Bayesian inference for plasmonic nanometrology

Joseph M. Lukens and Ali Passian

Phys. Rev. A **104**, 053501 — Published 1 November 2021

DOI: 10.1103/PhysRevA.104.053501

Bayesian inference for plasmonic nanometrology

Joseph M. Lukens* and Ali Passian†
Quantum Information Science Group, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA
(Dated: September 8, 2021)

We introduce a Bayesian method for the characterization of plasmonic nanoparticles, which is applicable to both near- and far-field problems. Designed to combine data generated from any photon-plasmon interaction experiment with physically motivated theoretical models, our approach leverages state-of-the-art Markov chain Monte Carlo sampling techniques and returns parameter estimates on nanometric scales. Simulated spectral datasets, describing resonant scattering of photons from ellipsoidal and toroidal nanoparticles, are explored as concrete examples of our approach, with the resulting Bayesian estimates showing excellent agreement with the ground truth, even under conditions of high statistical noise. By incorporating Bayes factors into the method as well, we reveal how model selection can determine which one of competing geometric shapes better explains the observed data. Our comprehensive nanometrology procedure can be tailored to a variety of light-particle interaction models, and its reliance on Bayesian inference furnishes automatic uncertainty quantification. In addition to applicability to a host of plasmonic configurations such as nanoparticle dimers, trimers, and array studies, it is proposed that the presented analysis can be extended to the quantum regime, where nonclassical photon statistics may provide additional insight for inference of scatterer properties.

I. INTRODUCTION

In atomic, molecular, and optical physics, statistical methods are ubiquitous for processing signals associated with photon detection. Light generated via radiative decay of excited states or via absorption and scattering of photons can carry information about the light-emitting or scattering system, which has motivated long-standing studies of inverse problems. For example, in applications based on dynamic light scattering [1], interest rests on inference of particle and aggregate size distributions, leading to a parameter estimation problem. Since validating estimates is difficult without invoking an independent measurement of the parameters, a range of estimates is often desired that can be scrutinized via accepted statistical procedures. In previous studies to this end, a monochromatic beam of light has been scattered from an ensemble of particles in liquid or aerosol medium, with scattered light measured at a given angle as a function of time. Due to stochastic processes (e.g., Brownian motion) in such media, an inhomogeneous Fredholm integral equation results from the time autocorrelation analysis of the measured intensity signal, and through equation inversion, one seeks the particle size distribution. It is generally recognized that such a procedure is mathematically ill-posed, so that there can be many parameter combinations or candidate aggregate size distributions that produce results consistent with the observed data.

To tackle ill-posed problems, several approaches have been proposed, including Bayesian inference as nicely described by Huber *et al.* in the context of characterizing aerosolized fractal aggregates of soot [2]. As a general procedure, Bayesian inference [3] provides a principled and conceptually straightforward approach for estimating any parameters $\mathbf{x}$ of interest given experimental observations $\boldsymbol{\mathcal{D}}$. A probability distribution over unknowns can be defined and updated with experimental observations via Bayes' theorem, $P(\mathbf{x}|\boldsymbol{\mathcal{D}}) = P(\boldsymbol{\mathcal{D}}|\mathbf{x})P(\mathbf{x})/P(\boldsymbol{\mathcal{D}})$, where the distribution $P(\boldsymbol{\mathcal{D}}|\mathbf{x})$ is based on a model of the physical process and $P(\mathbf{x})$ delineates any prior beliefs on the system. By furnishing a complete probability distribution for $\mathbf{x}$, Bayesian inference provides uncertainties automatically, while also revealing any assumptions explicitly via the prior $P(\mathbf{x})$. Moreover, the Bayesian mean estimator of any function $f(\mathbf{x})$, defined as $f_B = \int d\mathbf{x}\, P(\mathbf{x}|\boldsymbol{\mathcal{D}})f(\mathbf{x})$ is optimal in terms of attaining the minimum squared error on average [4].

Bayesian concepts are receiving increasing attention in metrology in general. This is noted for example by Hejazi *et al.* [5], who place their work on parameter estimation for single-wavelength photodetectors within the broader context of data analytics in condensed matter and materials sciences. Another example of how Bayesian inference may aid measurements on physical systems has been contributed by Bera *et al.* [6], who describe estimation of the parameters that control the dynamics of an optically trapped Brownian particle. In the work of Aihara *et al.* [7], the physical system on which measurements were carried out took the form of an evaporated bismuth thin film, and Bayesian estimates were obtained for the parameters characterizing the transient profile of coherent phonons. Similarly, a thin film superlattice was characterized via Bayesian analysis of the measured photoluminescence spectrum in the work of Iwamitsu *et al.* [8]. In a different direction, Bayesian optimization techniques [9] have also been applied in the design of subwavelength structures with engineered scattering patterns [10, 11]. Here, Bayes' theorem facilitates efficient search and optimization over a complicated design space, rather than

* lukensjm@ornl.gov
† passianan@ornl.gov

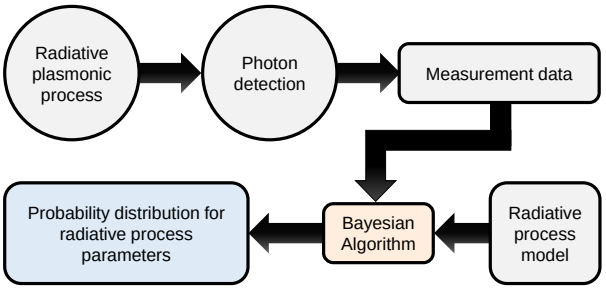

FIG. 1. Simplified workflow for the Bayesian inference of plasmonic nanoparticle properties.

being used for the inference of unknown parameters from experimental data.

In this work, we are interested in Bayesian inference associated with light scattering from plasmonic nanoparticles, as broadly depicted in Fig. 1. When an electron or photon interacts with a nanoparticle, under proper conditions of field polarization, wavelength, and momentum, as well as nanoparticle geometric shape, surface morphology, and dielectric properties, resonant photonic modes can be excited in the surface or bulk regions of the particle. In the case of metallic nanoparticles, these modes are associated with collective electronic oscillations, where the minimum involved energy is called a plasmon. The decay of these modes can occur radiatively or nonradiatively, with the former leading to emission of a photon and the latter resulting in nanoparticle heating [12]. Thus, calculation of emission probabilities and experimental measurement of photons are important in the nanometrology of the emitting particles.

In this Article, we extend the applicability of Bayesian inference to plasmonic systems in which field enhancement, energy confinement, and strong coupling to emitters are known to play important roles in the radiative processes. Leveraging advanced Markov chain Monte Carlo (MCMC) methods and able to incorporate any first-principles, phenomenological, or empirical model of interest, we test the method on simulated experimental datasets from ellipsoidal and ring-toroidal gold nanoparticles, demonstrating accurate estimation of specific geometric parameters. As an important additional capability, we also compute the ratio of posterior probabilities from alternative geometrical models, finding convincing evidence in favor of the ground truth in all datasets examined and thereby indicating our method's utility for both parameter estimation and model selection. While we concentrate explicitly on the estimation of geometric properties from nanoparticle scattering data, the presented approach is applicable to many other situations of interest in plasmonics, such as plasmonic tomography [13], qubit-particle interactions [14], and surface plasmon resonance sensors [15], thus providing a valuable complementary capability to standard near- or far-field techniques in nanometrology.

We have organized our presentation as follows. Section II briefly introduces how plasmon modes may be computed in a broad spectral band for an arbitrarily shaped nanoparticle of a given material. This section provides the needed background for the models to be used later in inference. In Sec. III, we overview the applied Bayesian method, including details of the MCMC sampling procedure that enables efficient numerical evaluation. Section IV describes the nanoparticle geometries considered for simulated datasets, with Sec. V summarizing the results obtained from these with Bayesian inference. Section VI considers the question of model selection, showing how Bayes factors can identify the more probable of competing models given observations. Concluding remarks are provided in Secs. VII and VIII.

II. NANOPARTICLE PLASMONICS

Light-matter interaction [16] in the form of photon scattering from nanoparticles [17] plays a central role in important applications ranging from chemical and biological sensing [18] to emerging experiments in quantum sensing [19]. With the development of new light sources (e.g., single photons [20]), as well as with the discovery of increasingly sensitive nanometrological techniques (e.g., scattering near-field scanning optical microscopy [21]) that can probe the electromagnetic environment of the nanoparticles, studies of coupled photon-particle systems are gaining further momentum. These include ion trapping [22], quantum plasmonics [23], and qubit entanglement control [24], to mention a few.

For nanoparticles that support resonant surface modes, significant absorption and inelastic scattering occurs, which may be revealed from a broadband excitation of the modes with specific polarization states. Under optimum resonance conditions, high degrees of field confinement may occur in the vicinity of the nanoparticles and their sub- and super-strates. Such electromagnetic environments can be probed via nearfield scattering to provide information that may not be recoverable in the farfield. Farfield detection of the scattered photons, however, offers the advantage of noninvasiveness, simplifying the light-particle interaction model. For example, in tip-enhanced or scattering-type nearfield measurements, any model must account for the convolution of the particle nearfield with the scattering tip. Statistical methods that can help deconvolute the data and infer information regarding the scatterer are therefore important.

Consider a typical measurement scenario for characterizing both radiative decay and nanoparticle attributes, as depicted in Fig. 2. The polarization state of an incoming beam of photons of flux Φ is controlled with a linear polarizer and a polarization rotator, and the beam is directed toward a single or a many-nanoparticle system. Although the probability of a photon-particle plasmon interaction does change with the particle's *geometric* cross-section, the differential *scattering* cross-section $d\sigma$,

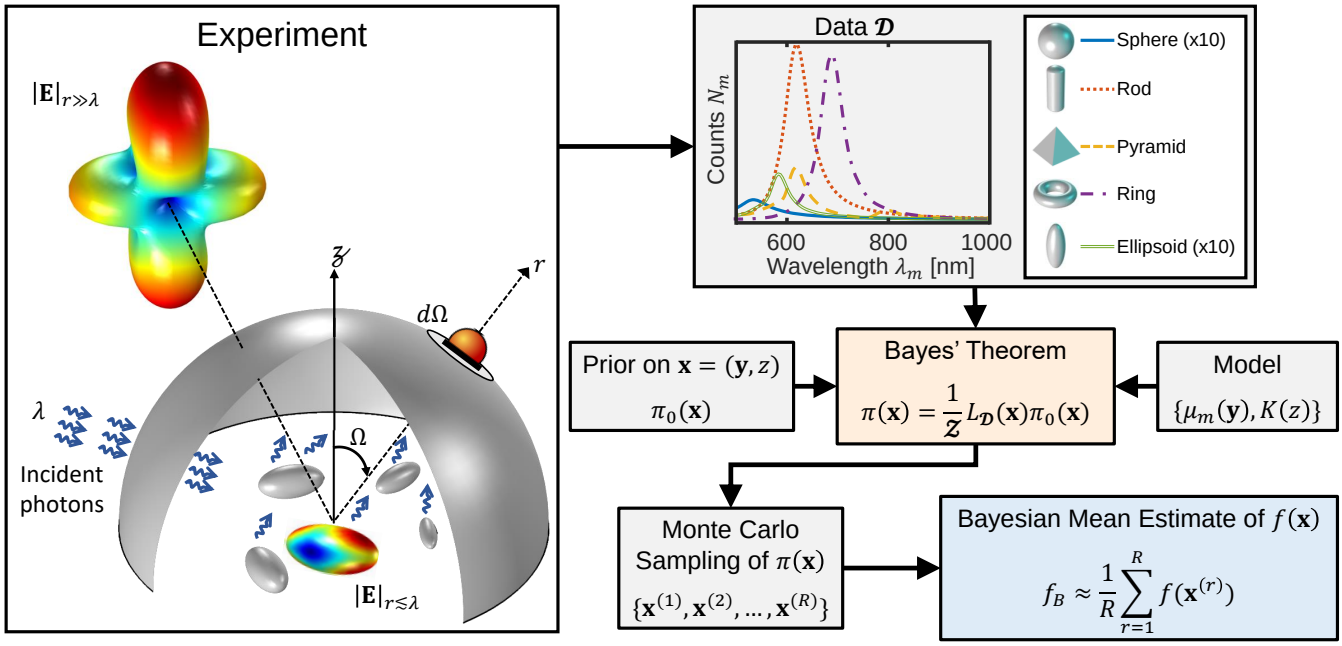

FIG. 2. Overview of plasmonic inference problem. Light emitted from a plasmonic system is detected in the near- or far-field regime, giving a dataset $\mathcal{D}$ of counts. Bayes' theorem combines this data with a model and prior information to define a probability distribution on the unknown parameters. Through Monte Carlo techniques, samples are obtained to estimate any function of interest. As visual examples, we show the norm of the near-field electric field distribution and the far-field (radiation pattern) computed at $\lambda = 400$ nm for a circular ellipsoid with major (minor) radius of 45 nm (10 nm). Example cross-sections for five nanoparticle shapes highlight the types of data expected. Other dimensions assumed were: 50 nm radius for the sphere; 25 nm radius and 100 nm height for the rod; 100 nm $\times$ 100 nm base and 100 nm height for the pyramid; and 40 nm (21 nm) major (minor) radius for the ring.

that is, the ratio of the scattered power at $d\Omega$ to Φ, is the quantity of interest since it properly accounts for the resonance properties of the excited surface modes, which are known to lead to significantly higher scattering than could be accounted for from geometric effects alone. Following this interaction with the system under test, subsequent plasmon excitation, and radiative decay, the emanating photons may be detected with a properly positioned photodetector at the apex of the solid angle $d\Omega$. In the left panel of Fig. 2, the plasmon excitation is exemplified for a single ellipsoidal nanoparticle, where the norm of the induced surface field is visualized. The displayed radiation pattern of the excited mode is obtained from near- to far-field transformation. The excited plasmon mode and its radiation pattern were obtained computationally in the frequency domain using the finite element method (FEM) [25]. To generate the datasets $\mathcal{D}$, shown in the right panel of Fig. 2, we employ the finite difference time domain (FDTD) method, as will be discussed in Sec. IV. Datasets $\mathcal{D}$ describe the photon scattering cross-sections for various nanoparticles, as annotated in Fig. 2.

Following a single $d\Omega$ measurement, the detector position may be changed incrementally in Ω to cover the upper half-space above the substrate. Accordingly, for each wavelength, the end result is a map of the angular distribution of the emitted photons, which can be kept as a function of Ω or integrated over all angles of interest (as in the examples in Figs. 2 and 4). We note that, instead of wavelength scanning, one may choose to probe with a broadband pulse and measure the transient response of the nanoparticles, performing an inverse Fourier transform to acquire the emission spectrum for a single angular position of the detector.

III. OVERVIEW OF BAYESIAN PLASMONIC CHARACTERIZATION METHOD

A. Preliminaries

As our prototypical physical system, we consider nanoparticles of various materials and morphologies, such as the well-studied systems of gold and silver spheroids. When performing measurements on the proposed physical system, a single event is defined as the detection of a photon of a particular wavelength and polarization state at a given point in the near- or far-field. The collection of all such events can be summarized as a dataset $\mathcal{D}$. In order to relate this dataset to the unknown parameters of interest—which can be defined generically as a vector $\mathbf{x}$—one must invoke a physical

model that specifies $P(\boldsymbol{\mathcal{D}}|\mathbf{x})$, i.e., the probability of observing the data $\boldsymbol{\mathcal{D}}$ given the parameters $\mathbf{x}$. Lastly, we define some prior distribution $P(\mathbf{x})$ that encompasses all beliefs about the system's parameters before performing experiments. $P(\mathbf{x})$ can be chosen as informative or as uninformative as desired to reflect the knowledge and objectives of the inference procedure. These three elements—data, model, and prior—then are united into a single probability distribution according to Bayes' theorem $[P(\mathbf{x}|\boldsymbol{\mathcal{D}}) = P(\boldsymbol{\mathcal{D}}|\mathbf{x})P(\mathbf{x})/P(\boldsymbol{\mathcal{D}})]$. This conditional probability of parameters $\mathbf{x}$, given observations $\boldsymbol{\mathcal{D}}$, formally solves the problem.

Yet, the theoretical simplicity of Bayes' rule belies the computational challenges associated with utilizing it in practice. For all but the simplest models, integrals of the form $f_B = \int d\mathbf{x}\, P(\mathbf{x}|\boldsymbol{\mathcal{D}}) f(\mathbf{x})$ cannot be performed analytically, and numerical approaches are hindered by the often high-dimensional nature of $\mathbf{x}$. However, if samples of $P(\mathbf{x}|\boldsymbol{\mathcal{D}})$ can be obtained reasonably efficiently, then the Bayesian mean estimator f_B can be computed directly. Monte Carlo techniques offer the possibility to obtain such samples, and the efficiency of these methods has improved significantly in recent years, opening a valuable opportunity for the application of Bayesian estimation in plasmonic characterization.

B. Formal problem formulation

Returning to the scenario of Fig. 2, we can represent the unknown geometric properties of the system under test by a vector of parameters $\mathbf{y}$. Then, for a particular measurement $m \in \{1, 2, ..., M\}$, this system is illuminated by a flux $\Phi_m(\lambda, q)$: the number of incident photons per unit time, area, and wavelength, with $q \in \{1, 2\}$ denoting possible projections onto two orthogonal polarization states. The scattered photons are detected according to an efficiency function $g_m(\lambda, q, \Omega) \in [0, 1]$, which specifies the probability of successfully detecting a scattered photon at wavelength λ, polarization q, and solid angle Ω. Both Φ_m and g_m are free to vary with m, reflecting a tunable excitation source or tunable measurement device, respectively.

The core of the theoretical model centers on the scattering cross-section (though it may be formulated for absorption cross-section as well) of the plasmonic nanostructure, expressed in differential form as $\partial\sigma(\mathbf{y}, \lambda, q, \Omega)/\partial\Omega$. We will assume a spectral region $\lambda \in \Lambda$, with Λ sufficiently broad to recover the resonance features of the nanoparticle excitation. Furthermore, as depicted in Fig. 2, we assume that the emitted photons can be probed only from the upper ($\mathfrak{z} > 0$) half-space encompassing angles Ω^+, which is reasonable for substrate-bound nanoparticles or nanostructures, although radiative plasmon decay can also be observed from the substrate domain, for example when used in prism-based experiments. The rate of detections for measurement setting m can then be written as:

$$\tilde{\mu}_m(\mathbf{y}) = \sum_q \int_\Lambda d\lambda \int_{\Omega^+} d\Omega\, \Phi_m(\lambda, q) \frac{\partial\sigma(\mathbf{y}, \lambda, q, \Omega)}{\partial\Omega} g_m(\lambda, q, \Omega). \tag{1}$$

This expression utilizes detailed information from all aspects of the experiment—source, nanoparticles, and detection—so that the absolute cross-section of the structure under test can in principle be harnessed to aid in inference. Experimentally, however, the level of calibration required for absolute cross-section determination can be quite difficult to attain. Uncertainties in the spot size of the probe beam, collection efficiency, and detector responsivity must all be minimized. While techniques such as spatial modulation spectroscopy [26–28] or calibration with a known reference nanostructure [29, 30] can be used to determine absolute efficiencies in practice, the majority of plasmonic scattering experiments provide results on a relative scale.

Accordingly, we remove the need to independently establish the absolute efficiencies in $\tilde{\mu}_m(\mathbf{y})$ by introducing a constant scale factor $K(z)$ that is a function of an additional to-be-inferred parameter z. Then, by defining $\tilde{\mu}_m(\mathbf{y}) = K(z)\mu_m(\mathbf{y})$, where $\mu_m(\mathbf{y})$ is normalized to sum to unity via

$$\mu_m(\mathbf{y}) = \frac{\tilde{\mu}_m(\mathbf{y})}{\sum_{m'=1}^{M} \tilde{\mu}_{m'}(\mathbf{y})}, \tag{2}$$

any reference to absolute efficiencies is avoided. Importantly, since $K(z)$ is of the same order as the total number of detection events, an initial estimate can be found easily from the data, making it a relatively straightforward addition to the Bayesian procedure. Thus, the experimental challenge of system efficiency calibration can be eliminated by incurring one extra parameter, extending the full parameter vector to

$$\mathbf{x} = (\mathbf{y}, z). \tag{3}$$

Of course, with this simplification, inference based on overall scatterer efficiency is no longer possible; two systems having cross-sections that differ by a constant scale factor ($\sigma_A \propto \sigma_B$), cannot be distinguished. Yet, this ambiguity reflects the experimental conditions we consider, and not the inference procedure, which could be adapted to the case of absolute rates by working with $\tilde{\mu}_m(\mathbf{y})$ directly.

We consider M experimentally obtained photon counts N_m, making up the dataset:

$$\boldsymbol{\mathcal{D}} = \{N_1, N_2, ..., N_M\}. \tag{4}$$

For each measurement m, the mean photon count according to the model is $K(z)\mu_m(\mathbf{y})$, making up the set:

$$\mathcal{M} = \{K(z)\mu_1(\mathbf{y}), K(z)\mu_2(\mathbf{y}), ..., K(z)\mu_M(\mathbf{y})\}. \tag{5}$$

We utilize a Poissonian model with means $\mathcal{M}_m = K(z)\mu_m(\mathbf{y})$, where each measurement is assumed independent of the others. In this way, the likelihood $L_{\boldsymbol{\mathcal{D}}}(\mathbf{x}) \propto P(\boldsymbol{\mathcal{D}}|\mathbf{x})$ can be written as:

$$L_{\boldsymbol{\mathcal{D}}}(\mathbf{x}) = e^{-K(z)} \prod_{m=1}^{M} \mathcal{M}_m^{N_m}. \tag{6}$$

Bayes' theorem for the posterior $\pi(\mathbf{x}) \equiv P(\mathbf{x}|\boldsymbol{\mathcal{D}})$ can then be expressed in the convenient form

$$\pi(\mathbf{x}) = \frac{1}{\mathcal{Z}} L_{\boldsymbol{\mathcal{D}}}(\mathbf{x})\pi_0(\mathbf{x}), \tag{7}$$

where $\mathcal{Z}$ is a normalizing constant and $\pi_0(\mathbf{x})$ the prior distribution on $\mathbf{x}$. Equation (7) is equivalent to the standard formula $P(\mathbf{x}|\boldsymbol{\mathcal{D}}) = P(\boldsymbol{\mathcal{D}}|\mathbf{x})P(\mathbf{x})/P(\boldsymbol{\mathcal{D}})$, yet we have chosen the $\pi(\mathbf{x})$ notation to emphasize that only the functional dependencies on $\mathbf{x}$ need be considered in the MCMC approach below, and the likelihood $L_{\boldsymbol{\mathcal{D}}}(\mathbf{x})$ has no requirement for normalization like a true probability distribution $P(\boldsymbol{\mathcal{D}}|\mathbf{x})$.

In lieu of direct integration, the inference procedure obtains R samples $\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(R)}\}$ from the distribution $\pi(\mathbf{x})$ such that the Bayesian mean estimator of any function $f(\mathbf{x})$ can be computed as:

$$f_B = \int d\mathbf{x}\, \pi(\mathbf{x}) f(\mathbf{x}) \approx \frac{1}{R} \sum_{r=1}^{R} f(\mathbf{x}^{(r)}). \tag{8}$$

To proceed, we must therefore consider a strategy for Monte Carlo sampling of $\pi(\mathbf{x})$ (next step in Fig. 2).

C. Preconditioned CrankNicolson algorithm and numerical procedure

Efficient methods to generate the samples $\mathbf{x}^{(r)}$ comprise a longstanding research program in Bayesian statistics [3, 4] and historically have posed a significant entry barrier to the adoption of Bayesian methods in a variety of fields. Markov chain Monte Carlo (MCMC) techniques attack this problem by implementing a stochastic process such that samples therefrom converge to the target distribution $\pi(\mathbf{x})$ for sufficiently long chains. Some common MCMC algorithms include Metropolis–Hastings [31, 32], Gibbs [33], slice sampling [34], hybrid Monte Carlo [35], and sequential Monte Carlo [36], each facing unique tradeoffs in its implementation. For our inference procedure, we adopt the relatively recent MCMC algorithm known as preconditioned Crank–Nicolson (pCN) [37], a modification to Metropolis–Hastings designed specifically for dimension-independent performance.

Consider a Markov chain in state $\mathbf{x}$; a possible next state $\mathbf{x}'$ is generated according to the proposal distribution $q(\mathbf{x}'|\mathbf{x})$. Metropolis–Hastings will accept $\mathbf{x}'$ with probability

$$A(\mathbf{x}', \mathbf{x}) = \frac{\pi(\mathbf{x}')}{\pi(\mathbf{x})} \frac{q(\mathbf{x}|\mathbf{x}')}{q(\mathbf{x}'|\mathbf{x})}. \tag{9}$$

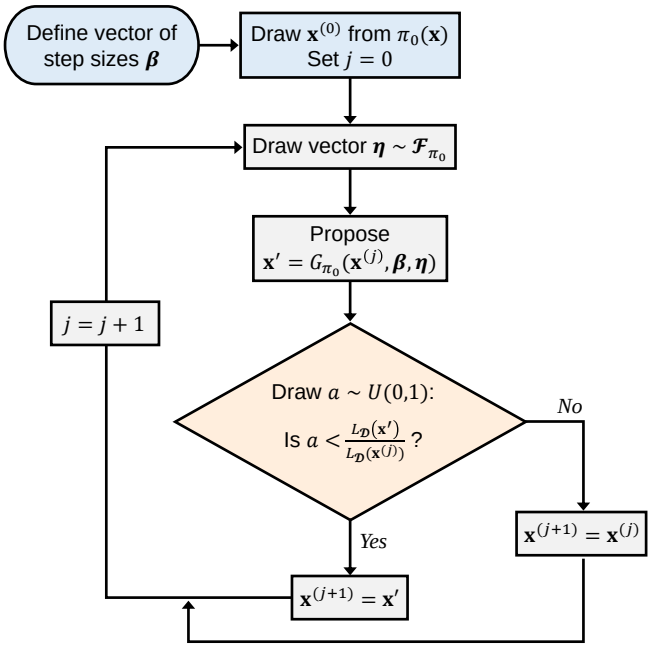

FIG. 3. Outline of steps required to generate samples $\mathbf{x}^{(j)}$ from the posterior $\pi(\mathbf{x}) \propto L_{\boldsymbol{\mathcal{D}}}(\mathbf{x})\pi_0(\mathbf{x})$. The unspecified distribution $\mathcal{F}_{\pi_0}$ and function G_{π_0} are assumed to be chosen such that the acceptance probability reduces to a ratio of likelihoods only (pCN condition).

In the case of pCN, the proposal distribution is chosen such that it is reversible with respect to the prior, i.e., it satisfies $q(\mathbf{x}|\mathbf{x}')\pi_0(\mathbf{x}') = q(\mathbf{x}'|\mathbf{x})\pi_0(\mathbf{x})$; thus, from Eq. (7) the acceptance probability reduces to $A(\mathbf{x}', \mathbf{x}) = \frac{L_{\boldsymbol{\mathcal{D}}}(\mathbf{x}')}{L_{\boldsymbol{\mathcal{D}}}(\mathbf{x})}$. This simplification can faciliate massive speed-ups in MCMC convergence, eliminating the step-size/acceptance rate tradeoff intrinsic to standard random-walk proposals [37, 38]. We recently applied pCN methods to develop a highly efficient Bayesian procedure for quantum state tomography [39], and here we expand these concepts to the types of likelihoods and priors encountered in plasmonic nanometrology.

The basic numerical procedure is outlined in the flowchart in Fig. 3. The vector $\boldsymbol{\beta}$ scales the step sizes for all parameters, while the randomly drawn $\boldsymbol{\eta}$ defines the individual jumps. The acceptance decision is made probabilistically by drawing $a \in [0, 1]$ from a uniform distribution and comparing to the likelihood ratio. In this flowchart, we have expressed the procedure as generically as possible, leaving distribution $\mathcal{F}_{\pi_0}$ and function G_{π_0} unspecified. Algorithmically speaking, one must determine the combination $(\mathcal{F}_{\pi_0}, G_{\pi_0})$ carefully such that reversibility with respect to the prior is maintained $[q(\mathbf{x}|\mathbf{x}')\pi_0(\mathbf{x}') = q(\mathbf{x}'|\mathbf{x})\pi_0(\mathbf{x})]$. For reference, we list known pCN proposals for three common prior distributions—normal, gamma, and uniform—in the Appendix, which can be applied to many situations encountered in practice, including the examples in the following

sections.

Computationally speaking, each step in the chain requires evaluation of $L_{\boldsymbol{\mathcal{D}}}(\mathbf{x}')$ for a new point $\mathbf{x}'$, so it is critical to be able to compute this efficiently with minimal cost. Therefore, rather than performing the integral expressed in Eq. (1) in real-time for each proposed point, an analytical or approximate expression would be preferred. If not available, we suggest precomputing $\mu_m(\mathbf{y}_n)$ for a sufficiently dense set of points $\mathbf{y}_n$, leaving the results in the form of a look-up table which can then be used to interpolate the value $\mu_m(\mathbf{y})$ at any given $\mathbf{y}$. This approach is particularly well suited to the use of computational electrodynamic methods, which can be executed to model the variety of configurations expected in the prior *before* Bayesian inference.

Together, Eqs. (1-8) and Fig. 3 complete the Bayesian plasmonic inference method we introduce here. In what follows, we apply this approach to several test cases, specializing the general framework to realistic experimental scenarios.

IV. EXAMPLE DATASETS

Except for nanoparticles with simple geometries, analytical calculation of elecrodynamic quantities is either not feasible or requires high degrees of approximations. Thus, numerical approaches are indispensable. To compute the scattering cross-sections we use the FDTD method [40], a versatile platform for solving partial differential equations [41]. Within the framework of FDTD, one defines a computational domain with sufficient spatial and temporal resolution, the proper and proportional choices of which assure numerical stability and minimal numerical dispersion. In the standard implementation of FDTD, one considers a rectilinear 3D mesh, where Maxwell's equations are discretized via approximation of the involved differential operators by finite differences, which can be developed to various orders of accuracy. Though FDTD is typically formulated in rectilinear coordinates, general coordinate systems also can be explored, albeit at the cost of additional complexity pertaining to discretization of the differential operators in curvilinear coordinates. For example, a spherical FDTD domain allows for natural modeling of particles with spherical symmetry, but particles without spherical symmetry will lack smooth boundaries. So rather than tailoring coordinate systems for each simulation to match the geometry under test, we consider the standard Cartesian 3D mesh in all cases.

The computational domain can be formulated to generate either the total field or the scattered field. Such total-field scattered-field formulation offers a number of advantages for plane-wave nanoparticle excitation. Various nanoparticles may then be modeled by occupying the computational domain such that each object will have (staircase) boundaries resulting from a voxelization (volume pixeling) process, which can be improved, especially for high curvature surfaces, using the technique of conformal meshing. The field components are then properly staggered across the unit cell (Yee cell) so as to satisfy Maxwell's equations. Each mesh point is ascribed a dielectric function which in our case is imported from Johnson and Christy's compilation for gold [42]. Unlike FEM-based methods, in FDTD the time domain representation of the dispersion, as opposed to the Fourier domain representation $D(\omega) = \epsilon(\omega)E(\omega)$, is implemented (typically, assuming a local dielectric function). An ultrashort excitation pulse covering the spectral range of 300–1000 nm, and a sufficiently dense sampling of frequencies, generates the scattered nearfields, which can then be transformed to the farfield via proper integration. Here, we compute the nearfields and generate the needed scattering cross-sections for a given excitation field polarization state.

The spatial resolution or mesh cell size δ_i $(i = 1, 2, 3)$ can be quantified with respect to the free-space wavelength λ. The time increments for evolving the fields will then be set, obeying the (Courant) stability criterion [40] $\Delta t \leq \left(\mu\epsilon/\sum_i \delta_i^{-2}\right)^{1/2}$, where $\mu = 1$ for nonmagnetic materials and ϵ is the permittivity. In our Λ broad simulation, the smallest wavelength is $\lambda = 300$ nm. To adequately resolve the fields, the (largest) mesh cell $\delta_i \lesssim a\lambda$, where a is a suitably small number, typically in the range $a \in [0.05, 0.1]$. For a more precise estimate, the dielectric properties of the materials involved can be taken into account. In such cases, $a \in [0.05, 0.1]/n_{\max}$, where $n_{\max}$ is the largest refractive index in the computational domain. Therefore, a sufficient number of mesh points both per λ and per plasmon decay length can be obtained. Since the skin depth for gold over Λ can be as small as $\sim$2 nm, cell sizes of $\delta_i \lesssim 2$ nm were employed to compute the fields inside the nanoparticles.

As specific examples, we consider an ellipsoid with minor radius of 10 nm and a ring with major radius 40 nm. These nanostructures can be conveniently modeled in three dimensions using available software packages [43]. By performing parametric studies involving the radius of each structure made of a given material, the wavelength, polarization, and incidence angle of the exciting field, the scattering cross-sections can be readily calculated from the computed nearfields, as shown in Fig. 4. Each is integrated over all angles in the upper hemisphere and normalized by the corresponding geometric cross-section to obtain the dimensionless "effective" cross-section σ_e. Since no resonances appear in the wavelength range 300–500 nm for the considered parameter ranges, we plot the cross-section over the domain $\Lambda = [500, 1000]$ nm. The parameters scanned in these studies—major radius of ellipsoid, denoted by $\mathcal{Y}$, and minor radius of ring, denoted by $\mathcal{Y}'$—represent the geometric features to be estimated from Bayesian inference; the other characteristic dimensions in each system are assumed fixed and known. Figure 4 presents, within the considered spectral and shape parameter windows, the complete plasmon dispersion for the nanoparticles without resorting to dipole or quasi-static approximation, for 50 evenly spaced values of $\mathcal{Y}$

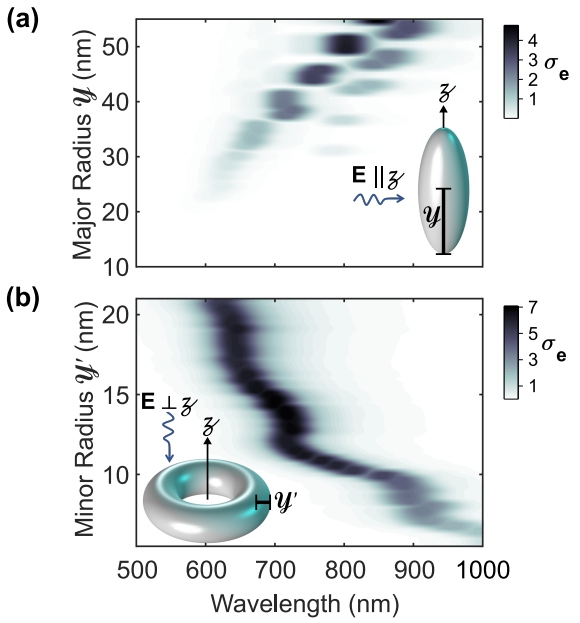

FIG. 4. Models to be employed in Bayesian inference. The effective scattering cross-sections σ_e for (a) gold ellipsoidal and (b) gold toroidal nanoparticles are obtained from registering the fields in the nearfield zone of the nanoparticles. The polarization states $\boldsymbol{E}_{\perp\mathcal{Z}}$ and $\boldsymbol{E}_{\parallel\mathcal{Z}}$ of the interacting photons are shown in the insets along with the characteristic nanoparticle dimensions $\mathcal{Y}$ and $\mathcal{Y}'$.

and $\mathcal{Y}'$. The island-like appearance in Fig. 4(a) reflects the fact that the computed solution contains contributions from all plasmonic modes. As the shape parameters change, the resonances of each mode vary at different rates, creating the observed alternating pattern.

Using the cross-section calculations in Fig. 4, we simulate scattering experiments for a variety of ground truth geometric parameters: major radii of $\mathcal{Y}_g \in \{10, 15, ..., 55\}$ nm for the ellipsoid, and minor (tube) radii of $\mathcal{Y}'_g \in \{6, 7, ..., 20\}$ nm for the ring. In what follows, the subscript g always denotes a ground truth value, to distinguish it from a result of Bayesian inference (subscript B). The computed cross-sections correspond to a fixed linear excitation polarization q and integration over a farfield hemisphere, so that the detection rate in Eq. (1) reduces to an integral over λ only:

$$\tilde{\mu}_m(\mathbf{y}) = \int_\Lambda d\lambda\, \Phi_m(\lambda)\sigma(\mathbf{y},\lambda)g_m(\lambda). \tag{10}$$

We assume each measurement setting m is centered at wavelength λ_m such that $\Phi_m(\lambda)g_m(\lambda) = h(\lambda - \lambda_m)$, where $h(\lambda)$ is a singly peaked positive function centered at $\lambda = 0$. If $\sigma(\mathbf{y},\lambda)$ is slowly varying with λ compared to $h(\lambda)$—the natural condition for a high-resolution measurement—we further obtain $\tilde{\mu}_m(\mathbf{y}) \propto \sigma(\mathbf{y},\lambda_m)$, so that the normalized theoretical rates are

$$\mu_m(\mathbf{y}) = \frac{\sigma(\mathbf{y},\lambda_m)}{\sum_{m'=1}^{M}\sigma(\mathbf{y},\lambda_{m'})}, \tag{11}$$

which can be computed for any $(\mathbf{y},\lambda_m)$ via interpolation of the results in Fig. 4. For simplicity in these examples, the spectral shape and relative efficiency of the flux/-collection combination are taken as uniform over m, but experimental variations could be readily incorporated via an m-dependent scale factor in Eq. (10).

We generate random data for $M = 100$ equispaced wavelengths in 500–1000 nm from Poisson distributions of mean $K_g\mu_m(\mathbf{y})$: the mean number of detected photons, chosen as $K_g \in \{10^2, 10^3, 10^4\}$, allows us to explore the impact of statistical noise. Considering the 10 ground truth values of $\mathcal{Y}_g$ and 15 values of $\mathcal{Y}'_g$ noted above, this produces 75 distinct datasets $\boldsymbol{\mathcal{D}}$, one for each $(\mathcal{Y}_g, K_g)$ or $(\mathcal{Y}'_g, K_g)$ pair; Fig. 5 plots six of these as examples, three each for the ellipsoid and ring. Appreciable statistical fluctuations appear in the $K_g = 100$ cases, which reduce significantly as K_g increases. The two ring examples with $\mathcal{Y}'_g = 12$ nm and $\mathcal{Y}'_g = 14$ nm have nearly identical scattering spectra. This effect can be seen directly in Fig. 4(a), which contains two intervals (roughly 12–14 nm and 17–19 nm) over which the cross-section varies minimally with $\mathcal{Y}'_g$. This offers an important opportunity for comparison, since inference of the tube radius would be expected to prove more challenging in these regions than others.

We have selected this case study in order to focus on the basics of the method with minimal technical distractions that may arise when studying more complex cases. For example, considering angle-resolved detection, as allowed in the general expression of Eq. (1), would lead to more complexities in the calculation of $\mu_m(\mathbf{y})$ but would also provide additional information that should reduce uncertainty in the inference process. As discussed in Sec. VII, no fundamental limitations prevent our method from handling significantly more complex systems and measurement conditions such as these, but the feasibility of any specific model will require its own investigation.

V. BAYESIAN INFERENCE

Armed with the datasets simulated above, we next apply the full method of Sec. III for Bayesian inference. This entails setting up a prior $\pi_0(\mathbf{x})$ and then calculating posterior probabilities $\pi(\mathbf{x})$ based on the prior and likelihood $L_{\boldsymbol{\mathcal{D}}}(\mathbf{x})$, as indicated in Fig. 2. Through this process, the prior probabilities are updated via data collection $\boldsymbol{\mathcal{D}}$.

The plasmonic geometries considered contain a single unknown, $\mathcal{Y}$ ($\mathcal{Y}'$), which we express in terms of the dimensionless parameter $y \in [-1, 1]$ as:

$$\mathcal{Y}(y) = \frac{(y+1)\mathcal{Y}_H - (y-1)\mathcal{Y}_L}{2}, \tag{12}$$

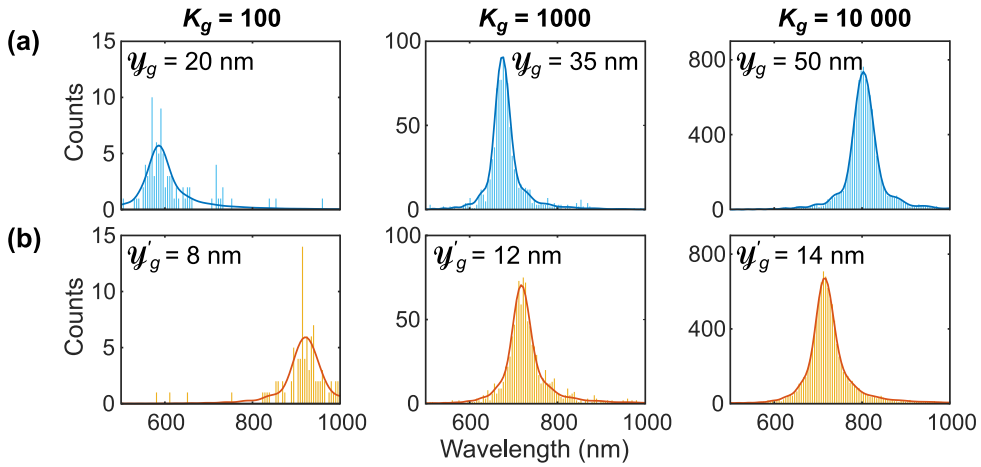

FIG. 5. Example datasets $\mathcal{D}$ from simulated experiments. (a) Scattering from ellipsoidal nanoparticle. (b) Scattering from ring. Histograms show the raw count data; solid lines provide the ground truth cross-section for the given geometry and counts.

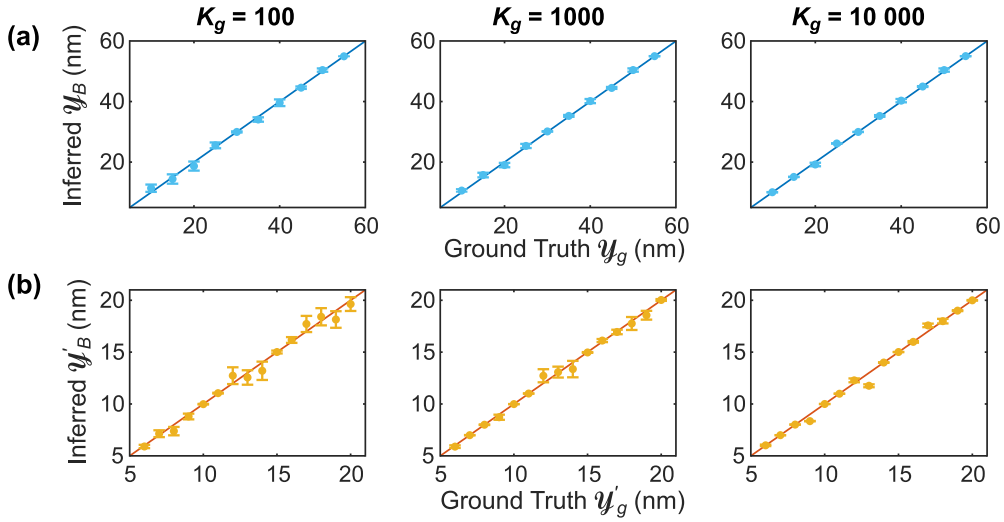

FIG. 6. Bayesian inference results for datasets from the (a) ellispoid and (b) ring. Points are the mean of the Bayesian posterior, and error bars the standard deviation. Each column denotes a specific ground truth average photon number K_g.

and analogously for $\mathcal{Y}'$. Here, $\mathcal{Y}_H$ ($\mathcal{Y}_L$) is the maximum (minimum) value in the simulated domain of Fig. 4. The scale factor K is parameterized as $K(z) = K_0(1 + kz)$, with K_0 and k constant. On this parameter set $\mathbf{x} = (y, z)$, one must then specify a prior distribution $\pi_0(\mathbf{x})$. In lieu of specific information which may heavily favor some regions of the parameter space over others, it is generally good practice to define a prior that is as "uninformative" as possible, giving appreciable weight to all possible parameter values. This results in a relatively conservative estimation procedure that returns low uncertainties only when justified by the amount of data gathered. In our specific case, we take y uniformly distributed in its domain $[y \sim \mathcal{U}(-1, 1)]$ and z following a standard normal distribution $[z \sim \mathcal{N}(0, 1)]$ so that the hyperparameters K_0 and k specify the mean and standard deviation of an arbitrary normal distribution on K. As an empirically convenient choice, we select $K_0 = \sum_m N_m$ (the sum over all observed counts) and $k = 0.1$ [44], which we have found gives a sufficiently uninformative prior for the numbers of counts in these test datasets. Thus, the total prior is $\pi_0(\mathbf{x}) \propto \mathbf{1}_{[-1,1]}(y)e^{-z^2/2}$, where $\mathbf{1}_{[-1,1]}(\cdot)$ is the indicator function. Both y and z distributions have pCN-compatible proposals (see Appendix) so that the conditions presupposed in Fig. 3 are satisfied.

For efficient computation of the likelihood in the sampling procedure, we precompute the cross-section at discrete values y_n and measured wavelengths λ_m; then during the algorithm, we obtain $\sigma(y, \lambda_m)$ for a specific

$y \in [y_n, y_{n+1}]$ via linear interpolation:

$$\sigma(y, \lambda_m) = \frac{\sigma(y_{n+1}, \lambda_m) - \sigma(y_n, \lambda_m)}{y_{n+1} - y_n}(y - y_n) + \sigma(y_n, \lambda_m), \tag{13}$$

from which the normalized rates follow per Eq. (11). This formulation is computationally efficient and reveals how the results of complex cross-section modeling can be incorporated seamlessly into our inference method.

We then perform the pCN MCMC algorithm for all 75 datasets, employing the appropriate model (ellipsoid or ring) in each case. As a slight enhancement over Fig. 3, we also adaptively tune the step sizes $\boldsymbol{\beta}$, which amounts to increasing or decreasing $\boldsymbol{\beta}$ at selected iterations j to maintain an acceptance rate in a desired range (0.1–0.3 in our case) [39, 45, 46]. The total MCMC chain length is RT: $R = 1024$ is the number of samples $\mathbf{x}^{(r)}$ kept for estimation, and T is a thinning parameter, which is doubled successively until the estimate of $\mathcal{Y}_g$ ($\mathcal{Y}'_g$) converges in both mean and standard deviation. We found $T = 2^{15}$ sufficient for all cases except $K_g = 10,000$ on the ring, where we sampled up to $T = 2^{17}$. For reference, the total wall clock time for running each $T = 2^{17}$ MCMC chain was approximately 24 min. on a CPU with 2.5 GHz clock. The final results appear in Fig. 6; points are the Bayesian means $\mathcal{Y}_B$ ($\mathcal{Y}'_B$) estimated from the R samples, error bars are the posterior standard deviations, and the solid lines trace the ground truth values for comparison.

Overall, the estimates are extremely close to the ground truth, even in the $K_g = 100$ cases with high statistical noise. However, lower accuracies and higher uncertainties were obtained in the two regions of the ring radius identified in Sec. IV (12–14 nm and 17–19 nm)—precisely where the cross-section varies slowly with $\mathcal{Y}'$. These uncertainties reduce with increasing photon number and generally hone in on the ideal curve, apart from $\mathcal{Y}'_g = 13$ nm, for which $\mathcal{Y}'_B = 11.8$ nm at $K_g = 10,000$. Close inspection of Fig. 4(b) at these two radii in particular show very similar shapes, indicating again that the observed deviation from ground truth reflects inherent ambiguities in the cross-section itself.

VI. MODEL SELECTION

In the above examples, the inference models employed match the actual plasmonic structures under test; i.e., the ellipsoid model was applied to datasets generated from a ground truth ellipsoid, and likewise for the ring. The defining parameters were unknown, but the model was assumed accurate. The specification of a single well-posed model is certainly the desired case in practice, but may not be feasible in all situations, in which case Bayesian methods can again be invoked for model comparison and selection. In this section, we outline through example how this can be implemented in our plasmonic nanometrology workflow.

Taking the datasets corresponding to $K_g = 1000$,

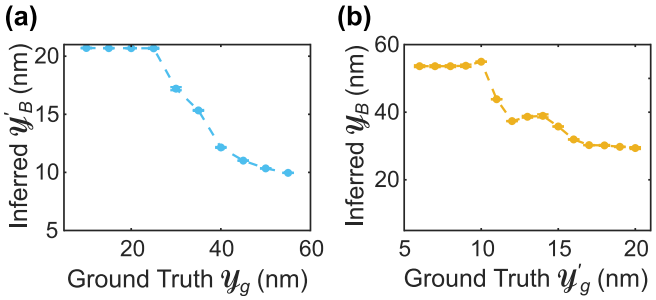

FIG. 7. Bayesian inference using a mismatched model on datasets corresponding to $K_g = 1000$. (a) Ellipsoid ground truth, ring inference model. (b) Ring ground truth, ellipsoid inference model.

we redo Bayesian inference utilizing mismatched cross-section models in the likelihood: the ring model [Fig. 4(b)] for the data actually generated by an ellipsoid, and the ellipsoid model [Fig. 4(a)] for data produced by a ring. Figure 7 plots the results obtained in these "cross-model" tests. The general negative slope of the inferred parameters stems from the opposite spectral trends for the two nanoparticles: the scattering peak of the ellipsoid redshifts with increasing dimension, and vice versa for the ring. The flat, almost clipped portions of both curves correspond to regions where the peak of the observed cross-section reaches the edge of the range supported by the model. Thus, these results can be viewed as competitors to those found earlier: Fig. 7(a) provides an alternative analysis of the same data that are utilized in the center plot of Fig. 6(a), while Fig. 7(b) serves as an alternative to the center plot of Fig. 6(b). Without knowing the ground truth already, how can one rank these competing models?

Bayes' rule again offers a principled solution. Let E (R) denote the hypothesis of ellipsoid (ring). Then the posterior odds in favor of E over R can be written as:

$$\frac{P(E|\boldsymbol{\mathcal{D}})}{P(R|\boldsymbol{\mathcal{D}})} = \frac{P(\boldsymbol{\mathcal{D}}|E)}{P(\boldsymbol{\mathcal{D}}|R)}\frac{P(E)}{P(R)} = B_{ER}\frac{P(E)}{P(R)}, \tag{14}$$

where B_{ER} is the Bayes factor [47], equal to the posterior odds whenever the prior model weights are equal. Across-model Bayesian strategies, such as reversible jump MCMC [48], deal with multiple models directly by sampling over a joint distribution of models and parameters. As computationally simpler alternatives, within-model approaches estimate the marginals $P(\boldsymbol{\mathcal{D}}|E)$ and $P(\boldsymbol{\mathcal{D}}|R)$ independently, through the integral

$$P(\boldsymbol{\mathcal{D}}|\mathcal{H}) = \int d\mathbf{x}_\mathcal{H}\, P(\boldsymbol{\mathcal{D}}|\mathbf{x}_\mathcal{H}, \mathcal{H}) P(\mathbf{x}_\mathcal{H}|\mathcal{H}), \tag{15}$$

for each hypothesis $\mathcal{H} \in \{E, R\}$, where the subscript $\mathcal{H}$ has been added to explicitly note conditioning on a specific model. Computing Eq. (15) is equivalent to determining the normalization factor $\mathcal{Z}$ in Eq. (7)—a significantly more challenging endeavor than sampling from the

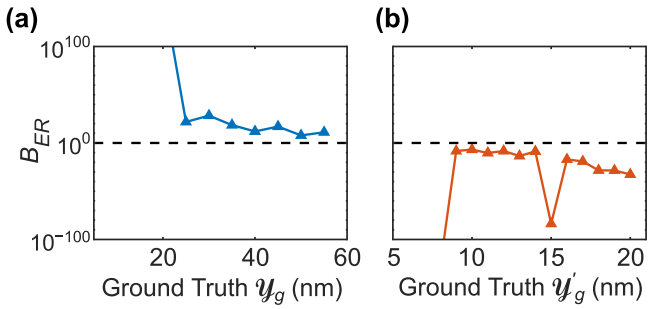

FIG. 8. Bayes factors for competing models. (a) Ellipsoid ground truth. (b) Ring ground truth. The horizontal line at 10^0 marks the threshold above (below) which the ellipsoid (ring) model is more probable.

unnormalized posterior as in conventional MCMC. Approaches to compute these marginal probabilities have been built on Gibbs sampling [49], simulated annealing [50], and power posteriors [51].

In our case, however, the small number of parameters involved allows us to use direct Monte Carlo integration: drawing samples $\mathbf{x}_{\mathcal{H}}^{(r)}$ from the prior $P(\mathbf{x}_{\mathcal{H}}|\mathcal{H})$ for each hypothesis gives the approximation [47]

$$P(\boldsymbol{\mathcal{D}}|\mathcal{H}) \approx \frac{1}{R}\sum_{r=1}^{R} P(\boldsymbol{\mathcal{D}}|\mathbf{x}_{\mathcal{H}}^{(r)},\mathcal{H}), \tag{16}$$

where $P(\boldsymbol{\mathcal{D}}|\mathbf{x}_{\mathcal{H}},\mathcal{H})$ is equal to the likelihood $L_{\boldsymbol{\mathcal{D}}}(\mathbf{x})$ defined in Eq. (6) multiplied by Poissonian factors $\prod_{m=1}^{M}(N_m!)^{-1}$ for normalization. The final results for $R = 2^{20}$ samples (no thinning) appear in Fig. 8. The horizontal dotted line corresponds to $B_{ER} = 1$: all points above (below) this line denote higher odds for the ellipsoid (ring) model. The Bayes factors for all datasets convincingly favor the ground truth model over its competitor; $B_{ER} > 10^7$ for all points in Fig. 8(a), while $B_{ER} < 10^{-6}$ for all points in Fig. 8(b). The Bayes factors for $\mathcal{Y}_g \leq 20$ nm and $\mathcal{Y}'_g \leq 8$ nm are so extreme that they do not even appear on this logarithmic scale.

As practical alternatives to the Bayes factor when estimating the marginal probabilities may be much more difficult than in the present examples, the computationally simpler Akaike [52] or Schwarz [53] information criteria could be invoked instead. In cases such as Fig. 8 where one model is unambiguously superior, the differences between all methods are expected to be minimal. However, the Bayes factor has the advantage as a directly interpretable metric of comparison for any amount of data, even when the differences in models may be small [$B_{ER} \sim \mathcal{O}(1)$].

VII. DISCUSSION

The excellent agreement obtained above suggests that the presented approach for Bayesian plasmonic nanometrology should apply to many systems for which a model can be formulated in terms of the parameters involved in the radiative processes. Ultimately though, the success of the inference process will hinge on the details of the model, here specifically the scattering model. For example, strong variation in the spatio-spectral distribution of scattered light with the parameters of interest is a prerequisite for low-uncertainty inference—the absence of which was highlighted here by the reduced accuracy of the ring model in spectral regions where the cross-section varies little with tube radius.

With respect to computational cost, the specific functional features of $\tilde{\mu}_m(\mathbf{y})$ and the efficiency with which it can be computed significantly impact MCMC speed, in terms of both the time per step and the total number of steps required to reach convergence. Unfortunately, theoretically predicting the operational cost (e.g., wall clock time or chain length) for a given model is extremely difficult, and must be determined empirically by performing MCMC. Indeed, we used a total chain length of $RT = 2^{25}$ for most cases here, a forebodingly large number given the fact we examine only two parameters.

Nevertheless, more parameters do not necessarily lead to longer chains; in a previous example, we applied a similar inference algorithm for quantum state tomography with 512 unknowns, and found $RT = 2^{22}$ sufficient for convergence—10 min. total runtime on the same desktop computer [54]. The fact a system with a 256-fold increase in parameters required shorter chains than those here highlights the unpredictability of forecasting MCMC performance in general. Importantly, the pCN sampling procedure we include in Fig. 3 is designed specifically to maintain high convergence rates as the number of parameters increases [37] so that our method furnishes comparative optimism to handle much more complicated systems. In the end, of course, the speed of Bayesian methods cannot compete with that of simpler inferential approaches such as least-squares fitting. Yet the ultimate motivation for adopting the Bayesian viewpoint is not computational in nature, but rather based on its optimality guarantees and uncertainty quantification, which our initial tests here have shown can be brought to bear on plasmonic scattering.

Finally, moving beyond the presented scattering use-case, it would be interesting to extend the Bayesian inference method to other contexts in plasmonic nanometrology, such as the emerging applications of quantum plasmonics and quantum sensing in general [55]. Nearfield scattering properties offer a window into the excited nanoparticle surface modes, but experimental measurements of these modes are notoriously difficult due to the presence of the probing structure, substrate coupling, and coupling to neighboring particles. Furthermore, engineering of the electromagnetic environment to control plasmon-plasmon/plasmon-emitter coupling and Purcell enhancement involves sensitive parameter dependencies with respect to particle dimensions, orientation, and density, as well as the decay characteristics of the excited

emitter states, all of which can alter the radiative decay channels behind the observation of emitted photons. In each of these potential applications, the Bayesian formalism's ability to handle models with arbitrary functional forms in a single consistent paradigm should make it a useful tool for highly interacting plasmonic systems.

VIII. CONCLUSION

In conclusion, we have introduced a Bayesian procedure for estimating plasmonic nanoparticle geometries from scattering data. Our self-contained method applies highly efficient pCN sampling and readily incorporates custom cross-section models. Utilizing simulated datasets from both ellipsoidal and toroidal geometries, we demonstrated accurate inference of unknown parameters under multiple levels of statistical noise. Finally, computation of Bayes factors was shown to reveal the ground truth nanoparticle morphology (ellipsoid or ring) in all cases examined, revealing how our method integrates with model selection as well as parameter estimation. Overall, the presented procedure establishes a valuable framework that can be specialized to a variety of problems in plasmonic nanometrology.

ACKNOWLEDGMENTS

We thank K.J.H. Law for introducing us to the pCN proposal for uniform priors. This work was performed at Oak Ridge National Laboratory, operated by UT-Battelle for the U.S. Department of Energy under contract no. DE-AC05-00OR22725. Funding was provided by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, through the Quantum Algorithm Teams Program.

Appendix: pCN Proposals for Selected Priors

For reference, here we list pCN proposals for three selected prior distributions. When each parameter (assumed independent of all others) is updated according these steps, $q(\mathbf{x}|\mathbf{x}')\pi_0(\mathbf{x}') = q(\mathbf{x}'|\mathbf{x})\pi_0(\mathbf{x})$ obtains, so that the acceptance probability follows a simple likelihood ratio. Any parameter distributed according to the following can be added to the model seamlessly. Other distributions have pCN proposals as well, but we have found these three to be particularly useful, since they cover domains encountered widely in practice: $(-\infty, \infty)$, $[-1, 1]$ (a finite interval), and $[0, \infty)$.

1. Normal Distribution

Consider parameter $x \sim \mathcal{N}(0, 1)$, so that the prior is

$$\pi_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \tag{A.1}$$

Define step size $\beta \in (0, 1)$ and draw $\eta \sim \mathcal{F}_{\pi_0} = \mathcal{N}(0, 1)$. Then the proposal x' is [37]

$$x' = G_{\pi_0}(x, \beta, \eta) = \sqrt{1 - \beta^2}x + \beta\eta. \tag{A.2}$$

2. Uniform Distribution

Consider parameter $x \sim \mathcal{U}(-1, 1)$, so that the prior is

$$\pi_0(x) = \frac{1}{2}\mathbf{1}_{[-1,1]}(x). \tag{A.3}$$

The step size is $\beta \in (0, 1)$; draw $\eta \sim \mathcal{F}_{\pi_0} = \mathcal{U}(-1, 1)$. Then the proposal x' is the reflected random walk [38]:

$$x' = G_{\pi_0}(x, \beta, \eta) = \begin{cases} -2 - (x + \beta\eta) & x + \beta\eta \leq -1 \\ x + \beta\eta & |x + \beta\eta| < 1 \\ 2 - (x + \beta\eta) & x + \beta\eta \geq 1. \end{cases} \tag{A.4}$$

3. Gamma Distribution

Consider parameter $x \sim \Gamma(\alpha, 1)$, so that the prior is

$$\pi_0(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x}. \tag{A.5}$$

Take step size $\beta \in (0, 1)$. Draw η_1 from the beta distribution $\eta_1 \sim \mathcal{B}((1-\beta)\alpha, \beta\alpha)$, and η_2 from the gamma distribution $\eta_2 \sim \Gamma(\beta\alpha, 1)$. Then take as the proposal [56]:

$$x' = G_{\pi_0}(x, \beta, \boldsymbol{\eta}) = \eta_1 x + \eta_2. \tag{A.6}$$

[1] A. P. Ramos, A. L. Da Róz, M. Ferreira, F. de Lima Leite, and O. N. Oliveira, Ch. 4 - Dynamic light scattering applied to nanoparticle characterization, in *Nanocharacterization Techniques* (William Andrew Publishing, 2017) pp. 99–110.

[2] F. J. T. Huber, S. Will, and K. J. Daun, J. Quant. Spectrosc. Radiat. Transf. **184**, 27 (2016).

[3] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, Cambridge, UK, 2003).

[4] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods* (Springer, New York, 1999).

[5] D. Hejazi, S. Liu, S. Ostadabbas, and S. Kar, ACS Appl. Nano Mater. **2**, 4075 (2019).

[6] S. Bera, S. Paul, R. Singh, D. Ghosh, A. Kundu, A. Banerjee, and R. Adhikari, Sci. Rep. **7**, 41638 (2017).
[7] S. Aihara, M. Hamamoto, K. Iwamitsu, M. Okada, and I. Akai, AIP Adv. **7**, 045107 (2017).
[8] K. Iwamitsu, Y. Furukawa, M. Nakayama, M. Okada, and I. Akai, J. Lumin. **197**, 18 (2018).
[9] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, Proc. IEEE **104**, 148 (2016).
[10] F. F. Qin, Z. Z. Liu, Q. Zhang, H. Zhang, and J. J. Xiao, Sci. Rep. **8**, 14033 (2018).
[11] F. Qin, D. Zhang, Z. Liu, Q. Zhang, and J. Xiao, Opt. Express **27**, 31075 (2019).
[12] A. L. Lereu, Nat. Photonics **1**, 368 (2007).
[13] A. C. Atre, B. J. M. Brenny, T. Coenen, A. García-Etxarri, A. Polman, and J. A. Dionne, Nat. Nanotechnol. **10**, 429 (2015).
[14] X.-H. Jiang, P. Chen, K.-Y. Qian, Z.-Z. Chen, S.-Q. Xu, Y.-B. Xie, S.-N. Zhu, and X.-S. Ma, Sci. Rep. **10**, 11503 (2020).
[15] M. Bocková, J. Slabý, T. Špringer, and J. Homola, Annu. Rev. Anal. Chem. **12**, 151 (2019).
[16] N. Rivera and I. Kaminer, Nat. Rev. Phys. **2**, 538 (2020).
[17] L. F. Guerra, T. W. Muir, and H. Yang, J. Phys. Chem. C **123**, 18565 (2019).
[18] L. Polavarapu, J. Pérez-Juste, Q.-H. Xu, and L. M. Liz-Marzán, J. Mater. Chem. C **2**, 7460 (2014).
[19] S. M. Sadeghi and C. Mao, J. Appl. Phys. **121**, 014309 (2017).
[20] I. I. Izhnin, K. A. Lozovoy, A. P. Kokhanenko, K. I. Khomyakova, R. M. H. Douhan, V. V. Dirko, A. V. Voitsekhovskii, O. I. Fitsych, and N. Y. Akimenko, Appl. Nanosci. (2021).
[21] Z. Yao, S. Xu, D. Hu, X. Chen, Q. Dai, and M. Liu, Adv. Optical Mater. **8**, 1901042 (2020).
[22] M. Salhi, A. Passian, and G. Siopsis, Phys. Rev. A **92**, 033416 (2015).
[23] Z.-K. Zhou, J. Liu, Y. Bao, L. Wu, C. E. Png, X.-H. Wang, and C.-W. Qiu, Prog. Quantum Electron. **65**, 1 (2019).
[24] F. Zhang, J. Ren, X. Duan, Z. Chen, Q. Gong, and Y. Gu, J. Phys. Condens. Matter **30**, 305302 (2018).
[25] COMSOL, Inc., "COMSOL Multiphysics," `https://www.comsol.com/comsol-multiphysics` (2021).
[26] A. Arbouet, D. Christofilos, N. Del Fatti, F. Vallée, J. R. Huntzinger, L. Arnaud, P. Billaud, and M. Broyer, Phys. Rev. Lett. **93**, 127401 (2004).
[27] H. Baida, P. Billaud, S. Marhaba, D. Christofilos, E. Cottancin, A. Crut, J. Lermé, P. Maioli, M. Pellarin, M. Broyer, N. D. Fatti, F. Vallée, A. Sańchez-Iglesias, I. Pastoriza-Santos, and L. M. Liz-Marzań, Nano Lett. **9**, 3463 (2009).
[28] P. Billaud, S. Marhaba, N. Grillet, E. Cottancin, C. Bonnet, J. Lermé, J.-L. Vialle, M. Broyer, and M. Pellarin, Rev. Sci. Instrum. **81**, 043101 (2010).
[29] A. Gaiduk, M. Yorulmaz, P. V. Ruijgrok, and M. Orrit, Science **330**, 353 (2010).
[30] J. R. G. Navarro and M. H. V. Werts, Analyst **138**, 583 (2013).
[31] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[32] W. K. Hastings, Biometrika **57**, 97 (1970).
[33] S. Geman and D. Geman, IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-6**, 721 (1984).
[34] R. M. Neal, Ann. Stat. **31**, 705 (2003).
[35] S. Duane, A. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).
[36] P. D. Moral, A. Doucet, and A. Jasra, J. R. Stat. Soc. Ser. B **68**, 411 (2006).
[37] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White, Stat. Sci. **28**, 424 (2013).
[38] S. J. Vollmer, SIAM/ASA J. Uncertain. Quantif. **3**, 535 (2015).
[39] J. M. Lukens, K. J. H. Law, A. Jasra, and P. Lougovski, New J. Phys. **22**, 063038 (2020).
[40] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed. (Artech House, 2005).
[41] P. Decleer, A. Van Londersele, H. Rogier, and D. Vande Ginste, J. Comput. Appl. Math. **381**, 113023 (2021).
[42] P. B. Johnson and R. W. Christy, Phys. Rev. B **6**, 4370 (1972).
[43] Ansys, Inc., "Ansys Lumerical FDTD," `https://www.ansys.com/products/photonics/fdtd` (2021).
[44] E. M. Simmerman, H.-H. Lu, A. M. Weiner, and J. M. Lukens, Opt. Lett. **45**, 2886 (2020).
[45] G. O. Roberts, A. Gelman, and W. R. Gilks, Ann. Appl. Prob. **7**, 110 (1997).
[46] G. O. Roberts and J. S. Rosenthal, J. Appl. Prob. **44**, 458 (2007).
[47] R. E. Kass and A. E. Raftery, J. Am. Stat. Assoc. **90**, 773 (1995).
[48] P. J. Green, Biometrika **82**, 711 (1995).
[49] S. Chib, J. Am. Stat. Assoc. **90**, 1313 (1995).
[50] R. M. Neal, Stat. Comput. **11**, 125 (2001).
[51] N. Friel and A. N. Pettitt, J. R. Stat. Soc. Ser. B **70**, 589 (2008).
[52] H. Akaike, IEEE Trans. Auto. Contr. **19**, 716 (1974).
[53] G. Schwarz, Ann. Stat. **6**, 461 (1978).
[54] J. M. Lukens, K. J. H. Law, and R. S. Bennink, arXiv:2012.08997 (2020).
[55] A. S. Clark, M. Chekhova, J. C. F. Matthews, J. G. Rarity, and R. F. Oulton, Appl. Phys. Lett. **118**, 060401 (2021).
[56] P. A. Lewis, E. McKenzie, and D. K. Hugus, *Gamma processes*, Tech. Rep. (Naval Postgraduate School, 1986).