Soft Matter

Automating Bayesian inference and design to quantify acoustic particle levitation

Journal:	*Soft Matter*
Manuscript ID	SM-ART-07-2021-001116.R1
Article Type:	Paper
Date Submitted by the Author:	23-Oct-2021
Complete List of Authors:	Dhatt-Gauthier, Kiran; Columbia University, Chemical Engineering Livitz, Dimitri; Columbia University, Chemical Engineering Bishop, Kyle; Columbia University, Chemical Engineering; Columbia University

Soft Matter

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Automating Bayesian inference and design to quantify acoustic particle levitation

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

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Self-propulsion of micro- and nanoparticles powered by ultrasound provides an attractive strategy for the remote manipulation of colloidal matter using biocompatible energy inputs. Quantitative understanding of particle motion and its dependence on size, shape, and composition requires accurate characterization of the acoustic field, which depends sensitively on the experimental setup. Here, we show how automated experiments based on Bayesian inference and design can accurately and efficiently characterize the acoustic field within resonant chambers used to propel acoustic nanomotors. Repeated cycles of observation, inference, and design (OID) are guided by a physical model that describes the rate at which levitating particles approach the nodal plane. Using video microscopy, we observe the relaxation of tracer particles to this plane following the application of the acoustic field. We use sequential Monte Carlo methods to infer model parameters such as the amplitude and frequency of the resonant chamber while accounting for particle-level measurement noise and population-level heterogeneity in the field. Guided by simulated outcomes, we select the optimal design for the next experiment as to maximize the information gain in the relevant parameters. We show how this iterative process serves to discriminate between competing hypotheses and efficiently converges to accurate parameter estimates using only few automated experiments. We discuss the need for model criticism to ensure the validity of the guiding model throughout automated cycles of observation, inference, and design. This work demonstrates how Bayesian methods can learn the parameters of nonlinear, hierarchical models used to describe video microscopy data of active colloids.

1 Introduction

Direct video imaging[1,2] of colloidal objects—from enzymes[3] and nanoparticles[4–6] to living cells[7] and rheological probes[8,9]—provides useful data with which to understand their properties,[6] dynamics,[3] interactions,[4,10] assemblies,[11] and microenvironments.[8] More than a century after Perrin's pioneering experiments,[12] the tracking of Brownian particles is now commonly used to quantify their diffusive motions, from which properties of the particle (e.g., size) and its local environment (e.g., fluid velocity, viscosity) are readily inferred. These methods are also useful in quantifying the propulsion mechanisms for a growing variety of active particles[13–16] powered by chemical fuels[17] and/or external fields.[18]

Whether active or passive, the analysis of particle motion relies on models that predict the observed data in terms of uncertain model parameters and user-specified design variables (Fig. 1a). To account for variability in experimental observations, these models are inherently probabilistic. They describe the odds of observing a particular outcome given certain assumptions about the validity of the model and the values of it parameters. Such probabilistic models may contain multiple sources of stochastic variation (i.e., "noise") due to thermal fluctuations, particle dispersity, heterogeneous environments, and measurement error among others. Faced with these many uncertainties, scientists conduct experiments in order to infer or "learn" model parameters from the observed data and also to evaluate model performance. Prior to each experiment, they tune the available design variables or "knobs" (e.g., the frequency of an applied signal) as to influence the measurement outcome. By carefully selecting these design variables, "good" scientists hope to learn model parameters efficiently using as few experiments as possible. The present work seeks to automate where possible these important processes of parameter estimation and experimental design.

The broader challenges of experimental design outlined above will be investigated here within the specific context of propelling micro- and nanoparticles using ultrasound.[20–24] The self-

Department of Chemical Engineering, Columbia University, New York, NY. E-mail: kyle.bishop@columbia.edu

† Electronic Supplementary Information (ESI) available: evidence supporting the linear relation between particle height and size; details on marginalization procedures used in inference and the estimation of the model evidence; discussion of the two-stage design process for hierarchical models; additional data from automated optimal experiments. See DOI: 10.1039/cXsm00000x.

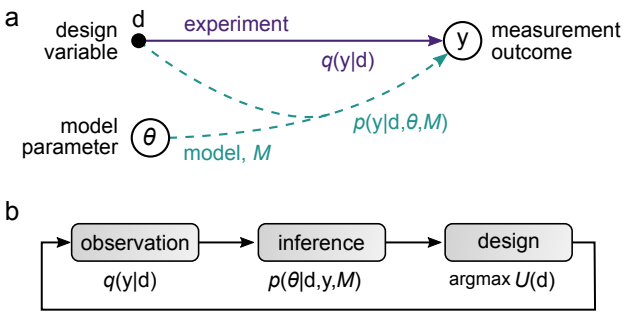

Fig. 1 Bayesian experimental design. (a) The relationship between an experimental design $\mathbf{d}$ and a measurement outcome $\mathbf{y}$ can be described by the conditional probability distribution $q(\mathbf{y}|\mathbf{d})$, which is unknown to the experimenter. The user-specified model M provides a parametric approximation to this relationship $q(\mathbf{y}|\mathbf{d}) \approx \int p(\mathbf{y}|\boldsymbol{\theta},\mathbf{d},M)p(\boldsymbol{\theta}|M)d\boldsymbol{\theta}$, where the distribution $p(\boldsymbol{\theta}|M)$ describes our uncertain knowledge of the model parameters $\boldsymbol{\theta}$, and the distribution $p(\mathbf{y}|\boldsymbol{\theta},\mathbf{d},M)$ describes the likelihood of the measurement outcome under model M with parameters $\boldsymbol{\theta}$. (b) Model parameters can be learned through iteration of a three-step cycle of observation, inference, and design (OID).[19] Upon conducting an experiment (observation), our knowledge of the model parameters is updated (inference) and used to design future experiments to maximize their expected utility (design).

propelled motions of solid particles levitating at the nodal plane of a standing acoustic wave are thought to derive from steady streaming flows[22,25] that depend on particle size,[22] shape,[21,22] and composition.[21] The propulsion speed further depends on the local acoustic field, which is often spatially heterogeneous due to imperfections in the resonant acoustic chamber. In order to understand the relationships between particle properties and their self-propelled motions, we first require quantitative knowledge of the acoustic field in which particles move. Resonant fields are often sensitive to small perturbations in experimental conditions, which can be difficult to control from one experiment to the next. It is therefore desirable to characterize the acoustic field efficiently and perhaps repeatedly during investigations aimed at quantifying the mechanisms of acoustic propulsion.

In this context, Bayesian data analysis[26,27] provides a principled approach for inferring model parameters and designing maximally informative experiments (Fig. 1b). Given a probabilistic model for the observed data, Bayes' rule describes how prior distributions for model parameters should be updated to infer their likely values conditioned on experiment outcomes. The trained model can then be used in designing future experiments as to maximize user-specified objectives. Commonly, one seeks designs that maximize the expected information gain about the model parameters of interest.[28] This cycle of observation, inference, and design (OID) can be repeated automatically to learn model parameters to a specified precision using the fewest number of experiments.[19] In practice, however, parameter estimation is complicated by intractable nonlinear models containing many unobserved variables. Even modern numerical techniques based on Hamiltonian Monte Carlo (HMC)[29]or variational inference (VI)[30] have difficulty in capturing multimodal distributions that can arise in nonlinear models of physicochemical systems. Bayesian design is even more computationally expensive to implement than inference, requiring expectations over all possible experiment outcomes.[31] Despite these challenges, thoughtful modeling choices combined with probabilistic programming tools (e.g., Stan,[32] PyMC3[33]) can enable the wide-spread application of automated experiments for parameter estimation using modest computational resources.

Here, we demonstrate an automated platform for quantifying the acoustic field within resonant chambers used in the study of micro- and nanomotors powered by ultrasound. A physical model of these chambers is characterized by four parameters that describe the amplitude, natural frequency, quality factor, and spatial heterogeneity of the resonant acoustic field. To infer these parameters, we observe the motion of colloidal tracer particles as they move to the nodal plane upon application of the field at a prescribed driving frequency. The data from video microscopy is analyzed using a full probability model that describes the noisy measurement of each particle within the spatially heterogeneous field. We use sequential Monte Carlo (SMC) methods[34] to sample parameters from the posterior distribution and simulate outcomes of future experiments. For the next experiment, we select the design—namely, the driving frequency and the video frame rate—that maximizes the expected information gain. The cycle of observation, inference, and design is implemented in an automated closed loop with computer-controlled actuation and data collection. We show that accurate parameter estimates can be achieved using a small number of carefully selected experiments. The methods developed here are immediately applicable to related problems in characterizing and controlling field-driven colloids.[35–39]

2 Observation-Inference-Design

2.1 Observation

Our experiments are performed in a cylindrical acoustic cell (height, $H = 220\ \mu$m; diameter, 4 mm) containing an aqueous dispersion of polystyrene (PS) spheres (radius, $R = 7.5\ \mu$m; Fig. 2a). The cell is actuated from below by a piezoelectric transducer subject to a sinusoidal voltage of magnitude V and angular frequency ω (see Appendix A for experimental details). Near the natural frequency, $\omega/2\pi \approx c/2H = 3.41$ MHz, a standing wave is created with wavelength equal to twice the cell height. Upon application of the acoustic field, the PS particles levitate from the chamber floor to the midplane of the cell due to acoustic radiation forces[40].

We quantify the magnitude of the radiation force by tracking the motion of particles as they come into focus at the nodal plane (Fig. 2b).[41] During each experiment, we first focus on a collection of particles levitating at the midplane of the cell (Fig. 2b, region i). The applied voltage is then switched off for a period of time $\tau \approx 7$ s, during which the particles sediment out of focus due to gravity (Fig. 2b, region ii). Reapplication of the acoustic field causes the particles to return to the nodal plane at a rate proportional to the magnitude of the radiation force (Fig. 2b, region iii). Using video microscopy, we quantify the apparent size y of each particle as it slowly increases during sedimentation and quickly recovers upon reapplication of the field (Fig. 2c; see Appendix

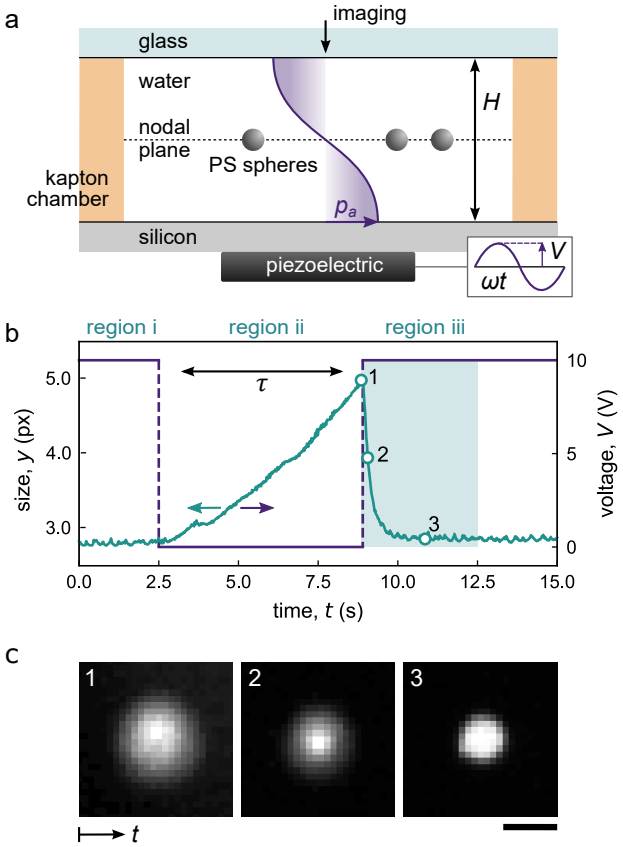

Fig. 2 Observation. (a) Schematic illustration of the acoustic cell containing polystyrene (PS) tracer spheres. (b) During each experiment, the voltage signal is switched off for a time τ and then reapplied (purple curve, right axis). The apparent size y of each particle increases steadily as it falls out of focus and then recovers rapidly up reapplication of the acoustic field (aqua curve, left axis). (c) Image sequence showing a single particle as it moves into focus at the nodal plane. The corresponding particle size in each image is plotted in (b) (open circles). The scale bar is 15 μm.

A).

For each experiment, the design variable **d** specifies the magnitude V and frequency ω of the applied signal, the sedimentation time τ, as well as the number N_t and frame rate f of images captured upon reapplication of the field. These features of the experiment—unlike the number of PS tracer particles, for example—are easily controlled by a computer and therefore useful in guiding sequences of automated experiments. For a given design, we use the captured images to generate the data **y** representing the size of the N_p tracked particles at the N_t time points. From the observed relationship between design **d** and data **y**, we infer the resonant acoustic field within the cell using a minimal number of automated experiments. To reduce the dimensionality of the design space, we fix the applied voltage to $V = 10$ V, the sedimentation time to $\tau = 7$ s, and the number of time points to $N_t = 100$. The driving frequency ω and frame rate f are selected for each experiment so as to maximize the expected information gain with respect to the model parameters as detailed below.

2.2 Inference

Generative Model

Bayesian inference and design requires a probabilistic model for the observed data **y** conditioned on the experimental design **d**. Here, the physical model describes the physics of acoustic particle levitation and its dependence on the resonant acoustic field within the cell. The motion of the PS spheres through the viscous fluid is approximated by the overdamped dynamics

$$6\pi\eta R\frac{dh}{dt} = F_g + F_a \tag{1}$$

where h is the height of the particle from the floor, η is the fluid viscosity, F_g is the gravitational force, and F_a is the acoustic radiation force on the sphere. In the absence of the acoustic field (Fig. 2b, region ii), the particle sediments at a constant velocity (here, ca. 6 μm/s) determined by the balance of the viscous drag and the gravitational force. In the acoustic field (Fig. 2b, region iii), the particle experiences an additional radiation force of the form

$$F_a = \frac{\pi^2\Phi R^3 p_a^2}{\rho c^2 H}\sin\left(\frac{2\pi h}{H}\right) \tag{2}$$

where p_a is the amplitude of the acoustic pressure wave, $\rho = 998$ kg/m^3 and $c = 1480$ m/s are, respectively, the density and speed of sound in water, and $\Phi = 0.11$ is the acoustophoretic contrast factor for PS in water.[40] Near the midplane, equation (2) is well approximated by a linear force-displacement relationship like that of an elastic spring. The linearized dynamics can be integrated to obtain the following approximation for the transient height of the particle following the application of the acoustic field

$$h(t) = h_\infty + (h_0 - h_\infty)e^{-\lambda t} \tag{3}$$

where h_0 is the initial height at time zero, and $\lambda = \pi^2\Phi R^2 p_a^2/3\eta\rho c^2 H^2$ is the rate at which the particle approaches its asymptotic height h_∞. In our experiments, the magnitude of the acoustic force is typically much larger than the gravitational force, and the asymptotic height of the particle is indistinguishable from that of the nodal plane, $h_\infty \approx H/2$.

To describe the experimental data, we must further specify how the observed size y depends on the height of the particle h. For small particle displacements (i.e., $\Delta h \approx R$), the size y can be approximated by a linear function of the height h as confirmed by control experiments on sedimenting particles (see Supplementary Figure S2†). With these assumptions, the model predicts that the size evolves in time as

$$y(t) = a + be^{-\lambda t} \tag{4}$$

where the parameters a and b depend on the initial and asymptotic height of the particle and on the approximate linear relationship between height and size. These quantities do not depend on the characteristics of the acoustic field—namely, the pressure p_a—and their values are of little interest. We therefore treat them as nuisance parameters which are later discarded by marginalization.[26]

The rate of particle motion λ depends on the frequency ω of

the driving signal by way of the acoustic pressure p_a, which determines the magnitude of the radiation force. Near the natural frequency, we approximate this relationship as

$$\lambda(\omega) = \frac{\alpha\omega^2}{\beta^2 + (\omega-\beta)^2/\gamma^2} \quad (5)$$

where the parameters α, β, and γ characterize the magnitude, peak location, and width of the frequency response, respectively.[42] This expression is appropriate when the quality factor of the resonance is high (i.e., $Q = 1/2\gamma \gg 1$) and when the applied frequency ω is close to the resonant value β. Together, equations (4) and (5) describe how the observed size evolves in time as specified by particle-level parameters a, b and by cell-level parameters α, β, γ.

To describe the deviations between our experimental observations and the deterministic model outlined above, we consider two sources of stochastic variation or "noise". First, at the level of the individual particles, we assume that our measurements of the size are subject to additive white Gaussian noise with standard deviation $s \approx 0.1$ pixels. Second, at the level of the acoustic cell, we note that the acoustic field exhibits spatial variations within the nodal plane—so-called nodal structure—due to excitation of acoustic modes directed normal to the chamber thickness. As a result, the rate of particle motion λ can differ between particles at different locations in the nodal plane. As this rate is strictly positive, we assume that logarithm of λ is subject to additive Gaussian noise with standard deviation σ. For simplicity, the noise parameters s and σ are assumed to be constant and independent of design variables such as the applied frequency.

With these preliminaries, we can now write the full probability model to describe the observed data $\mathbf{y} = \{y_{ijk}\}$ for N_e experiments containing N_p particles at N_t time points (see Appendix B for details). Figure 3a shows a graphical representation of the model and the relationships among the different parameters and the design variables. We are primarily interested in the cell-level parameters—denoted $\theta = \{\alpha, \beta, \gamma, \sigma\}$—that describe the magnitude, peak location, and width of the frequency response as well as the magnitude of any nodal structure. Knowledge of these parameters along with the design variables $\mathbf{d}_i$ for experiment i determine the probability distribution for the rate parameter λ_{ij} for particle j. Other particle-level parameters—denoted $\phi_{ij} = \{a_{ij}, b_{ij}\}$—influence the distribution of the observed data y_{ijk} at time point k; however, their values are not of immediate interest.

Parameter Estimation

Given the observed data $\mathbf{y}$, we use Bayesian inference to estimate the posterior probability distribution for the cell-level parameters θ—namely, the magnitude α, peak location β, width γ, and heterogeneity σ of the acoustic resonance. Following each experiment, we use sequential Monte Carlo (SMC) methods[34] to sample parameter values from the posterior distribution, $p(\theta \mid \mathbf{y}, \mathbf{d})$ (see Appendix B). The required likelihood function $p(\mathbf{y} \mid \theta, \mathbf{d})$ is obtained by analytical marginalization of the latent rate parameters λ_{ij} and the nuisance parameters ϕ_{ij} for each experiment i and particle j (see ESI†).

Figures 3b-d show the analysis of particle tracking data from two experiments collected at two different driving frequencies ω. For each frequency, the analysis of the $N_p = \{18, 15\}$ particles provides as many estimates of the rate parameter λ (Fig. 3b); variations in this parameter from particle to particle reflect the spatial heterogeneity in the acoustic field. Conditioned on the data from these experiments, the posterior distribution for the cell parameters is bimodal suggesting two qualitatively different interpretations: one in which the natural frequency β lies between the two driving frequencies, and another in which it lies above them. These two modes are clearly visible in the marginal distribution for the resonance amplitude α and natural frequency β (Fig. 3c). We emphasize that the posterior is not analytically tractable owing to the nonlinear dependence of the rate λ on the cell-level parameters. Figure 3d shows the predicted resonant response $\lambda(\omega)$ of equation (5) using parameters sampled from the posterior (purple curves). These curves illustrate the two competing hypotheses for the natural frequency, both of which are compatible with data from the two experiments (markers). In designing the next experiment, we seek to discriminate between these competing scenarios as to eliminate one possibility or the other with high confidence.

To quantify the intuitive notion that some experiments are better than others, we define a utility function $u(\mathbf{y}, \mathbf{d})$ that depends on the observed data $\mathbf{y}$ collected under design $\mathbf{d}$. The distinction between "better" and "worse" experiments is necessarily subjective and reflects the objectives of the experimenter (e.g., to maximize knowledge, minimize cost, etc.). Here, we equate the utility of the data to the information gained with respect to the cell-level parameters θ

$$u(\mathbf{y}, \mathbf{d}) = \int p(\theta \mid \mathbf{y}, \mathbf{d}) \ln\left[\frac{p(\theta \mid \mathbf{y}, \mathbf{d})}{p(\theta)}\right] d\theta \quad (6)$$

This utility function is equal to the Kullback–Leibler divergence between the posterior and prior distributions for the model parameters and measures the amount of information provided by the data in reducing our uncertainty in these quantities.[28] Better experiments provide more information (here, measured in nats) about the cell-level parameters of interest. In practice, we estimate the utility using Monte Carlo integration with respect to the sampled posterior distribution (see Appendix B).[43] For example, the two experiments shown in Figure 3d provide 6.58 nats (9.49 bits) of information about the cell parameters. To interpret this information theoretic result, we can imagine dividing the space of possible parameter values into $e^{6.58} \approx 720$ regions with equal prior probabilities for enclosing the "true" value. The information provided by these two experiments is analogous to that of discovering which one of the 720 regions contains the true value.

2.3 Design

Given data $\mathbf{y}$ from previous experiments collected under designs $\mathbf{d}$, the design $\tilde{\mathbf{d}}$ for the next experiment is selected so as to maximize the expected utility of future outcomes $\tilde{\mathbf{y}}$. The expected utility $U(\tilde{\mathbf{d}})$ is obtained by integrating over possible outcomes $\tilde{\mathbf{y}}$

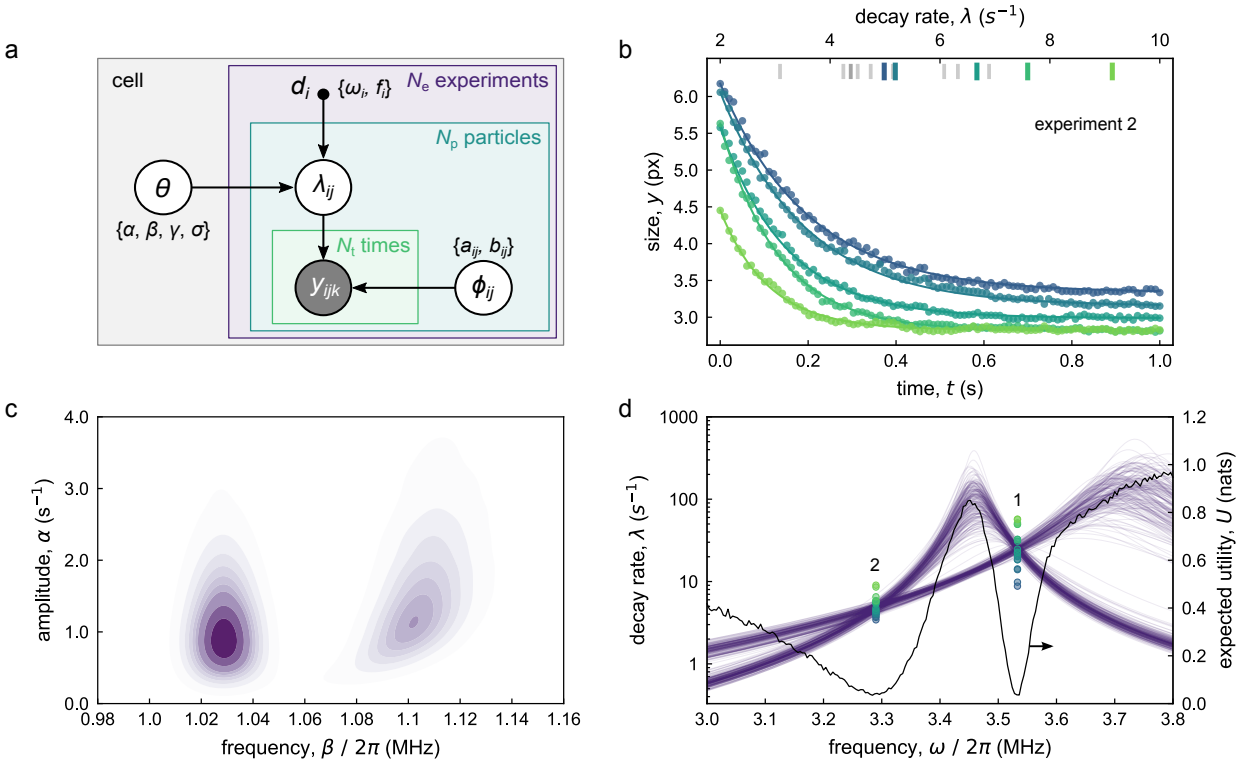

Fig. 3 Inference & Design. (a) Graphical representation of the full probability model for acoustic levitation experiments. Circles denote random variables with observed variables shaded grey. Colored boxes—so-called plates—indicate replication for $i = 1 \ldots N_e$ experiments, $j = 1 \ldots N_p$ particles, and $k = 1 \ldots N_t$ time points. (b) During a particular experiment $i = 2$, the observed sizes of the $N_p = 15$ particles decreases exponentially with different rates λ_{ij} for $j = 1, \ldots, N_p$ (rectangular markers). Circular markers denote experimental observations for five selected particles; solid curves denote model predictions using the maximum likelihood parameter estimates for λ_{ij} and ϕ_{ij} (see Appendix B). (c) Posterior distribution for the resonance amplitude α and the natural frequency β conditioned on data from two experiments. (d) Predicted rate parameter $\lambda(\omega)$ as a function of frequency ω based on data from two experiments. Each purple curve is evaluated using equation (5) with parameter values α, β, γ drawn from the posterior distribution in (c). Markers show estimates of the rate parameter λ_{ij} for each experiment i and particle j. The solid black curve represents the expected utility U of a subsequent experiment as function of the driving frequency; here, the optimal design maximizing U occurs at the right boundary of the frequency window.

weighted by the predictive distribution $p(\tilde{\mathbf{y}} \mid \tilde{\mathbf{d}}, \mathbf{y}, \mathbf{d})$ conditioned on the previous data

$$U(\tilde{\mathbf{d}}) = \int u(\tilde{\mathbf{y}}, \tilde{\mathbf{d}}) p(\tilde{\mathbf{y}} \mid \tilde{\mathbf{d}}, \mathbf{y}, \mathbf{d}) d\tilde{\mathbf{y}} \tag{7}$$

In other words, $U(\tilde{\mathbf{d}})$ describes the average utility of simulated outcomes conditioned on actual outcomes of past experiments. The optimal experimental design $\mathbf{d}^\star$ is that which maximizes this quantity: $\mathbf{d}^\star = \arg\max U(\tilde{\mathbf{d}})$. Using the utility of equation (6), the expected utility is equal to the mutual information between the cell-level parameters θ and the experimental outcome $\tilde{\mathbf{y}}$. With this choice, optimal experimental designs serve to maximize the amount of information shared between the parameters and the data. As a result, knowledge of the experimental outcome $\tilde{\mathbf{y}}$ acts reduce our uncertainty in the model parameters θ.

Computing the expected utility is arguably the biggest challenge in the practical implementation of Bayesian designs based on analytically intractable models. In addition to the integral over possible outcomes $\tilde{\mathbf{y}}$, equation (7) requires the evaluation of additional integrals over the parameters θ to compute both the utility and the posterior predictive distribution. As detailed in Appendix C, we use simulated experiments along with Monte Carlo importance sampling to estimate the expected utility of each candidate design $\tilde{\mathbf{d}}$.[31,44] To mitigate additional challenges due the hierarchical model, we decompose the optimization problem into two steps. We first identify the optimal frequency $\omega^\star$ neglecting the uncertainty in the particle rates $\tilde{\lambda}_{ij}$; we then optimize the frame rate $f^\star$ as to minimize that uncertainty. Details of this multistage approximation and its accuracy are discussed in Appendix C.

Continuing the example above, Figure 3d shows the expected information gain U as a function of the driving frequency ω conditioned on data from two experiments. The expected information gain is smallest at the two driving frequencies investigated previously and larger at unexplored frequencies reaching a maximum value at the upper bound of the frequency window. This optimal design maximizing U corresponds to the frequency at which the model predictions are most uncertain.[19] In light of the two hypotheses implicit in the bimodal posterior (Fig. 3c), the design algorithm selects the frequency at which the respective hypotheses make the most divergent predictions. Upon conducting the recommended experiment, the processes of observation, inference, and design can be repeated in an iterative fashion to efficiently

learn model parameters as we now show.

3 Results & Discussion

3.1 Validation of the Physical Model

The Bayesian approach to experimental design outlined in the previous section assumes that the experimental data is in fact generated by the proposed model. While it accounts for uncertainty in the model parameters, the algorithm does not question the validity of the model itself (unless, of course, alternatives are explicitly provided). It is therefore imperative that the physical model provide a reasonably effective approximation to the real data generating process (i.e., the experiments) if we are to trust in and benefit from the recommended designs. To evaluate the ability of the proposed model to describe the experimental data, we performed a series of experiments on a set of $N_p \approx 10$ particles at 14 frequencies equally spaced from 3.74 to 4.00 MHz using the maximum frame rate of 1000 fps (Fig. 4). From these data, we inferred the model parameters and compared the model predictions with the experimental observations.

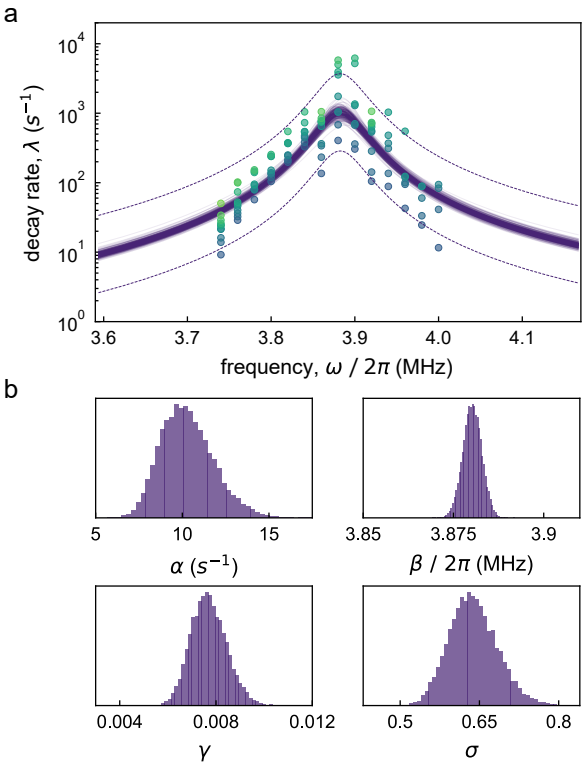

Fig. 4 Model Validation. (a) Experimental estimates of the rate parameter λ (markers) for a set of $N_p \approx 10$ particles at different frequencies ω agree with the posterior predictive distribution $p(\lambda \mid \mathbf{y}, \mathbf{d})$ conditioned on the experimental data $\mathbf{y}$ (curves). Solid curves denote the noise-free resonant response $\lambda(\omega)$ sampled from the posterior; the dashed curves denote 95% credible intervals for the rate λ. The $N_e = 14$ experiments were conducted at regular intervals of 20 kHz using a constant frame rate of 1000 fps. Agreement between the experimental data and the model predictions is strong at low frequencies (conducted earlier) but diminishes at high frequencies (conducted later) due to particle aggregation and cell aging. (b) Marginal distributions for the four cell-level parameters—the resonance amplitude α, natural frequency β, width γ, and heterogeneity σ—conditioned on the experimental data in (a).

Figure 4a shows the posterior predictive distribution for the rate parameter λ (curves) and the values inferred from the experimental data (markers) as a function of the driving frequency ω. The fact that the experimental data fall within the high probability regions of the fitted model indicates that the latter provides an effective and concise description of the former. Consistent with the inferred value for σ, the rate parameters λ show variation spanning from ca. three times less to three times more than the median value predicted by equation (5). Notably, data collected at latter times after sustained activation of the cell showed higher than expected variation in the inferred rate parameter λ (see the high frequency region of Fig. 4a). With repeated activation of the cell, the levitated particles accumulate within secondary pressure nodes within the nodal plane thereby reducing the number of free particles available for analysis. Other forms of cell aging—for example, solvent evaporation or mechanical relaxation of cell components—may also contribute to a decrease in data quality with sustained use. Below, we discuss how anomalous experimental results inconsistent with model expectations (i.e., "outliers") can be detected using posterior predictive checks (PPCs), thereby alerting the experimenter to possible failures of the model.

Figure 4b shows the marginal distributions for the cell-level parameters conditioned on the experimental data. The parameters are all estimated to within 15% of their mean values; however, some are inferred with considerably greater precision. In particular, the natural frequency β of the acoustic cell is estimated to be 3.8802 ± 0.0026 MHz; the width of the resonance is $\gamma = (7.71 \pm 0.74) \times 10^{-3}$, confirming the assumption $\gamma \ll 1$ that underlies equation (5). By contrast, the amplitude $\alpha = 10.3 \pm 1.5$ s^{-1} of the resonant response is considerably less certain, owing to spatial heterogeneity in the pressure field characterized by the noise parameter $\sigma = 0.639 \pm 0.046$. Overall, this set of 14 experiments provides a total information gain of 10.8 nats (15.6 bits) in the cell-level parameters with respect to the weakly informative priors detailed in Appendix B. Further experiments provide little additional information about the parameters owing to the heterogeneity in the acoustic field.

We emphasize that the experimental data used to characterize the acoustic field and validate our probabilistic model are unique to the particular cell on which they are collected. Despite efforts to fabricate cells with reproducible characteristics, uncontrolled variations—for example, in the position of the glass reflector slide—lead to differences in the cell parameters and thereby differences in the acoustic field for a common driving voltage and frequency. This variability motivates the need for automated experiments with which to quickly learn model parameters and characterize acoustic cells.

3.2 Automated Cycles of Observation-Inference-Design

Starting from a state of uncertainty, the cycle of observation-inference-design can be automated to accurately infer model parameters using a relatively small number of experiments selected algorithmically. Here, our initial state of knowledge—as quantified by our choice of priors for the model parameters—is one of significant uncertainty but not complete ignorance. The charac-

teristic magnitudes of the parameters are known (e.g., $\beta/2\pi \approx c/2H = 3.4$ MHz for the natural frequency); however, their precise values may be ca. three times smaller or larger depending on the particular acoustic cell in question (see Appendix B).

We consider the design space $\mathbf{d}$ specified by the driving frequency ω and the video frame rate f, holding all other variables constant. The driving frequency is constrained to lie on a finite range $\omega_{\min} < \omega < \omega_{\max}$ with $\omega_{\min} = 3.4$ MHz and $\omega_{\max} = 4.0$ MHz for the present experiments. Outside of this range, acoustic radiation forces are too weak to levitate particles to the nodal plane during the observation time. The frame rate f is discretized into five possible values spaced evenly on a logarithmic scale from 100 to 1000 fps. This small set of discrete options helps to accelerate the design process while capturing the range of rate parameters λ encountered in the experiments. During each iteration of the cycle, we select the design $\mathbf{d}$ that maximizes the expected information gain U as estimated by the model conditioned on previously collected data.

Figure 5 illustrates the first three iterations of the automated OID cycle. Following the first experiment, analysis of the observed data serves to reduce our uncertainty in the resonant response $\lambda(\omega)$ at the prescribed frequency but less so at higher or lower frequencies (Fig. 5a, purple curves). The expected information gain U for the next experiment is greatest at these previously unexplored frequencies (Fig. 5a, black curve). We note that the optimal frequency identified will differ somewhat from the true value due to noise inherent in the Monte Carlo estimates for the expected information gain. Such errors can be reduced as necessary by increasing the number of simulated experiments at the expense of added computation. After accepting the recommended frequency (Fig. 5a, vertical line), we select the optimal frame rate with which to capture the anticipated particle motion in the next experiment (see Appendix C).

Following the second experiment, the posterior predictions of the resonant response $\lambda(\omega)$ exhibit the same bimodality encountered above in our discussion of inference (cf. Figs. 5b & 3d). The optimal frequency for the subsequent experiment is selected at the lower frequency limit $\omega_{\min}$ to best discriminate between the two competing hypotheses implicit in the bimodal posterior. After performing this third experiment, the posterior distribution for the cell-level parameters converges on a single mode; however, the precise values of the magnitude, natural frequency, and quality of the resonance remain uncertain (Fig. 5c). The designs for subsequent experiments are selected as to systematically reduce the remaining uncertainty in the model parameters despite significant noise in the observed particle dynamics.

Figure 6a shows the posterior prediction for the resonant response $\lambda(\omega)$ conditioned on data from an automated sequence of ten optimal experiments (purple curves; see also Fig. S8). The estimated decay rates for the observed particles (colored markers) are scattered above and below the median response due to heterogeneities in the acoustic field (dashed curves). The expected information gain for further experiments is small and approximately independent of the applied frequency frequency (black curve). Figure 6b shows how the cumulative information gain relative to the prior increases with successive experiments (purple markers). Here, the information gain $u(\mathbf{y}, \mathbf{d})$ of equation (6) is computed using posterior parameter samples conditioned on all data $\mathbf{y}$ collected prior to and including the n^{th} experiment.[43] Notably, the information gain computed after each experiment agrees favorably with the predictions of simulated experiments used to guide experimental design (grey distributions).

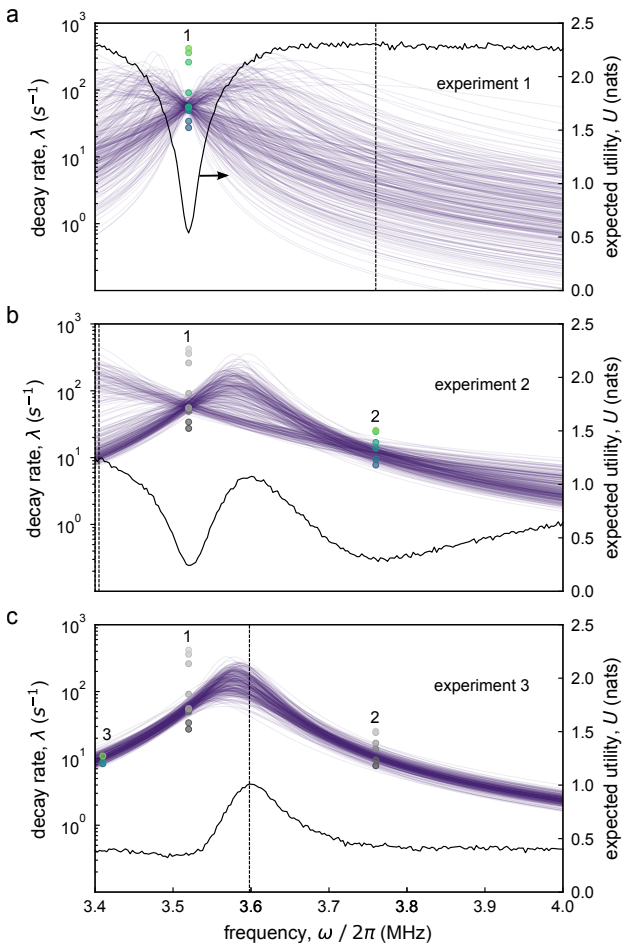

Fig. 5 Three automated experiments illustrating the observation-inference-design cycle. In each plot, the circular markers denote the experimental estimates for the rate parameters λ_{ij} for each particle j in experiment i; colored markers denote the results of the current experiment (Observation). The purple curves are posterior samples of the noise-free resonant response $\lambda(\omega)$ conditioned on the experimental data (Inference). The black curves show the expected information gain U of a subsequent experiment as a function of the driving frequency ω; the vertical line shows the optimal frequency $\omega^\star$ that maximizes U (Design). See Supplementary Figures S7 and S8 for the full data set.†

Together, these results demonstrate that the cell-level parameters can be learned using only five optimal experiments within the automated OID cycle. Here, these few experiments provide ca. 10 nats of information about the parameters, which is comparable to that provided by the more comprehensive data set illustrated in Figure 4. Following these initial experiments, the posterior distribution becomes unimodal and can be approximated by a multivariate normal distribution. Further experiments aimed at reducing parameter uncertainty can then make use of simplifying approximations based on linearized models to accelerate infer-

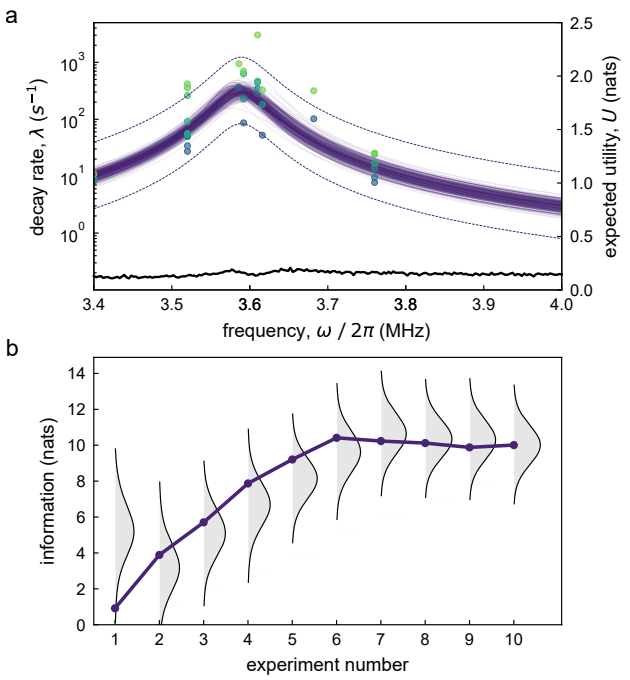

Fig. 6 Results of ten optimal experiments. (a) Rate parameter λ as function of the driving frequency ω (left axis). Markers represent maximum likelihood estimates of λ_{ij} for each particle j and experiment i; purple curves denote the noise-free resonance $\lambda(\omega)$ sampled from the posterior distribution; the dashed curves show the 95% credible interval for $\lambda(\omega)$. The black curve shows the expected information gain U for a subsequent experiment (right axis). (b) Cumulative information gain as a function of the experiment number n. Purple markers denote the information $u(\mathbf{y},\mathbf{d})$ in the cell parameters provided by the observed data $\mathbf{y}$ relative to the prior distribution. The grey distributions denote 99% credible intervals for the information gain predicted prior to the n^{th} experiment using simulated experiments; the expected value of this distribution U is use in designing the n^{th} experiment.

ence (e.g., Laplace approximation[45]) and design (e.g., D-optimal design[46]).

Throughout the automated cycle, we repeatedly assess the ability of the fitted model to describe the observed data and provide warnings of possible failures in data collection or analysis. We perform posterior predictive checks (PPCs) at both the particle and cell levels comparing experimental observations to stimulated data from the fitted model. At the particle level, we assess the exponential decay in the particle size (see Fig. 3b) using a chi-squared test for the sum-squared-error between the observed data and the predictions of the fitted model. Particles that fail this test (p-value $< 5\%$) are excluded from our analysis. At the cell level, we compare the estimated decay rate for each particle to posterior predictions of the fitted model and generate warnings of potentially anomalous behaviors (p-value $< 1\%$). When significant deviations between the observed data and the model predictions arise, the fitted model may lead to false conclusions, and it cannot be relied upon to design effective experiments. For this reason, some form of model criticism[47] is essential to the reliable automation and execution of the observation-inference-design cycle.

4 Conclusions

We demonstrate how automated experiments guided by physical models and Bayesian algorithms can accurately characterize resonant acoustic chambers using only few measurements. The hierarchical model developed here accounts for uncertainties at the particle level due to measurement noise and at the population level due to heterogeneities in the resonant field. Such models arise naturally in the study of active colloids by video imaging, where noisy measurements of multiple particles must be incorporated to draw quantitative conclusions about system-level properties or dynamics. We show how Bayesian inference using sequential Monte Carlo sampling can be used to describe multimodal distributions for unknown parameters that arise from nonlinear models applied to limited data. In general, the different modes of the posterior distribution describe the different qualitatively distinct hypotheses by which to explain the current data. We show how Bayesian designs aimed at maximizing the information gain can identify experimental conditions that best discriminate between such competing hypotheses.

While the present example is comparatively simple, the ability to entertain multiple hypotheses and to systematically evaluate them through an iterative process of experimentation is an essential component of the scientific method. Given a strong model for the relevant phenomena, computational methods of Bayesian inference and design can be applied algorithmically and automatically to learn model parameters using a minimal number of carefully selected experiments. Throughout this process, one must be careful to ensure that experimental observations are consistent with the limited "imagination" of the prescribed model. To this end, posterior predictive checks (PPCs) comparing experimental and simulated data can help to maintain confidence in model predictions and the resulting designs.

In the context of acoustic nanomotors, the methods developed here for *learning* model parameters can be extended to *engineering* particle propulsion directed by shape.[22] Consider the problem of identifying which particle shape satisfies—or perhaps even optimizes—a given functional objective (e.g., swim fast in circles). One approach to solve this problem relies on predictive models to first simulate particle motions for different candidate shapes and then select the best one. Unfortunately, existing models of acoustic propulsion[22,25,48] are inaccurate and/or computationally demanding thereby limiting their utility for engineering particle motions. Alternatively, one can use the observation-inference-design cycle to conduct experiments on different particle shapes in order to learn and validate simpler parametric models of shape-directed propulsion. The primary purpose of such models is to engineer particle performance rather than to understand acoustic propulsion. This difference in motivation should be reflected in the utility function used in the OID cycle as to balance the competing demands of "exploring" new designs and "exploiting" promising candidates. Bayesian optimization[49] based on surrogate models such as Gaussian processes[50] combined with acquisition functions such as expected improvement[51] provides one popular implementation of this general framework of observation, inference, and design.[19]

Looking forward, there remain significant opportunities for accelerating inference and design involving multimodal distributions.[52] In the present implementation, each OID cycle requires an average time of 1, 2, and 1 minutes, respectively, on an Intel i5-8400 processor; the code used to collect the present results is available on Github.[53] The extension of these methods to models containing more unknown parameters and/or design variables is likely to require more computational resources, additional simplifying approximations, and/or new methods for inference and design. In closing, we note that the methods described here rely on a strong guiding model, which incorporates knowledge of the relevant physics to create an efficient approximation of the data generating process (i.e., one with few parameters). Alternative approaches based on black-box models (e.g., neural networks) are unlikely to be competitive in the "small data" regime described here. Future applications of Bayesian inference and design are likely to benefit from hybrid methods[54] that incorporate physical knowledge where available while retaining the flexibility to describe nonlinear relationships between experimental designs and outcomes.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported as part of the Center for Bio-Inspired Energy Science, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award DE-SC0000989.

Appendix A: Experimental Details

Acoustic Cell

The acoustic cell was constructed from a silicon wafer covered with several layers of Kapton tape (CS Hyde Co. 18-1S-W) containing a cylindrical hole of height 220 μm and diameter 4.0 mm. A piezoelectric transducer (Steiner & Martins, Inc. SMD12T06R412WL) with a natural frequency of 3.4 MHz $\pm$ 5% was was attached to the opposite side of the wafer with epoxy. Prior to the experiments, the chamber was filled with a dilute dispersion of 15 μm fluorescent polystyrene (PS) spheres (Thermo-Scientific, FluoSpheres F8844) in deionized water. The chamber was capped by a glass coverslip and pressed firmly to remove excess liquid.

Automated Experiments

A sinusoidal voltage $V(t)$ with an amplitude of 10 V and a prescribed frequency ω was applied to the piezoelectric transducer using a function generator (Keithley 3390). Near the natural frequency, the applied signal caused the particles to levitate near the mid-plane of the chamber as observed by fluorescence microscopy (Nikon Plan Fluor 10$\times$ objective). The focal plane was positioned 15 μm above the nodal plane to achieve the desired linear relationship between particle size and height (Fig. S2). During each experiment, the signal was switched off for $\tau = 7$ s and then reapplied. Particle motions following the reapplication of the acoustic field were captured by a high speed camera (Phantom v710) at a frame rate of 1000 fps, which was later downsampled to the selected frame rate f prior to particle tracking. The number of frames analyzed during tracking was held constant at $N_t = 100$. Computer automation of the function generator and the high speed camera was achieved using an Arduino UNO microcontroller, which synchronized the function generator and high speed camera at the desired frequency ω and frame rate f for each experiment.

Particle Tracking

The microscopy videos were analyzed using TrackPy[55] to determine the (x,y) position of each particle as a function of time t. TrackPy uses an implementation of the Crocker-Grier centroid-finding algorithm[1] to locate the positions of Gaussian-like blobs of specified size within each frame. The algorithm uses a spatial bandpass filter to smooth the image and subtract off the background thereby transforming the imaged particles into Gaussian blobs on a black background. The quantity y used throughout the main text is the size parameter returned by the tracking algorithm, which measures the radius of gyration of these Gaussian blobs. To avoid complications due to particle interactions, particles separated by less than two diameters (30 μm) were excluded from the analysis. Any particle aggregates with eccentricity greater than 0.15 were also excluded.

Appendix B: Bayesian Inference

Full Probability Model

Figure 3a shows the generative model used to infer the unknown parameters from the observed data, $\mathbf{y} = \{y_{ijk}\}$, for all experiments i, particles j, and time points k. The joint probability distribution can be written as

$$p(\theta,\lambda,\phi,\mathbf{y}\,|\,\mathbf{d}) = p(\theta)\prod_{i=1}^{N_e}\prod_{j=1}^{N_p} p(\lambda_{ij}\,|\,\theta,\mathbf{d}_i)p(\phi_{ij})\prod_{k=1}^{N_t} p(y_{ijk}\,|\,\lambda_{ij},\phi_{ij},\mathbf{d}_i) \tag{8}$$

Here, the cell-level parameters, $\theta = \{\alpha,\beta,\gamma,\sigma\}$, include the amplitude, location, width, and heterogeneity of the resonance; the particle-level parameters include the relaxation rate λ_{ij} and the nuisance parameters, $\phi_{ij} = \{a_{ij}, b_{ij}\}$, that describe the initial and asymptotic values of the measured size.

Prior Distributions

We assume independent lognormal priors for the four global parameters

$$p(\theta) = p(\alpha)p(\beta)p(\gamma)p(\sigma) \tag{9}$$

where $\alpha \sim \text{Lognormal}(\mu_\alpha,\sigma_\alpha^2)$, $\beta \sim \text{Lognormal}(\mu_\beta,\sigma_\beta^2)$, $\gamma \sim \text{Lognormal}(\mu_\gamma,\sigma_\gamma^2)$, and $\sigma \sim \text{Lognormal}(\mu_\sigma,\sigma_\sigma^2)$. The choice of lognormal priors reflects the fact that each parameter is positive and that its logarithm lies within a known range. Similarly, we use independent normal priors with zero mean for the parameters ϕ_{ij}

$$p(\phi_{ij}) = p(a_{ij})p(b_{ij}) \tag{10}$$

where $a_{ij} \sim \mathcal{N}(0,\sigma_a^2)$ and $b_{ij} \sim \mathcal{N}(0,\sigma_b^2)$ for all particles j and experiments i. The hyperparameters μ_α, σ_α, etc. are summarized in

Table 1.

Parameter	Units	μ	σ
α	s^{-1}	0	0.5
β / 2π	MHz	0	0.5
γ	–	$\ln 0.1$	0.5
σ	–	$\ln 0.1$	0.5
a	pixels	0	2
b	pixels	0	2

Table 1 Hyperparameters for prior distributions

Likelihood Functions

As described in the main text, the distributions for the relaxation rate $p(\lambda_{ij} \mid \theta, \mathbf{d}_i)$ and the measured size $p(y_{ijk} \mid \lambda_{ij}, \phi_{ij}, \mathbf{d}_i)$ are given by

$$p(\lambda_{ij} \mid \theta, \mathbf{d}_i) = \text{Lognormal}\left(\lambda_{ij} \mid \mu_i, \sigma^2\right) \tag{11}$$

$$p(y_{ijk} \mid \lambda_{ij}, \phi_{ij}, \mathbf{d}_i) = \mathcal{N}\left(y_{ijk} \mid a_{ij} + b_{ij} e^{-\lambda_{ij} t_{ik}}, s^2\right) \tag{12}$$

Here, the parameter μ_i represents the logarithm of the median relaxation rate for experiment i

$$\mu_i = \ln \frac{\alpha \omega_i^2}{\beta^2 + (\omega_i - \beta)^2/\gamma^2} \tag{13}$$

The measurement error is specified as $s = 0.1$ pixel, and $t_{ik} = (k-1)/f_i$ is the time of the k^{th} frame of the i^{th} experiment following the reapplication of the field at $k = 1$. Together, the equations (8)–(12) specify the full probability model used in analyzing the experimental data and designing future experiments.

Marginal Likelihood

Given data $\mathbf{y}$ collected under design $\mathbf{d}$, we infer the probability distribution for the cell-level parameters θ by first marginalizing over the nuisance parameters ϕ and the latent parameters λ and then sampling the posterior $p(\theta \mid \mathbf{y}, \mathbf{d})$ a sequential Monte Carlo (SMC) inference engine. Exact marginalization over the particle parameters ϕ can be conducted analytically as detailed in the ESI†. By contrast, exact marginalization over the particle rates λ is not analytically tractable owing to the nonlinear dependence of the size on the rate in equation (4). However, for reasonably effective designs, the size data $\mathbf{y}_{ij}$ for particle j in experiment i provides an accurate estimate of the rate λ_{ij} such that the likelihood $p(\mathbf{y}_{ij} \mid \lambda_{ij})$ is well approximated by a lognormal distribution for λ_{ij}. Using this approximate likelihood, we marginalize over each λ_{ij} analytically to obtain the marginal likelihood $p(\mathbf{y} \mid \theta, \mathbf{d})$ for the cell-level parameters (see ESI†).

Posterior Sampling

Using this marginal likelihood, we use a sequential Monte Carlo (SMC) sampler implemented in PyMC3 (version 3.8) to sample cell-level parameters from the posterior $p(\theta \mid \mathbf{y}, \mathbf{d})$.[33,34] The PyMC3 implementation captures multimodal posteriors by using a population of Markov chains to sample from a sequence of probability distributions that gradually approach the desired posterior. We use 5000 independent chains with 5000 samples drawn from the posterior; other algorithm parameters are set to their default values.

Information Gain

To estimate the utility $u(\mathbf{y}, \mathbf{d})$ from the posterior parameter samples, we first apply Bayes' theorem to write equation (6) as

$$u(\mathbf{y}, \mathbf{d}) = \int p(\theta \mid \mathbf{y}, \mathbf{d}) \ln p(\mathbf{y} \mid \theta, \mathbf{d}) d\theta - \ln p(\mathbf{y}) \tag{14}$$

We use Monte Carlo integration based on N_θ posterior samples in $\{\theta_m\}$ to estimate the first term. We estimate the second term—the negative log-evidence—using the Monte Carlo algorithm of Heavens *et al.*[43]

$$u(\mathbf{y}, \mathbf{d}) \approx \frac{1}{N_\theta} \sum_m \ln p(\mathbf{y} \mid \theta_m, \mathbf{d}) - \ln E_o \tag{15}$$

where E_o is the MAP estimate of the evidence $p(\mathbf{y})$ (see ESI†).

Appendix C: Bayesian Design

Two-step Design

The design $\tilde{\mathbf{d}} = \{\tilde{\omega}, \tilde{f}\}$ of the next experiment is selected using a two-step optimization procedure, in which we first identify the applied frequency $\tilde{\omega}$ and then the frame rate $\tilde{f}$ that maximize the expected utility $U(\tilde{\mathbf{d}})$. Here, the expected utility of equation (7) is precisely the mutual information $I(\tilde{\mathbf{y}}; \theta)$ between the data $\tilde{\mathbf{y}}$ of the next experiment and the cell-level parameters θ. From the properties of the mutual information, this quantity can be expressed in terms of the latent rate parameter $\tilde{\lambda}$ as

$$U(\tilde{\mathbf{d}}) = I(\tilde{\lambda}; \theta) - I(\tilde{\lambda}; \theta \mid \tilde{\mathbf{y}}) + \cancel{I(\theta; \tilde{\mathbf{y}} \mid \tilde{\lambda})} \tag{16}$$

The last term represents the mutual information between the cell parameters θ and the data $\tilde{\mathbf{y}}$ conditioned on knowledge of the rate parameters $\tilde{\lambda}$; the structure of the graphical model implies that this quantity is identically zero (Fig. 3a). As the mutual information is non-negative, the remaining terms satisfy the inequalities, $I(\tilde{\lambda}; \theta) \geq I(\tilde{\lambda}; \theta \mid \tilde{\mathbf{y}}) \geq 0$. For the designs of interest, we make the simplifying assumption that the rate parameters $\tilde{\lambda}$ are accurately estimated from the data $\tilde{\mathbf{y}}$ such that $I(\tilde{\lambda}; \theta) \gg I(\tilde{\lambda}; \theta \mid \tilde{\mathbf{y}})$. The expected utility is then approximately equal to $I(\tilde{\lambda}; \theta)$, which depends on the frequency $\tilde{\omega}$ but not on the frame rate $\tilde{f}$. To maximize the expected utility, we first select the frequency $\omega^\star$ that maximizes $I(\tilde{\lambda}; \theta)$ and then select the frame rate $f^\star$ that maximizes $I(\tilde{\mathbf{y}}; \tilde{\lambda})$ which approximates the exact procedure of minimizing $I(\tilde{\lambda}; \theta \mid \tilde{\mathbf{y}})$. For further information on this approximation, see the ESI†.

Optimal Frequency $\omega^\star$

We use nested Monte Carlo integration[31,44,56] to estimate the mutual information $I(\tilde{\lambda}; \theta)$ and particle swarm optimization[57] to identify the frequency $\omega^\star$ that maximizes this estimate. Briefly, we use $N_{\text{out}} \times N_{\text{in}}$ samples of the cell-level parameters $\{\theta_{nm}\}$ drawn from the posterior distribution $p(\theta \mid \mathbf{y}, \mathbf{d})$ conditioned on data $\mathbf{y}$ from past experiments collected under design $\mathbf{d}$. We then

generate N_{out} samples of the rate parameters $\{\tilde{\lambda}_n\}$ from the distribution $p(\tilde{\lambda} \mid \theta_{n0}, \tilde{\omega})$, which is specified by equation (11) of the model. From these samples, the mutual information $I(\tilde{\lambda};\theta)$ is estimated as

$$\hat{I}(\tilde{\omega}) = \frac{1}{N_{\text{out}}} \sum_n \left[\ln p(\tilde{\lambda}_n \mid \theta_{n0}) - \ln \left(\frac{1}{N_{\text{in}}} \sum_m p(\tilde{\lambda}_n \mid \theta_{nm}) \right) \right] \quad (17)$$

This Monte Carlo estimate is both noisy and biased: the variance scales as $\mathcal{O}(N_{\text{out}}^{-1})$, the positive bias as $\mathcal{O}(N_{\text{in}}^{-1})$, and the total mean squared error as $\mathcal{O}(N_{\text{out}}^{-1} + N_{\text{in}}^{-2})$.[44] It is asymptotically optimal to set $N_{\text{out}} \propto N_{\text{in}}^2$ to achieve to the overall convergence rate $\mathcal{O}(T^{-1/3})$ in the total number of samples $T = N_{\text{out}} N_{\text{in}}$.[56] We set $N_{\text{in}} = N_{\text{out}} = 10^3$ which serves as a balance between computational speed and accuracy.[31] We use particle swarm optimization[57] with 100 iterations of 100 particles to identify the frequency $\omega^\star$ that maximizes the estimate $\hat{I}(\tilde{\omega})$.

Optimal Frame Rate $f^\star$

Rather than minimize directly the second component of equation (16) for the expected utility with respect to the frame rate $\tilde{f}$, we instead maximize the mutual information $I(\tilde{\mathbf{y}};\lambda)$ between the data $\tilde{\mathbf{y}}$ and the rate parameters λ. The approximate equivalence of these two procedures is explained in the ESI†. The objective function $I(\tilde{\mathbf{y}};\lambda)$ to be maximized is estimated as

$$\hat{I}(\omega^\star, \tilde{f}) = \frac{1}{N_{\text{out}}} \sum_n \left[\ln p(\tilde{\mathbf{y}}_n \mid \tilde{\lambda}_{n0}) - \ln \left(\frac{1}{N_{\text{in}}} \sum_m p(\tilde{\mathbf{y}}_n \mid \tilde{\lambda}_{nm}) \right) \right] \quad (18)$$

Here, the sampled rate parameters $\{\tilde{\lambda}_{nm}\}$ are drawn from the distribution $p(\tilde{\lambda} \mid \theta_{nm})$ using the optimal frequency $\omega^\star$; the sampled particle sizes $\{\tilde{\mathbf{y}}_n\}$ are drawn from the distribution $p(\tilde{\mathbf{y}} \mid \tilde{\lambda}_{n0})$. As each of the N_{p} particles is assumed the be independent, we sample measurement outcomes $\tilde{\mathbf{y}}$ for a single particle with N_{t} time points.

References

1 J. C. Crocker and D. G. Grier, *J. Colloid Interface Sci.*, 1996, **179**, 298–310.

2 K. A. Rose, M. Molaei, M. J. Boyle, D. Lee, J. C. Crocker and R. J. Composto, *J. Appl. Phys.*, 2020, **127**, 191101.

3 M. Xu, J. L. Ross, L. Valdez and A. Sen, *Phys. Rev. Lett.*, 2019, **123**, 128101.

4 Q. Chen, H. Cho, K. Manthiram, M. Yoshida, X. Ye and A. P. Alivisatos, *ACS Cent. Sci.*, 2015, **1**, 33–39.

5 Z. Ou, C. Liu, L. Yao and Q. Chen, *Acc. Mater. Res.*, 2020, **1**, 41–52.

6 K. S. Silmore, X. Gong, M. S. Strano and J. W. Swan, *ACS Nano*, 2019, **13**, 3940–3952.

7 J. Deng, M. Molaei, N. G. Chisholm and K. J. Stebe, *Langmuir*, 2020, **36**, 6888–6902.

8 J. C. Crocker, M. T. Valentine, E. R. Weeks, T. Gisler, P. D. Kaplan, A. G. Yodh and D. A. Weitz, *Phys. Rev. Lett.*, 2000, **85**, 888.

9 E. M. Furst and T. M. Squires, *Microrheology*, Oxford University Press, 2017.

10 J. C. Crocker and D. G. Grier, *Phys. Rev. Lett.*, 1994, **73**, 352.

11 X. Tang, B. Rupp, Y. Yang, T. D. Edwards, M. A. Grover and M. A. Bevan, *ACS Nano*, 2016, **10**, 6791–6798.

12 J. Perrin, *Brownian movement and molecular reality*, Taylor and Francis, London, 1910.

13 A. E. Patteson, A. Gopinath and P. E. Arratia, *Curr. Opin. Colloid Interface Sci.*, 2016, **21**, 86–96.

14 C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe and G. Volpe, *Rev. Mod. Phys.*, 2016, **88**, 045006.

15 W. Fei, Y. Gu and K. J. Bishop, *Curr. Opin. Colloid interface Sci.*, 2017, **32**, 57–68.

16 P. Illien, R. Golestanian and A. Sen, *Chem. Soc. Rev.*, 2017, **46**, 5508–5518.

17 A. M. Brooks, M. Tasinkevych, S. Sabrina, D. Velegol, A. Sen and K. J. Bishop, *Nat. Commun.*, 2019, **10**, 1–9.

18 Z. Zhang, H. Yuan, Y. Dou, M. O. de la Cruz and K. J. Bishop, *Phys. Rev. Lett.*, 2021, **126**, 258001.

19 T. J. Loredo, *AIP Conf. Proc.*, 2004, **707**, 330–346.

20 W. Wang, L. A. Castro, M. Hoyos and T. E. Mallouk, *ACS Nano*, 2012, **6**, 6122–6132.

21 S. Ahmed, W. Wang, L. Bai, D. T. Gentekos, M. Hoyos and T. E. Mallouk, *ACS Nano*, 2016, **10**, 4763–4769.

22 S. Sabrina, M. Tasinkevych, S. Ahmed, A. M. Brooks, M. Olvera de la Cruz, T. E. Mallouk and K. J. Bishop, *ACS Nano*, 2018, **12**, 2939–2947.

23 C. Zhou, L. Zhao, M. Wei and W. Wang, *ACS Nano*, 2017, **11**, 12668–12676.

24 L. Ren, N. Nama, J. M. McNeill, F. Soto, Z. Yan, W. Liu, W. Wang, J. Wang and T. E. Mallouk, *Sci. Adv.*, 2019, **5**, eaax3084.

25 F. Nadal and E. Lauga, *Phys. Fluids*, 2014, **26**, 082001.

26 D. Sivia and J. Skilling, *Data Analysis: a Bayesian Tutorial*, Oxford University Press, 2006.

27 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.

28 D. V. Lindley, *Ann. Math. Statist.*, 1956, **27**, 986–1005.

29 M. Betancourt, 2017, arXiv:1704.02434v2.

30 D. M. Blei, A. Kucukelbir and J. D. McAuliffe, *J. Am. Stat. Assoc.*, 2017, **112**, 859–877.

31 X. Huan and Y. M. Marzouk, *J. Comp. Phys.*, 2013, **232**, 288–317.

32 B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li and A. Riddell, *J. Stat. Softw.*, 2017, **76**, 1–32.

33 J. Salvatier, T. Wiecki and C. Fonnesbeck, *PeerJ Comput. Sci.*, 2016, **2**, e55.

34 P. Del Moral, A. Doucet and A. Jasra, *J. R. Stat. Soc. Series B*, 2006, **68**, 411–436.

35 K. Han, C. W. Shields IV and O. D. Velev, *Adv. Funct. Mater.*, 2018, **28**, 1705953.

36 Y. Yang and M. A. Bevan, *ACS Nano*, 2018, **12**, 10712–10724.

37 M. A. Fernandez-Rodriguez, F. Grillo, L. Alvarez, M. Rathlef, I. Buttinoni, G. Volpe and L. Isa, *Nat. Commun.*, 2020, **11**,

4223.
38 W. Fei, P. M. Tzelios and K. J. Bishop, *Langmuir*, 2020, **36**, 6977–6983.
39 Y. Dou, P. M. Tzelios, D. Livitz and K. J. Bishop, *Soft Matter*, 2021, **17**, 1538–1547.
40 H. Bruus, *Lab Chip*, 2012, **12**, 1014–1021.
41 R. Barnkob, P. Augustsson, T. Laurell and H. Bruus, *Lab Chip*, 2010, **10**, 563–570.
42 H. Bruus, *Lab Chip*, 2012, **12**, 20–28.
43 A. Heavens, Y. Fantaye, A. Mootoovaloo, H. Eggers, Z. Hosenie, S. Kroon and E. Sellentin, 2017, arXiv:1704.03472.
44 K. J. Ryan, *J. Comput. Graph. Stat.*, 2003, **12**, 585–603.
45 J. P. McMullen and K. F. Jensen, *Org. Process Res. Dev.*, 2011, **15**, 398–407.
46 E. G. Ryan, C. C. Drovandi, J. M. McGree and A. N. Pettitt, *Int. Stat. Rev.*, 2016, **84**, 128–154.
47 D. M. Blei, *Annu. Rev. Stat. Appl.*, 2014, **1**, 203–232.
48 J. Voß and R. Wittkowski, *Nanoscale Advances*, 2020, **2**, 3890–3899.
49 E. Brochu, V. M. Cora and N. De Freitas, 2010, arXiv:1012.2599.
50 C. E. Rasmussen and C. K. I. Williams, *Gaussian processes in machine learning*, MIT Press, 2006.
51 J. Močkus, Optimization techniques IFIP technical conference, 1975, pp. 400–404.
52 Y. Yao, A. Vehtari and A. Gelman, 2020, arXiv:2006.12335.
53 *https://github.com/bishopgroup/acousticBayes*.
54 M. Von Stosch, R. Oliveira, J. Peres and S. F. de Azevedo, *Comput. Chem. Eng.*, 2014, **60**, 86–101.
55 D. Allan, T. Caswell, N. Keim and C. van der Wel, *Zenodo*, 2016, **60550**, 2019.
56 T. Rainforth, R. Cornish, H. Yang, A. Warrington and F. Wood, 2017, arXiv:1709.06181.
57 Tisimst, *Pyswarm*, `https://github.com/tisimst/pyswarm`, 2015.