Approximate Bayesian computation techniques for optical characterization of nanoparticle clusters

OZAN BURAK ERICOK,[1] ● ALI TAYLAN CEMGIL,[2] ● AND HAKAN ERTURK[1],* ●

[1]*Department of Mechanical Engineering, Bogazici University, 34342 Bebek, Istanbul, Turkey*
[2]*Department of Computer Engineering, Bogazici University, 34342 Bebek, Istanbul, Turkey*
Corresponding author: hakan.erturk@boun.edu.tr

Characterization of nanoparticle aggregates from observed scattered light leads to a highly complex inverse problem. Even the forward model is so complex that it prohibits the use of classical likelihood-based inference methods. In this study, we compare four so-called *likelihood-free* methods based on approximate Bayesian computation (ABC) that requires only numeric simulation of the forward model without the need of evaluating a likelihood. In particular, rejection, Markov chain Monte Carlo, population Monte Carlo, and adaptive population Monte Carlo (APMC) are compared in terms of accuracy. In the current model, we assume that the nanoparticle aggregates are mutually well separated and made up of particles of same size. Filippov's particle-cluster algorithm is used to generate aggregates, and discrete dipole approximation is used to estimate scattering behavior. It is found that the APMC algorithm is superior to others in terms of time and acceptance rates, although all algorithms produce similar posterior distributions. Using ABC techniques and utilizing unpolarized light experiments at 266 nm wavelength, characterization of soot aggregates is performed with less than 2 nm deviation in nanoparticle radius and 3–4 deviation in number of nanoparticles forming the monodisperse aggregates. Promising results are also observed for the polydisperse aggregate with log-normal particle size distribution.   © 2017 Optical Society of America

OCIS codes: (000.5490) Probability theory, stochastic processes, and statistics; (290.0290) Scattering; (290.3200) Inverse scattering; (160.4236) Nanomaterials.

https://doi.org/10.1364/JOSAA.35.000088

1. INTRODUCTION

Developments in nanotechnology have brought new opportunities to various fields from medicine [1] to combustion diagnosis [2,3]. Synthesizing nanostructures with unique optical and physical properties necessitates new and improved characterization tools, as manufacturing unique structures with desired properties is only possible through using proper characterization.

There are different techniques for characterization of the size, shape, and configuration of the nanostructures. Although the direct methods relying on microscopy such as transmission electron microscopy (TEM), scanning electron microscopy (SEM), and atomic force microscopy (AFM) are capable of producing high-quality images and are considered as the most reliable means of characterization, they usually are not suitable for in-line monitoring. Recently, characterization of the nanostructures based on their emission profiles via time-resolved laser-induced incandescence (TR-LII) and

scattering profiles via light scattering have been gaining interest. While these methods have *in situ* characterization capability, the light-scattering methods might be preferable over TR-LII methods, as they are not only limited to absorbing materials.

There are different methods to predict scattering behavior of aggregates in the literature. Rayleigh–Debye–Gans (RDG) approximation [4] and RDG for fractal aggregates (RDG-FA) [5,6] are often used when the scatterer is irregularly shaped and smaller than the wavelength. RDG-FA is valid if $|m-1| \ll 1$ and $k|m-1|d_p \ll 1$, where m, k, and d_p are the refractive index, wavenumber, and primary particle diameter of aggregate. These approximations are often used to estimate the optical properties of combustion-generated particles, as in Refs. [7,8]. However, Klusek *et al.* [9] stated that soot aggregates cannot be treated as Rayleigh scatterers and Köylü and Faeth [10] stated that RDG-FA theory cannot be used for soot aggregates with large refractive index. Besides these, they

neglect the multiple and self-induced scattering effects and the error in angular scattering properties such as S_{11} and S_{12}, which can be as high as 50% for large aggregates [11]. The T-matrix method [12,13,14], on the other hand, is a fast and accurate method to model scattering of light by aggregates. Although there is no limitation on primary particle size, the T-matrix method is limited to nonoverlapping spheres. In reality, aggregates may overlap, and overlapping has an impact on scattering properties [15]. Discrete dipole approximation (DDA), proposed by Purcell and Pennypacker [16], is another method that is often used in the literature. DDA is a semianalytical volume integral method that can handle the most complex geometries by representing them with a finite set of interacting dipoles. Through the interactions of these dipoles, DDA can easily calculate the optical behavior of any geometry with any morphology. Though it can handle very complex geometries accurately, it is slow compared to approximate methods [17].

Applications of light-scattering methods for soot characterization is a well-studied problem in the literature. RDG-FA theory is one of the most frequently used light-scattering methods. For example, Reimann *et al.* [18] combined elastic light-scattering techniques with LII, whereas Huber *et al.* [19] combined wide angle light-scattering methods with the TR-LII method to estimate the primary particle radius and radius of gyration in the aggregate. Using the Mueller matrix for characterization of nanoparticle aggregates is another approach. Mengüç and Manickavasagam [20] outlined a procedure to obtain the number of nanoparticles in the aggregate using polarization information stored in elements such as S_{12}, S_{33}, and S_{34} if the primary particle diameter is known via an *ex situ* measurement such as SEM. In different studies, Charnigo *et al.* used the normalized versions and derivatives of S_{11}, S_{12}, S_{33}, and S_{34} to estimate the agglomeration level [21] and nanoparticle diameter [22]. The characterization problem is also formulated as a parameter estimation problem of identifying the average size and the number of nanoparticles forming the aggregate that can be estimated through the solution of the optimization problem. More recently, Ericok and Erturk [23] used only the nonpolarization information, S_{11}, to estimate the number and radius of the nanoparticles in the aggregate using the tabu search algorithm, while their analysis showed that characterization is possible for aggregates larger than 20 nm using a light source at 266 nm. However, in all of these deterministic methods, it is not possible to validate the results, since the analysis focuses on matching the measured scattering behavior rather than the characterized parameters directly. Therefore, presenting the range of existence of the parameters with a desired confidence level would be more appropriate for characterization unless an independent experiment is designed to measure these parameters directly.

Statistical inversion based on Bayesian inference is a probabilistic approach that is capable of predicting the probability distribution of the unknown parameters based on measurements. The basic idea of the Bayesian inference relies on updating the information about the unknown parameters based on observations. Once the final probability distribution of the unknown parameters is obtained, credible intervals with a desired confidence level can be constructed.

The Bayesian inference approach is relatively new for nanoparticle characterization. Burr *et al.* [24] used Bayesian inference to recover soot aggregate size distribution from multiangle elastic light-scattering data using maximum *a posteriori* (MAP) inference. Sipkens *et al.* [25] used the TR-LII method to characterize gas-borne silicon nanoparticles and analyzed their measurements with the Bayesian approach to find the most probable nanoparticle size distribution parameters. Otero *et al.* [26] studied the potentials of parametric and nonparametric Bayesian schemes on the inverse problems of light scattering. In another study, Clementi *et al.* [27] used the Bayesian inversion method based on multidirectional dynamic light-scattering measurements to estimate the particle size distribution of latexes. Charnigo *et al.* [28] used different elements of the Mueller matrix and their derivatives to construct credible intervals for nanoparticle diameter and agglomeration level. They introduced both systematic and stochastic error in their analysis and utilized polarization information via normalized S_{11}, S_{12}, S_{33}, and S_{34} for different cases. In a more recent work, Hadwin *et al.* [29] used Bayesian inference to calculate the MAP and credible intervals of soot volume fraction and peak temperature for different likelihoods and prior densities, whereas Huber *et al.* [30] used Bayesian inference to estimate the size and morphology parameters of soot-laden aerosols. Ericok and Erturk [31] used Bayesian inference for characterization of the radius and number of nanoparticles of the soot aggregates with different credible interval levels.

Computing the likelihood function is vital for Bayesian inference. However, it is not always available in a closed analytical form or feasible to compute due to complex physical models that require too many parameters. In general, there are three ways to overcome this problem: (1) nonparametric estimation of the likelihood. In this method, one can run M auxiliary experiments based on the parameters that affect the outcome. The auxiliary experiments must be run at discrete points for a single parameter, and intermediate points can be interpolated. However, if the number of parameters is not small, this method suffers from what is known as the "curse of dimensionality." (2) Parametric estimation of likelihood. One can assume an analytical form, e.g., Gaussian or Poisson distributions, for likelihood. Although this method is very practical, it may misrepresent the physical reality. (3) Parametric estimation after dimension reduction. This method does not use the whole outcome of the system, but only its sufficient statistics. If a sufficient statistics such as mean, standard deviation, etc., exists that represents the outcome well, one may reduce the number of unknowns and use parametric estimation on the reduced unknowns. More information can be found in Ref. [28].

One alternative way to cope with this problem is not using likelihood altogether. If likelihood is not available or cannot be obtained by the methods described above, Bayesian inference is performed via what is known as approximate Bayesian computation (ABC) techniques. ABC consists of methods that do not require the calculation of the likelihood function to obtain the posterior distribution. The idea is to sample points from the posterior distribution directly, if a model exists that relates the parameters of interest to observable quantities. While ABC methods are widely used in different fields such as

genetics [32], epidemiology [33], cosmology [34], and signal detection theory [35], etc., they have not been used for solution of inverse radiative transfer and optical characterization problems.

In radiative transfer and optical characterization, Bayesian inference methods are generally performed by assuming an analytical form to the likelihood function as in Refs. [24–31]. However, most of the problems considered are complex, where the physical behavior of the system cannot be represented by an analytical expression, necessitating the use of numerical modeling. For the analysis of such problems, using ABC methods is a more reasonable approach. In this study, some of the ABC methods will be introduced (for the first time, to our best knowledge) and applied rigorously for characterization of nanoparticle aggregates to demonstrate the strengths and limitations of the proposed methods.

2. ABC

The very fabric of the statistical inversion is to reconstruct the solution of the inverse problem in terms of probability distributions. The inverse problems generally comprise identifying the parameters of interest that are related to the measured quantities through a physical model. Likewise, the statistical inversion theory takes all the information about the system that is known to the observer prior to the measurement, as well as the measurement itself, and estimates the probability distributions of the parameters of interest. Therefore, the solution of the inverse problem is not a single set of parameters, but a probability distribution referred to as the posterior probability distribution. The fundamental principles and some details of the concepts can be found in Refs. [36,37].

In Bayesian statistical inference, all the variables are treated as random variables, which are denoted by capital letters. The observable random variable $Y \in R^m$, the measurement, and the unobservable random variable $\Theta \in R^n$, the parameters of interest, are related to each other by

$$Y = f(\theta, E), \tag{1}$$

where $E \in R^k$ is the noise.

The conditional probability of θ given $Y = y_{\text{observed}}$, $\pi(\theta|y_{\text{observed}})$, is the solution of the statistical inverse problem, as it represents the probability distribution of $\Theta = \theta$ based on the measurement y_{observed}:

$$\pi_{\text{post}}(\theta) = \pi(\theta|y_{\text{observed}}) = \frac{\pi_{\text{pr}}(\theta)\pi(y_{\text{observed}}|\theta)}{\int_{R^n} \pi_{\text{pr}}(\theta)\pi(y_{\text{observed}}|\theta)d\theta}, \tag{2}$$

where $\pi_{\text{pr}}(\theta)$ is the prior density that contains all the knowledge about θ before the measurement. The conditional probability of random variable Y given $\Theta = \theta$, $\pi(y|\theta)$, is called the likelihood function.

Equation (2) suggests that the posterior probability distribution, $\pi_{\text{post}}(\theta)$, can easily be obtained if the analytical forms of both the prior and the likelihood are available. The easiest part in Eq. (2) is the prior density, since it is either known or assumed by the researcher apart from the model. However, representing the likelihood analytically can be impossible for some cases, as in most of the computational or simulation-based models. In these cases, the standard methods of Bayesian

estimation cannot be applied. As stated earlier, ABC methods overcome this problem by not calculating the likelihood function but by sampling enough points from the posterior distribution itself. In this study, four ABC algorithms—rejection sampler, Markov chain Monte Carlo (MCMC), population Monte Carlo (PMC), and adaptive population Monte Carlo (APMC)—are considered.

A. Rejection Sampler

The first ABC algorithm introduced by Tavaré *et al.* [38] was a simple rejection algorithm to estimate the common ancestor times for DNA sequence data. Tavaré's idea was extended by Pritchard *et al.* [39]. First, a point, θ^*, is generated from the prior density. Then, the artificial data set, x, associated with the generated point θ^* is created using the simulation model, f. They proposed that if the generated data set and the actual data set are close enough, $\rho(x, y) < \epsilon$, then the point is sampled from the posterior. In this inequality, ρ is a scalar metric, and ϵ is the user-defined tolerance. If these steps are repeated until N points are drawn from the posterior, a good approximate of the posterior can be obtained without calculating the likelihood function. Algorithm 1, presented below, is the rejection algorithm proposed by Pritchard. Although it is simple to implement, the rejection algorithm suffers from high rejection rates—in other words, low acceptance rates of sampled points.

Algorithm 1: ABC rejection sampler

1:   $count = 0$
2:   **for** $i = 1...N$ **do**
3:       **while** $\rho(x, y) > \epsilon$ **do**
4:          $count = count + 1$
5:          generate $\theta^* \sim \pi_{pr}(\theta)$
6:          simulate $x \sim f(x|\theta^*)$
7:       store $\theta_i = \theta^*$
8:   $r_{accept} = N/count$
9:   return.

B. MCMC

The main disadvantage of the ABC rejection sampler algorithm is the high rejection rates caused by sampling points from the prior distribution. Rejection sampler algorithm samples the new point without using any information from the previously sampled points. However, it is possible, in principle, to reduce the rejection rates by utilizing the information from the previously sampled points. The MCMC algorithm samples a new point by making a jump from the current point with the assumption that the new point will possibly be near the last sampled point. Thus, it eliminates the complete randomness and makes use of available information.

The MCMC algorithm is similar to the Metropolis–Hasting algorithm. The only difference is that there are no likelihood calculations involved in MCMC. Like the Metropolis–Hasting algorithm, MCMC starts with an initial point, $\theta^{(0)}$, sampled from the prior. Then, the new points are generated by a transition kernel, q, which relates the point from the previous MC step, $\theta^{(i-1)}$, to the new one, θ^*. Data set x is generated by using θ^*, and $\rho(x, y)$ is calculated. If $\rho(x, y) > \epsilon$, then $\theta^{(i-1)}$ is a better

approximation than θ^*; therefore no replacement is made. Otherwise, θ^* is accepted with probability, h, where h is a uniformly distributed random number between 0 and 1, $U(0, 1)$. It should be noted that the MCMC algorithm only uses the last sampled point, not the ones that are sampled earlier. Algorithm 2, given below, is the MCMC algorithm proposed by Marjoram *et al.* [40]. More detailed discussions about the algorithm can be found in Refs. [41,42].

Algorithm 2: ABC MCMC

1:   $accepted = 0$
2:   generate $\theta^{(0)} \sim \pi_{pr}(\theta)$
3:   **for** $i = 1 \dots N$ **do**
4:      generate $\theta^* \sim q(\theta^*|\theta^{(i-1)}, \Sigma)$
5:      simulate $x \sim f(x|\theta^*)$
6:      **if** $\rho(x, y) > \epsilon$ **then**
7:         $\theta^{(i)} = \theta^{(i-1)}$
8:      **else**
9:         calculate $h = \frac{\pi_{pr}(\theta^*)}{\pi_{pr}(\theta^{(i-1)})} \times \frac{q(\theta^{(i-1)}|\theta^*, \Sigma)}{q(\theta^*|\theta^{(i-1)}, \Sigma)}$
10:        **if** $U(0, 1) < h$ **then**
11:           $accepted = accepted + 1$
12:           store $\theta^{(i)} = \theta^*$
13:        **else**
14:           store $\theta^{(i)} = \theta^{(i-1)}$
15:   $r_{accept} = accepted/N$
16:   **return**

Although the acceptance rates of MCMC are higher than those of the rejection algorithm, there are several disadvantages. First, sampled points are highly correlated to each other due to a transition kernel. This is not a major drawback unless noncorrelated points are needed. Second, small ϵ together with high correlation property might cause MCMC to stack in a low probability region, resulting in a harder transition from one point to another. Third, a covariance matrix should be defined by the user *a priori* to the algorithm. This has a great impact on the performance of MCMC, since the covariance matrix determines the nature of the transition kernel. Finally, it is hard to parallelize MCMC chains [43].

C. PMC

Using the points that are successfully sampled previously and adjusting the tolerance level ϵ are two good ways to increase the acceptance rate. The partial rejection control (PRC) algorithm that Sisson *et al.* [44] proposed uses weighted resampling from the points sampled previously and a decreasing tolerance level in each iteration. Beaumont *et al.* [45] corrected the bias noted in the original algorithm proposed by Sisson *et al.* and suggested the PMC algorithm. Algorithm 3, given below, is the PMC algorithm suggested by Beaumont *et al.*

PMC divides the calculations into $t = T$ iterations and the corresponding tolerance levels, ϵ_t's, are defined by the user with $\epsilon_1 > \epsilon_t > \epsilon_T$. Thus, more points are accepted at the beginning, since a relatively relaxed tolerance is applied. The main advantage of PMC is that it samples new points using the points that represent the posterior well, which results in the elimination of the points that represent the posterior poorly.

Algorithm 3: ABC PMC

1:   $count = 0$
2:   $t = 1$
3:   **for** $i = 1 \dots N$ **do**
4:      **while** $\rho(x, y) < \epsilon_1$ **do**
5:         $count = count + 1$
6:         draw $\theta_i^{(t)} \sim \pi_{pr}(\theta)$
7:         simulate $x \sim f(x|\theta_i)$
8:      let $w_i^{(t)} = 1/N$
9:   let Σ twice the empirical covariance of $\{\theta_i^{(t)}\}_{i=1,2,\dots N}$
10:   **for** $t = 2 \dots T$ **do**
11:      **for** $i = 1 \dots N$ **do**
12:         **while** $\rho(x, y) < \epsilon_t$ **do**
13:           $count = count + 1$
14:           pick $\theta_i^* \sim \theta_j^{(t-1)}$ with probabilities $w_j^{(t-1)}$
15:           sample $\theta_i^{(t)} \sim q(\theta_i^{(t)}|\theta_i^*, \Sigma)$
16:           simulate $x \sim f(x|\theta_i^{(t)})$
17:         calculate and normalize $w_i^{(t)} \propto \frac{\pi_{pr}(\theta_i^{(t)})}{\sum_j w_j^{(t-1)} q(\theta_i^{(t)}|\theta_j^{(t-1)}, \Sigma)}$
18:      let Σ twice the empirical covariance of $\{\theta_i^{(t)}\}_{i=1,2,\dots N}$
19:   $r_{accept} = N/count$
20:   **return**.

At $t = 1$, the algorithm samples N points satisfying $\rho(x, y) < \epsilon_1$ from the prior distribution and assigns a weight, $w_i^{(1)}$, to each of them. Thus, the algorithm has a set of points, $\{\theta_i^{(1)}\}_{i=1,2,\dots N}$, whose elements satisfy the corresponding tolerance level. The set at iteration t, $\{\theta_i^{(t)}\}_{i=1,2,\dots N}$, is formed by selecting a point, θ_i^*, from the previous set $\{\theta_i^{(t-1)}\}_{i=1,2,\dots N}$ with probability $w_i^{(t-1)}$ and sampling a new point, $\theta_i^{(t)}$, by using a transition kernel, $q(\theta_i^{(t)}|\theta_i^*, \Sigma)$. Unlike MCMC, the covariance matrix at each iteration, Σ, is calculated by taking twice the empirical covariance of the set $\{\theta_i^{(t-1)}\}_{i=1,2,\dots N}$.

The major drawback of PMC is the user-defined tolerance levels. The user may define the tolerances too sharply and too shallowly, which leads to an increase or decrease in the computation time required.

D. APMC

The problem of user-defined tolerances is solved by defining online tolerance at each step as suggested by Wegmann *et al.* [46], Drovandi and Pettitt [47], and Del Moral *et al.* [48]. The APMC algorithm proposed by Lenormand *et al.* [49] is designed to overcome this problem and is based on the basic principles of these works and that of Beaumont *et al.*

APMC goes a step further than PMC by adjusting tolerance levels adaptively. Algorithm 4 requires three parameters: N, α, and $p_{acc_{min}}$, where N is the total number of points that will be sampled, $\alpha \in [0, 1]$ is the ratio of points that will be kept in each iteration to the total number of points, and $p_{acc_{min}}$ is the minimal acceptance rate. The number of points to keep in each iteration, N_α, is αN.

At $t = 1$, the algorithm draws N points from the prior distribution and calculates the $\rho_i^{(0)}$ and sets $w_i^{(0)} = 1$. Then, the first tolerance level, ϵ_1, is taken as the first α quantile of $\{\rho_i^{(0)}\}_{i=1,2,\dots N}$. After the first tolerance level is determined, APMC constructs the new set, $\{\theta_i^{(1)}, w_i^{(1)}, \rho_i^{(1)}\}$, from the elements of the old set, where $\rho_i^{(0)}$ is smaller than ϵ_1. Thus, a total

of N_α elements are obtained at the first iteration. Since the new set consists of the points with the best metrics, $\rho_i^{(1)}$, the covariance matrix is taken as twice the empirical covariance of this set. At this point, the rest of the iterations start by setting p_{acc} to 1.

From this point, the algorithm tries to find the best $N - N_\alpha$ points to complete the set at each iteration. At iterations $t > 1$, a new point is generated by drawing a point from the transition kernel, $q(\theta_i^{(t-1)}|\theta_i^*, \Sigma)$, around the point θ_i^*, which is picked from the previous set, $\theta_j^{(t-1)}$, with normalized weight (i.e., probabilities), where $j = 1, 2, \ldots N_\alpha$. After the new point is generated, the data set is simulated, and the corresponding metric and the weight of the old point, $\theta_i^{(t-1)}$, and the point itself, are replaced with the new one.

When the process of replacing old points with new ones is finished, APMC calculates the new tolerance level for the next iteration as the first α quantile of $\{\rho_i^{(t-1)}\}_{i=1,2,\ldots N}$ and determines the minimal acceptance rate, p_{acc}. The algorithm repeats these steps until $p_{acc} \leq p_{acc_{min}}$.

The main advantage of this algorithm is that the tolerance levels are calculated and updated in each step, resulting in utilization of more optimized tolerance levels.

Algorithm 4: ABC APMC

```
1:   count = 0
2:   N_α = αN
3:   for t = 1 do
4:      for i = 1…N do
5:         count = count + 1
6:         draw θ_i^(0) ~ π_pr(θ)
7:         simulate x ~ f(x|θ_i^(0))
8:         calculate ρ_i^(0) = ρ(x, y)
9:         let w_i^(0) = 1
10:     define ε_1 = Q_ρ^(0)(α) the first α-quantile of {ρ_i^(0)}_{i=1,2,…N}
11:     {θ_i^(1), w_i^(1), ρ_i^(1)} = {(θ_i^(0), w_i^(0), ρ_i^(0))|ρ_i^(0) ≤ ε_1}_{i=1,2,…N}
12:     let Σ twice the empirical covariance of {θ_i^(1)}_{i=1,2,…N_α}
13:     define p_acc = 1
14:  while p_acc > p_acc_min do
15:     t = t + 1
16:     for i = N_α + 1…N do
17:        count = count + 1
18:        pick θ_i^* ~ θ_j^(t-1) with probability {w̄_j^(t-1)}_{j=1…N_α}
19:        draw θ_i^(t-1) ~ q(θ_i^(t-1)|θ_i^*, Σ)
20:        simulate x ~ f(x|θ_i^(t-1))
21:        calculate and update ρ_i^(t-1) = ρ(x, y)
22:        update w_i^(t-1) ∝ π_pr(θ_i^(t-1)) / (Σ_j w_j^(t-1) q(θ_i^(t-1)|θ_j^(t-1), Σ))   j = 1…N_α
23:        calculate and update p_acc = 1/(N-N_α) Σ_{k=N_α+1}^N 1_{ρ_i^(t-1) < ε_{t-1}}
24:     let ε_t = Q_ρ^(t-1)(α) the first α-quantile of {ρ_i^(t-1)}_{i=1,2,…N}
25:     {θ_i^(t), w_i^(t), ρ_i^(t)} = {(θ_i^(t-1), w_i^(t-1), ρ_i^(t-1))|ρ_i^(t-1) ≤ ε_t}_{i=1…N}
26:     let Σ twice the empirical covariance of {θ_i^(t)}_{i=1,2,…N_α}
27:  r_accept = N_α/count
28:     return.
```

3. OPTICAL SCATTERING

Stokes showed that four elements, $\mathbf{K} = [IQUV]^T$, are enough to fully describe the intensity and polarization state of an electromagnetic wave, where I, Q, U, and V represent the total

intensity, the difference between horizontally and vertically polarized intensities, the difference between $+45°$ and $-45°$ polarized intensities, and the difference between the right-handed and left-handed circularly polarized intensities, respectively, with $I^2 = Q^2 + U^2 + V^2$ [50].

Incident, $\mathbf{K}_i$, and scattered, $\mathbf{K}_s$, intensities are related to each other by a 4×4 matrix known as the Mueller matrix, $[S_{ij}]$, by $\mathbf{K}_s = [S_{ij}]\mathbf{K}_i$. The Mueller matrix is the complete representation of the scattering event. Therefore, it can be used for characterization of nanoparticles. In this work, mainly the first element of the Mueller matrix, S_{11}, is used for characterization purposes. S_{11} is known as the differential scattering cross section and it relates the incident total intensity, I_i, to scattered total intensity, I_s, if unpolarized light is incident upon the particle.

4. REPRESENTATION OF AGGREGATES

It is required to represent the structure of the nanoparticle aggregates, since real-time light-scattering experiments are replaced with numerical experiments in this study. The aggregates are mathematically modeled with the fractal equation,

$$N_p = k_f \left(\frac{R_g}{r_p}\right)^{D_f}, \tag{3}$$

where N_p is the number of particles in the aggregate, r_p is the primary particle radius, R_g is the radius of gyration, and D_f and k_f are the fractal dimension and fractal prefactor, respectively.

There are different techniques to generate fractal-like aggregates such as ballistic aggregation, diffusion-limited aggregation (DLA) and reaction-limited aggregation (RLA) [51]. The most commonly used one is Filippov's sequential algorithm [52]. Filippov's algorithm can be implemented by two alternative approaches. The sequential algorithm is a particle-cluster aggregation method that adds spherical particles to each one by one, while satisfying the fractal equation at every step. Generation of fractal-like aggregates is also possible with cluster-cluster aggregation. This method first generates small aggregates and adds them to each other to form a larger aggregate. However, the particle-cluster algorithm is preferred to the cluster-cluster algorithm due to its implementation and tuning. Moreover, Filippov's algorithm was originally designed to generate monodisperse aggregates; it is easy to implement polydispersity, as illustrated in Ref. [53].

5. PROBLEM STATEMENT

Characterization of soot aggregates using ABC techniques is presented in this study to illustrate the strengths and weaknesses of the proposed ABC algorithms via numerical light-scattering experiments. It must be noted that although this work focuses on a specific problem, ABC methods are flexible enough to be applied to characterization of different nanoparticle aggregates.

This work considers dilute suspensions of identical, but randomly oriented soot aggregates that are well separated from each other. The total scattering behavior of the whole measurement volume comprises the scattering behavior of individual aggregates and can be approximated by the scattering behavior of a single aggregate averaged over many orientations [17,54]. Orientation averaging compensates for the effects of random

orientations of many aggregates as well as movement of these aggregates during the measurement time. Both monodisperse and polydisperse aggregates with log-normal particle size distribution of nonoverlapping spheres are considered.

The fractal-like aggregates are generated by Filippov's algorithm [52], and the implementation of Filippov's particle-cluster algorithm given by Skorupski *et al.* [53] is followed in this work. The particle-cluster algorithm relies on random numbers that affect the aggregate geometry. Therefore, different geometries will be obtained at each time when the particle-cluster algorithm is executed with exactly the same parameters. Finally, although the particle-cluster algorithm fails to produce the expected slope of pair-correlation function [52,53], the particle-cluster algorithm is preferred, since the generated aggregates would be more similar to aggregates in real ensembles.

The values of the parameters D_f and k_f are assumed to be known *a priori* following the results of two independent experiments conducted by Yon *et al.* [55] and taken as 1.81 and 1.37, respectively, as inputs to the particle-cluster algorithm. The wavelength of the incident light is considered as 266 nm, which lies in the wavelength range of a commercial device [56], as suggested by [20] in all of the calculations, and the optical properties of soot at 266 nm are adopted from [55].

Tian *et al.* [57] showed experimentally that the radius of the primary particles of soot aggregates is in the range of 6–25 nm, and the probability of aggregates with $N_p < 30$ is higher than those with $N_p > 30$. Moreover, in our previous study [23], it was shown that the characterization of N_p and r_p highly depends on the effective radius, r_{eff}, given in Eq. (4). When r_{eff} is smaller than 20 nm, the deviations in estimated values for r_p and N_p begin to increase, since order of magnitude of S_{11} at 266 nm becomes comparable with that of the uncertainty of the commercial device. Therefore, we have selected test cases according to these observations. Table 1 shows the cases selected to test the proposed methods. The effective radius is defined as

$$r_{\text{eff}} = \left(\sum_{i=1}^{N_p} r_i^3 \right)^{1/3}. \qquad (4)$$

6. SOLUTION METHODOLOGY

Obtaining the scattering behavior of a given aggregate with known geometry and morphology requires the solution of the direct problem by solving Maxwell's equations. Since an analytical solution of the Maxwell's equation for a complex geometry such as an aggregate is not available, use of numerical methods is required. In this work, Fortran implementation of DDA, DDSCAT, developed by Draine and Flatau [58], is used for the solution of the direct problem.

A. Synthetic Measurements

Since the algorithms considered are evaluated by the numerical experiments, the actual experimental measurements are replaced with synthetic measurements that are obtained by adding random measurement error to the simulated scattering behavior of known aggregates. The scattering behavior of

Table 1. Cases Studied in This Work

Cases	N_p	r_p [nm]	σ_r [nm]	r_{eff} [nm]
Case 1	34	4	0	13
Case 2	4	10	0	16
Case 3	10	10	0	21
Case 4	16	16	0	40
Case 5	34	16	0	52
Case 6	40	22	0	75
Case 7	30	15	1.25	50

the entire measurement volume is approximated by the orientation-averaged scattering profile of a single aggregate, and 490 orientations are considered as recommended in Ref. [20].

The effects of the randomness in Filippov's algorithm should be taken into account while constructing the synthetic measurements. Every fractal geometry generated with exactly the same parameters will be different due to the randomness of the algorithm. Therefore, 10 different fractal geometries are created for monodisperse cases (Cases 1–6), presented in Table 1. Moreover, the measurement error may also have some impact on the final results. Therefore, 10 different random error vectors are introduced to the scattering profile of each geometry created for each case. The measurement error is assumed to have a normal distribution with zero mean ($\mu = 0$) and a known standard deviation ($\sigma = 10^{-4}$), considering the uncertainty of a commercial device [56]. Therefore, the synthetic measurements are created according to the formula below:

$$y_{i,j} = S_{11,i}(N_p, r_p, \sigma_r) + e_{i,j}(\mu = 0, \sigma = 10^{-4}), \qquad (5)$$

where i and j are the indices of geometry and error, respectively. As a result, 100 different synthetic measurements (10 different measurement errors to each of 10 geometries) are obtained with Eq. (5) for each monodisperse case.

The synthetic measurement for a polydisperse case (Case 7) is created in a slightly different way. Since the primary particle radii vary significantly among aggregate realizations generated with the same fractal parameters, the scattering behavior of the ensemble cannot be assumed as the scattering behavior of a single aggregate. Therefore, an average over aggregate realizations is also required for Case 7. Although our preliminary analysis showed that taking an average over 40 realizations is sufficient to represent the scattering behavior of the ensemble, 100 aggregate realizations are created. After random error vectors are added to the scattering behavior of these realizations, the synthetic measurement is created by taking the average of these 100 simulated measurements.

The physical limitations of the actual experimental setup must be considered while creating synthetic measurements. The measurements at $\phi = 0°$ and $\phi = 180°$ are not possible in the same plane. Moreover, measurements near these angles are highly disturbed by the incident light. Considering these limitations, the initial and final polar angles are taken as $\phi_i = 5°$ and $\phi_f = 165°$, respectively. In our previous study [23], it was observed that the increment between these angles may not affect the results of characterization much if $\Delta\phi < 16°$. Therefore, the scattering profiles of every parameter set contain measurements between $\phi_i = 5°$ and $\phi_f = 165°$ with $\Delta\phi = 8°$.

B. Metric Function

The main objective of the ABC methods is to sample points from the posterior to approximate the actual posterior distribution of the parameters of interest. This is achieved by defining a metric between the scattering behavior of the actual parameters, y, and that of the simulated parameters, x. Therefore, the proposed method utilizes the first element of the Mueller matrix of a given aggregate at N_m polar angles, ϕ_i, where $i = 1 \dots N_m$. Then, a small distance between $x = [S_{11}(\phi_1) \dots S_{11}(\phi_m)]^T$ and $y = [S_{11}^*(\phi_1) \dots S_{11}^*(\phi_m)]^T$ means that the parameters leading to x may be the parameters of interest, where $*$ indicates the synthetic measurements. The ratio of the norm of the difference between x and y to the norm of y is used as a metric. Since the ABC methods rely on the distance between the simulated estimate, x, and synthetic measurement, y, a threshold value, ϵ, must be defined to decide whether a point is sampled from the posterior or not. A point is considered to be sampled from the posterior if the metric is smaller than the user-defined tolerance,

$$\rho(x, y) = \frac{\|x - y\|}{\|y\|} < \epsilon. \tag{6}$$

The definition of the distance metric is the same for all the algorithms considered in this work. Second, all algorithms assume a uniform prior density, meaning that there is no knowledge about the parameters of interest before the experiments. Choice of uniform prior density is a common practice in the literature, such as the cases in Refs. [25,28]. Third, a preliminary study is conducted to determine how many points an algorithm should sample in order to have results that do not depend on sample size. Algorithm 1 is executed with different numbers of samples, and it is found that sampling 100 points or more from the posterior distribution for a single synthetic measurement yields almost identical results. Therefore, all algorithms except for Algorithm 2 sample 100 points, which results in 10^4 points (100×100) in total for each case; a sample of 100 points is sufficient. Since Algorithm 2 is MCMC, the number of points that will be sampled cannot be controlled with given implementation. Moreover, the transition kernel, q, is assumed as Gaussian with a known covariance matrix for Algorithms 2, 3, and 4. However, some parameters are different from each other for different algorithms.

The main purpose of the transition kernel is to propose a new point based on the previous one, and the very nature of the transition kernel must be chosen considering the physics of the problem. While we have tried different transition kernels before selecting the one used in this study, the choice is not arbitrary. Since all algorithms start from the points that are randomly selected within the domain, the initial point might be very far away from the actual posterior distribution region. In such cases, a transition kernel with small covariance might not propose acceptable jumps, whereas the rate of acceptance of a too large covariance will still be small. Therefore, an ideal kernel should have covariance that is not too small or too large. The covariance matrix is then taken as $[36 - 10; -10\ 36]$ for Algorithm 2. Finally, the negative sign of the off-diagonal terms is because r_p and N_p are inversely proportional to each other for

a constant effective radius through Eq. (4). The covariance matrices in Algorithms 3 and 4 are calculated empirically based on the previously sampled points.

The threshold value is taken as $\epsilon = 0.03$ for Algorithms 1 and 2, whereas the interval [0.1 0.03] is divided into $T = 9$ equal spaces for Algorithm 3, where $T = 9$ is the total number of iterations in Algorithm 3. The minimum acceptable tolerance considering the normal measurement error is calculated with Eq. (6) for a given parameter set and found around 0.023. The threshold value of $\epsilon = 0.03$ considers both the measurement error and the interpolation error within the database. Besides, larger values of tolerance would result in larger prediction zones for parameters of interest. While this increases the probability of an accurate or acceptable prediction, which can be identified in the test cases by checking if the actual parameters are contained within the predicted zone or not, predicting a smaller zone is also desired. Therefore, $\epsilon = 0.03$ would yield a smaller prediction zone while encapsulating both the measurement and interpolation error. Finally, $\alpha = 0.5$ and $p_{acc_{min}} = 0.01$ are taken as inputs for Algorithm 4.

It should be noted that although Algorithms 3 and 4 have their stopping criteria, the covariance matrix may also cause these algorithms to stop. The covariance matrix is symmetric and positive-definite by definition. If all the sampled points are identical or similar to each other, then, covariance among the sampled points would be equal to or very close to 0. Algorithms 3 and 4 terminate as the covariance matrix is no more positive-definite, and it is assumed that the algorithm converges to that particular point.

C. Database

The presented algorithms require the solution of the direct problem of calculating the scattering profile of a given set of parameters. The computational time to obtain the orientation-averaged scattering profile of a given parameter set is demanding, as 490 orientations are considered. Solving the direct problem for a parameter set $N_p = 15$, $r_p = 8$ nm, and $\sigma_r = 0$ with the parallel implementation of DDSCAT using 3393 dipoles and averaged over 490 orientations takes 653 s in an 8-core system with 2.50 GHz frequency. The computational time required would increase if the number or size of nanoparticles in the aggregate increases. Therefore, constructing a database and interpolating the solutions of the intermediate parameter sets would be a feasible approach to sample points from the posterior distribution.

Another challenge that must be considered while creating a database is the randomness of fractal geometries. The geometries generated with same parameters according to Eq. (3) would be all different due to the randomness in Filippov's algorithm. Therefore, two separate aggregates generated based on the exact same parameters would have different geometries, and hence, their scattering profile would be different without error being introduced. If a database is to be constructed, scattering behaviors of different geometries with the same particle size and number must be averaged to reduce geometry dependency.

The database is constructed as follows: Fifteen different numbers of particles ranging from 3 to 45 with an increment

of 3, and 15 different particle radii ranging from 3 nm to 45 nm with an increment of 3 nm are considered. Therefore, the database comprises a total of 225 sets of parameters (15 for particle size and number, each). Twelve geometries are simulated with Filippov's algorithm for each of these parameter sets, assuming all the nanoparticles in the aggregate have the same radius. The direct problem is solved for each of twelve geometries and parameter sets and the orientation-averaged S_{11} values are obtained with 1° increments in the polar angle. This resolution in polar angle allows us to use any subinterval with any integer increment in calculations. Cubic splines are used to interpolate the intermediate values of parameter sets. Extending the limits of the database for a desired resolution is very simple and straightforward.
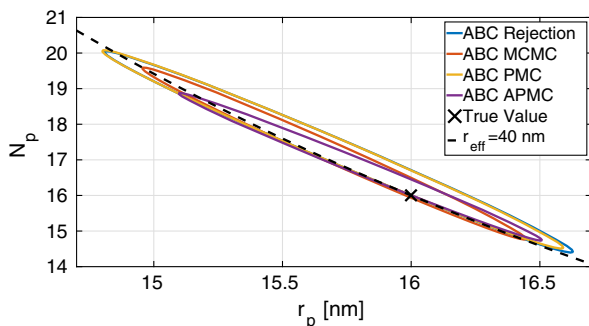
7. RESULTS AND DISCUSSION

A verification study is conducted at the beginning of this study by reproducing the results presented in Ref. [23]. The parameters of DDSCAT regarding the number of dipoles, orientation averaging, etc., are also studied to obtain accurate results.

The results for Case 4, shown in Table 1, are presented rigorously, first comparing the performance of the four algorithms. Following the identification of the algorithm with the superior performance considering estimation accuracy and efficiency, the selected algorithm is tested for all cases to determine the limits of characterization.
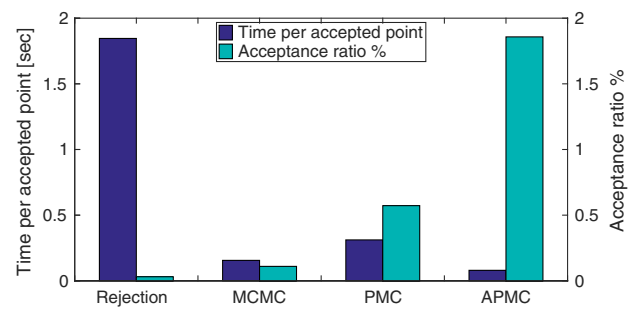
A. Case 4

The ellipsoids are plotted with mean and covariance matrix of the points sampled from the posterior for each algorithm. Figure 1 shows the results for Case 4. It can be observed that the sampled points for all algorithms follow the corresponding constant effective radius curve calculated with Eq. (4), as expected. The APMC algorithm gives the narrowest ellipsoid among others while still capturing the true value.

Figure 2 shows the time elapsed to accept a point for all algorithms and the acceptance ratio, which is defined as the ratio of total number of sampled points to the total number of trials. It can be observed that the rejection algorithm has the highest time with the lowest acceptance rate, as low as 0.03%, whereas the APMC algorithm has the lowest time with the highest acceptance rate of 1.8%. Since the APMC

Fig. 2. Time and acceptance rate of all algorithms for Case 4.

algorithm gives accurate results in a reasonable amount of time with high acceptance rates, the other cases are only investigated with the APMC method. As the time and acceptance rate results are similar for the other cases, they are not presented here explicitly.
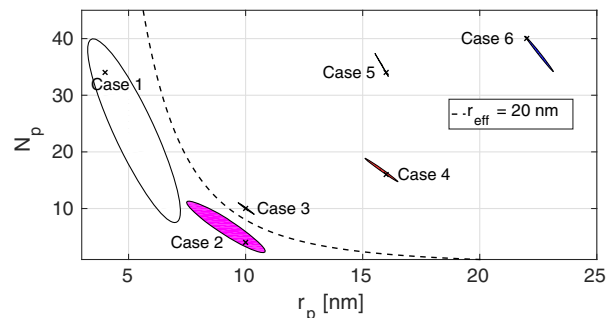
B. Monodisperse Cases (Cases 1–6)

Figure 3 shows the results of Cases 1–6. Again, the ellipsoids are plotted with the mean and covariance matrix of the corresponding sampled points. All of the ellipsoids include corresponding true values that are defined in Table 1 and are marked with a cross sign. When the effective radius is less than 20 nm, the ellipsoids grow larger, and the deviations begin to increase. For the smaller aggregates, the order of magnitude of measurement error becomes comparable to, or larger than, the measured S_{11} values. Considering these results, we can safely say that the characterization is reliable above the 20 nm effective radius curve for a 266 nm light source.

Figure 3 also suggests that the deviations in r_p are less than 2 nm and that of N_p is around 3–4. Finally, each case follows its constant effective radius curve, though these are not shown here explicitly.

C. Polydisperse Case (Case 7)

In reality, it is well known that the radii of the soot aggregates follow a log-normal distribution. A preliminary analysis is carried out to understand the applicability of the proposed method for such aggregates. For that, the monodispersity assumption for nanoparticle radii is omitted and an aggregate of $N_p = 30$ particles with a log-normal size distribution with

Fig. 1. Results of Case 4 ($N_p = 16$, $r_p = 16$ nm, $\sigma_r = 0$) for all algorithms.

Fig. 3. Results of all monodisperse cases combined. Cross signs indicate the corresponding true values that are defined in Table 1.
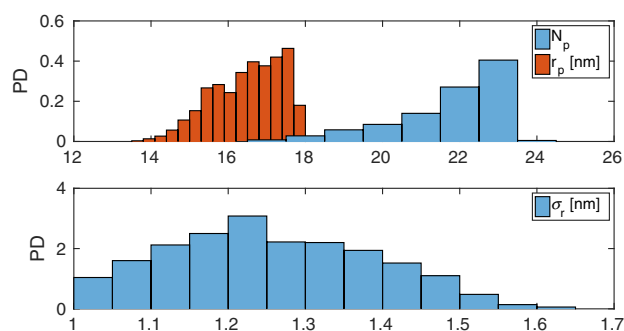
Fig. 4. APMC results of Case 7 ($N_p = 30$, $r_p = 15$ nm, $\sigma_r = 1.25$ nm). PD stands for probability density values.

$r_p = 15$ nm and $\sigma = 1.25$, which is typical for soot aggregates [59] and [60], is considered for Case 7.

The results plotted in Fig. 4 are obtained by sampling 1000 points using the APMC algorithm with $p_{acc_{min}} = 0.01$. Figure 4 shows the probability density values of N_p, r_p, and σ_r. It can be observed that the peak probabilities occur near the expected parameters, except for N_p. Although the range for results is larger than desired, a rigorous analysis on other S_{ij} profiles along with ABC methods may provide significant improvement on solution accuracy.

8. CONCLUSION

In this study, ABC techniques are introduced for optical characterization of monodisperse and polydisperse nanoparticle aggregates. The aggregates are considered as similar and well separated from each other and generated with Filippov's particle-cluster algorithm. The scattering behavior of the ensemble is approximated by a scattering behavior of a single aggregate averaged over multiple orientations for monodisperse aggregates, whereas 100 different aggregates realizations, in addition to orientation averaging, are also averaged for polydisperse aggregates.

The capability and limitations of the four ABC methods, rejection, MCMC, PMC, and APMC, are investigated for characterization of both monodisperse and polydisperse soot aggregates based on numerical light-scattering measurements. Seven cases with effective radii ranging from 13 to 75 nm are considered. It is shown that similar results can be found by all methods investigated. However, the APMC algorithm is superior to others in terms of the high acceptance rates and low computation time it requires.

The measurements of S_{11} values of the Mueller matrix allow characterization with unpolarized light without the use of *ex situ* particle size or number measurements. For the monodisperse cases, it is shown that a nanoparticle radius can be found with a deviation less than 2 nm, and the number of particles forming the aggregate can be found with a deviation of 3–4 when the effective radius is greater than 20 nm when a wavelength of 266 nm is considered. Moreover, promising results are also obtained for the polydisperse case. The results presented in this work are also consistent with the results that are obtained by following the conventional Bayesian inference methodology described in Ref. [31]. The limitations would vary if other

wavelengths are considered. Effects of the use of different wavelengths and other Mueller matrix elements on solution sensitivity and limitations should also be investigated.

ABC methods are fast and robust computational tools that can be used to approximate the posterior distributions of n-dimensional unknown spaces. In a more challenging characterization problem, morphological parameters and refractive indices are also unknowns that need to be recovered. For example, if fractal parameters are also considered as unknowns constructing the likelihood function with 15 nodes per parameter, it would take years to calculate. In such cases, use of conventional Bayesian inference methods is not feasible, whereas ABC methods do not require the calculation of likelihood function.

Finally, a more challenging characterization problem briefly investigated here should be examined in more detail. For example, the effects of polarization information and distribution parameters on solution accuracy should be investigated. Another distribution for the number of nanoparticles can also be assumed instead of taking a single aggregate. Moreover, necking and overlapping may occur while formation of soot aggregates. A model that takes these parameters into account would be more realistic and may need to use polarization information stored in other elements of the Mueller matrix.

REFERENCES

1. S. S. Agasti, S. Rana, M.-H. Park, C. K. Kim, C.-C. You, and V. M. Rotello, "Nanoparticles for detection and diagnosis," Adv. Drug Delivery Rev. **62**, 316–328 (2010).
2. C. L. Myung, A. Ko, and S. Park, "Review on characterization of nanoparticle emissions and pm morpology from internal combustion engines: part 1," Int. J. Automotive Technol. **15**, 203–218 (2014).
3. S. Choi, C. L. Myung, and S. Park, "Review on characterization of nano-particle emissions and pm morpology from internal combustion engines: part 2," Int. J. Automotive Technol. **15**, 219–227 (2014).
4. M. Kerker, *The Scattering of Light and Other Electromagnetic Radiation*, Vol. **16** of Physical Chemistry (Academic, 1969).
5. C. M. Sorensen, "Light scattering by fractal aggregates: a review," Aerosol Sci. Technol. **35**, 648–687 (2001).
6. G. M. Faeth and Ü. Ö. Köylü, "Soot morphology and optical properties in nonspreximed turbulent flame environments," Combust. Sci. Technol. **108**, 207–229 (1995).
7. T. L. Farias, Ü. Ö. Köylü, and M. G. Carvalho, "Range of validity of the Rayleigh–Debye–Gans theory for optics of fractal aggregates," Appl. Opt. **35**, 6560–6567 (1996).
8. J. Kim, H. Bauer, T. Dobovičnik, R. Hitzenberger, D. Lottin, D. Ferry, and A. Petzold, "Assessing optical properties and refractive index of combustion aerosol particles through combined experimental and modeling studies," Aerosol Sci. Technol. **49**, 340–350 (2015).
9. C. Klusek, S. Manickavasagam, and M. Mengüç, "Compendium of scattering matrix element profiles for soot agglomerates," J. Quantum Spectrosc. Radiat. Transfer **79**, 839–859 (2003).
10. U. O. Koylu and G. M. Faeth, "Radiative properties of flame-generated soot," J. Heat Transfer **115**, 409–417 (1993).
11. Y. Zhao and L. Ma, "Applicable range of the Rayleigh–Debye–Gans theory for calculating the scattering matrix of soot aggregates," Appl. Opt. **48**, 591–597 (2009).
12. B. Peterson and S. Ström, "T-matrix for electromagnetic scattering from an arbitrary number of scatterers and representations of E(3)," Phys. Rev. D **8**, 3661–3678 (1973).
13. M. I. Mishchenko, J. W. Hovenier, and L. D. Travis, *Light Scattering by Nonspherical Particles: Theory, Measurements and Applications* (Academic, 1999).
14. L. Liu, M. I. Mishchenko, and W. Patrick Arnott, "A study of radiative properties of fractal soot aggregates using the superposition T-matrix

method," J. Quant. Spectrosc. Radiat. Transfer **109**, 2656–2663 (2008).

15. N. Doner and F. Liu, "Impact of necking and overlapping on radiative properties of coated soot aggregates," Aerosol Sci. Technol. **51**, 532–542 (2017).

16. E. M. Purcell and C. R. Pennypacker, "Scattering and absorption of light by nonspherical dielectric grains," Astrophys. J. **186**, 705–714 (1973).

17. A. Karlsson, T. Yi, and P.-E. Bengtsson, "Absorption and scattering of light from ensembles of randomly oriented aggregates," J. Opt. Soc. Am. A **30**, 316–324 (2013).

18. J. Reimann, S. A. Kuhlmann, and S. Will, "2D aggregate sizing by combining laser-induced incandescence (LII) and elastic light scattering (ELS)," Appl. Phys. B **96**, 583–592 (2009).

19. F. J. T. Huber, M. Altenhoff, and S. Will, "A mobile system for a comprehensive online-characterization of nanoparticle aggregates based on wide-angle light scattering and laser-induced incandescence," Rev. Sci. Instrum. **87**, 053102 (2016).

20. M. Mengüç and S. Manickavasagam, "Characterization of size and structure of agglomerates and inhomogeneous particles via polarized light," Int. J. Eng. Sci. **36**, 1569–1593 (1998).

21. R. Charnigo, M. Francoeur, M. P. Mengüç, A. Brock, M. Leichter, and C. Srinivasan, "Derivatives of scattering profiles: tools for nanoparticle characterization," J. Opt. Soc. Am. A **24**, 2578–2589 (2007).

22. R. Charnigo, M. Francoeur, P. Kenkel, M. P. Mengüç, B. Hall, and C. Srinivasan, "Estimating quantitative features of nanoparticles using multiple derivatives of scattering profiles," J. Quantum Spectrosc. Radiat. Transfer **112**, 1369–1382 (2011).

23. O. B. Ericok and H. Erturk, "Inverse characterization of nanoparticle clusters using unpolarized optical scattering without ex-situ measurements," J. Quantum Spectrosc. Radiat. Transfer **198**, 117–129 (2017).

24. D. W. Burr, K. J. Daun, O. Link, K. A. Thomson, and G. J. Smallwood, "Determination of the soot aggregate size distribution from elastic light scattering through Bayesian inference," J. Quantum Spectrosc. Radiat. Transfer **112**, 1099–1107 (2011).

25. T. A. Sipkens, R. Mansmann, K. J. Daun, N. Petermann, J. T. Titantah, M. Karttunen, H. Wiggers, T. Dreier, and C. Schulz, "In situ nanoparticle size measurements of gas-borne silicon nanoparticles by time-resolved laser-induced incandescence," Appl. Phys. B **116**, 623–636 (2014).

26. F. A. Otero, H. R. Barreto Orlande, G. L. Frontini, and G. E. Eliçabe, "Bayesian approach to the inverse problem in a light scattering application," J. Appl. Stat. **42**, 994–1016 (2015).

27. L. A. Clementi, J. R. Vega, L. M. Gugliotta, and H. R. Orlande, "A Bayesian inversion method for estimating the particle size distribution of latexes from multiangle dynamic light scattering measurements," Chemom. Intell. Lab. Syst. **107**, 165–173 (2011).

28. R. Charnigo, M. Francoeur, P. Kenkel, M. P. Menguc, B. Hall, and C. Srinivasan, "Credible intervals for nanoparticle characteristics," J. Quantum Spectrosc. Radiat. Transfer **113**, 182–193 (2012).

29. P. J. Hadwin, T. A. Sipkens, K. A. Thomson, F. Liu, and K. J. Daun, "Quantifying uncertainty in soot volume fraction estimates using Bayesian inference of auto-correlated laser-induced incandescence measurements," Appl. Phys. B **122**, 1–16 (2016).

30. F. J. T. Huber, S. Will, and K. J. Daun, "Sizing aerosolized fractal nanoparticle aggregates through Bayesian analysis of wide-angle light scattering (WALS) data," J. Quantum Spectrosc. Radiat. Transfer **184**, 27–39 (2016).

31. O. B. Ericok and H. Erturk, "Characterization of nanoparticle aggregates using Bayesian inference via light scattering experiments," in *ASME International Mechanical Engineering Congress and Exposition* (Phoenix, 2016), V008T10A026.

32. C. Leuenberger and D. Wegmann, "Bayesian computation and model selection without likelihoods," Genetics **184**, 243–252 (2010).

33. M. M. Tanaka, A. R. Francis, F. Luciani, and S. A. Sisson, "Using approximate Bayesian computation to estimate tuberculosis transmission parameters from genotype data," Genetics **173**, 1511–1520 (2006).

34. A. Weyant, C. Schafer, and W. M. Wood-Vasey, "Likelihood-free cosmological inference with type Ia supernovae: approximate Bayesian computation for a complete treatment of uncertainty," Astrophys. J. **764**, 116–130 (2013).

35. B. M. Turner and T. Van Zandt, "Hierarchical approximate Bayesian computation," Psychometrika **79**, 185–209 (2014).

36. D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*, 2nd ed. (Athena Scientific, 2008).

37. J. Kaipio, E. Somersalo, E. S. S. Antman, J. E. M. L. Sirovich, A. J. K. Hale, P. H. J. Keener, J. K. B. J. Matkowsky, and A. M. C. S. Peskin, *Statistical and Computational Inverse Problems*, Vol. **160** of Applied Mathematical Sciences (Springer, 2006).

38. S. Tavaré, D. J. Balding, R. C. Griffiths, and P. Donnelly, "Inferring coalescence times from DNA sequence data," Genetics **145**, 505–518 (1997).

39. J. K. Pritchard, M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman, "Population growth of human Y chromosomes: a study of Y chromosome microsatellites," Mol. Biol. Evol. **16**, 1791–1798 (1999).

40. P. Marjoram, J. Molitor, V. Plagnol, and S. Tavare, "Markov chain Monte Carlo without likelihoods," Proc. Natl. Acad. Sci. USA **100**, 15324–15328 (2003).

41. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. (CRC Press, 2014).

42. C. Robert and G. Casella, *Monte Carlo Statistical Methods* (Springer, 2004).

43. B. M. Turner and T. Van Zandt, "A tutorial on approximate Bayesian computation," J. Math. Psychol. **56**, 69–85 (2012).

44. S. A. Sisson, Y. Fan, and M. M. Tanaka, "Sequential Monte Carlo without likelihoods," Proc. Natl. Acad. Sci. USA **104**, 1760–1765 (2007).

45. M. A. Beaumont, J. Cornuet, J. Marin, and C. P. Robert, "Adaptive approximate Bayesian computation," Biometrika **96**, 983–990 (2009).

46. D. Wegmann, C. Leuenberger, S. Neuenschwander, and L. Excoffier, "ABCtoolbox: a versatile toolkit for approximate Bayesian computations," BMC Bioinf. **11**, 116 (2010).

47. C. C. Drovandi and A. N. Pettitt, "Estimation of parameters for macroparasite population evolution using approximate Bayesian computation," Biometrics **67**, 225–233 (2011).

48. P. Del Moral, A. Doucet, and A. Jasra, "An adaptive sequential Monte Carlo method for approximate Bayesian computation," Stat. Comput. **22**, 1009–1020 (2012).

49. M. Lenormand, F. Jabot, and G. Deffuant, "Adaptive approximate Bayesian computation for complex models," Comput. Statist. **28**, 2777–2796 (2013).

50. C. E. Bohren and D. R. H. Wiley, *Absorption and Scattering of Light by Small Particles* (Wiley, 1984).

51. P. Meakin, "A historical introduction to computer models for fractal aggregates," J. Sol-Gel Sci. Technol. **15**, 97–117 (1999).

52. A. Filippov, M. Zurita, and D. Rosner, "Fractal-like aggregates: relation between morphology and physical properties," J. Colloid Interface Sci. **229**, 261–273 (2000).

53. K. Skorupski, J. Mroczka, T. Wriedt, and N. Riefler, "A fast and accurate implementation of tunable algorithms used for generation of fractal-like aggregate models," Physica A: Stat. Mech. Appl. **404**, 106–117 (2014).

54. M. J. Berg and C. M. Sorensen, "Internal fields of soot fractal aggregates," J. Opt. Soc. Am. A **30**, 1947–1955 (2013).

55. J. Yon, R. Lemaire, E. Therssen, P. Desgroux, A. Coppalle, and K. F. Ren, "Examination of wavelength dependent soot optical properties of diesel and diesel/rapeseed methyl ester mixture by extinction spectra analysis and LII measurements," Appl. Phys. B **104**, 253–271 (2011).

56. Hinds Instruments, "Technology for Polarization Measurement Exicor 150XT," technical report, 2013.

57. K. Tian, F. Liu, K. A. Thomson, D. R. Snelling, G. J. Smallwood, and D. Wang, "Distribution of the number of primary particles of soot aggregates in a nonpremixed laminar flame," Combust. Flame **138**, 195–198 (2004).

58. B. T. Draine and P. J. Flatau, "Discrete-dipole approximation for scattering calculations," J. Opt. Soc. Am. A **11**, 1491–1499 (1994).

59. K. J. Daun, B. J. Stagg, F. Liu, G. J. Smallwood, and D. R. Snelling, "Determining aerosol particle size distributions using time-resolved laser-induced incandescence," Appl. Phys. B **87**, 363–372 (2007).

60. F. Liu, B. J. Stagg, D. R. Snelling, and G. J. Smallwood, "Effects of primary soot particle size distribution on the temperature of soot particles heated by a nanosecond pulsed laser in an atmospheric laminar diffusion flame," Int. J. Heat Mass Transfer **49**, 777–788 (2006).