contribution to the magnetic torque. With this one added parameter ($\Delta\alpha$), the particle dynamics is still described by equation (1).

Gravity model: Effect of gravity

In the gravity model, we consider the gravitational force on the particle as well as the magnetic torque. Subject to the capillary constraint, the angular velocity of the particle is given by

$$\boldsymbol{\Omega} = \left(\frac{\boldsymbol{\delta} - \mathbf{nn}}{8\pi\eta a^3 + 6\pi\eta a R^2} + \frac{\mathbf{nn}}{8\pi\eta a^3}\right)\cdot\left[(\mathbf{m}\times\mathbf{B}) + RM(\mathbf{n}\times\mathbf{g})\right] \quad (2)$$

as derived in the ESI, Section 1.2†. The gravitational field $\mathbf{g}$ is antiparallel to the imaging direction and approximately parallel to the axis of the precessing field $\mathbf{B}(t)$. While the field magnitude is known, the buoyant mass of the particle M is treated as an uncertain parameter with an estimated value of $M \approx 2\times10^{-11}$ g based on the composition of the particle and the surrounding fluids. Owing to the heavy metal coating on one hemisphere, the particle also experiences a gravitational torque; however, its contribution is expected to be much smaller than that of the force considered here (by a factor of $a/R \ll 1$). We therefore neglect this additional torque.

Correlated Gaussian noise

The deterministic models outlined above are augmented by additive Gaussian noise $w(t)$ with zero mean and covariance

$$\langle w(t)w(t')\rangle = \sigma^2\exp(-|t-t'|/\tau) \quad (3)$$

where σ is the noise magnitude, and τ is the correlation time. Such noise is added to the computed particle positions $x(t)$ and $y(t)$ to describe the observed particle trajectories projected onto the xy-plane. With the addition of noise, we obtain a full probability model for the observed tracking data that describes the likelihood of experiment outcomes conditioned on the model parameters. This probabilistic description of the observed data is a prerequisite for applying methods of Bayesian parameter estimation and model selection.[19,20] Our use of the correlated noise model (3) (as opposed to white noise) is motivated by analysis of the residuals between the experimental data and predictions of the deterministic model.[19] The autocorrelation of the residuals decays in time at a characteristic rate comparable to the slower of the two intrinsic rates of the system: the precession frequency ω and the magnetic relaxation rate[15] $k_\mathrm{m} = mB/6\pi\eta a R^2$. We specify the correlation rate as $\tau^{-1} = \min(k_\mathrm{m}, \omega)$ and the noise magnitude as $\sigma = 1$ pixel (0.58 μm). In the present experiments, the imaging frame rate is typically much faster than the relaxation rate resulting in highly correlated measurements. By accounting for these correlations, the noise model (3) gives conservative estimates for parameter uncertainty that depend on the *effective* number of data points t_max/τ independent of the frame rate.

2.3 Parameter Estimation & Model Criticism

Bayesian inference

We use Bayesian inference to estimate the most probable parameter values for each of the candidate models and their combinations using data from 20 experiments conducted at different precession frequencies ω and angles φ (Fig. 1c, multi-color). An additional 10 experiments are set aside for use in validating the fitted models (Fig. 1c, gray). The duration of each time series is truncated to three times the respective correlation time to facilitate comparison of parameter estimates between experiments. The posterior distribution for the parameters (denoted collectively as θ) conditioned on the data $\mathcal{D}$ and model M is evaluated using Bayes theorem

$$p(\theta \mid \mathcal{D}, M) = \frac{p(\mathcal{D} \mid \theta, M)\, p(\theta \mid M)}{p(\mathcal{D} \mid M)} \tag{4}$$

Here, the likelihood function $p(\mathcal{D} \mid \theta, M)$ describes the probability of observing the data $\mathcal{D}$ under model M given knowledge of the parameter values θ. The prior distribution $p(\theta \mid M)$ describes knowledge of the uncertain parameter values before considering the data. For example, the prior for the magnetic moment m is approximated by a log-normal distribution with a median of 3×10^{-14} A m^2 and an interquartile range spanning 1.4 times more or less.[15] Other priors are detailed in the ESI, Section 2†.

We use numerical optimization to identify the most probable parameter estimates that maximize the logarithm of the posterior distribution (4). Starting from many (∼100) initial parameter guesses, we use automatic differentiation[27] to evaluate function gradients and the L-BFGS algorithm to ascend to local maxima. In pursuit of the global optimum, we use Latin hypercube sampling to distribute the initial parameter guesses over the high probability regions of the prior distribution. This procedure reliably converges to the same posterior mode for different realizations of the initial guesses, suggesting that the global optimum is identified.

We then approximate the posterior by a multivariate normal distribution centered on the most likely parameter estimates. To determine the covariance matrix of this distribution, we use automatic differentiation to compute the Hessian matrix of the log-posterior evaluated at the mode. Under the so-called Laplace approximation,[19] the covariance matrix is equal to the negative inverse of this Hessian matrix. When the true posterior has a single well defined mode, this approximation provides an accurate and concise description of the likely parameter values and their respective uncertainties.

To accelerate the process of parameter estimation, we divide the data into 5 batches of 4 experiments and analyze each batch independently of the others. By focusing on 4 experiments at a time rather than 20, the dimensionality of the parameter space is reduced significantly (e.g., from 84 to 20 in the base model) thereby facilitating numerical optimization. Moreover, the respective posteriors conditioned on each batch are more likely to have a single mode to which the optimization algorithm reliably converges. As detailed below, these distributions can be compared and combined to produce the final parameter estimates. In this way, batching provides a scalable approach to analyzing large amounts of experimental data with large numbers of models parameters. Here, we assign experiments to batches at random as illustrated in Figure 1c. The batch size of 4 experiments is selected to accelerate computation and produce unimodal posteriors well described by the Laplace approximation.

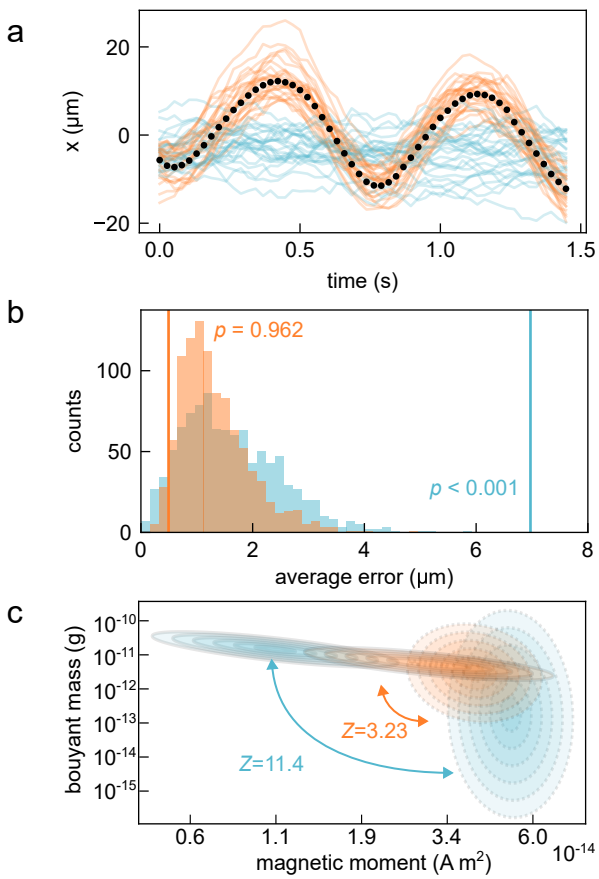

Fig. 3 Model criticism. (a) Posterior predictions for the x-component of particle position for the base model (blue curves) and the angle+gravity model (orange curves) as compared to experimental data (black markers). The experiment is from batch 2 with precession angle $\varphi = 1.0$ rad and frequency $\omega/2\pi = 0.28$ Hz (Fig. 1b). (b) Histogram of average absolute error for simulated trajectories like those shown in (a); the measured error for the experimental trajectories is plotted as a vertical line. The right-tailed p-value for the base model is much less than 0.05 indicating a poor description of the data. (c) Marginal posteriors for the magnetic moment m and the buoyant mass M for two batches (solid vs. dashed contours) as estimated using the paramagnetic+gravity model (blue) and the angle+gravity model (orange). The distributions for each model are annotated by the pairwise z-score of equation (5).

Posterior predictive checks

The fitted models are criticized to assess (1) the model description of the experimental data, (2) the consistency of parameter estimates across batches, and (3) the accuracy of model predictions for unseen data. First, at the batch level, we perform posterior predictive checks (PPCs) that compare observed data to simulated data produced by the fitted model. Using the Laplace approximation for the posterior, we sample ∼1000 parameter values and simulate noisy particle trajectories to produce an ensem-

ble of model predictions. Figure 3a illustrates such predictions for one component of the particle position generated using two different models. Visually, the predictions of the angle+gravity model "look like" the experimental data, while those of the base model do not. To quantify this distinction, we first compute the average absolute error between each simulated trajectory and the noise-free trajectory computed using the most probable parameter values. Figure 3b shows histograms of these errors for the two different models measured in units of particle diameters. We then compute the analogous error for the experimental trajectory and compare it to the distribution of simulated errors. Failure of the experiment error to conform to expectations of the fitted model is quantified using a p-value, which describes the probability of obtaining an error greater than or equal to that observed in experiment assuming that the model is correct. A small p-value (often $p < 0.05$) indicates that the fitted model does not provide a self-consistent description of the experimental data. Continuing the example above, the small p-value for the base model in Figure 3b suggests that this model is unable to describe the experimental data at the level of accuracy specified by the noise model. By contrast, the angle+gravity model passes the PPC to provide one consistent description of the data—but not the only one.

Model consistency between batches

Further support for a candidate model is obtained by comparing the batch-level parameter estimates for consistency. When a model is sufficiently accurate, parameter estimates from different batches of experimental data should agree with one another within the uncertainty of the posterior distributions. Figure 3c shows examples of parameter estimates for two batches using two models that are more or less consistent. The degree of consistency can be quantified by a pairwise z-score

$$z_{ij} = \left[(\mu_j - \mu_i)^T (\Sigma_i + \Sigma_j)^{-1} (\mu_j - \mu_i)\right]^{1/2} \tag{5}$$

where μ_i and Σ_i are the posterior mean and covariance matrix of batch i. The z-score measures the distance between the two parameter estimates scaled by the uncertainty. Small z-scores of order one suggest consistent parameter estimates between batches while large z-scores indicate inconsistencies.

Figure 3c shows the marginal posteriors for the the magnetic moment m and the buoyant mass M for two batches analyzed using two different models. For the angle+gravity model, the overlapping distributions produce a z-score of 3.23, while the paramagnetic+gravity model leads to non-overlapping distributions with $z_{12} = 11.4$. The large z-score of the paramagnetic+gravity model suggests that inferences drawn from the two batches are inconsistent with one another. Such inconsistencies may indicate that the model is wrong—for example, it uses the descriptive ability of one physical effect (e.g., paramagnetic anisotropy) to describe a different effect (e.g., particle tilting). Alternatively, inconsistencies between batches may indicate that the model is overfitting each batch, in which case larger batch sizes may resolve the inconsistency.

Cross-validation

The candidate models are further criticized using cross-validation in which parameters inferred from one batch are used to predict data in other batches. We use the most probable parameter estimates from each batch to make noise-free predictions of experiments in other batches and compute the root mean square (RMS) error for each prediction. Ideally, a good model should make accurate predictions of unseen data not used in model training. Moreover, if a model is consistent from batch to batch, we can combine the batch-level parameter distributions into a global posterior of the form

$$p(\theta \mid \mathcal{D}, M) = \frac{p(\theta \mid M)}{p(\mathcal{D} \mid M)} \prod_{b=1}^{B} p(\mathcal{D}_b \mid \theta, M) \tag{6}$$

where $p(\mathcal{D}_b \mid \theta, M)$ is the likelihood for data $\mathcal{D}_b$. When the batch-level posterior $p(\theta \mid \mathcal{D}_b, M)$ is approximated by the Laplace approximation and the prior $p(\theta \mid M)$ is normally distributed, the combined posterior of equation (6) can be evaluated analytically (see ESI, Section 3 for details†). We use the mean of this posterior approximation to initialize one final optimization of the log-posterior given all the data. We take the resulting optimum as our final parameter estimate, which is used to make predictions for all experiments, including batches and long-time data excluded during inference (Fig. 1c, gray).

Bayesian model selection

Finally, we use Bayesian model selection to evaluate the (relative) probability of each candidate model based on the global posterior (6) conditioned on data from the combined batches. Using this multivariate normal approximation, we evaluate the model likelihood $p(\mathcal{D} \mid M)$ analytically by integrating equation (6) with respect to the parameters θ (see ESI, Section 3 for details†). Assuming that each model is equally probable a priori, the posterior probability for any model M_i is directly proportional to its likelihood $p(\mathcal{D} \mid M_i)$. The most probable model is that with the largest likelihood.

3 Results and Discussion

Posterior predictive checks (PPCs) show that the simplest models fail to describe the experimental data from one or more batches (Fig. 4a). The p-values computed for the base, angle, paramagnetic, and gravity models are all much smaller than one for at least one of the five batches considered. These models lack the descriptive power necessary to reproduce data from multiple experiments simultaneously. By contrast, composite models that combine multiple effects—for example, angle+gravity—pass the PPC for all batches and remain viable candidates for further criticism and comparison. We find that the effects of gravity on particle motion cannot be neglected—particularly, for low driving frequencies $\omega \ll Mg/6\pi\eta aR \sim 5$ rad/s. However, the addition of either a paramagnetic contribution to the particle's magnetic moment (paramagnetic) or a tilt to its orientation at the interface (angle) is sufficient to reproduce the experimental observations. In other words, there are three models that describe the data: angle+gravity, paramagnetic+gravity, and an-

gle+paramagnetic+gravity. It remains to determine which of these candidate models is favored over the others.

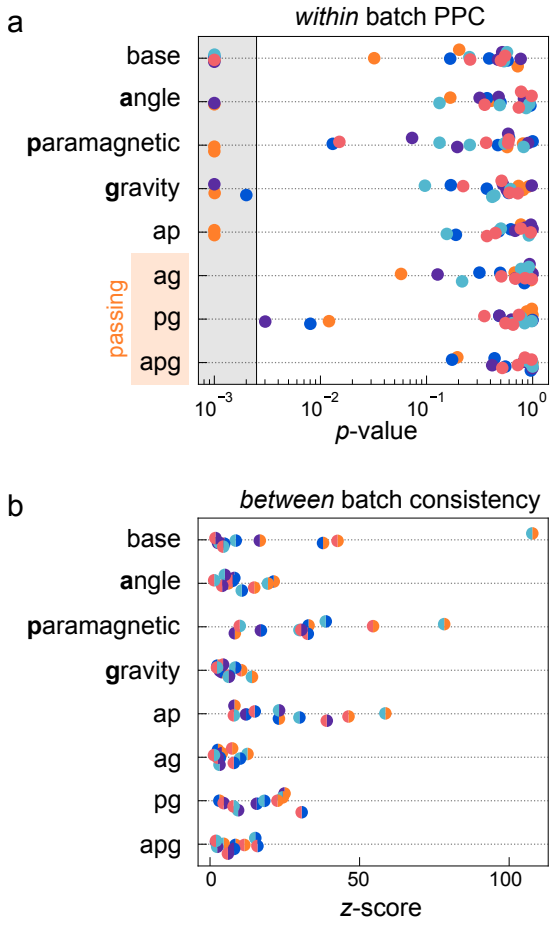

Fig. 4 (a) Results of posterior predictive checks (PPCs) for each of the eight models and five batches of four experiments. Each marker corresponds to one experiment; colors denote different batches. The vertical line shows the 5% significance threshold accounting for multiple comparisons—namely, $p = 0.05/20$.[28] The last three models pass the PPC with no p-values smaller than the threshold. (b) Between batch z-scores of equation (6) for each of the eight models and 20 batch-batch pairs. Colors denote different batch pairs.

Of the three candidate models that pass the PPCs, the angle+gravity (ag) model shows the highest degree of consistency between parameters inferred from different batches (Fig. 4b). The average z-score is 4.35 as compared to 14.6 and 6.75 for the paramagnetic+gravity (pg) and angle+paramagnetic+gravity (apg) models, respectively. In other words, by using the ag-model to analyze different subsets of the experimental data, we are led to similar conclusions about common parameter values such as the tilt angle μ or the magnetic moment m. Such consistency provides further support for this model as well as other aspects of the analysis such as the Laplace approximation and the choice of batch size. For the pg-model, the higher z-scores indicating lower levels of consistency between batches suggests that this model is less capable of describing the data from all experiments.

The apg-model shows less consistency between batches despite containing the ag-model as a limiting case (i.e., when $\Delta\alpha = 0$). This observation is indicative of overfitting whereby the added parameter—here, the polarizability difference $\Delta\alpha$—adopts different values from batch to batch as to reduce the batch-level error. In addition to inconsistent parameter estimates, this type of overfitting leads also to inaccurate predictions as detailed below. The chosen batch size of 4 experiments is sufficient to infer as many global parameters from the data (e.g., the moment m, angle μ, mass M, drop radius R of the ag-model). However, the analysis of more complicated models with more global parameters would likely benefit from larger batch sizes.

Based on cross-validation, the ag-model shows the smallest average prediction error (0.33 particle diameters) across all experiments including unseen data excluded during parameter estimation (Fig. 5a). Only the apg-model which contains the ag-model as a limiting case performs as well. Competing models such as pg and g exhibit nearly twice the error in the predicted particle trajectories. Figures 5b and 5c illustrate the average prediction errors for the ag- and pg-models as a function of the design variables ω and φ. The ag-model is characterized by small prediction errors over the entire design space, including unfamiliar conditions excluded during parameter estimation (gray markers). By contrast, the pg-model exhibits comparatively large errors for conditions near the transition region between the pole-orbiting and zig-zag behaviors. These results highlight the challenge of discriminating between competing models using limited data. Despite the different physics they contain, both the ag- and pg-models can provide accurate descriptions of the pole-orbiting and zig-zag behaviors. Only by considering all the data—in particular, the transition region between the two behaviors—do the deficiencies of the pg-model become clear.

Further support for the ag-model is provided by Bayesian model selection, which favors this model decisively over the other candidates considered. In particular, the ag-model is ca. 10^9 times more probable than the next best apg-model. The latter is judged to be less likely despite containing the favored ag-model as a limiting case. To understand this result, we decompose the model likelihood as the product of two components

$$p(D \mid M) = p(D \mid \hat{\theta}, M) \times \frac{p(\hat{\theta} \mid M)}{p(\hat{\theta} \mid D, M)} \tag{7}$$

The first term is simply the likelihood function evaluated at the most probable parameter values $\hat{\theta}$. The addition of more parameters tends to increase this quantity by enhancing the descriptive capacity of the model. The second term describes the ratio between prior and posterior densities evaluated at the parameter estimate $\hat{\theta}$. This quantity is typically less than one and becomes smaller still with the addition of superfluous parameters. Because it favors simpler models, the logarithm of this ratio is sometimes referred to as the Occam factor. In comparing the ag and apg models, the added parameter does not alter the log-likelihood, but it does decrease the Occam factor thereby favoring the simpler ag model (Tab. 1). By contrast, the pg model is strongly disfavored on the basis of both contributions.

To summarize, the ag-model provides an accurate, self-consistent description of all experimental data using physically

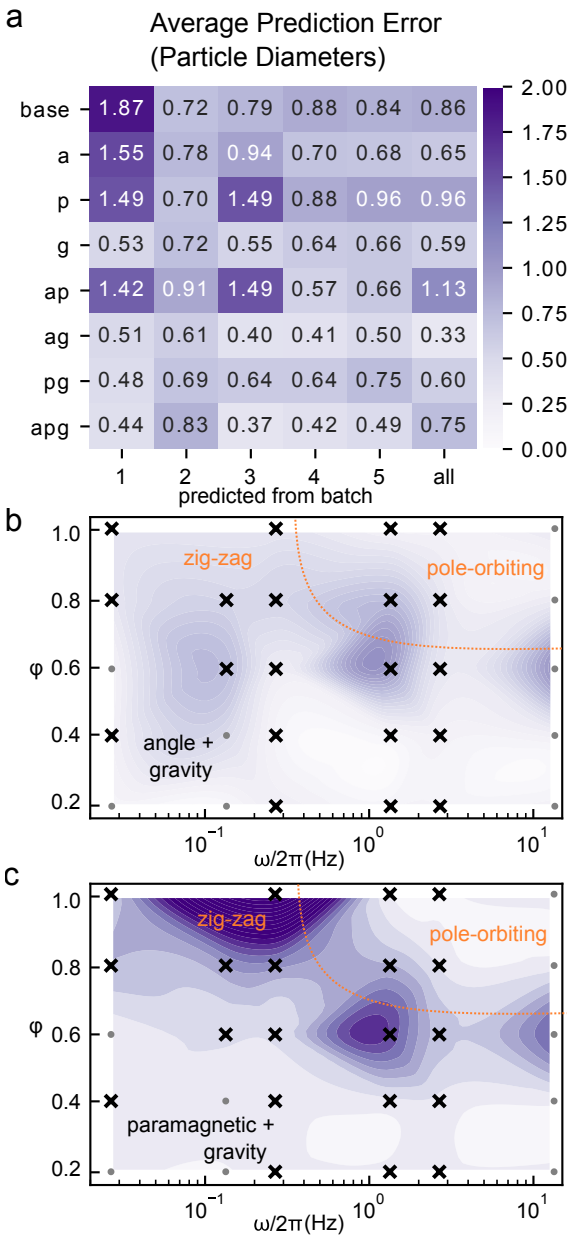

Fig. 5 Cross-validation. (a) Table of average prediction errors across all 30 experiments using each of the eight candidate models with parameters inferred from different batches (1 to 5) and their combination (all). The reported errors denote the average absolute error between the experimental data and noise-free model predictions using the most probable parameter estimates. (b,c) Prediction errors as a function of the frequency ω and angle φ of the precessing field for the (b) angle+gravity and (c) parametric+gravity models. Parameter values are inferred from the 20 experiments marked with black ×'s; gray circles denote experiments excluded during parameter estimation. The contour plot uses bilinear interpolation between experiments with the colormap from (a).

reasonable parameter estimates. The buoyant mass of the particle is inferred to be $M = 0.70 \times 10^{-11}$ g ($\pm 20\%$), as compared to the prior estimate of 1.2×10^{-11} g based on the stated size and composition of the Janus sphere. The inferred magnetic moment is $m = 2.2 \times 10^{-14}$ A m^2 ($\pm 10\%$), which is close a previous estimate of 2.9×10^{-14} A m^2 for similar particles.[15] Importantly,

model	$\ln p(M \mid \mathcal{D})$	$\ln p(M \mid \hat{\theta}, \mathcal{D})$	Occam factor
ag	0.0	104.2	−104.2
apg	−21.6	104.2	−125.8
pg	−139.4	31.4	−170.8

Table 1 Bayesian model selection. The top three candidate models are ranked by the logarithm of the model likelihood $\ln p(\mathcal{D} \mid M)$. This quantity is further discomposed into two additive contributions: log-likelihood $\ln p(\mathcal{D} \mid \hat{\theta}, M)$ evaluated at the most probable parameter estimates $\hat{\theta}$ and the Occam factor, defined as $\ln[p(\hat{\theta} \mid M)/p(\hat{\theta} \mid \mathcal{D}, M)]$. All quantities are shifted by a common factor such that the model log-likelihood is zero for the favored ag model.

these estimates for M and m depend on the accuracy of our estimate for the hydrodynamic resistance to translation $\lambda_t = 6\pi\eta a$ because the dynamics depends on the ratios m/λ_t and M/λ_t. Enhanced (diminished) resistance at the interface would lead us to underestimate (overestimate) the bouyant mass M and the magnetic moment m; however, the predictions of the model would remain unaffected. The moment is tilted relative to the interface by an angle $\mu = -8.8^\circ$ ($\pm 10\%$), where the negative value implies a contribution directed into the drop center. The origin of this tilt, however, remains unclear. In one interpretation, the moment is aligned parallel to the equator of the Janus sphere, and the particle itself tilts at the interface due to pinning of the three-phase contact line.[29] Alternatively, the Janus sphere may adsorb in its preferred orientation while its internal magnetic moment tilts with respect to the Janus equator.

The ag-model predicts new types of particle motion which are borne out by additional experiments on particles with large tilt angles. For tilted particles, the zig-zag trajectories found at lower precession angles and/or frequencies are displaced from the drop equator towards the lower (upper) pole when the moment is directed into (out from) the drop. While we cannot control the tilt angle of Janus spheres at the drop interface, additional experiments like those of Figure 1d revealed flower-shaped trajectories anticipated by the ag-model (Fig. 6). Such trajectories were not predicted previously for particles with zero tilt angle;[16] however, they are well explained by the present ag-model. Using same the bouyant mass and magnetic moment as estimated above, we infer a tilt angle of 51° directed into the interface for the data in Figure 6; the average error is 0.79 particle diameters. Such asymmetric, zig-zag trajectories provide further evidence for the ag-model, which was selected above on the basis of more subtle quantitative evidence. These trajectories are potentially useful for droplet mixing[16] and/or propulsion in time-varying fields; however, further work is needed to control the orientation of the particle's magnetic moment relative to the interface.

In addition to physical insights into magneto-capillary particle dynamics, the present analysis provides useful lessons in applying Bayesian methods to analyze time-series data using dynamical models. Posterior predictive checks (PPCs) are necessary to ensure that candidate models have the descriptive capacity to reproduce the data on which they are trained. Inferences made using models that fail such PPCs are unreliable and may lead to false conclusions. Model parameters should be inferred using data from multiple experiments simultaneously to increase the statisti-

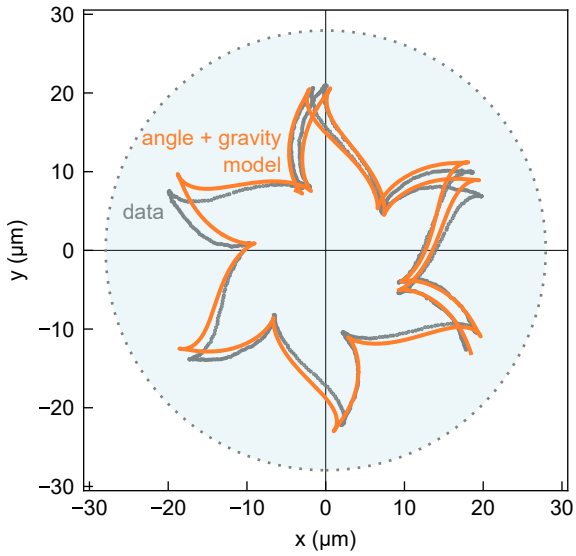

Fig. 6 Flower-shaped trajectory (gray) of a tilted particle in a precessing field with angular frequency, $\omega/2\pi = 0.73$ Hz, and precession angle, $\varphi = 0.4$ rad. The data is well described by the predictions (orange) of the ag-model using previous estimates for the magnetic moment and the buoyant mass. The estimated drop position and radius is shaded in blue.

cal power of model criticism and avoid overfitting. Grouping experiments into batches allows for scaling inference to larger data sets and enables further criticism based on cross-validation. The combination of efficient integration algorithms and automatic differentiation greatly accelerates the search for optimal parameter estimates that maximize the posterior probability. Given multiple interpretations of the experimental data, the Bayes model ratio provides a useful metric by which to discriminate between competing models. Finally, correlated noise models with user-specified parameters can provide useful approximations of missing physics that help to identify predictive descriptions that are good enough.[30]

4 Conclusions

We have developed and experimentally validated a dynamical model that accurately predicts the field-driven motion of a magnetic Janus sphere on the curved interface of a liquid drop. The particle dynamics is governed by a balance of magnetic, viscous, and gravitational forces/torques subject to rigid capillary constraints imposed by the interface. Through quantitative comparison of model predictions and experimental observations, we accurately infer unobserved quantities such as the magnitude and orientation of the particle's magnetic moment. The analysis of this specific problem offers general lessons for the analysis of time series data using dynamical models for the purpose of parameter estimation, model criticism, and model selection.

Looking forward, predictive models of magneto-capillary particle dynamics provide opportunities for designing future experiments that deepen our knowledge or enhance functional capabilities. By carefully choosing the experimental conditions, one can accurately infer model parameters using limited data from few experiments (see reference [31] for details). This process of Bayesian experimental design is guided by model predictions that anticipate the likely outcomes of hypothetical experiments and identify designs that advance a user-specified objective—for example, to maximize the information gain about uncertain model parameters.[32] In the present context, the different batches of experiments provide different amounts of information about the model parameters: as large as 8680 bits for batch 4 and as small as 676 bits for batch 3 (see ESI, Sec. 4†). Bayesian design guides the selection of better—that is, more informative—experiments. Ideally, this process can be automated within iterative cycles of observation, inference, and design that converge rapidly to accurate parameter estimates.[31,33] Alternatively, by changing the objective function, the same iterative process can be used to identify design conditions (e.g., the time-varying field) that achieve desired capabilities such as drop mixing and propulsion despite uncertainty in the model parameters. In this way, predictive models serve to accelerate the design of active colloids[34] and microrobots[35,36] with increasingly autonomous capabilities such as self-propulsion and self-guided navigation.[37]

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1935228.

Notes and references

1 W. Fei, Y. Gu and K. J. M. Bishop, *Curr. Opin. Colloid Interface Sci.*, 2017, **32**, 57–68.
2 F. Martínez-Pedrero, *Adv. Colloid Interface Sci.*, 2020, **284**, 102233.
3 J. Deng, M. Molaei, N. G. Chisholm, T. Yao, A. Read and K. J. Stebe, *Curr. Opin. Colloid Interface Sci.*, 2022, 101629.
4 Z. Lin, Z. Zhang, Y. Li and Y. Deng, *Chem. Eng. J.*, 2016, **288**, 305–311.
5 C.-Y. Xie, S.-X. Meng, L.-H. Xue, R.-X. Bai, X. Yang, Y. Wang, Z.-P. Qiu, B. P. Binks, T. Guo and T. Meng, *Langmuir*, 2017, **33**, 14139–14148.
6 S. Lam, E. Blanco, S. K. Smoukov, K. P. Velikov and O. D. Velev, *J. Am. Chem. Soc.*, 2011, **133**, 13856–13859.
7 E. Blanco, S. Lam, S. K. Smoukov, K. P. Velikov, S. A. Khan and O. D. Velev, *Langmuir*, 2013, **29**, 10019–10027.
8 E. Kim, K. Stratford and M. E. Cates, *Langmuir*, 2010, **26**, 7928–7936.
9 A. Snezhko, M. Belkin, I. S. Aranson and W.-K. Kwok, *Phys. Rev. Lett.*, 2009, **102**, 118103.
10 A. Snezhko and I. S. Aranson, *Nat. Mater.*, 2011, **10**, 698–703.
11 G. Lumay, N. Obara, F. Weyer and N. Vandewalle, *Soft Matter*, 2013, **9**, 2420–2425.
12 G. Grosjean, M. Hubert and N. Vandewalle, *Adv. Colloid Interface Sci.*, 2018, **255**, 84–93.
13 Y. Collard, G. Grosjean and N. Vandewalle, *Commun. Phys.*, 2020, **3**, 112.
14 Q. Xie, G. B. Davies and J. Harting, *ACS Nano*, 2017, **11**, 11232–11239.
15 W. Fei, M. M. Driscoll, P. M. Chaikin and K. J. M. Bishop, *Soft*

Matter, 2018, **14**, 4661–4665.
16 W. Fei, P. M. Tzelios and K. J. M. Bishop, *Langmuir*, 2020, **36**, 6977–6983.
17 A. Sukhov, M. Hubert, G. Grosjean, O. Trosman, S. Ziegler, Y. Collard, N. Vandewalle, A.-S. Smith and J. Harting, *Eur. Phys. J. E*, 2021, **44**, 59.
18 E. T. Jaynes, *Probability Theory: the Logic of Science*, Cambridge University Press, 2003.
19 D. Sivia and J. Skilling, *Data Analysis: a Bayesian Tutorial*, Oxford University Press, 2006.
20 A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin, *Bayesian Data Analysis*, CRC press, 3rd edn, 2013.
21 A. Gelman, X.-L. Meng and H. Stern, *Stat. Sin.*, 1996, 733–760.
22 D. M. Blei, *Annu. Rev. Stat. Appl.*, 2014, **1**, 203–232.
23 B. J. Park, T. Brugarolas and D. Lee, *Soft Matter*, 2011, **7**, 6413–6417.
24 J. Diebel, *Matrix*, 2006, **58**, 1–35.
25 C. Rackauckas and Q. Nie, *Journal of Open Research Software*, 2017, **5**, year.
26 A. Dörr, S. Hardt, H. Masoud and H. A. Stone, *J. Fluid Mech.*, 2016, **790**, 607–618.
27 J. Revels, M. Lubin and T. Papamarkou, *arXiv preprint arXiv:1607.07892*, 2016.
28 J. P. Shaffer, *Annu. Rev. Psychol.*, 1995, **46**, 561–584.
29 H.-M. Gao, Z.-Y. Lu, H. Liu, Z.-Y. Sun and L.-J. An, *The Journal of Chemical Physics*, 2014, **141**, 134907.
30 G. E. P. Box, *J. Am. Stat. Assoc.*, 1976, **71**, 791–799.
31 K. Dhatt-Gauthier, D. Livitz and K. J. M. Bishop, *Soft Matter*, 2021, **17**, 10128–10139.
32 D. V. Lindley, *Ann. Math. Stat.*, 1956, **27**, 986–1005.
33 T. J. Loredo, AIP Conference Proceedings, 2004, pp. 330–346.
34 K. J. M. Bishop, S. L. Biswal and B. Bharti, *Annu. Rev. Chem. Biomol. Eng.*, 2023, **14**, 1–30.
35 K. Dhatt-Gauthier, D. Livitz, Y. Wu and K. J. M. Bishop, *JACS Au*, 2023, **3**, 611–627.
36 J. G. Lee, R. R. Raj, N. B. Day and C. W. Shields, *ACS Nano*, 2023, **17**, 14196–14204.
37 Y. Dou, P. M. Tzelios, D. Livitz and K. J. M. Bishop, *Soft Matter*, 2021, **17**, 1538–1547.