A Bayesian approach to analyzing holograms of colloidal particles

Thomas G. Dimiduk[1] and Vinothan N. Manoharan[2,1*]

[1] *Department of Physics, Harvard University, 17 Oxford St, Cambridge MA 02138, USA*
[2] *Harvard John A. Paulson School of Engineering and Applied Sciences, 29 Oxford St, Cambridge MA 02138, USA*
[*] *vnm@seas.harvard.edu*

Abstract: We demonstrate a Bayesian approach to tracking and characterizing colloidal particles from in-line digital holograms. We model the formation of the hologram using Lorenz-Mie theory. We then use a tempered Markov-chain Monte Carlo method to sample the posterior probability distributions of the model parameters: particle position, size, and refractive index. Compared to least-squares fitting, our approach allows us to more easily incorporate prior information about the parameters and to obtain more accurate uncertainties, which are critical for both particle tracking and characterization experiments. Our approach also eliminates the need to supply accurate initial guesses for the parameters, so it requires little tuning.

OCIS codes: (090.0090) Holography; (090.1995) Digital holography; (100.3190) Inverse problems; (100.3200) Inverse scattering; (150.6910) Three-dimensional sensing.

References and links

1. F. C. Cheong, S. Duarte, S.-H. Lee, and D. G. Grier, "Holographic microrheology of polysaccharides from *Streptococcus mutans* biofilms," Rheol. Acta. **48**, 109–115 (2009).
2. R. W. Perry, G. Meng, T. G. Dimiduk, J. Fung, and V. N. Manoharan, "Real-space studies of the structure and dynamics of self-assembled colloidal clusters," Faraday Discuss. **159**, 211–234 (2012).
3. J. Fung, K. E. Martin, R. W. Perry, D. M. Kaz, R. McGorty, and V. N. Manoharan, "Measuring translational, rotational, and vibrational dynamics in colloids with digital holographic microscopy," Opt. Express **19**, 8051–8065 (2011).
4. D. M. Kaz, R. McGorty, M. Mani, M. P. Brenner, and V. N. Manoharan, "Physical ageing of the contact line on colloidal particles at liquid interfaces," Nat. Mater. **11**, 138–142 (2012).
5. B. Ovryn, "Three-dimensional forward scattering particle image velocimetry applied to a microscopic field-of-view," Exp. Fluids **29**, S175–S184 (2000).
6. F. C. Cheong, B. S. Rémi Dreyfus, J. Amato-Grill, K. Xiao, L. Dixon, and D. G. Grier, "Flow visualization and flow cytometry with holographic video microscopy," Opt. Express **17**, 13071–13079 (2009).
7. S.-H. Lee, Y. Roichman, G.-R. Yi, S.-H. Kim, S.-M. Yang, A. van Blaaderen, P. van Oostrum, and D. G. Grier, "Characterizing and tracking single colloidal particles with video holographic microscopy," Opt. Express **15**, 18275–18282 (2007).
8. C. Wang, H. Shpaisman, A. D. Hollingsworth, and D. G. Grier, "Celebrating Soft Matter's 10th Anniversary: Monitoring colloidal growth with holographic microscopy," Soft Matter **11**, 1062–1066 (2015).
9. C. Wang, X. Zhong, D. B. Ruffner, A. Stutt, L. A. Philips, M. D. Ward, and D. G. Grier, "Holographic Characterization of Protein Aggregates," J. Pharm. Sci. **105**, 1074–1085 (2016).
10. Y. Pu and H. Meng, "Intrinsic aberrations due to Mie scattering in particle holography," J. Opt. Soc. Am. A **20**, 1920–1932 (2003).
11. B. Ovryn and S. H. Izen, "Imaging of transparent spheres through a planar interface using a high-numerical-aperture optical microscope," J. Opt. Soc. Am. A **17**, 1202–1213 (2000).
12. A. Wang, T. G. Dimiduk, J. Fung, S. Razavi, I. Kretzschmar, K. Chaudhary, and V. N. Manoharan, "Using the discrete dipole approximation and holographic microscopy to measure rotational dynamics of non-spherical colloidal particles," J. Quant. Spectrosc. Radiat. Transfer **146**, 499–509 (2014).
13. J. Fung, R. W. Perry, T. G. Dimiduk, and V. N. Manoharan, "Imaging Multiple Colloidal Particles by Fitting Electromagnetic Scattering Solutions to Digital Holograms," J. Quant. Spectrosc. Radiat. Transfer **113**, 2482–2489 (2012).
14. L. Denis, D. Lorenz, E. Thiébaut, C. Fournier, and D. Trede, "Inline hologram reconstruction with sparsity constraints," Opt. Lett. **34**, 3475–3477 (2009).
15. F. Soulez, L. Denis, C. Fournier, É. Thiébaut, and C. Goepfert, "Inverse-problem approach for particle digital holography: accurate location based on local optimization," J. Opt. Soc. Am. A **24**, 1164–1171 (2007).
16. A. Yevick, M. Hannel, and D. G. Grier, "Machine-learning approach to holographic particle characterization," Opt. Express **22**, 26884–26890 (2014).

17. T. Latychevskaia, F. Gehri, and H.-W. Fink, "Depth-resolved holographic reconstructions by three-dimensional deconvolution," Opt. Express **18**, 22527–22544 (2010).
18. M. Seifi, C. Fournier, L. Denis, D. Chareyron, and J. L. Marié, "Three-dimensional reconstruction of particle holograms: a fast and accurate multiscale approach," J. Opt. Soc. Am. A **29**, 1808–1817 (2012).
19. T. G. Dimiduk, R. W. Perry, J. Fung, and V. N. Manoharan, "Random-subset fitting of digital holograms for fast three-dimensional particle tracking," Appl. Opt. **53**, G177–G183 (2014).
20. D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," J. Soc. Ind. Appl. Math. **11**, 431–441 (1963).
21. P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support* (Cambridge University, 2005).
22. J. Goodman and J. Weare, "Ensemble samplers with affine invariance," Comm. Appl. Math. Comp. Sci. **5**, 65–80 (2010).
23. D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "emcee: The MCMC Hammer," Publ. Astron. Soc. Pac. **125**, 306–312 (2013).
24. D. J. Earl and M. W. Deem, "Parallel tempering: Theory, applications, and new perspectives," PCCP **7**, 3910–3916 (2005).
25. T. G. Dimiduk, J. Fung, R. W. Perry, and V. N. Manoharan, "HoloPy - Hologram Processing and Light Scattering in Python," https://github.com/manoharan-lab/holopy (2016).

1. Introduction

Digital in-line holography has emerged as a powerful tool for both tracking and characterizing colloidal particles. Such particles are a few hundred nanometers to a few micrometers in diameter and dispersed in a fluid. In a typical tracking experiment, one measures the position of the particle in three dimensions as a function of time by analyzing a time-series of holograms. During the experiment, the particle might be subject to Brownian motion, interactions, or external forces. Tracking measurements are important to a number of different experiments in soft-matter physics and fluid mechanics, including microrheology [1], the study of phase transitions [2], measuring interactions in colloidal systems [3,4], and particle image velocimetry [5,6]. In a characterization experiment, one measures the size, refractive index, and shape of individual colloidal particles [7,8]. This type of measurement is important for quality control of formulations [9] and studies of polymerization kinetics [8], among other applications.

Many of these experiments have been enabled by a method of analyzing holograms based on scattering theory. A hologram is a two-dimensional interference pattern that encodes information about the three-dimensional (3D) positions and optical properties of a particle or set of particles. An inline hologram results from the interference of the scattered fields from the sample and the undiffracted beam, which acts as a reference wave (Fig. 1). Traditionally, the 3D information is recovered through reconstruction—effectively, shining light back through the hologram to generate a real 3D image. However, when the particles are comparable to the wavelength, reconstruction produces a number of imaging artifacts [10]. An alternative to reconstruction was first demonstrated by Ovryn and Izen [5,11], who showed that inline holograms could be directly predicted from Lorenz-Mie scattering theory and a model for the propagation and interference of the fields in the optical train of the microscope. Later work by Lee and coworkers [7], using a simpler model of the optical train, showed that one could fit a scattering model based on Lorenz-Mie theory to both track and characterize spherical colloidal particles. Further work from our research group has shown that other models, based on different scattering theories, can be fit to holograms of non-spherical particles [12] and clusters of spherical particles [2,13].

With these techniques, the optical field is never reconstructed; instead, a forward model—a model that can calculate a hologram based on scattering theory and propagation along the optical train— is fit to the data. In this method, which we will refer to as "least-squares fitting," an iterative minimization algorithm is used to find the parameters in the model that minimize the sum of the squared differences, pixel-by-pixel, between the measured hologram and the model. Although this process is computationally more expensive than reconstruction, it yields much more accurate and precise measurements: for tracking experiments, analysis of the mean-square

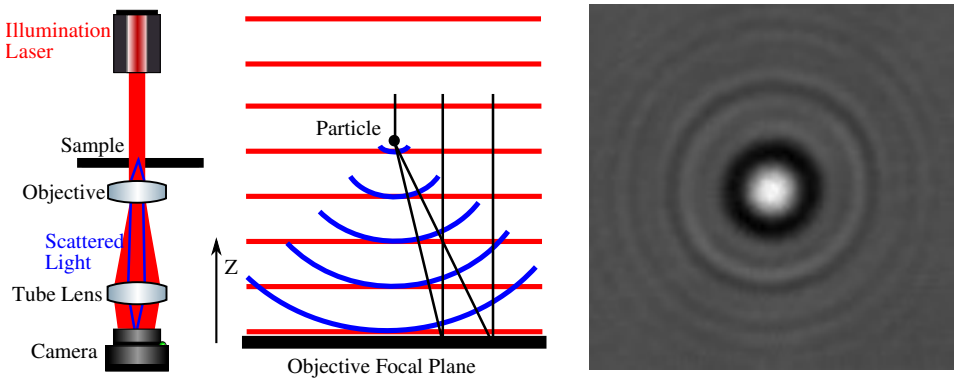

Fig. 1. A typical optical setup used to record inline digital holograms. We image with a collimated laser beam and record the hologram on a CMOS camera. The coordinate z is defined as the distance from the focal plane of the objective to the particle ($z > 0$ if the particle lies further from the objective than the focal plane). The reference wave is shown in red and the scattered wave in blue. x and y, not shown, represent the position of the particle in the plane normal to the optical axis. On the right, a measured, background-divided hologram of a 1-μm-diameter silica particle in water, taken using a 660 nm laser (hologram courtesy of Anna Wang).

displacement shows that the uncertainty in the position can be on the order of a few nanometers or less [4].

However, there are two problems with the least-squares technique. First, it is difficult to obtain accurate uncertainties on parameters. In a characterization experiment, for example, one would like to know the uncertainty in the measured size or refractive index. The uncertainties on the fit parameters are critical not only for characterization experiments, but for any experiment in which the data will later be compared to a physical theory—for example an equation of motion for the particles—since the agreement with theory can only be judged when uncertainties are known. It is possible to obtain some estimates of uncertainty from the parameter covariance matrix, which is usually an output of the minimization algorithm, but these uncertainties cannot be easily marginalized, as we explain below. Second, the iterative minimization schemes used in least-squares fitting require good initial guesses for and constraints on all the parameters [2,3,12]. Determining the initial guesses and the constraints often requires a great deal of effort, particularly for more complex scattering models.

Here we demonstrate a Bayesian analysis method that overcomes these problems. Bayesian approaches to parameter estimation differ from classical frequentist approaches such as least-squares fitting in that we are allowed to speak of the probability distribution *of a parameter*. As in the least-squares technique, we start with a forward model, but instead of maximizing the likelihood, which is the probability of the data given parameter values, we maximize the posterior probability, which is the probability of the parameter values given the data. There are other Bayesian approaches to analyzing holograms [14], but none that we are aware of use forward models based on scattering solutions.

There are two principle advantages of the Bayesian approach over least-squares fitting: first, any prior information about the particle, such as its position, size, or composition, can be formalized as a prior probability distribution, rather than implicitly encoded into initial guesses or constraints. Second, because probabilities are assigned to parameters, we can marginalize, or integrate out, the uncertainties on some of the parameters to obtain the uncertainties on the others. For example, in the characterization problem, we are interested in the uncertainty on the size or refractive index after taking into account the uncertainties on the positions. This type of analysis is not

easily done using classical frequentist methods, and hence the parameters that are not of interest (but required by the model) are often called "nuisance parameters."

As an example of the utility of this technique, we show the results from a Bayesian analysis of a hologram of a 1-μm-diameter silica particle in water in Fig. 2, compared to the results obtained from least-squares fitting. The Bayesian approach yields nearly the same "best fit" values for the parameters, but it also gives us much more information about the parameter uncertainties. From the Bayesian results, we see that there are strong correlations between the estimated particle refractive index, size, and axial position. The covariances between these parameters are difficult to extract in a least-squares approach. Because we have access to these covariances in the Bayesian approach, we can integrate them out, or "marginalize," to estimate the uncertainties on any one parameter, as shown by the distributions along the diagonal.

The rest of this paper explains how we obtain these results and how we interpret them. In practice, the Bayesian approach involves three steps: specifying a forward model, specifying priors on all the parameters, and calculating the posterior probability distribution of all the parameters. The last step, calculating the posterior, is most effectively done using Markov-chain Monte Carlo (MCMC) techniques. In what follows, we limit our discussion to the measurement of a single microsphere, and we use a simple model for the propagation and interference along the optical train, equivalent to that of Lee and coworkers [7]. Nonetheless, the approach can easily be extended to more complex models of scattering or propagation. After describing the model and the Bayesian approach to calculating the posterior probability distribution, we apply our method to both the particle tracking problem and the particle characterization problem. We also show how to speed up the approach—and obviate the need for initial starting guesses for our MCMC chains—by a type of tempering based on choosing random subsets of pixels.

2. Background

Here we describe the forward model for hologram formation from a single spherical colloidal particle and the least-squares method used to fit this model to the data. The data are obtained from an in-line holographic setup as shown in Fig. 1. We follow the work of Lee and coworkers [7] closely. In their model, a hologram $\mathbf{H}$ is computed as a function of six parameters:

$$\mathbf{H}(x, y, z, n, r, \alpha) = |\alpha E_{\text{scat}}(x, y, z, r, n) + E_{\text{ref}}|^2 \tag{1}$$

where x, y, and z are the coordinates of the center of the particle, n is its (possibly complex) refractive index, r is its radius, and α is an auxiliary parameter that rescales the scattered field E_{scat}. E_{ref} is the reference field. We use boldface type for the hologram $\mathbf{H}$ to indicate that it is a matrix consisting of measurements of the intensity at each pixel (i, j) on our detector. The parameter α is needed to compensate for a number of different effects along the optical train and to produce satisfactory agreement between the model and measured holograms. Although a single scalar parameter is likely insufficient to account for all of these effects, the Bayesian analysis framework we develop below can easily be extended to more complex models, including ones that model the propagation along the optical train.

We divide the measured hologram $\mathbf{H}$ by a measured background, usually from a hologram captured without any particles in the field of view. The measured background should be very close to the intensity of the reference wave, $|E_{\text{ref}}|^2$, differing only by the light scattered by the particle of interest. We therefore approximate the background as the reference intensity to obtain an expression for the normalized hologram $\mathbf{h}$:

$$\mathbf{h} = \frac{\mathbf{H}(x, y, z, r, n, \alpha)}{|E_{\text{ref}}|^2} = \left| \frac{\alpha E_{\text{scat}}(x, y, z, r, n)}{E_{\text{ref}}} + 1 \right|^2 . \tag{2}$$

To infer the parameters x, y, z, r, n, and α, we need a model for E_{scat} and an inference method. For a single sphere, the Lorenz-Mie theory can be used to calculate E_{scat}; other scattering

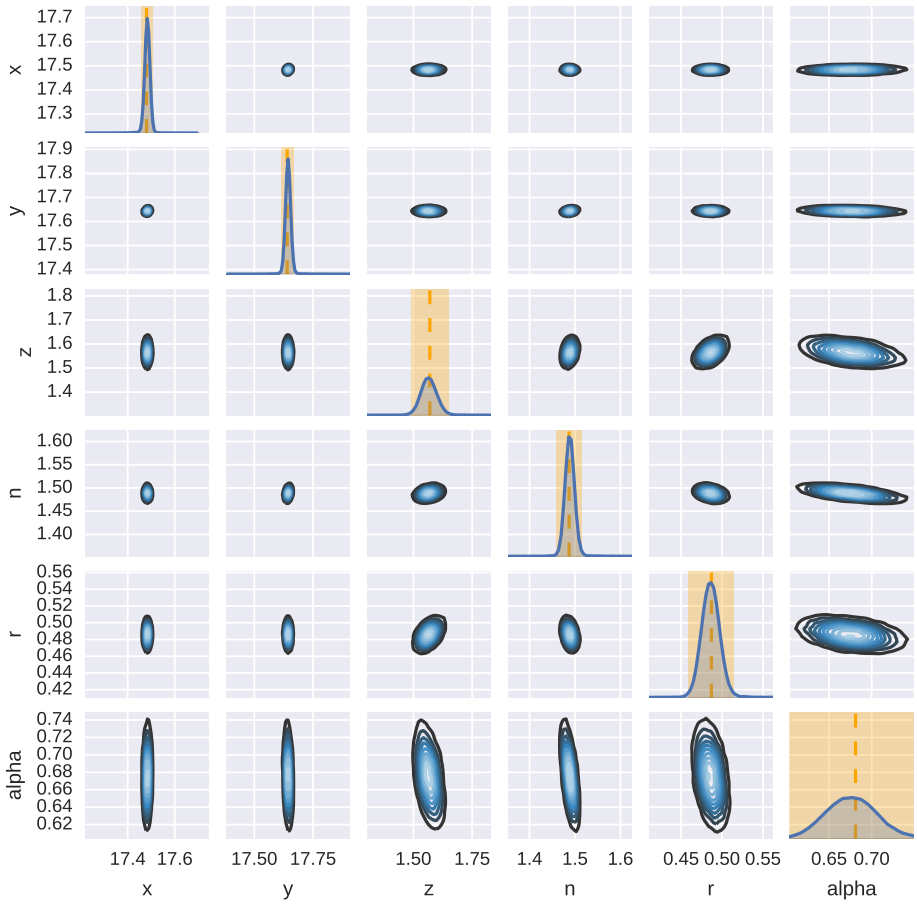

Fig. 2. Comparison of results obtained from least-squares fitting using the Levenberg-Marquardt method, as described in section 2, to those obtained from the Bayesian inference method described in this paper. Units of x, y, z and r are micrometers. The results from fitting are shown by the dashed orange lines with orange regions showing one-sigma confidence intervals calculated from the parameter covariance matrix. For α ("alpha") the width of the interval exceeds the width of the plot. Fully marginalized distributions obtained from Markov-chain Monte Carlo (MCMC) sampling are shown along the diagonal, while the off-diagonal contour plots show the joint distributions of each pair of parameters, represented as Gaussian kernel density estimates. We explain the Bayesian approach and the MCMC technique later in the text.

theories [3, 12, 13] can be used for different types of particles or clusters of particles. In the inference method used by Lee and coworkers [7] and found in many subsequent papers, the parameters are inferred by minimizing the sum (over pixels) of the squared differences between model and data. Letting θ be the set of parameters $\{x, y, z, n, r, \alpha\}$, we have

$$\theta = \arg\min_{\theta} \chi^2 = \arg\min_{\theta} \sum_{i,j} \left[h_{ij} - h_{ij}^{M}(\theta) \right]^2 \tag{3}$$

where h_{ij} is the intensity measured at pixel (i, j) in the normalized recorded hologram and h_{ij}^{M} is the same for the hologram calculated from the forward model M. Using this least-squares formulation is equivalent to maximum-likelihood estimation when the noise in the measurements has a Gaussian distribution.

Because evaluating the Mie fields involves many computationally expensive special-function evaluations over a large number of pixels, several alternatives and variations on this method have been developed. Soulez and coworkers [15] used a local refinement approach to find particle positions more efficiently. Yevick and coworkers [16] used a machine-learning method employing a trained support-vector machine, instead of the full Lorenz-Mie solution, to infer the parameters. It is also possible to obtain particle positions from reconstructed volumes using deconvolution [17] or successive resampling [18]. In a recent paper our research group showed that good fits could be obtained by modeling only a small, randomly chosen subset of the pixels [19]. We later discuss how the random-subset technique can speed up the Bayesian approach as well.

Least-squares fitting is generally done using an iterative non-linear algorithm such as Levenberg-Marquardt [20], which requires the user to specify an initial guess for the parameters. Good initial guesses are required for the algorithm to converge. To keep the algorithm from wandering off into unphysical regions of parameter space, one can fix certain parameters or allow them to vary only over a certain range. One can also add arbitrary functions to the χ^2 function in Eq. (3) to penalize certain regions of parameter space. However, fixed values for parameters might not accurately reflect our knowledge of the system. We might have some prior information about the particle size (for example, from the manufacturer) but not trust it enough to set the size to a certain value. A penalty function can more accurately capture our state of knowledge, but as we shall see, Bayesian priors offer an elegant alternative to these various methods of limiting the parameter space.

3. Bayesian approach

The problem of estimating parameters from a measured hologram and a forward model is well suited to Bayesian inference, which allows us to replace guesses, penalties, and constraints with a single prior probability function. Here we describe our approach using standard Bayesian terminology. Specifically, we look for what is called the *posterior probability* or, more simply, the "posterior": the probability distribution of the parameters, given the data we have measured. We can compute the posterior using Bayes' theorem, which relates the posterior to the *prior probability* or "prior"—which represents our knowledge of the parameters before the data is accounted for—and the *likelihood*, a function expressing the probability that the measured hologram would be observed, given particular values for the parameters and a model for the noise. For a more detailed discussion of Bayesian inference and incorporation of priors, see chapter 3 of reference [21].

We are interested in the posterior probability $p(\theta|\mathbf{h}, M, I)$ of a set of parameters θ given the recorded hologram $\mathbf{h}$, the forward model M, and known prior information I. I might include information about the apparatus, such as the magnification, wavelength, and observation volume, and the fact that we know we are looking at a single sphere. Applying Bayes' theorem, we can express the posterior in terms of a prior probability density $p(\theta|M, I)$, the likelihood function

$p(\mathbf{h}|\theta, M, I)$, and a normalization factor $p(\mathbf{h}|M, I)$:

$$p(\theta|\mathbf{h}, M, I) = \frac{p(\theta|M, I)p(\mathbf{h}|\theta, M, I)}{p(\mathbf{h}|M, I)}. \tag{4}$$

Because the normalization factor $p(\mathbf{h}|M, I)$ has no effect on the measured parameters or their uncertainties, we work with the unnormalized posterior probability distribution:

$$\begin{aligned} p'(\theta|\mathbf{h}, M, I) &= p(\theta|\mathbf{h}, M, I)p(\mathbf{h}|M, I) \\ &= p(\theta|M, I)p(\mathbf{h}|\theta, M, I). \end{aligned} \tag{5}$$

As shown by Eq. (5), we need to specify the prior $p(\theta|M, I)$ and the likelihood function $p(\mathbf{h}|\theta, M, I)$ to proceed.

3.1. Priors

The prior $p(\theta|M, I)$ is simply a mathematical statement of what we know about the parameters before doing the inference calculation. For example, priors for r and n might come from manufacturer's data sheets for commercially sourced particles. If we know the material composition of the particle, we can obtain a prior for n from a refractive index table and some estimated uncertainty. Similarly, we might obtain a prior probability distribution for r from the size distribution of the particles, as measured by the manufacturer or by some other technique, such as light scattering. In general, the prior distribution reflects our state of knowledge about a parameter, with wider distributions indicating greater uncertainty. We discuss specific choices of priors in section 4.1.

3.2. Likelihood function

To evaluate the likelihood $p(\mathbf{h}|\theta, M, I)$, we must model the noise in the measured holograms. There are a variety of noise sources, including photon shot noise, electronic readout noise at the camera, and stray fringes from other scatterers in the optical train. We consider only the shot noise and readout noise in what follows. We then write the measured hologram as the sum of the model hologram plus noise:

$$h_{ij} = h^M_{ij} + u_{ij} \tag{6}$$

where the model hologram $\mathbf{h}^M$ (the elements of which are h^M_{ij}) is calculated from the forward model (Eq. (2)) and the Lorenz-Mie theory. The noise in the intensity at pixel (i, j) is represented by u_{ij}.

In writing this equation, we are assuming that the noise can be represented by a single additive term that characterizes pixel-by-pixel variations in the intensity. We now further assume that the noise $\mathbf{u}$ is delta correlated in space and time and that each u_{ij} is Gaussian distributed with a variance σ^2_{ij} at each pixel. With these assumptions, we can write the likelihood as

$$p(\mathbf{h}|\theta, M, I) = \prod_{i,j} p(u_{ij}|\theta, M, I) \tag{7}$$

where

$$p(u_{ij}|\theta, M, I) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left\{ \frac{-\left[h_{ij} - h^M_{ij}(\theta)\right]^2}{2\sigma^2_{ij}} \right\}. \tag{8}$$

Combining Eq. (7) with Eq. (8), we obtain the full likelihood

$$p(\mathbf{h}|\theta, M, I) = \frac{1}{(2\pi)^{N/2} \prod_{i,j} \sigma_{ij}} \exp\left\{ -\sum_{i,j} \frac{\left[h_{ij} - h^M_{ij}(\theta)\right]^2}{2\sigma^2_{ij}} \right\} \tag{9}$$

where N is the total number of pixels in the hologram. We can estimate σ_{ij}^2 from a series of background holograms where no particles are present.

To justify this form of the likelihood, we first note that at the intensities we measure, the number of photons and readout electrons are large enough that the both shot noise and readout noise are approximately Gaussian, and hence their sum is also Gaussian. Also, the non-linear least squares fitting methods used in previous work implicitly assume that the noise is Gaussian, so our approach is consistent with these methods. Other noise models that include the effect of correlations (that is, models for speckle) are a topic for future work.

3.3. Putting it all together

We recognize the term in the exponent of the likelihood in Eq. (9) as $\chi^2/2$, where

$$\chi^2(\theta) = \sum_{i,j} \frac{\left[h_{ij} - h_{ij}^M(\theta)\right]^2}{\sigma_{ij}^2}. \tag{10}$$

In this work we assume that the standard deviation of the noise at each pixel is identical and equal to a constant σ, so that

$$\prod_{i,j} \sigma_{ij} = \sigma^N. \tag{11}$$

Substituting the above into Eq. (5) yields an equation for the unnormalized posterior:

$$p'(\theta|\mathbf{h}, M, I) = \frac{p(\theta|M, I)}{(2\pi)^{N/2}\sigma^N} \exp\left[-\frac{\chi^2(\theta)}{2}\right]. \tag{12}$$

Because the χ^2 term depends on the parameters θ through the forward model, it is difficult if not impossible to calculate the posterior directly from Eq. (12). We therefore use a Markov-chain Monte Carlo technique as described in section 4.1.

For the specific problems of particle tracking and particle characterization, we are interested only in subsets of the parameters. In particle tracking, we want to infer the position as a function of time. The position is described by one or more of the parameters x, y, and z along with their uncertainties, which can be obtained by marginalizing over n, r, and α (and over the position variables that are not of interest). In particle characterization, we want to infer the index of refraction n and radius r of the particle along with their associated uncertainties, which can be obtained by marginalizing over x, y, z, and α. Formally, marginalization involves integration of the posterior distribution over the parameters to be marginalized. The marginalized distribution for the particle tracking problem is

$$p'(x, y, z|\mathbf{h}, M, I) = \int dr\, dn\, d\alpha\, p'(x, y, z, r, n, \alpha|\mathbf{h}, M, I) \tag{13}$$

and that for the particle characterization problem is

$$p'(n, r|\mathbf{h}, M, I) = \int dx\, dy\, dz\, d\alpha\, p'(x, y, z, n, r, \alpha|\mathbf{h}, M, I) \tag{14}$$

In practice, we do not have to evaluate these integrals. As we describe in section 4.1, we can sample the posterior using MCMC techniques and then marginalize by binning samples or by kernel density estimation.

3.4. Time-series

In both particle tracking and particle characterization we work with time-series of holograms. There are two different categories of parameters within time-series: those that change over time and those that do not. In most cases, the particle radius r and refractive index n do not change as a function of time, whereas x, y, and z do. Therefore when calculating the marginalized posterior for the characterization problem, $p'(n, r|\mathbf{h}, M, I)$, we can use information from the entire time-series to minimize the uncertainty on the parameters. When computing $p'(x, y, z|\mathbf{h}, M, I)$ we analyze each frame individually.

There are at least two procedures to calculate the posterior $p'(n, r|\mathbf{h}, M, I)$ using information from the entire time-series. One is to evaluate it for a single frame and then use the result as the prior for the next frame. This procedure is similar to using a Kalman filter. It should yield the best estimate of n and r when the last frame is analyzed. The other procedure is to do a global calculation over the entire time-series. This procedure would involve expanding the model to include parameters for x, y, and z in each frame, then marginalizing all of these position parameters. Here we use the first approach because it is simpler to implement and computationally less demanding.

4. Results and discussion

4.1. Markov-chain Monte Carlo technique

Because Eq. (12) does not admit an analytic solution, we must calculate the posterior probability distribution numerically. Brute force evaluation of the posterior is prohibitively expensive owing to the six parameters. Thus we use a sampling technique based on a Markov-chain Monte Carlo algorithm.

Specifically, we use an affine-invariant ensemble sampler [22] as implemented in the Python library emcee [23]. This sampler uses an ensemble of "walkers" to explore the posterior probability distribution. After a sufficient number of steps, the ensemble will converge to a steady-state distribution that is, by construction, equal to the posterior probability density. Thus, in the long-time limit, the method yields a set of samples directly from the posterior. Each sample consists of values for all six parameters in our forward model. We can plot these samples as in Fig. 2 to visualize the joint distributions between the parameters, or we can marginalize simply by binning the samples as a function of only the parameters we wish to infer.

We demonstrate these techniques on holograms of 1-μm-diameter silica spheres in water (index 1.333), illuminated with a 660 nm laser. We assign priors as follows. For the refractive index n, we choose a Gaussian prior with a mean of 1.5 and a standard deviation of 0.1, chosen based on the typical variation of refractive index from particle to particle measured in holographic microscopy. Based on manufacturer's specifications, we choose a Gaussian prior for the particle radius with mean of 0.5 μm and standard deviation of 0.05 μm. We estimate the in-plane particle position (x, y) using a Hough-transform based algorithm [6], and we assign Gaussian priors with standard deviations of 0.1 μm, chosen based on our prior experience with this algorithm. For α, we choose a Gaussian prior centered at 0.7 with width 0.05, based on prior experience from least-squares fitting. We have also used a uniform prior ranging from 0 to 1, with no significant effect on the results.

Choosing a prior distribution for z is the most difficult task; our knowledge of z is limited before we do the full inference calculation (indeed, one of the main goals of the model-based approach to holography is to make it possible to extract precise information about z). We discuss various choices for the prior on z below. In the end, we are able to use a uniform prior from 0 to 100 μm from the focal plane. We chose this width based on the sample height, which is on the order of 100 μm. The prior can be widened or narrowed with little effect on the final results.

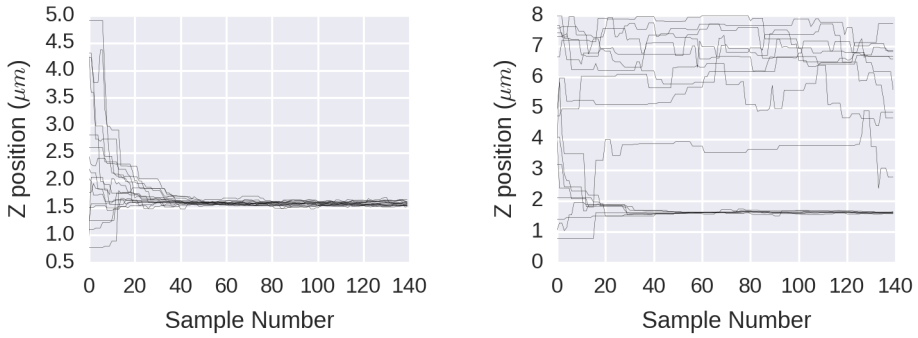

Fig. 3. Plots of walker position as a function of MCMC time step. In both cases, the walker positions are chosen from a uniform distribution with lower bound at zero. In the plot on the left, the upper bound is 5 μm, whereas on the right it is 8 μm. As can be seen from the plot on the right, walkers that are further away from the maximum in the posterior distribution get stuck and do not reach steady state.

We compute the noise (σ in Eq. 12) from a background image $\mathbf{b}$ as

$$\sigma = \text{std}(\mathbf{b})/\text{mean}(\mathbf{b}). \tag{15}$$

Typically, $\sigma \approx 0.1$ for holograms with no other particles present. For the results presented here, $\sigma = 0.119$.

4.1.1. Equilibration of the MCMC chain

With these choices for priors, we run our MCMC algorithm and examine the period of time required to reach steady state, called the "burn-in" period. Evaluating the burn-in period is important because the samples obtained within it are not representative of the posterior and therefore must be discarded.

We find that the number of samples needed to reach steady state is strongly affected by the initial positions of the ensemble of walkers. Fig. 3(a) shows that if the initial positions are not too far from the maximum of the posterior of z, the burn-in is quick. However, if the initial positions are slightly further from the maximum, the MCMC walkers can get stuck in local minima of the posterior distribution and take a long time to equilibrate (Fig. 3(b)).

4.1.2. Random-subset tempering

The slow equilibration poses a chicken-and-egg problem for the MCMC method: without a good guess for the position of the particle, the method does not yield a reliable estimate in any reasonable time, but it is difficult to obtain a sufficiently precise guess (within a few micrometers of the maximum of the posterior) without knowing the position ahead of time. One way to address this challenge is parallel tempering [24], in which several different MCMC chains are run in parallel at different "temperatures" (a parameter used to vary the sharpness of the probability distributions). Higher-temperature chains can more easily explore the entire probability distribution.

Inspired by this approach, we use a tempering scheme, but one that is simpler to implement and computationally less expensive. Our technique is based on previous work [19] showing that only a small, randomly selected subset of the pixels in a hologram are needed to extract information about the parameters in the model. We therefore start by choosing a small fraction $f = 0.0005$ of pixels (10 pixels out of 20,000 total in the hologram). We then initialize 500 walkers with

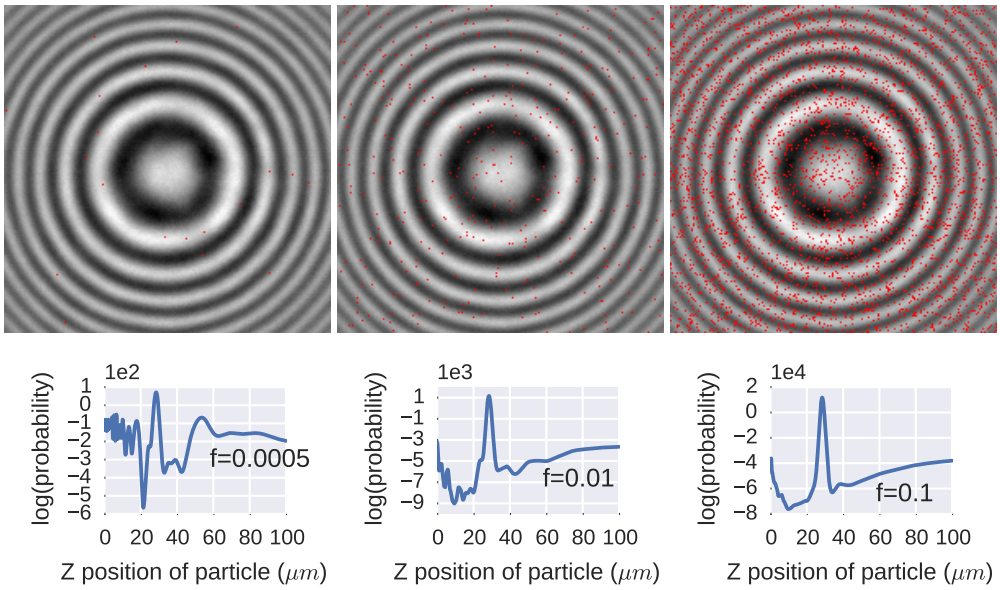

Fig. 4. Comparison of posterior probability distributions of the parameter z, as obtained through MCMC calculations at three subset fractions. Top row shows a hologram of a 1-μm-diameter polystyrene sphere in water (same hologram in all three columns). The red pixels are the randomly chosen pixels used in the calculations. The bottom row shows the posterior probabilities at the different random-subset fractions f. The peak in the probability gets sharper as the fraction increases, but it is present even at the lowest fraction, which represents only 10 pixels in the hologram.

positions sampled from the prior and run the MCMC algorithm for a short period (30 samples after burn-in, which takes 15 seconds or less on a modern CPU at these low fractions). We then calculate the marginalized probability distribution of z based on this MCMC chain, select a Gaussian distribution around the most likely value of z, and use this distribution to select the walker positions for the next iteration at higher fraction. We then repeat this procedure at successively higher fractions. At each stage, we change only the initial distribution of walkers, and not the priors. For holograms of single spheres, we find that three preliminary stages (at $f = 0.00025$, $f = 0.00116$, and $f = 0.00539$) are sufficient to obtain a narrow distribution for z, which we then use for a final run at $f = 0.025$. We also find that the posterior distribution does not change significantly at fractions above $f = 0.025$ for the holograms we analyze in this paper.

To illustrate this approach, we analyze a hologram of a 1-μm-diameter polystyrene sphere in water (index 1.33). Fig. 4 shows that there is a peak in the posterior probability of z even at the lowest pixel fraction. The probability distributions show that the fraction of pixels is analogous to inverse temperature in a tempering scheme; as the fraction increases, the peak in the posterior distribution becomes sharper, and the other modes subside. The prior and posterior distributions at each stage for a 1-μm-diameter silica sphere in water are shown in Fig. 5. The plots show that one can start with a very wide prior in z (a range of 0 to 100 μm above the focal plane, representing nearly complete ignorance about the position of the particle within the sample volume), and refine the uncertainty in the position down to a fraction of the particle size.

This tempering approach allows us to start with very little information and—in under a minute—obtain an excellent starting point for a final MCMC calculation that requires only a short burn-in. We speed up the computation by doing this final run at a fraction $f = 0.025$, for

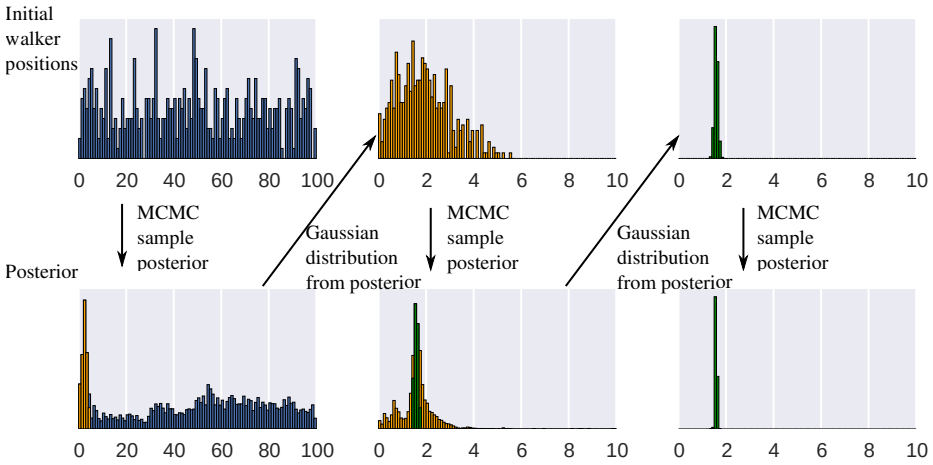

Fig. 5. The random-subset tempering procedure involves refining position estimates at successively higher fractions, starting from coarse guesses. Top row shows the initial walker z positions for each fraction and bottom row the posterior, as obtained from MCMC ensemble sampling. The intial position distribution for the second and third stage are chosen from the peak in the posterior for the previous stage. The final posterior distribution is much narrower than the distribution used for the first stage; note the change in scale in the horizontal axis, which shows the z position in micrometers.

which the calculation is about 40 times faster than that for the full hologram. We use 500 walkers in the final run. Under these conditions, it takes about 20 minutes on a modern desktop CPU to obtain 25000 samples, roughly 5000 of which are independent. Although the method takes longer than least-squares fitting, it does not require a good initial guess for z (or α). Thus there is less manual intervention required to accurately estimate the parameters, and the procedure can more easily be parallelized.

4.2. Particle tracking

With the tempering scheme in hand, we proceed to the particle-tracking problem. We present results from the analysis of a trajectory of a 1-μm silica particle, as described in section 4.1. First, we note that our estimates for the position, as calculated from the median of the posterior obtained for a single frame, are close to those obtained from least-squares fitting, as shown in Fig. 6, in which we have marginalized over n, r, and α. The Bayesian estimates, with boundaries of the 68% credible interval shown in subscript and superscript, are $x = 17.48^{+0.024}_{-0.026}$ μm, $y = 17.64^{+0.027}_{-0.028}$ μm, and $z = 1.56^{+0.077}_{-0.072}$ μm. There is no appreciable covariance between x and y and only slight covariance between the in-plane coordinate (x, y) and the height z. By comparison, a least-squares fit yields $x = 17.48 \pm 0.024$ μm, $y = 17.64 \pm 0.028$ μm, and $z = 1.57 \pm 0.079$ μm, where the uncertainties represent one-sigma confidence intervals, as calculated from the parameter covariance matrix.

In this case, the values and uncertainties for the least-squares fit and the Bayesian approach agree well. However, the Bayesian approach yields an estimate of the entire probability distribution of each parameter, whereas the least-squares method yields only the mean and uncertainty. Also, we note that the uncertainties have different interpretations. The confidence interval from the least-squares fit is a frequentist measure of uncertainty: it tells us that in 68% of a hypothetically large number of identical experiments, the parameter will lie in the computed confidence interval (which can differ for each experiment). The Bayesian credible interval tells us that there is a 68% probability that the true value of the parameter lies in the interval, given the data and the model.

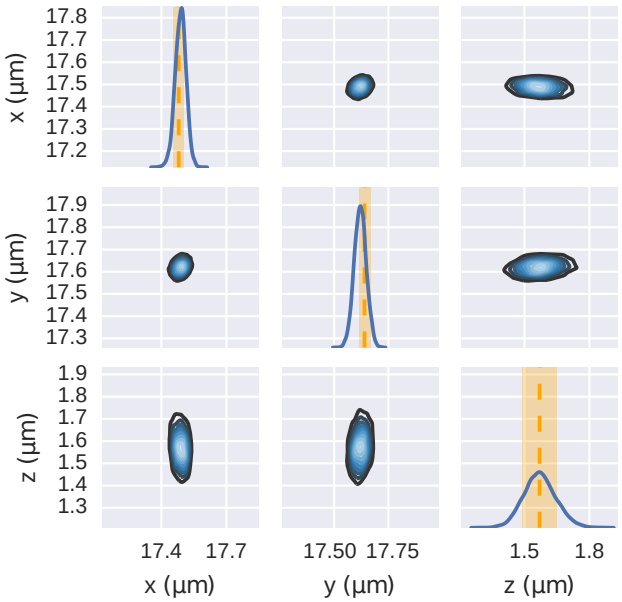

Fig. 6. Results for particle position estimation from a single frame of a time-series. The marginalized posteriors for x, y, and z obtained from our MCMC scheme are shown along the diagonal, while the off-diagonal contour plots show the covariances (Gaussian kernel density estimates). Orange lines are results from least-squares fitting.

With the posterior distributions obtained for each frame in a time-series, we construct the trajectory of the particle along with the uncertainty by sequential MCMC calculations, in which we use the posterior from one frame in the time-series to inform priors for the subsequent frame. Specifically, we approximate the posterior from one frame as a Gaussian and use it as a prior for n, r, and α for the next frame. For x, y, z, we use a Gaussian prior centered on the median of the posterior of the previous frame and assign it a width of 0.1 μm to allow for diffusion. The results of our procedure are shown in Fig. 7.

The uncertainty in the trajectory is important for hypothesis testing or model selection. In this data set the particle is subject to Brownian motion and forces from an optical trap and gravity. In other, similar experiments, one might wish to compare the actual trajectory to an equation of motion derived from a model of the forces. To quantitatively determine whether a particular equation of motion accurately models the data, one needs to know the uncertainty on the trajectory.

4.3. Particle characterization

We apply the same tempered MCMC scheme to the particle characterization problem, using the same data set as for the tracking analysis. We show the results from a single frame and those from 10 frames in Fig. 8. For the 10-frame analysis, we use the posterior for one frame as the prior for the subsequent frame in the time-series, then we combine the MCMC samples from the 10 individual frames. The resulting posterior distribution is significantly narrower (by a factor of about 10 for n and 26 for r), thus yielding much better estimates for the refractive index and radius than could be obtained from a single frame.

The refinement of the particle parameters is important in applications such as particle sizing.

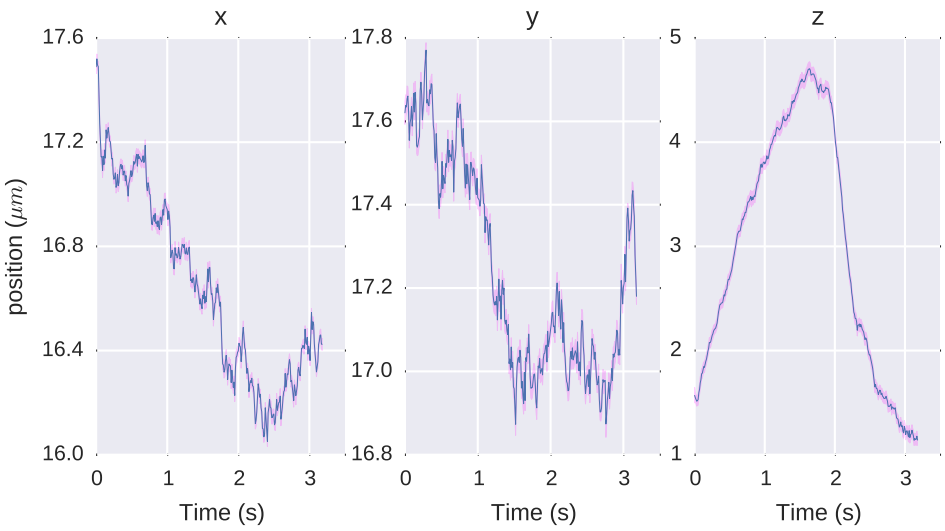

Fig. 7. Particle trajectory inferred from analysis of a 300-frame time-series (3 seconds total). The blue curve shows the position estimate (determined from the median of the posterior) at each frame, and the shaded magenta region indicates the uncertainty (68% credible interval).

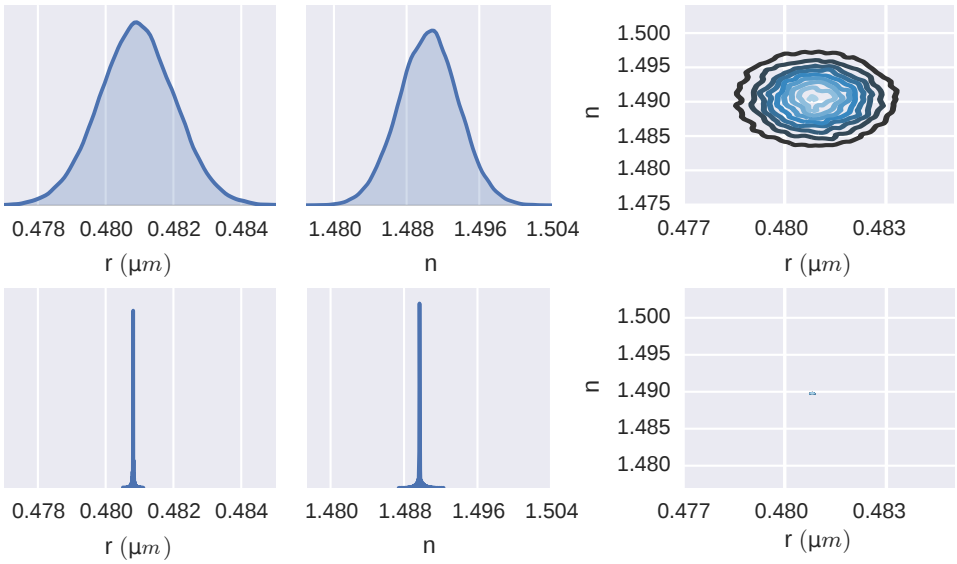

Fig. 8. Posterior distributions of particle properties (refractive index *n* and radius *r*) inferred from analysis of one frame (top row) and ten successive frames (bottom row) of a time-series of holograms. Plots are Gaussian kernel density estimates of MCMC samples.

In holographic particle sizing, the size distribution is measured by combining measurements for individual particles. It is therefore important that the uncertainty on the each measured particle be as small as possible, so that narrow size distributions and multimodal distributions can be resolved. In the 10-frame measurement that we show in Fig. 8, the relative uncertainty in the radius is 0.016%. Thus, as desired, the uncertainty on a single particle is much smaller than the width of even a monodisperse colloidal size distribution.

5. Conclusions

The Bayesian approach we have demonstrated is well-suited to the problem of extracting information from holograms of colloidal particles. One reason is that we understand well how particles scatter light, and there are accurate ways to calculate the scattering. Another is that the technique captures essential details of the scattered field, such as the phase information that is encoded in the interference fringes. All of this means that accurate forward models for hologram formation can be developed from first principles. With a good forward model, MCMC schemes like the one we demonstrate here can be brought to bear on parameter estimation.

The main disadvantage of the Bayesian approach compared to least-squares fitting is that the computations take much longer. This disadvantage is outweighed, in our opinion, by the many advantages of the approach. First, prior information can be specified explicitly at the outset, and need not be spread over initial guesses, constraints, and penalty functions. Second, the tempered MCMC scheme we have developed does not require accurate initial guesses, thus eliminating much of the manual effort required to tune the guesses. Third, the analysis yields the entire joint probability distribution of the parameters, so that the uncertainty can be quantified accurately and marginalized to account for uncertainties in parameters that are not of interest. Thus the method can be applied, as we have shown, to extract realistic uncertainties on the trajectory of a particle or on its properties, taking into account data over an entire time-series.

But we have not demonstrated what could be the most important advantage of the Bayesian approach: its adaptability to other, more complex forward models. In the future, we aim to apply this technique to more complex scattering models such as the superposition solution for multiple spheres, which allows us to infer the structure and dynamics of colloidal clusters [2]. We would also like to model correlated noise (speckle) in the holograms as well as other optical effects arising from propagation along the optical train of the microscope. A more detailed description of noise and imaging effects should allow us to model the hologram more accurately than the current formulation, which lumps these effects into a single parameter, α.

All of these extensions to the forward model will require more parameters, which is problematic for least-squares analysis but much less so for the Bayesian approach. In least-squares techniques, most of the computational time is spent evaluating the Hessian, which scales as the number of parameters squared. Furthermore, adding parameters complicates the chi-squared surface, producing local minima corresponding to unphysical solutions. It becomes more and more difficult to tune the least-squares algorithm as the number of parameters increases, even when there is no risk of overfitting. In contrast, the MCMC techniques used in a Bayesian approach do not rely on gradients and are designed to handle models with hundreds or even thousands of parameters. Also, tempering methods allow one to explore the entire probability distribution, including local maxima. Thus we believe that tempered MCMC approaches like the one we describe in this paper will scale better to more complex models.

Finally, although the approach we describe is by no means simple, it is easy to use because it requires very little tuning. We therefore hope that it will be useful to other researchers considering the use of holographic microscopy. The subset tempering algorithm makes it straightforward for a scientist who is not experienced with holography to make measurements. To facilitate the adoption of these techniques, we have implemented them in our open-source hologram analysis library HoloPy [25].

Funding

We acknowledge support from the National Science Foundation through grant number DMR-1306410.

Acknowledgments

We thank Anna Wang for providing the raw holograms that we analyze in the paper.