Journal of Quantitative Spectroscopy & Radiative Transfer 362 (2026) 110034

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy and Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

An optimization tool for inverse problems of multilayered spherical particles

Alexandru Doicu [a], Dmitry S. Efremenko [b],[*], Maxim A. Yurkin [c]

[a] Independent researcher, Germering, Germany
[b] Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Münchener Str., 20, Oberpfaffenhofen, 82234, Wessling, Germany
[c] Université Rouen Normandie, INSA Rouen Normandie, CNRS, CORIA UMR 6614, Rouen, France

ARTICLE INFO

Dataset link: https://github.com/DmitryEfremenko/MLSI

Keywords:
Multilayered spheres
T-matrix method
Mie scattering
Optimization algorithms
Regularization

ABSTRACT

We present an optimization framework for retrieving the layer thicknesses and refractive indices of multilayered spherical particles from either angular scattering patterns or spectral scattering/extinction data. The forward problem is solved using the T-matrix formalism, with multiple recurrence schemes and analytical derivatives enabling efficient and accurate sensitivity analysis. The inverse problem is addressed within a global optimization setting that combines bound-constrained parameter estimation with optional data transformations for enhanced stability and reduced dimensionality. A MultiStart strategy is employed to explore the parameter space and identify distinct candidate solutions, followed by local refinement. As a point estimate, the minimum-residual solution is determined, while the associated uncertainties are computed using the covariance matrix approximation at the minimum-residual solution, and the marginal posterior expectation. The developed optimization tool is publicly available at https://github.com/DmitryEfremenko/MLSI.

1. Introduction

Light scattering is ubiquitously used for non-invasive characterization of various objects, constituting a broad field of inverse scattering [1]. This inversion is the easiest when the corresponding particle model allows fast direct simulations. The obvious example of the latter is the Lorenz–Mie theory for homogeneous spheres. For such shapes, plenty of methods have been developed in three broad classes: low-dimensional inversion, nonlinear regression and neural networks (machine learning). The first class incurs measurement of only a few signals, e.g., in flow systems [2,3], or compressing the quasi-continuous signal, e.g., through the parameters of its Fourier spectrum [4,5]. These methods are very fast and typically robust, but do not have explicit error control.

Nonlinear regression is conceptually the simplest method, as it is based on fitting theoretical simulations to 1D or 2D experimental signals [6,7]. However, the corresponding inverse problem is typically ill-posed and nonconvex, requiring one to use global optimization and seek compromise between speed and robustness. Estimating uncertainties of model parameters is integral to the regression; however, it may become ambiguous if the real particle differs from the idealized model [8,9]. Machine learning can be considered as enhancement to the above approaches, naturally fitting the scenarios where large training datasets are readily available. Specifically, the corresponding surrogates have been trained on both angular and spectral forward models and validated on real measurements [10,11].

Ongoing development of computational capabilities made the inversion perfectly viable for non-spherical models as well, although the forward simulation may take minutes for a single set of input parameters. It is based on large precomputed datasets which are either used directly, e.g., with the nearest-neighbor interpolation [12,13] or through training a neural-network surrogate [14,15]. However, the main challenge in this case is not the computational time for a single particle shape per se, thanks to efficient methods like the discrete dipole approximation (DDA) [16], but rather due to the increased dimensionality which hampers both building a dense enough forward dataset and robust inversion. Thus, only relatively simple shape models, described by a few parameters, have been considered so far.

In this paper, we focus on multilayered spheres—a compromise between generality and computational feasibility. This shape has a lot of relevant applications in aerosol science, nano- and biophotonics. The forward map—from layer radii and complex refractive indices to measured observables—can be instantiated in several ways: angular intensity patterns, their Fourier spectra, and spectral responses such as scattering or extinction cross sections. All inversion methods for spheres allow natural generalization to such case. An example of nonlinear regression is Strokotov et al. [8], who used a coated-sphere model to characterize lymphocytes from light-scattering patterns measured with the scanning flow cytometry. Spectral methods do not aim to retrieve all shape parameters, but can provide one or two

---

[*] Corresponding author.
E-mail address: dmitry.efremenko@dlr.de (D.S. Efremenko).

https://doi.org/10.1016/j.jqsrt.2026.110034
Received 6 May 2026; Received in revised form 5 June 2026; Accepted 8 June 2026

sizes [17]. Deep learning has accelerated spectral inversions for multilayered nanospheres, either by accelerating forward model [18] or training the inverse network after preliminary compression of extinction spectrum [19]. However, despite explosive development of deep learning, such methods are still "black-box" in the sense that outliers (large errors) are always possible in practical applications [20]. In principle, the same concerns apply, albeit in a lesser extent, to any method based on global optimization. However, its careful consideration for multilayered spheres, including stability of forward models, noise-aware regularization, robust global-to-local optimization, data richness (angle, wavelength), and good initialization, is still missing.

In the following, we present an optimization algorithm for determining the layer thicknesses and refractive indices of a multilayered spherical particle from either (i) the angular dependence of the scattered intensity or (ii) the spectral dependence of the scattering or extinction cross section. The algorithm is designed around three guiding themes. First, we compute angular intensities and spectral cross sections using stable recurrence relations and differentiate them analytically with respect to the layer radii and refractive indices. Second, we enforce box constraints, which can be viewed as a simple form of regularization, and provide optional data transformations (dimensionality reduction or low-pass filtering). Third, we solve the inverse problem using a MultiStart strategy, ensuring both basin discovery and accurate refinement. As a point estimate, we determine the minimum-residual solution, while the associated uncertainties are computed using the covariance matrix approximation at the minimum-residual solution, and the marginal posterior expectation.

2. Scattering by a multilayered particle

The electromagnetic scattering theory for a coated sphere was originally formulated by Aden and Kerker [21], whose approach became the foundation for most later numerical studies. Bohren and Huffman [22] subsequently presented a clear description of this problem and even provided an accompanying computational program. Despite the apparent simplicity of the formulation, numerical evaluations based on these classical expressions often encounter difficulties: round-off errors can accumulate, and for large or strongly absorbing spheres, some quantities may reach magnitudes that cause numerical overflow. To mitigate these issues, Toon and Ackerman [23] reformulated the standard expressions for a single-layer sphere in a manner better suited to stable numerical computation. Later, Bhandari [24] derived a complete set of scattering coefficients for multilayered spheres, introducing a recursive scheme that connects the coefficients of an ℓ-layer sphere to those of a $(\ell-1)$-layer configuration. Wu and Wang [25] subsequently proposed an elegant recursive method, which was later refined by Wu et al. [26] to handle efficiently both plane-wave and Gaussian-beam illumination. This improved algorithm offers a compact representation and enhanced numerical stability. Johnson [27] also developed a related recurrence formulation of different structure, while Mackowski et al. [28] applied recursive techniques to study the thermal behavior of particulate systems.

For finely stratified or highly absorbing spheres, Kai and Massoli [29] introduced a method based on Taylor-series expansions to compute ratios of Riccati–Bessel functions more accurately, while Yang [30] proposed a simplified and numerically robust recurrence scheme, particularly effective for large particles with thin absorbing shells. Finally, Lock [31] developed a parallel iteration procedure for computing scattering by multilayered spheres. The procedure employs a successive-doubling strategy applied to four sets of multiple-scattering amplitudes and is reminiscent of the fast Fourier transform algorithm.

For spheres with many concentric layers, it is often advantageous to reformulate the problem in terms of transfer matrices. In this approach, each interface is represented by a 2×2 matrix that relates the inward and outward field amplitudes across that interface, and the response of the entire particle is obtained by matrix multiplication from the core outward. Moroz [32] presented a general recursive transfer-matrix solution that is both numerically stable and able to compute internal fields, stored energy, and emission. Qiu et al. [33] applied this method to the design and optimization of multilayered plasmonic nanoparticles, making it popular in nanophotonics. More recently, Rasskazov et al. [34] proposed a compact recursive transfer-matrix algorithm and released the open-source STRATIFY code, offering robust results even for highly absorbing or finely stratified particles. Further refinements aimed at improved stability for extremely thin shells have been introduced by Zhang [35].

Besides forward modeling, many applications require sensitivities of the scattering quantities with respect to size parameter or refractive index — for example, in inverse problems or uncertainty quantification. Grainger et al. [36] derived analytic derivatives of the Mie coefficients and scattering efficiencies, while Li and Bowler [37] extended these results to derivatives with respect to complex permittivity and permeability, proposing numerically stable expressions using logarithmic derivatives.

Among the various methods available for modeling light scattering by multilayered spheres, we employ the T-matrix formalism as the principal solution approach. This method is particularly suitable for the present application, and numerically stable recurrence relations for evaluating the T-matrix elements were derived by Doicu et al. [38]. These relations were originally formulated for multilayered axisymmetric particles and subsequently specialized to the case of a sphere. Appendix A summarizes the theoretical background for computing the scattered intensity and cross sections, along with their derivatives with respect to layer radii and refractive indices. In particular, it presents three alternative sets of recurrence relations for the diagonal elements of the transition matrix. The first set (Eqs. (A.27)–(A.28)) corresponds to the formulation given in Doicu et al. [38], while the other two sets (Eqs. (A.36)–(A.37) and (A.41)–(A.42)) represent equivalent variants. The last set, in particular, closely resembles the recurrence scheme proposed by Wu and Wang [25]. For completeness, Appendix A also includes a brief description of an alternative algorithm based on the transfer-matrix method, which is used solely for computing cross sections and is not part of the main solution procedure.

Our numerical simulations, shown in Figs. 1 and 2, indicate that all methods produce nearly identical results for the computation of the forward model (intensity or scattering cross section). However, when evaluating derivatives, slight differences were observed between the first method and the other two methods, which themselves yield almost identical outcomes.

3. The inversion algorithm

The inversion algorithm presented in this work is designed as a general framework for the solution of nonlinear ill-posed inverse problems. It combines statistical modeling, regularization techniques, and global optimization strategies in a unified approach that is applicable to a broad class of problems beyond the specific application considered here. In particular, the algorithm is able to incorporate data-error correlations, prior information, and multiple sources of uncertainty in a consistent Bayesian setting.

In the following, we present the main features of this general algorithm, including the formulation of the inverse problem, the modeling of data covariance, the use of data transformations, the maximum-likelihood and Bayesian formulations, the whitening strategy, residual-based uncertainty quantification, and the global optimization procedure based on a MultiStart strategy.

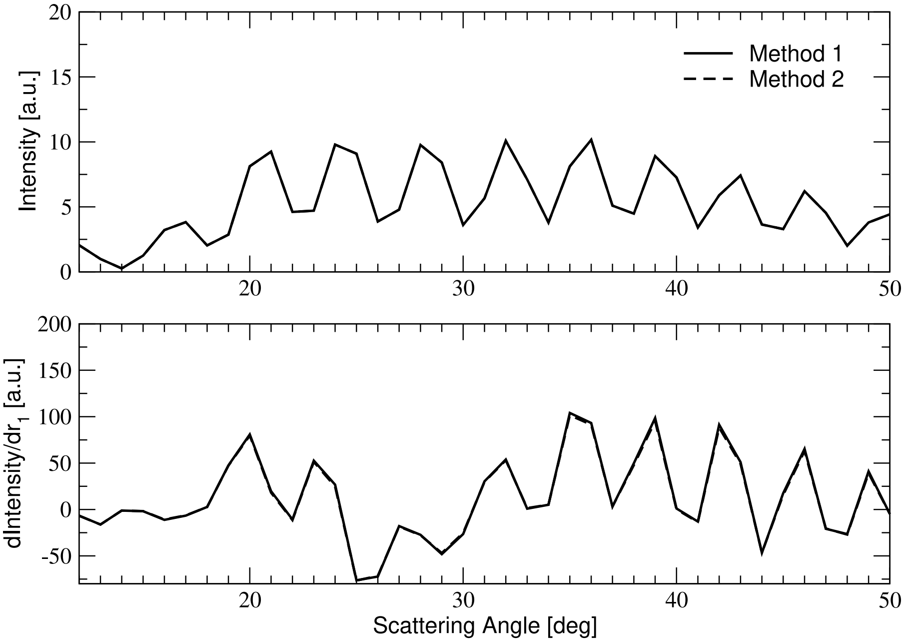

Fig. 1. Weighted intensity and its derivative with respect to the outer radius of the sphere r_1 as functions of the scattering angle. The results correspond to a coated sphere with a core radius of 3.0 μm and an outer radius of 3.65 μm, placed in buffered saline solution with a refractive index of 1.337. The relative refractive indices of the core and the shell in vacuum are 1.44 and 1.38, respectively. The wavelength in vacuum is 0.488 μm, and the angular domain of analysis ranges from 12° to 50° with an angular spacing of 0.5°.

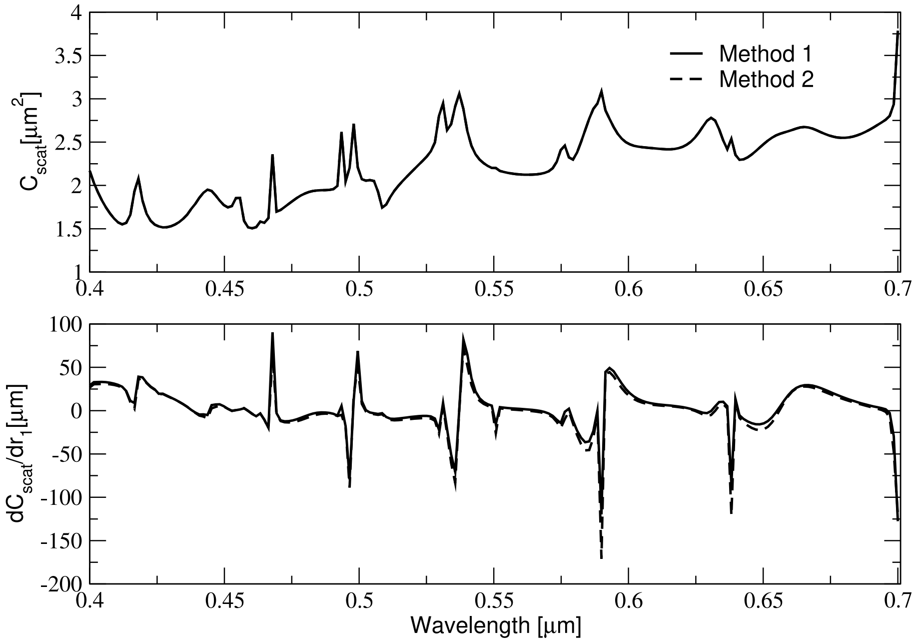

Fig. 2. Scattering efficiency and its derivative with respect to the outer radius of the sphere r_1 as functions of the wavelength. The results correspond to an eight-layered sphere with individual layer thicknesses of 0.07 μm and alternating refractive indices of titanium dioxide and silica. The refractive index of silica is fixed at 1.428, while the refractive index of titanium dioxide is wavelength dependent and given by $m_{\mathrm{TiO}_2}(\lambda) = \sqrt{5.913 + 0.2441/(\lambda^2 - 0.0803)}$, where λ denotes the wavelength. The scattering cross section is normalized by a reference radius of 1.0 μm. A total of 200 spectral points were chosen in a wavelength interval ranging from 0.4 μm to 0.7 μm.

3.1. Inverse problem formulation

The inversion algorithm is designed for nonlinear inverse problems of the form

$$\mathbf{F}(\mathbf{x}) = \mathbf{y}^\delta,$$

where $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ is the forward model, $\mathbf{x} \in \mathbb{R}^n$ is the unknown state vector, and $\mathbf{y}^\delta \in \mathbb{R}^m$ is the measured data vector. The measurements are assumed to be contaminated by additive Gaussian noise with zero mean and covariance matrix C_δ, i.e., $\mathbf{y}^\delta = \mathbf{y} + \boldsymbol{\delta}$, where $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \mathrm{C}_\delta)$ and $\mathbf{y}$ denotes the true data vector. The objective is to estimate the unknown parameter vector, quantify the uncertainty of the solution, and incorporate prior information and data-error correlations in a consistent statistical setting.

A key element of the formulation is the decomposition $\mathrm{C}_\delta = \sigma^2 \mathscr{C}_\delta$ with $\mathrm{trace}(\mathscr{C}_\delta) = m$, where σ^2 represents the overall noise variance (i.e., the magnitude of the data errors), and $\mathscr{C}_\delta$ is the shape covariance matrix, which describes the correlation structure of the errors independently of their scale. This decomposition separates the noise level from the correlation pattern of the measurement errors and is particularly convenient for whitening, regularization, residual analysis, and discrepancy-based stopping criteria.

3.2. Data covariance model and synthetic measurements

Several practical situations are distinguished according to the available information on the data covariance.

1. When the true covariance matrix is known, it is normalized directly to obtain the shape covariance matrix.
2. When only the covariance shape is known up to a scale factor, the missing scale is inferred from a prescribed signal-to-noise ratio.
3. If no covariance information is available, a diagonal covariance model is adopted, again with the variance estimated from the signal-to-noise ratio.
4. For synthetic tests, a covariance model can be generated artificially and then scaled so that its overall variance is compatible with a prescribed noise level.

For inverse problems involving multilayered spheres, in addition to measurement noise, uncertainties in the forward model parameters also contribute to the overall data uncertainty. These include, for example, (i) deviations of the layer boundaries from perfect spherical geometry and (ii) offsets of the layer centers relative to the particle center. In principle, such uncertainties can be quantified by employing a scattering model for nonspherical particles and performing a Monte Carlo analysis under the assumption that the statistics of the model parameters are known. The outcome of this procedure is a covariance matrix of the forward model parameter errors in the data space, which can serves as input for the inversion algorithm (Case 1 above).

For synthetic problems, the shape covariance matrix may be chosen as $\mathscr{C}_\delta = \mathsf{D}\mathsf{P}(\boldsymbol{\phi})\mathsf{D}$, where D is a diagonal scaling matrix derived from pointwise standard deviations and $\mathsf{P}(\boldsymbol{\phi})$ is a correlation matrix defined through a parametric autocorrelation function. The correlation structure may be modeled, for example, by exponential, Gaussian, damped cosine, Matérn, AR(1), Wendland, or rational quadratic kernels. In this context, synthetic measurements are generated by adding Gaussian noise with the prescribed covariance to the exact forward data. If $\mathscr{C}_\delta = \mathsf{R}\mathsf{R}^\top$ is the Cholesky factorization of the normalized covariance shape matrix, then a realization of the data error is obtained as $\boldsymbol{\varepsilon} = \overline{\sigma}\,\mathsf{R}\mathbf{z}$, where $\overline{\sigma}^2$ is an a priori estimate of the data-error variance and $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathrm{I}_m)$. The measurement vector is then computed as $\mathbf{y}^\delta = \mathbf{y} + \boldsymbol{\varepsilon}$.

3.3. Forward model and data transformations

The forward model considered in this work is representative of inverse problems arising in Mie-scattering applications. Depending on the measurement configuration, the data may consist either of weighted scattering intensities sampled over the scattering angle or of scattering cross sections sampled over wavelength. In the intensity-based setting, a weighting function can be introduced to emphasize specific angular regions of interest.

To improve robustness or reduce dimensionality, several data transformations may be applied before solving the inverse problem.

1. Principal Component Analysis (PCA) projects the data onto a low-dimensional subspace spanned by dominant eigenvectors of a sample covariance matrix.
2. Low-pass filtering based on the Discrete Cosine Transform (DCT) retains only the dominant low-frequency components of the data.
3. A differential model with polynomial detrending removes a low-order polynomial trend from both simulated and measured data.

In each case, the transformed inverse problem retains the same statistical structure, but with a modified covariance matrix and a reduced effective data dimension. A detailed description of the data transformation techniques is provided in Appendix B.

3.4. Maximum-likelihood and Bayesian formulation

Under the Gaussian data-error model $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathscr{C}_\delta)$, the likelihood leads to the generalized least-squares functional

$$\mathcal{F}_{\mathrm{x}}(\mathbf{x}) = \frac{1}{2}(\mathbf{y}^\delta - \mathbf{F}(\mathbf{x}))^\top \mathscr{C}_\delta^{-1}(\mathbf{y}^\delta - \mathbf{F}(\mathbf{x})).$$

Its minimizer provides the maximum-likelihood estimate when the covariance shape matrix is known.

A Bayesian extension is obtained by assigning a Gaussian prior to the unknown vector, $\mathbf{x} \sim \mathcal{N}(\mathbf{x}_{\mathrm{a}}, \mathsf{C}_{\mathrm{x}})$, where $\mathsf{C}_{\mathrm{x}} = \sigma_{\mathrm{x}}^2\mathscr{C}_{\mathrm{x}}$ is the a priori covariance matrix, σ_{x}^2 represents the overall variance (i.e., the magnitude of the prior uncertainty in the state vector), and $\mathscr{C}_{\mathrm{x}}$ is the shape covariance matrix describing the correlation structure of the prior errors independently of their scale. This leads to the regularized objective function

$$\mathcal{F}_{\mathrm{x}\alpha}(\mathbf{x}) = \frac{1}{2}\left[(\mathbf{y}^\delta - \mathbf{F}(\mathbf{x}))^\top \mathscr{C}_\delta^{-1}(\mathbf{y}^\delta - \mathbf{F}(\mathbf{x})) + \alpha(\mathbf{x} - \mathbf{x}_{\mathrm{a}})^\top \mathscr{C}_{\mathrm{x}}^{-1}(\mathbf{x} - \mathbf{x}_{\mathrm{a}})\right],$$

where $\alpha = \sigma^2/\sigma_{\mathrm{x}}^2$ is the regularization parameter. This establishes the connection between Bayesian estimation and Tikhonov regularization.

In addition to point estimates, e.g., $\mathbf{x}^\delta = \arg\min_{\mathbf{x}} \mathcal{F}_{\mathrm{x}}(\mathbf{x})$, marginal posterior expectations may be approximated by importance sampling or Metropolis–Hastings sampling. In the present work, the marginal posterior expectation constitutes a central component of the inversion methodology, as it provides a statistically consistent estimate of the solution and its uncertainty. In particular, it enables the computation of posterior mean values and covariance matrices that incorporate both data noise and model-induced variability. To account for correlated residuals, an effective number of degrees of freedom is introduced, leading to a modified posterior distribution that better reflects the true information content of the data. The resulting expectations are efficiently evaluated by importance sampling (IS), with proposal distributions constructed around the local minima of the objective function, or by Metropolis–Hastings (MH) sampling, in which the mean and variance are estimated from samples generated by a Markov chain targeting the posterior density. A detailed description of this approach is provided in Appendix C.

The objective functions $\mathcal{F}_{\mathrm{x}}(\mathbf{x})$ and $\mathcal{F}_{\mathrm{x}\alpha}(\mathbf{x})$ are minimized subject to simple bound constraints on the variables. Constraining the solution to the box $X = [a_1, b_1] \times [a_2, b_2] \ldots \times [a_n, b_n]$, i.e., $a_i \le x_i \le b_i$ for $i = 1, \ldots, n$, provides control over the magnitude of the state vector and reduces instabilities caused by data errors. In maximum-likelihood estimation, such bounds can be viewed as a simple form of regularization: by limiting the admissible domain, they prevent unbounded solutions and implicitly penalize large parameter values, in a manner analogous to an $\|\mathbf{x}\|^2$-type regularization. Although not a substitute for explicit regularization schemes, box constraints represent a practical option when no prior information on the solution is available.

3.5. Whitening and least-squares reformulation

A major computational simplification is achieved through whitening. If $\mathscr{C}_\delta = \mathsf{R}\mathsf{R}^\top$,then the whitening matrix $\mathsf{W} = \mathsf{R}^{-1}$ transforms the generalized least-squares problem into an ordinary one,

$$\mathcal{F}_{\mathrm{x}}(\mathbf{x}) = \frac{1}{2}\|\widehat{\mathbf{y}}^\delta - \widehat{\mathbf{F}}(\mathbf{x})\|^2,$$

where $\widehat{\mathbf{y}}^\delta = \mathsf{W}\mathbf{y}^\delta$ and $\widehat{\mathbf{F}}(\mathbf{x}) = \mathsf{W}\mathbf{F}(\mathbf{x})$. In the whitened space, the transformed data error $\widehat{\boldsymbol{\delta}} = \mathsf{W}\boldsymbol{\delta}$ is Gaussian with covariance $\sigma^2\mathrm{I}_m$, so the problem becomes a classical least-squares problem with uncorrelated homoscedastic errors. A similar transformation may be applied to the regularization term: if the a priori shape covariance matrix is factorized as $\mathscr{C}_{\mathrm{x}} = \mathsf{R}_{\mathrm{x}}\mathsf{R}_{\mathrm{x}}^{\mathrm{T}}$, then the regularization matrix is defined as $\mathsf{L} = \mathsf{R}_{\mathrm{x}}^{-1}$, which leads to an augmented least-squares formulation that combines data fidelity and prior information in a single residual vector.

3.6. Discrepancy principle and regularization strategy

In Bayesian estimation, the discrepancy principle and the choice of the a priori state vector, regularization matrix, and regularization parameter play a central role.

1. Discrepancy principle. Given an a priori estimate $\overline{\sigma}^2$ of the data-error variance, the squared error level is defined by $\Delta^2 = m\overline{\sigma}^2$. A computed solution is accepted if its whitened residual norm $R(\mathbf{x}^\delta) = \|\hat{\mathbf{r}}\|^2$ satisfies $R(\mathbf{x}^\delta) \le \tau\Delta^2$, with τ a safety factor close to one. Since the expected residual norm after fitting n parameters is $(m-n)\sigma^2$, this criterion is conservative. In practice, when the a priori estimate of the noise level is uncertain, the ratio between the achieved residual norm and the prescribed error level is best interpreted as a diagnostic indicator.
2. A priori information and regularization matrix. When only bounds $a_i \le x_i \le b_i$ are available, the a priori vector is chosen as the midpoint of the admissible interval, $x_{\mathrm{a}i} = (a_i + b_i)/2$, and the prior standard deviations are estimated by $\sigma_i = (b_i - a_i)/3.92$, which corresponds to interpreting the bounds as an approximate 95% confidence interval of a Gaussian prior. This yields a diagonal prior covariance matrix and a regularization matrix that penalizes strongly constrained parameters more heavily than weakly constrained ones. Specifically, the diagonal elements of the regularization matrix are $[\mathsf{L}]_{ii} = \sqrt{\sigma_{\mathrm{x}}^2}/\sigma_i$, where $\sigma_{\mathrm{x}}^2 = (1/n)\sum_{i=1}^{n}\sigma_i^2$ is the average of the component-wise prior variances.
3. Regularization parameter. The regularization parameter α can be kept constant during the iteration, or it can be varied. In the latter case, it is chosen in a practical manner: relatively large values are used at the beginning of the iteration to stabilize the problem and suppress undesirable local minima, and α is then decreased gradually during the iteration. This strategy is closely related to iteratively regularized Gauss–Newton methods.

3.7. MultiStart strategy

Because the objective function may possess several local minima, the inversion algorithm uses a MultiStart strategy to explore the admissible domain more broadly [39–46]. The method is organized into a sequence of configurations. At each configuration, a set of candidate points is generated, the objective function is evaluated at these points, and only the most promising candidates are retained for local optimization. The resulting local minimizers are then compared with the set of previously accepted solutions to determine whether a genuinely new solution has been found. In this way, the method seeks not only the minimum-residual estimate but also the collection of distinct admissible solutions compatible with the data.

3.7.1. Generation and validation of starting points

Candidate starting points may be generated by pseudo-random number generators [47,48], chaotic methods [49–51], low discrepancy methods (Halton, Sobol, and Faure sequences), Latin hypercube sampling [52], quasi-oppositional differential evolution [53,54], and centroidal Voronoi tessellation [55]. In the present implementation, pseudo-random generators and low-discrepancy sequences are used.

At a given configuration, a set of $N_{0\mathrm{C}}$ candidate points is generated in the admissible domain. The objective function is evaluated at these points. If the global accumulation option is enabled, the newly generated samples and their objective values are added to the collection accumulated from previous configurations. Otherwise, only the points generated in the current configuration are used. The objective values are sorted in ascending order, the candidate points are reordered consistently, and only the best N_0 points are retained as effective starting points for local optimization.

Before launching a local optimization, candidate starting points may be screened in order to avoid repeated descents into already known basins of attraction. Several validation tests are available in the framework [43]. In the present implementation, only a basin-exclusion test is activated. A candidate starting point $\mathbf{x}_{0k}$ is rejected if there exists a previously found solution $\mathbf{x}_{\mathrm{S}n}$ such that $\|\mathbf{x}_{0k} - \mathbf{x}_{\mathrm{S}n}\| < R_{\mathrm{SB}}(n)$, where $R_{\mathrm{SB}}(n)$ is an estimate of the basin radius associated with $\mathbf{x}_{\mathrm{S}n}$.

3.7.2. Residual statistics and uncertainty quantification

The statistics of the residuals provide the basis for variance estimation and uncertainty analysis. Under a local linearization of the forward map around the accepted solution $\mathbf{x}_{\mathrm{S}}$, the whitened residual vector $\hat{\mathbf{r}} = \mathsf{W}\mathbf{r} = \hat{\mathbf{y}}^\delta - \hat{\mathbf{F}}(\mathbf{x}_{\mathrm{S}}$ satisfies $\hat{\mathbf{r}} \approx (\mathrm{I}_m - \hat{\mathsf{A}})\hat{\boldsymbol{\delta}}$, where $\hat{\mathsf{A}} = \hat{\mathsf{K}}(\hat{\mathsf{K}}^\top\hat{\mathsf{K}})^{-1}\hat{\mathsf{K}}^\top$ is the influence matrix, provided that the whitened Jacobian $\hat{\mathsf{K}}$ has full column rank. Hence, $\mathbb{E}\{\|\hat{\mathbf{r}}\|^2\} = (m-n)\sigma^2$, and, in the unregularized case, the data-error variance can be estimated by

$$\hat{\sigma}^2 = \frac{1}{m-n}\|\hat{\mathbf{r}}\|^2. \tag{1}$$

When regularization is present the influence matrix becomes $\hat{\mathsf{A}}_\alpha = \hat{\mathsf{K}}(\hat{\mathsf{K}}^\top\hat{\mathsf{K}}+\alpha\mathsf{L}^\top\mathsf{L})^{-1}\hat{\mathsf{K}}^\top$, and the unbiased residual-based variance estimator involves the denominator $m - 2\,\mathrm{trace}(\hat{\mathsf{A}}_\alpha) + \mathrm{trace}(\hat{\mathsf{A}}_\alpha^2)$.

The same linearized framework yields expressions for the covariance matrix of the estimation error. In the unregularized case,

$$\hat{\mathsf{C}}_{\mathrm{x}} = \hat{\sigma}^2(\mathsf{K}^\top\mathscr{C}_\delta^{-1}\mathsf{K})^{-1}, \tag{2}$$

whereas in the regularized case the Laplace approximation gives $\hat{\mathsf{C}}_{\mathrm{x}} = \hat{\sigma}^2(\mathsf{K}^\top\mathscr{C}_\delta^{-1}\mathsf{K} + \alpha\mathsf{L}^\top\mathsf{L})^{-1}$.

3.7.3. Local optimization and distinct-solution test

Each validated starting point is refined by a local optimizer. Two local solvers are available. The first is the routine DRN2GB from the PORT library [56,57], a derivative-based algorithm for bounded nonlinear least-squares problems that combines Gauss–Newton steps with quasi-Newton corrections. The second is the limited-memory BFGS method [58], which approximates the inverse Hessian by a small number of correction vectors and is well suited for medium- and large-scale problems.

After each local search, the candidate minimizer is compared with the set of already accepted solutions. A newly computed solution $\mathbf{x}_{\mathrm{S}}$ is considered distinct if

$$\|\mathbf{x}_{\mathrm{S}} - \mathbf{x}_{\mathrm{S}n}\| > \tau \max\left(\|\mathbf{x}_{\mathrm{S}}\|, \|\mathbf{x}_{\mathrm{S}n}\|\right), \qquad n = 1, \ldots, N_{\mathrm{S}},$$

where $\tau > 0$ is a prescribed tolerance. If this condition is satisfied for all n, the candidate is inserted into the solution set; otherwise, it is regarded as numerically identical to a previously found solution.

3.7.4. Hybrid variants and algorithmic structure

In addition to the pure MultiStart strategy, one may combine a global exploration method with a second-order local refinement. Possible global methods include DIRECT [59], JADE [60], LSHADE [61], and LJADE [62]. Their purpose is to reduce the sensitivity of the algorithm to the initial sampling strategy. In the present implementation, however, the logical variables controlling the hybrid stage are set internally to `false`, so the reported results correspond to the pure MultiStart variant.

The pseudocode for the multistart algorithm is presented in Algorithm 1. The input parameters of the inversion procedure are the noisy data vector $\mathbf{y}^\delta$, the number of effective starting points per configuration

N_0, the total number of candidate points generated at each configuration $N_{0\mathrm{C}}$, the number of configurations N_C, and the tolerance τ used in the distinct-solution test. The algorithm initializes the number of distinct solutions $N_\mathrm{S} = 0$, the set of distinct solution X_S, the corresponding set of objective-function values $\mathcal{F}_\mathrm{S}$, the set of residual vectors R_S, the set of data error variance estimators V_{yS}, the set of the diagonal elements of the covariance matrix of the estimation error V_{xS}, the set of basin radii associated with each distinct solution R_{SB}, and the minimum of the objective function over all configurations and starting points $\mathcal{F}_{\min}$. For each successful local search, the algorithm stores the minimizer $\mathbf{x}_\mathrm{S}$, the objective value $\mathcal{F}_\mathrm{x}(\mathbf{x}_\mathrm{S})$, the whitened residual vector $\hat{\mathbf{r}}(\mathbf{x}_\mathrm{S})$, the estimator of the data-error variance $\hat{\sigma}^2(\mathbf{x}_\mathrm{S})$, the diagonal elements of the covariance matrix of the estimation error $\boldsymbol{\sigma}_\mathrm{x}^2(\mathbf{x}_\mathrm{S}) = \mathrm{diag}\big(\hat{\mathrm{C}}_\mathrm{x}(\mathbf{x}_\mathrm{S})\big)$, and an estimate $R_\mathrm{B}(\mathbf{x}_\mathrm{S})$ of the basin radius. The final output is the set of admissible solutions together with their residuals, variance estimators, covariance information, and the minimum-residual estimate. In particular, the minimum-residual solution is selected as the point estimate of the unknown parameter vector.

Algorithm 1 Multistart algorithm. The meaning of the variables is described in the text.

```
1. Inputs: $\mathbf{y}^\delta$, $N_0$, $N_{0C}$, $N_C$, and $\tau$.
2. Initialization: Set $N_S = 0$, $X_S = \emptyset$, $\mathcal{F}_S = \emptyset$, $R_S = \emptyset$, $V_{yS} = \emptyset$, $V_{xS} = \emptyset$, $R_{SB} = \emptyset$, and $\mathcal{F}_{\min} = 10^{30}$.
3. Configuration Loop:
For each configuration $k$, $k = 1, \ldots, N_C$ do:
  • Generate Starting Points: Generate a random set of $N_{0C}$ points and store them in $S_{0C}$. At each point, evaluate the objective function and store the results in $\mathcal{F}_{0C}$. If the global accumulation option is enabled, append this set to the global collections: $N_{0A} = kN_{0C}$, $S_{0A} = S_{A0} \cup S_{0c}$, $\mathcal{F}_{0A} = \mathcal{F}_{0A} \cup \mathcal{F}_{0C}$. Otherwise (independent treatment per configuration), set $N_{0A} = N_{0C}$, $S_{0A} = S_{0c}$, $\mathcal{F}_{0A} = \mathcal{F}_{0C}$. Sort $\mathcal{F}_{0A}$ in increasing order, applying the same permutation to $S_{A0}$. Select the first $N_0$ columns of $S_{A0}$ to form the effective set $S_0$.
  • Solution Search:
    – For each starting point $\mathbf{x}_{0j} \in S_0$, $j = 1, \ldots, N_0$ do:
        ∗ Validity check: If the MultiStart algorithm is selected, accept $\mathbf{x}_{0j}$ only if it lies outside all known basin $R_{SB}(N_S)$ around $X_S$; otherwise, accept $\mathbf{x}_{0j}$ directly.
        ∗ If $\mathbf{x}_{0j}$ is valid:
            · Optional preliminary global steps: If the DIRECT algorithm or a stochastic algorithm is selected, run the algorithm from $\mathbf{x}_{0j}$ to obtain $\mathbf{x}_S$ and set $\mathbf{x}_{0j} = \mathbf{x}_S$.
            · Second-order refinement: Run PORT or BFGS, with or without regularization, starting from $\mathbf{x}_{0j}$ to obtain $\mathbf{x}_S$, $\mathcal{F}_x$, $\hat{\mathbf{r}}$, $\hat{\sigma}^2$, $\boldsymbol{\sigma}_x^2$, and $R_B$. If $\mathbf{x}_S \notin X_S$ with tolerance $\tau$, (i) set $N_S = N_S + 1$, $X_S = X_S \cup \{\mathbf{x}_S\}$, $\mathcal{F}_S(N_S) = \mathcal{F}_x$, $R_S = R_S \cup \{\hat{\mathbf{r}}\}$, $V_{yS}(N_S) = \hat{\sigma}^2$, $V_{xS} = V_{xS} \cup \{\boldsymbol{\sigma}_x^2\}$, and $R_{SB}(N_S) = R_B$, and (ii) track the minimum-residual solution so far: if $\mathcal{F}_x < \mathcal{F}_{\min}$, update $\mathcal{F}_{\min} = \mathcal{F}_x$ , $\mathbf{x}_{S\min} = \mathbf{x}_S$, $\hat{\mathbf{r}}_{\min} = \hat{\mathbf{r}}$, $\hat{\sigma}^2_{\min} = \hat{\sigma}^2$, and $\boldsymbol{\sigma}^2_{x\min} = \boldsymbol{\sigma}^2_x$.
4. Outputs: Return the sets $X_S$, $\mathcal{F}_S$, $R_S$, $V_{yS}$, and $V_{xS}$ corresponding to all solutions, and the quantities $\mathbf{x}_{S\min}$, $\mathcal{F}_{\min}$, $\hat{\mathbf{r}}_{\min}$, $\hat{\sigma}^2_{\min}$, and $\boldsymbol{\sigma}^2_{x\min}$ corresponding to the minimum-residual solution.
```

4. Numerical simulations

We consider a multilayered sphere with L concentric layers of thicknesses d_ℓ and relative refractive indices m_ℓ with respect to the ambient medium. The radii of the layers are computed as

$$r_\ell = \begin{cases} r_{\ell+1} + d_\ell, & \ell = 1, \ldots, L-1, \\ d_L, & \ell = L, \end{cases}$$

meaning that the core of the particle has radius $r_L = d_L$, and the radius of the outermost layer (i.e., of the entire sphere) is r_1.

The inverse problem based on the angular distribution of the scattered intensity is referred to as Problem 1, whereas that based on the spectral dependence of the scattering cross section is Problem 2. In Problem 1, we aim to reconstruct both the layer thicknesses and refractive indices. The wavelength of the incident radiation is $\lambda = 0.488\,\mu\mathrm{m}$, the particle is immersed in liquid, and the angular range is from $\theta_{\min} = 12^\circ$ to $\theta_{\max} = 50^\circ$. Problem 2 involves only the reconstruction of the layer thicknesses for a particle in vacuum. In this case, we consider $N_\lambda = 200$ spectral points in a range from $\lambda_{\min} = 0.4\,\mu\mathrm{m}$ to $\lambda_{\max} = 0.7\,\mu\mathrm{m}$, and choose alternating refractive indices $m_{\mathrm{TiO_2}}(\lambda) = \sqrt{5.913 + 0.2441/(\lambda^2 - 0.0803)}$ (Titanium dioxide) and $m_{\mathrm{SiO_2}} = 1.428$ (Silica).

The inverse problem is solved using the maximum-likelihood estimate, and the PORT optimizer. For the Multistart algorithm, the number of effective starting points per configuration is set to $N_0 = 50$, and the total number of starting points generated at each configuration step to $N_{0\mathrm{C}} = 1000$.

4.1. Data affected by forward model errors

In the first test we consider a coated-sphere model, i.e. a specific case of Problem 1, to characterize lymphocytes from light-scattering patterns measured with the scanning flow cytometry [8]. We consider the displacement of the core (nucleus) relative to the particle center as an example of forward model errors. Simulations of light scattering by the nonsymmetric optical model of a lymphocyte are performed with the ADDA computer code (v. 1.5.0-alpha3). This is an efficient DDA implementation on a cluster of computers, parallelizing a single DDA computation [63,64]. Default code parameters were used.

The parameters of the coated sphere, which are the components of the true state vector $\mathbf{x}^\dagger$, are chosen as $d_1^\dagger = 0.65\,\mu\mathrm{m}$, $d_2^\dagger = 3.0\,\mu\mathrm{m}$, $m_1^\dagger = 1.38$, and $m_2^\dagger = 1.44$. The coated sphere is placed in a buffered saline solution with a refractive index of 1.337. The core is shifted along the x-axis (only positive direction, since negative is equivalent for the computed quantity), and along the z-axis (both positive and negative) in steps of 0.1 μm. For a beam propagating along the positive direction of the z-axis, the scattered intensity is computed at a set of discrete angles in the interval $[12^\circ, 50^\circ]$ with a step of $\Delta\theta = 0.5^\circ$, averaged over the azimuthal scattering angle, and weighted by the angular function

$$w(\theta) = \frac{1^\circ}{\theta^\circ} \exp\left(-2\ln^2\left(\frac{\theta^\circ}{54^\circ}\right)\right). \tag{3}$$

These intensities, illustrated in Fig. 3, represent the synthetic measurements $\mathbf{y}^\delta$ for the inverse problem, while $\mathbf{y} = \mathbf{F}(\mathbf{x}^\dagger)$ denotes the true data vector.

The inverse problem is solved assuming an isotropic shape covariance matrix, i.e., $\mathscr{C}_\delta = \mathrm{I}_m$, and the following simple bounds on the variables: $d_1 \in [0.6, 0.7]$ μm, $d_2 \in [1.4, 3.8]$ μm, $m_1 \in [1.34, 1.41]$, and $m_2 \in [1.41, 1.58]$. The minimum-residual solution $\mathbf{x}^\delta$ is computed using the MultiStart algorithm, and the posterior expectation of $\mathbf{x}$, i.e., $\mathbb{E}\{\mathbf{x}\} = \langle \mathbf{x} \rangle$ is evaluated using marginal posterior estimation via IS and MH sampling. Referring to Appendix C, the latter is computed as (cf. Eq. (C.15))

$$\mathbb{E}\{\mathbf{x}\} = \frac{\int \mathbf{x}\, p(\mathbf{x} \mid \mathbf{y}^\delta)\, d\mathbf{x}}{\int p(\mathbf{x} \mid \mathbf{y}^\delta)\, d\mathbf{x}},$$

where (cf. Eq. (C.14))

$$p(\mathbf{x} \mid \mathbf{y}^\delta) \propto [R(\mathbf{x})]^{-\kappa/2}$$

is the marginal posterior, $R(\mathbf{x}) = \|\mathbf{r}(\mathbf{x})\|^2$ is the residual energy, $\mathbf{r}(\mathbf{x}) = \mathbf{y}^\delta - \mathbf{F}(\mathbf{x})$ is the whitened residual at $\mathbf{x}$,

$$\kappa = \frac{m^2}{m + 2\sum_{k=1}^{m-1}(m-k)\rho_k^2}$$

is the effective number of independent observations (cf. Eq. (C.10)), and ρ_k is the autocorrelation function at lag k, computed using Eqs. (C.37), (C.38), and (C.42). The autocorrelation function is computed from the residuals at $\mathbf{x}^\delta$, i.e., $\mathbf{r}(\mathbf{x}^\delta) = \mathbf{y}^\delta - \mathbf{F}(\mathbf{x}^\delta)$. Tapering is not applied, as the autocorrelation function remains significant over the entire range. Typical examples of this function are shown in Fig. 4. The expectation $\mathbb{E}\{\mathbf{x}\}$ is computed using (i) diagonal-covariance importance sampling with $N_\mathrm{s} = 10^4$ samples and (ii) Metropolis–Hastings sampling with a total chain length of $N = 3 \times 10^4$, a burn-in length of $N_\mathrm{burn} = 10^4$, and $N_\mathrm{keep} = 10^4$ retained samples (see Appendix C).

Under the Laplace approximation, the posterior distribution is locally approximated by a Gaussian distribution centered at the minimum-residual solution. Accordingly, the posterior covariance matrix can be approximated by the covariance matrix of the estimation error. To estimate this covariance matrix, we consider both the minimum-residual covariance approximation and the marginal posterior expectation.

1. Minimum-residual solution. Let $\mathbf{x}^\delta$ be the minimum-residual solution, and assume that the forward map is locally linearized around $\mathbf{x}^\delta$. Then the data-error variance can be estimated by (cf. Eq. (1))

$$\hat{\sigma}^2 = \frac{1}{m-n}\|\hat{\mathbf{r}}\|^2, \tag{4}$$

   and the posterior covariance matrix by (cf. Eq. (2)) $\widehat{\mathrm{C}}_\mathrm{x} = \hat{\sigma}^2(\widehat{\mathrm{K}}^\top\widehat{\mathrm{K}})^{-1}$, where $\widehat{\mathrm{K}}$ is the whitened Jacobian matrix at $\mathbf{x}^\delta$. This expression is valid provided that $\widehat{\mathrm{K}}$ has full column rank.
2. Marginal posterior expectation with effective degrees of freedom. When the solution is defined in terms of the marginal posterior expectation, the posterior covariance matrix is computed directly from the posterior second central moment (Eq. (C.15)): $\widehat{\mathrm{C}}_\mathrm{x} = \mathbb{E}\{(\mathbf{x} - \langle\mathbf{x}\rangle)(\mathbf{x} - \langle\mathbf{x}\rangle)^\mathrm{T}\}$.

The results in Table 1 illustrate the effect of forward-model errors on the reconstruction accuracy when scattering data generated by a nonsymmetric optical model (a displaced core) are inverted using a concentric coated-sphere model. The displacement introduces a model discrepancy that propagates into the reconstructed parameters. The table reports the empirical signal-to-noise ratio, defined as the ratio between the signal energy and the energy of the realized noise vector

$$\mathrm{SNR}_\mathrm{e} = \frac{\|\mathbf{y}\|^2}{\left\|\mathbf{y}^\delta - \mathbf{y}\right\|^2}, \tag{5}$$

and the relative errors obtained using the minimum-residual solution and the marginal posterior expectation, namely,

$$\varepsilon_\mathrm{x} = \frac{\left\|\mathbf{x}^\delta - \mathbf{x}^\dagger\right\|}{\left\|\mathbf{x}^\dagger\right\|} \text{ and } \varepsilon_{\langle\mathrm{x}\rangle} = \frac{\left\|\langle\mathbf{x}\rangle - \mathbf{x}^\dagger\right\|}{\left\|\mathbf{x}^\dagger\right\|}, \tag{6}$$

respectively. Note that this definition gives more weight to the sizes in comparison to the refractive indices, since the ratio of full width of admissible parameter range to its true value is 15%–80% vs. 5%–12%, respectively. To address this imbalance and to assess the statistical calibration of the uncertainty estimates, we introduce the normalized errors.

$$z_i = \frac{x_i^\delta - x_i^\dagger}{s_i}, \quad \langle z_i\rangle = \frac{\langle x_i\rangle - x_i^\dagger}{\langle s_i\rangle} \qquad i = 1, \ldots, n, \tag{7}$$

where s_i are the standard deviations $s_i = \sqrt{(\widehat{\mathrm{C}}_\mathrm{x})_{ii}}$ obtained from the minimum-residual covariance matrix $\widehat{\mathrm{C}}_\mathrm{x} = \hat{\sigma}^2(\widehat{\mathrm{K}}^\top\widehat{\mathrm{K}})^{-1}$, whereas $\langle s_i\rangle$ are the same, but for the marginal posterior covariance matrix.

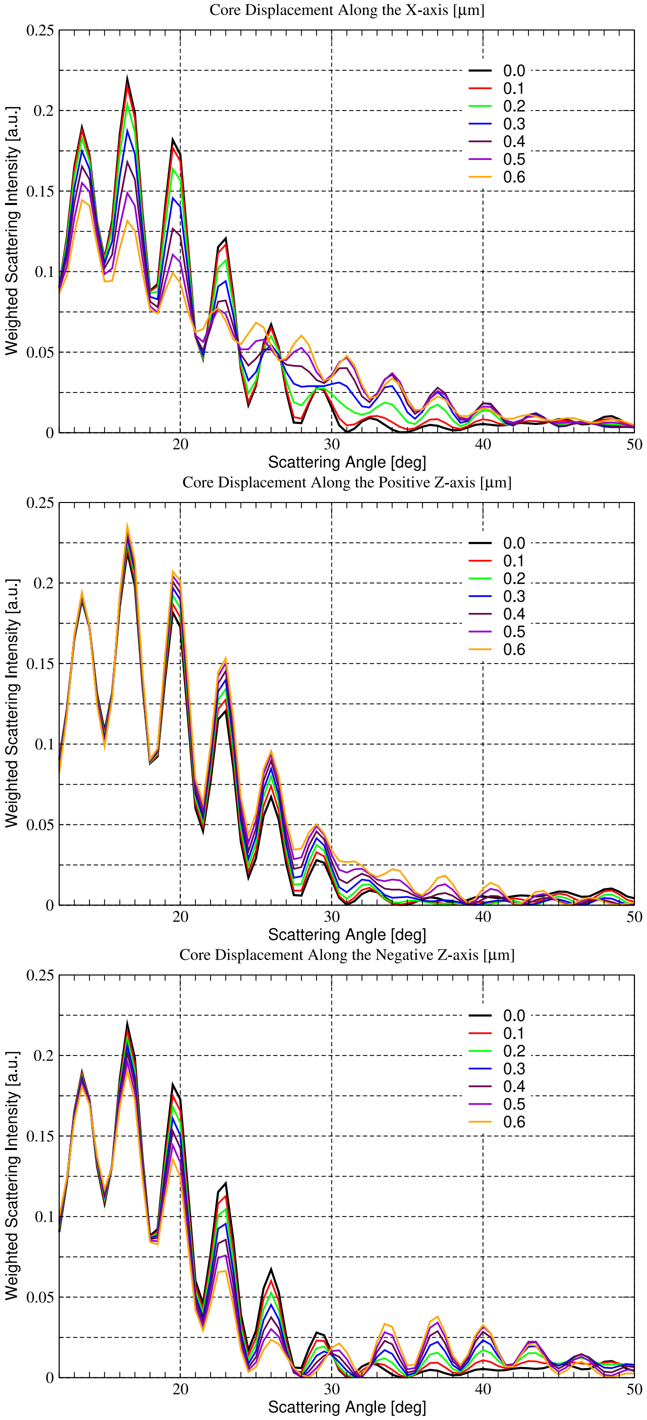

Fig. 3. Scattering intensities of a core–shell particle for different core displacements. The shell and core thicknesses are 0.65 µm and 3.0 µm, respectively, and the corresponding refractive indices are 1.38 and 1.44. The coated sphere is embedded in a buffered saline solution with a refractive index of 1.337.

We also introduce the combined indicators

$$Q_\mathrm{x} = \left(\sum_{i=1}^{n} z_i^2\right)^{1/2}, \qquad Q_{\langle\mathrm{x}\rangle} = \left(\sum_{i=1}^{n}\langle z_i\rangle^2\right)^{1/2}. \tag{8}$$

Let $\mathbf{e} = \mathbf{x}^\delta - \mathbf{x}^\dagger$ denote the estimation error and assume that $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathrm{C}_\mathrm{x})$, where C_x is the (unknown) true covariance matrix. If the

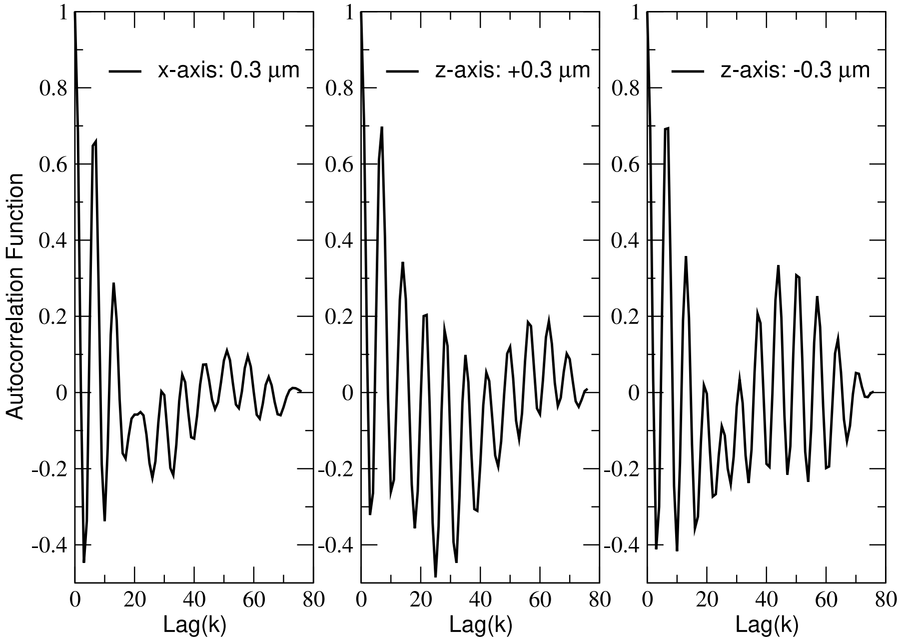

Fig. 4. Residual autocorrelation functions for different core displacements. The geometrical and optical parameters of the coated sphere are the same as in Fig. 3.

Table 1
Empirical signal-to-noise ratio, relative reconstruction errors, and uncertainty calibration indicators for the minimum-residual solution and the marginal posterior expectation. The marginal posterior expectation is evaluated using importance sampling (IS) and Metropolis–Hastings (MH) sampling. The geometrical and optical parameters of the coated sphere are the same as in Fig. 3.

Axis	Displ. [µm]	SNR	Minimum residual		Marginal posterior expectation			
					IS		MH	
			ε_{x} [%]	Q_{x}	$\varepsilon_{\langle \mathrm{x}\rangle}$ [%]	$Q_{\langle \mathrm{x}\rangle}$	$\varepsilon_{\langle \mathrm{x}\rangle}$ [%]	$Q_{\langle \mathrm{x}\rangle}$
x	0.1	1271	0.076	5.3	0.054	1.1	0.072	1.4
	0.2	103	0.42	8.1	0.31	2.1	0.27	1.4
	0.3	29	1.3	12	0.43	2.6	0.32	1.4
	0.4	14	15	82	14	39	14	23
	0.5	9	15	93	15	43	15	24
	0.6	6	15	99	15	36	15	21
$+z$	0.1	681	0.47	12	0.47	4.7	0.47	3.6
	0.2	180	0.91	13	0.91	4.4	0.88	3.7
	0.3	87	1.3	14	1.1	4.9	1.1	3.5
	0.4	49	1.5	12	1.1	4.3	1	2.4
	0.5	31	1.6	11	1.4	6.4	0.89	1.7
	0.6	22	1.7	12	1.5	9	0.91	1.6
$-z$	0.1	516	0.53	10	0.52	3.2	0.53	2.9
	0.2	124	1.1	10	0.92	3.1	0.94	2.8
	0.3	54	1.5	9.1	1.1	4.4	0.93	2.1
	0.4	30	1.5	6.5	0.71	1.4	0.78	1.5
	0.5	19	1.5	5.4	0.5	1.3	0.55	1.3
	0.6	14	1.5	4.9	0.39	1.4	0.9	1.8

components z_i are independent, i.e., if C_x is diagonal, then $z_i \sim \mathcal{N}(0,1)$ and $Q_{\mathrm{x}}^2 = \sum_{i=1}^{n} z_i^2 \sim \chi_n^2$. In the general case of correlated errors, this distribution does not apply. Assuming Gaussian distribution of these correlated errors, the Mahalanobis distance $Q_{\mathrm{M}}^2 = \mathbf{e}^{\top}\mathrm{C}_{\mathrm{x}}^{-1}\mathbf{e}$ does satisfy $Q_{\mathrm{M}}^2 \sim \chi_n^2$. However, after replacing the practically unknown C_{x} by its estimate $\hat{\mathrm{C}}_{\mathrm{x}}$, the corresponding statistic

$$\hat{Q}_{\mathrm{M}}^2 = \mathbf{e}^{\top}\hat{\mathrm{C}}_{\mathrm{x}}^{-1}\mathbf{e} \tag{9}$$

is only approximately χ_n^2-distributed. Nevertheless, for practical purposes, even Q_{x} remains a useful diagnostic indicator: for $n = 4$, a value of Q_{x} close to $\sqrt{n} = 2$ indicates a well-calibrated covariance estimate; values not exceeding $\sqrt{\chi_{4,0.95}^2} \approx 3.08$ are still statistically consistent, while values larger than $\sqrt{\chi_{4,0.99}^2} \approx 3.64$ indicate underestimated uncertainty. The results in Table 1 show the effect of forward-model errors on the reconstruction accuracy and uncertainty calibration, and allow a comparison between importance sampling (IS) and Metropolis–Hastings (MH) sampling for the marginal posterior expectation. The effect of forward-model errors depends strongly on the direction of the displacement.

1. For displacements along the x-axis, the empirical signal-to-noise ratio decreases rapidly as the displacement increases. For small displacements (0.1–0.3 µm), the marginal posterior expectation yields small reconstruction errors, comparable to or slightly smaller than those of the minimum-residual solution. However, for larger displacements ($\geq$ 0.4 µm), the relative error increases abruptly and saturates at approximately 15% for all methods. This value is determined by the parameter ranges and definition of Eq. (6). Such large model error can be explained by symmetry-breaking shift combined with further azimuthal averaging, which greatly reduces the contrast (oscillation amplitude) at intermediate scattering angles. This cannot be, in principle, reproduced by a symmetric model within the specified parameter ranges.
2. For displacements along the z-axis, the reconstruction errors remain significantly smaller (< 2%), which can be explained by preservation of the axial symmetry. In this regime, the marginal posterior expectation continues to yield improved accuracy compared with the minimum-residual solution.
3. A direct comparison between IS and MH shows that both methods produce similar relative errors, indicating that they approximate the same posterior distribution consistently. Some differences are observed for large axial shifts, but they are hard to generalize.

The indicators Q_{x} and $Q_{\langle \mathrm{x}\rangle}$ provide additional information beyond the relative errors by assessing the statistical calibration of the estimated uncertainty. We stress that the latter is even more important for the overall reliability of inverse method in practical applications than the parameter errors themselves.

1. For the minimum-residual solution, the values of Q_{x} are well above the reference thresholds, indicating a systematic underestimation of uncertainty.
2. For the marginal posterior expectation, both IS and MH lead to substantially smaller Q-values, indicating improved calibration. In particular, for $-z$ displacements, the values approach or fall below the 95% threshold, and in some cases are close to the ideal value $\sqrt{n}$.
3. Comparing IS and MH, the MH-based estimates tend to yield smaller Q-values, especially for moderate to large displacements. This suggests that MH sampling provides a more reliable characterization of the posterior spread.
4. A notable exception is large axial shifts, where all methods fail. In this case, large model error leads to large effect of assumed parameter boundaries on the obtained results and confidence intervals [8]. Even elaborate methods that we use do not account for that.

Overall, Table 1 leads to the following conclusions.

1. The reconstruction accuracy is much more sensitive to transverse displacements than to axial ones, probably due to breaking of axial symmetry.
2. The marginal posterior expectation significantly improves uncertainty calibration compared with the minimum-residual solution, especially when MH sampling is used.
3. All methods conceptually fail for large model deviations, when all of the admissible idealized models lead to comparable deviation from the data.

4.2. *Forward model transformation*

In the next test, we consider synthetic data and analyze the accuracy of four forward models defined as follows:

1. Forward Model 1: based on the weighted scattering intensities sampled over angle or scattering cross sections sampled over wavelength,
2. Forward Model 2: based on data reduction using Principal Component Analysis (PCA),
3. Forward Model 3: based on low-pass filtering using the Discrete Cosine Transform (DCT),
4. Forward Model 4: based on a differential representation obtained by removing a smoothing polynomial.

In the simulations, an approximation error $\varepsilon_{\mathrm{PCA}} = 5 \times 10^{-2}$ is used for the PCA model, energy fraction $\rho_{\mathrm{FFT}} = 0.99$ is used for the DCT model, and a quadratic function is employed as a smoothing polynomial (see Appendix B).

An a priori estimate of the data-error variance is defined in terms of the signal-to-noise ratio (SNR) as the ratio of signal power to noise power:

$$\overline{\sigma}^2 = \frac{\|\mathbf{y}_{\mathrm{ref}}\|^2}{m\,\mathrm{SNR}},$$

where SNR = 500, and $\mathbf{y}_{\mathrm{ref}} = \mathbf{F}(\mathbf{x}_{\mathrm{a}})$ is a reference spectrum corresponding to the a priori state $\mathbf{x}_{\mathrm{a}}$. The entries of the a priori vector are chosen as the midpoints of the admissible intervals of variation of the state vector components.

For Problem 1 we consider four layers ($L = 4$) and the following simple bounds on the variables: $d_1 \in [0.6, 0.7]$ μm, $d_2 \in [0.2, 0.5]$ μm, $d_3 \in [0.2, 0.3]$ μm, $d_4 \in [1.0, 3.0]$ μm, $m_1 \in [1.3570, 1.3574]$, $m_2 \in [1.368, 1.427]$, $m_3 \in [1.38, 1.48]$, and $m_4 \in [1.41, 1.58]$. The multilayered sphere is immersed in a buffered saline solution with a refractive index of 1.337. The shape covariance matrix is generated as $\mathscr{C}_\delta = \mathsf{D}\mathsf{P}(\boldsymbol{\phi})\mathsf{D}$, where D is a diagonal matrix with entries

$$[\mathsf{D}]_{ii} = \frac{\sigma_i}{\sqrt{\frac{1}{m}\sum_{k=1}^{m}\sigma_k^2}}, \qquad i = 1, \dots, m, \tag{10}$$

Table 2

Relative state vector errors and uncertainty calibration indicators for Problems 1 and 2 and different forward models using the minimum-norm solution and marginal posterior expectation. The marginal posterior expectation is evaluated using importance sampling (IS) and Metropolis–Hastings (MH) one.

Problem	Forward model	Minimum residual		Marginal posterior expectation			
				IS		MH	
		ε_{x} [%]	Q_{x}	$\varepsilon_{\langle \mathrm{x}\rangle}$ [%]	$Q_{\langle \mathrm{x}\rangle}$	$\varepsilon_{\langle \mathrm{x}\rangle}$ [%]	$Q_{\langle \mathrm{x}\rangle}$
1	1	1.7	1.6	1.1	2.5	0.85	2.6
	2	1.9	1.6	1.1	2.5	1.1	2.7
	3	1.9	6.6	1.1	2.7	0.85	2.1
	4	0.89	1.1	0.34	2.7	0.45	2.6
2	1	4.8	3.8	4.7	3.1	4.5	2.8
	2	4.3	3.6	4.5	2.7	4.1	2.2
	3	2.5	13	2.5	4.6	1.4	1.4
	4	4.9	3.5	4.8	2.8	4.6	2.5

and $\mathsf{P}(\boldsymbol{\phi})$ is a correlation matrix defined by an autocorrelation function ρ depending on parameters $\boldsymbol{\phi}$. Specifically, the variances σ_i^2 are defined as

$$\sigma_i^2 = \mathrm{Var}_1 + \mathrm{Var}_2 y_{\mathrm{ref}i}^2, \qquad i = 1, \dots, m, \tag{11}$$

where

$$\mathrm{Var}_1 = \frac{w_\sigma}{m\,\mathrm{SNR}}\|\mathbf{y}_{\mathrm{ref}}\|^2, \;\; \mathrm{Var}_2 = \frac{1 - w_\sigma}{\mathrm{SNR}}, \tag{12}$$

and $w_\sigma = 0.8$. The correlation matrix is defined on the sample grid $\theta_i = \theta_{\min} + (i-1)\Delta\theta$ as $[\mathsf{P}(\boldsymbol{\phi})]_{ij} = \rho(\theta_{ij}; \boldsymbol{\phi})$, where $\theta_{ij} = |i - j|\Delta\theta$ for $i, j = 1, \dots, m$. The autocorrelation function is the damped cosine

$$\rho(\theta; \Theta, \omega) = \exp\left(-\frac{\theta}{\Theta}\right)\cos(\omega\theta),$$

with $\Theta = 10\Delta\theta$ and $\omega = 2\pi/(30\Delta\theta)$.

For Problem 2 we consider eight layers ($L = 8$) and the bounds $d_\ell \in [0.03, 0.07]$ μm for $\ell = 1, \dots, 8$. The shape covariance matrix is the identity matrix, i.e., $\mathscr{C}_\delta = \mathrm{I}_m$.

In the simulations, we assume a solution pair $(\mathbf{x}^\dagger, \mathbf{y} = \mathbf{F}(\mathbf{x}^\dagger))$, generate a noisy version of the true data vector $\mathbf{y}$ as $\mathbf{y}^\delta = \mathbf{y} + \boldsymbol{\delta}$ with $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \mathsf{C}_\delta = \overline{\sigma}^2\mathscr{C}_\delta)$, and solve the inverse problem for the synthetic measurement vector $\mathbf{y}^\delta$. For Problem 1, we solve a generalized least-squares problem (since the shape covariance matrix is not the identity matrix), whereas for Problem 2, we solve an ordinary least-squares problem. The true state vector is given by $d_1^\dagger = 0.605$ μm, $d_2^\dagger = 0.428$ μm, $d_3^\dagger = 0.243$ μm, $d_4^\dagger = 1.898$ μm, $m_1^\dagger = 1.3572$, $m_2^\dagger = 1.3788$, $m_3^\dagger = 1.3997$, $m_4^\dagger = 1.5157$ for Problem 1, and $d_1^\dagger = 0.049$ μm, $d_2^\dagger = 0.063$ μm, $d_3^\dagger = 0.031$ μm, $d_4^\dagger = 0.052$ μm, $d_5^\dagger = 0.039$ μm, $d_6^\dagger = 0.05$ μm, $d_7^\dagger = 0.059$ μm, $d_8^\dagger = 0.033$ μm for Problem 2.

The true, measured, and retrieved data vectors are shown in Fig. 5 for Problem 1, and in Fig. 6 for Problem 2. For these simulations, the relative errors obtained using the minimum-residual solution and the marginal posterior expectation are reported in Table 2. The marginal posterior expectation is evaluated using IS and MH sampling. For both problems, $n = 8$, so values of Q_{x} close to $\sqrt{n} \approx 2.83$ indicates a well-calibrated covariance estimate, values exceeding $\sqrt{\chi^2_{8,0.99}} \approx 4.48$ suggest underestimated uncertainty, whereas values below $\sqrt{\chi^2_{8,0.05}} \approx 1.56$ indicate overestimated uncertainty. In practice, values significantly below about 2 already point to a tendency toward overestimation. The results can be summarized as follows:

1. Problem 1. Forward Models 1–3 yield comparable accuracy. The marginal posterior expectation reduces the relative error from about 1.7–1.9% to approximately 1.1%, with similar accuracy for IS and MH. The uncertainty indicators also show a clear improvement of posterior-based estimates over the minimum-residual one with respect to the ideal range.

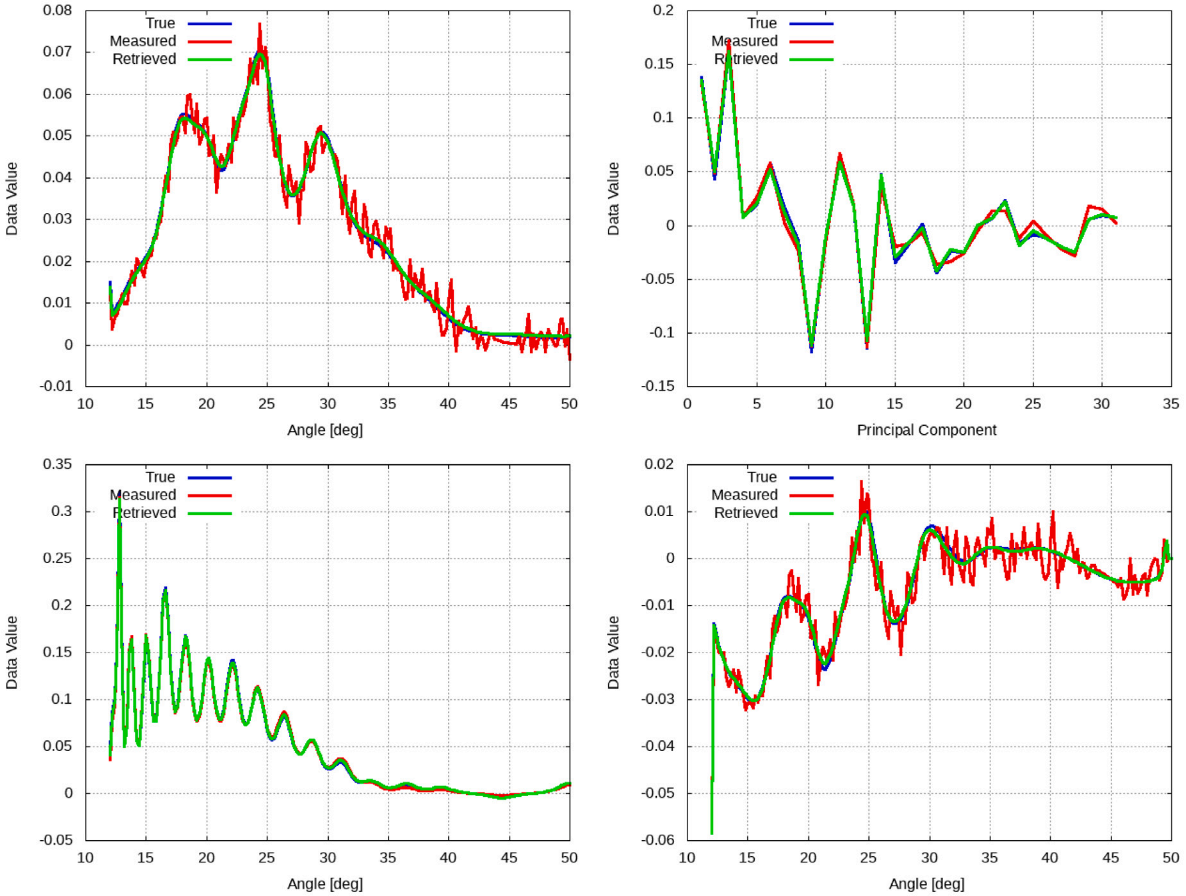

Fig. 5. True, measured, and retrieved data vectors for Problem 1 using Forward Models 1 (upper left), 2 (upper right), 3 (lower left), and 4 (lower right).

2. Problem 2. The reconstruction errors are larger (about 4%–5%). The marginal posterior expectation provides only limited improvement in accuracy (except for MH for forward model 3), indicating that the posterior is strongly concentrated around the minimum. However, the uncertainty calibration is still improved, with noticeable superiority of MH over IS.

4.3. *Statistical analysis of retrieval errors*

In the third test we consider the same scenario as in the second test and assess the accuracy of Forward Model 1 over a set of noisy data vector realizations. In the simulations, we assume a solution pair $(\mathbf{x}^\dagger, \mathbf{y} = \mathbf{F}(\mathbf{x}^\dagger))$, produce $N_\mathrm{y} = 200$ noisy versions $\mathbf{y}_k^\delta$ of $\mathbf{y}$ with $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \mathrm{C}_\delta = \overline{\sigma}^2 \mathscr{C}_\delta)$, and solve the inverse problem for each measurement vector $\mathbf{y}_k^\delta$. For the minimum-residual solution $\mathbf{x}_k^\delta$, we compute the (absolute) relative errors in $\mathbf{x}$, $\varepsilon_{\mathrm{x}k} = \left\|\mathbf{x}_k^\delta - \mathbf{x}^\dagger\right\| / \left\|\mathbf{x}^\dagger\right\|$ for $k = 1, \ldots, N_\mathrm{y}$. The histograms of the relative state vector errors are illustrated in Fig. 7. For Problem 1, the mean relative error is 0.35% with a standard deviation of 0.16%, whereas for Problem 2, the mean relative errors is 5.6% with a standard deviation of 5.1%.

The results highlight a clear difference in the distribution of reconstruction errors between the two problems. For Problem 1, the histogram is strongly concentrated near zero, with small mean error and low variability, indicating that, under the considered noise realizations, the reconstructed solutions remain close to the true state. In contrast, Problem 2 exhibits a much broader error distribution with a pronounced right tail, reflecting a higher variability of the reconstruction errors and the occurrence of larger deviations for some noise realizations. These differences in absolute error levels are primarily influenced by the specific characteristics of the problem, including parameter scaling and sensitivity of the forward model.

4.4. *Uncertainty analysis and calibration of confidence intervals*

In contrast to the previous test, where the data are affected by forward-model errors, we now consider Problem 2, in which the measurements are corrupted by additive Gaussian noise with an identity shape covariance structure. This setting allows us to assess the performance of the uncertainty estimation methods in a regime where the statistical assumptions underlying the inversion are satisfied and the forward model is consistent with the data-generating process.

4.4.1. *Confidence intervals*

In Fig. 8, we illustrate the 99% confidence intervals obtained using the minimum-residual covariance approximation, and the marginal posterior expectation. All methods produce intervals that contain the true solution for most components, indicating a qualitatively consistent description of uncertainty.

4.4.2. *Normalized-error analysis*

To assess the statistical calibration of the uncertainty estimates, we consider the normalized errors z_i defined in Eq. (7), the combined indicator Q_x defined in Eq. (8), and the Mahalanobis-distance indicator $\widehat{Q}_\mathrm{M}$ defined by Eq. (9). As before, values of Q near to $\sqrt{n} \approx 2.83$ indicate good calibration of the covariance estimate, with [1.48,4.19] being a 95% equal-tailed confidence interval. The computed values of Q_x are 2.4 for the linearization method, 1.9 for the marginal posterior expectation with IS, and 1.8 for that with MH sampling. Computation of Mahalanobis-distance indicator $\widehat{Q}_\mathrm{M}$, is sensitive to potential ill-conditioning of $\widehat{\mathrm{C}}_\mathrm{x}$, since its inverse can then strongly amplify errors along (noisy) directions associated with small eigenvalues. To mitigate this effect, we compute $\widehat{Q}_\mathrm{M}$ both with and without the contribution of

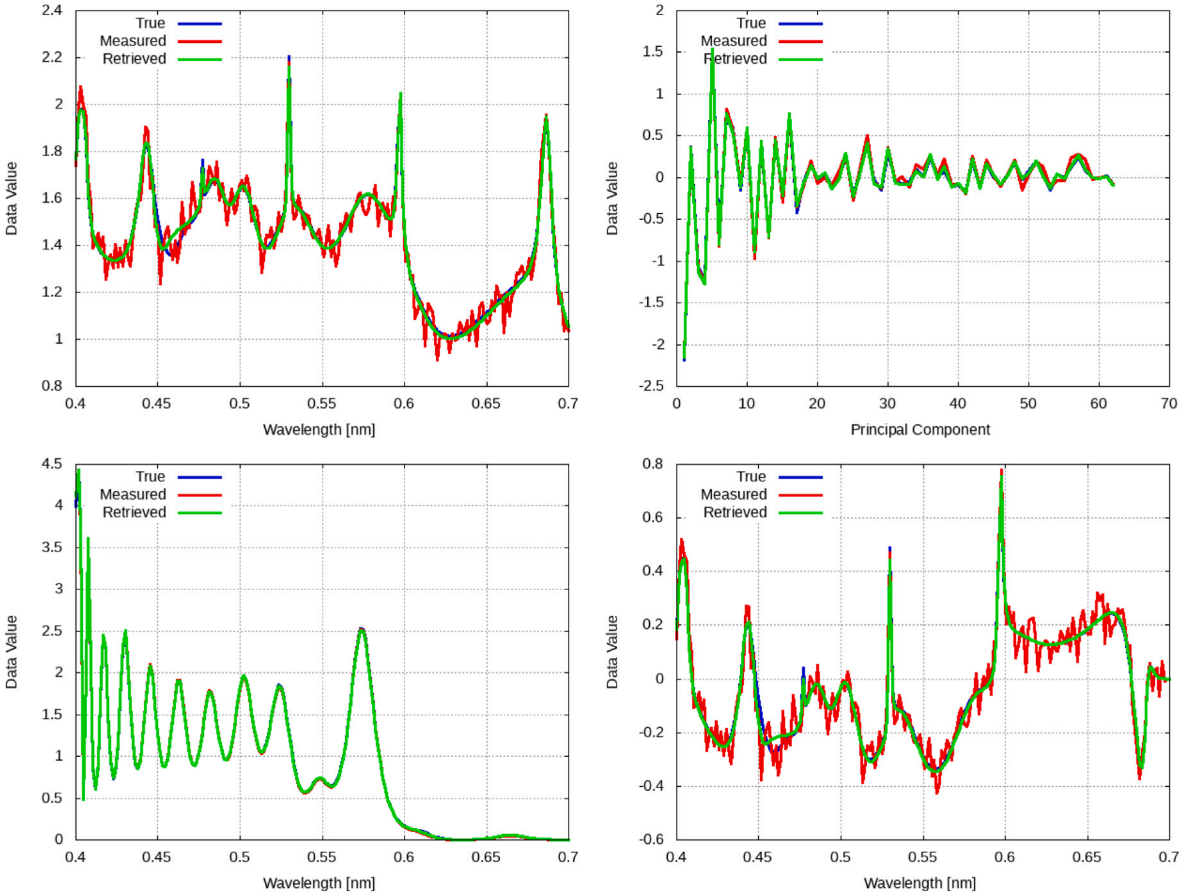

Fig. 6. True, measured, and retrieved data vectors for Problem 2 using Forward Models 1 (upper left), 2 (upper right), 3 (lower left), and 4 (lower right).

the smallest eigenvalue (of order $10^{-10} - 10^{-11}$). The resulting values are 12 and 3.1 for the linearization method, 4.6 and 2.1 for the marginal posterior expectation with IS, and 2.0 and 1.9 for that with MH sampling. Fig. 9 shows the absolute normalized errors $|z_i|$ together with the reference levels 1 and 2, corresponding approximately to the two-sided 68% and 95% quantiles of the standard normal distribution, respectively. The results can be summarized as follows:

1. The obtained values of Q_{x} are all below the reference value $\sqrt{n} \approx 2.83$, but within the 95% confidence interval. Specifically, the posterior-based methods exhibit slight overestimation of uncertainties.
2. Mahalanobis-distance indicators better agree with the expected χ^2 behavior, but only if negligibly small eigenvalues of the covariance matrix are discarded during the inversion.
3. The normalized errors $|z_i|$ confirm these findings: most values lie below the threshold 2, indicating that all methods provide a qualitatively consistent uncertainty description, while the systematically smaller values for IS and MH reflect wider confidence intervals and thus a tendency to overestimate uncertainty.

5. Conclusions

We developed an optimization tool to recover the layer thicknesses and refractive indices of a multilayered spherical particle from either (i) angular scattering patterns or (ii) spectral scattering/extinction data. The scattering by a multilayered spherical particle is computed using the T-matrix formalism. Three alternative recurrence schemes are implemented for the computation of the T-matrix elements, and analytical derivatives with respect to the layer radii and refractive indices are provided. The resulting inverse problem is solved using an optimization tool that is publicly available.

The proposed approach integrates accurate forward simulations with a robust global optimization strategy. The admissible parameter space is controlled through bound constraints, and optional data transformations are employed to improve stability and reduce dimensionality. The MultiStart framework enables the identification of multiple candidate solutions and their subsequent refinement. The final estimate and associated uncertainties are obtained from the minimum-residual solution and corresponding covariance-matrix approximation or through the marginal posterior expectation.

In the numerical analysis, several tests were performed:

1. Data affected by forward model errors. In the presence of forward-model errors, the reconstruction accuracy was found to depend strongly on the structure of the model discrepancy. Transverse perturbations lead to a rapid degradation of the solution due to structural inconsistency of the model, whereas axial perturbations preserve the effective symmetry and allow accurate parameter recovery. In this regime, the residuals are correlated and non-Gaussian, and the minimum-residual covariance approximation systematically underestimates the uncertainty. By contrast, the marginal posterior expectation, which accounts for the effective number of independent observations, provides improved robustness and more reliable uncertainty estimates.
2. Forward model transformation. The analysis of forward model transformations showed that PCA- and DCT-based reduced models provide efficient data compression while maintaining reconstruction accuracy comparable to the full forward model.
3. Statistical analysis of retrieval errors. The statistical analysis over multiple noisy realizations highlighted a clear difference

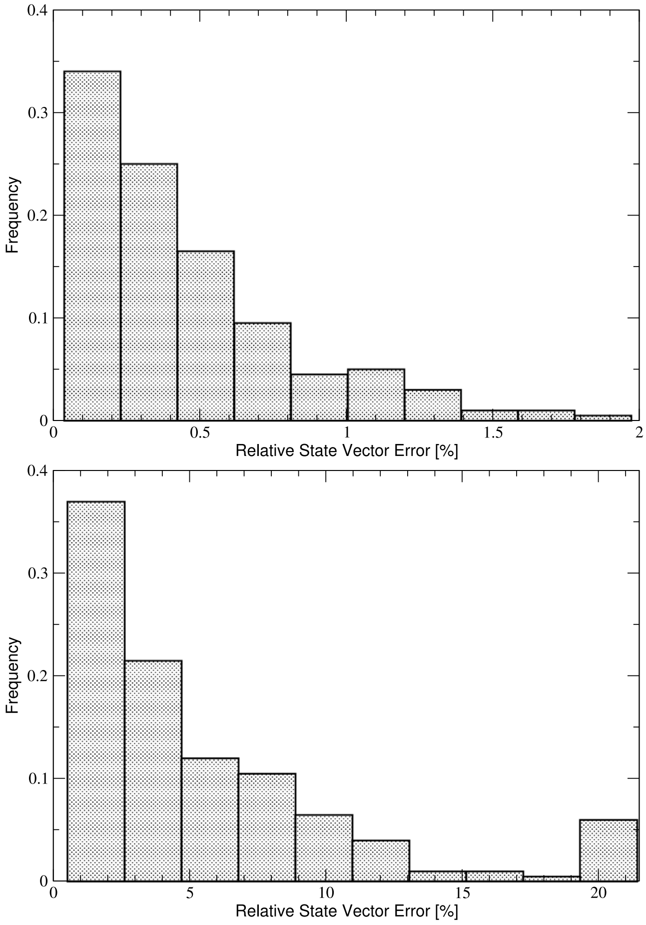

Fig. 7. Histograms of the relative state vector errors (based on minimum-residual solution) obtained from 200 noisy data vector realizations for Problem 1 (upper) and Problem 2 (lower).

between the two problem settings. Problem 1 exhibits stable and accurate reconstructions with small variability, whereas Problem 2 is more sensitive to noise, resulting in larger errors and increased variability.

4. Uncertainty analysis. The uncertainty analysis demonstrated that all methods provide qualitatively consistent confidence intervals containing the true solution. However, the calibration of these intervals depends on the method and the error structure. For Problem 2, where the statistical assumptions are satisfied, the minimum-residual covariance approximation yields well-calibrated uncertainty estimates, while the marginal posterior expectation tends to slightly overestimate it, as indicated by values of Q_{x} below the reference level $\sqrt{n}$, but still within 95% confidence interval. The overestimation is reduced when using the Mahalanobis-distance indicators, but only if negligibly small eigenvalues of the covariance matrix are discarded during the inversion.

Overall, the results indicate that the choice of uncertainty quantification method should be guided by the underlying error structure. Posterior averaging is advantageous in the presence of model discrepancies and correlated residuals, whereas the linearization-based covariance approximation remains reliable when the forward model is consistent with the data and the noise is Gaussian. The commonly used χ-type indicators provide useful qualitative diagnostics, but their quantitative interpretation must be treated with care, as they are sensitive to covariance ill-conditioning and may vary significantly depending on the noise realization when only a single measurement is available.

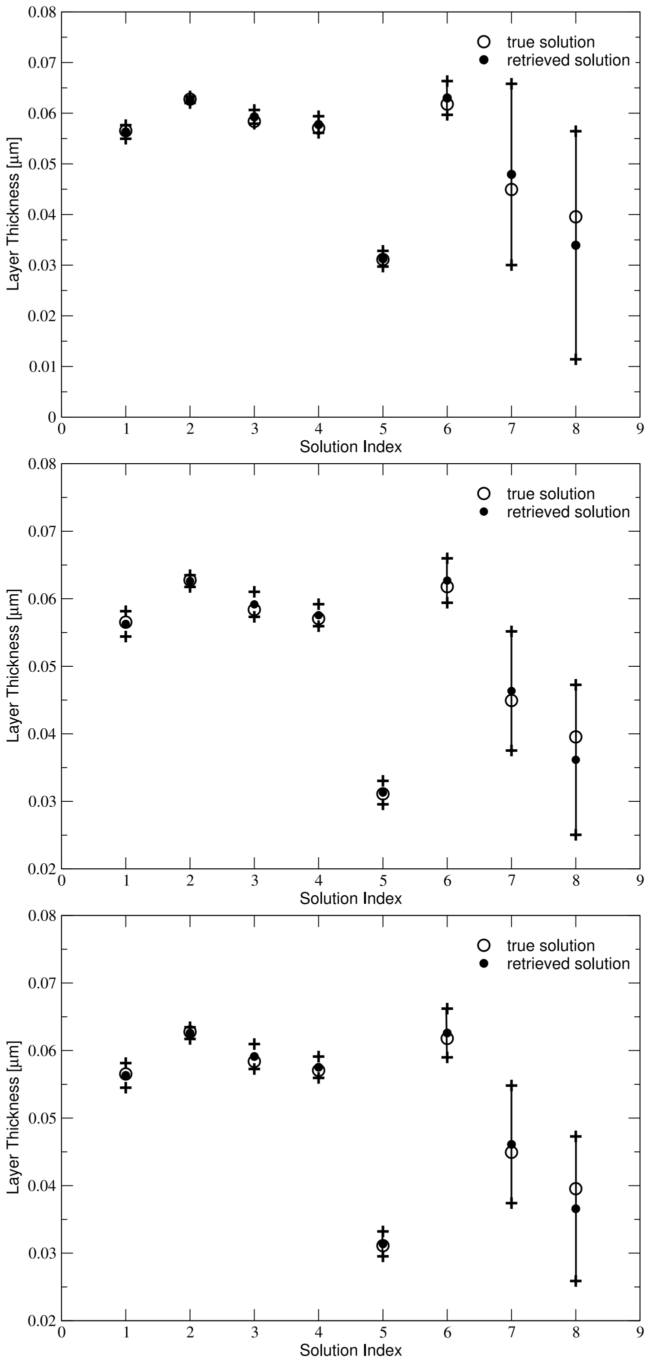

Fig. 8. 99% confidence intervals for the solutions of Problem marginal posterior expectation with either importance sampling (middle) or Metropolis–Hastings one (lower).

The developed optimization tool can be applied to a wide range of inverse problems of various dimensionality, given a routine that evaluates the forward model and its Jacobian at a given state vector. In future, we plan to integrate modern first-order adaptive gradients (e.g., Adam-family), momentum/accelerated schemes (Nesterov), proximal methods for sparse/structured priors (FISTA–Fast Iterative Shrinkage-Thresholding Algorithm), and scaled conjugate-gradient algorithms.

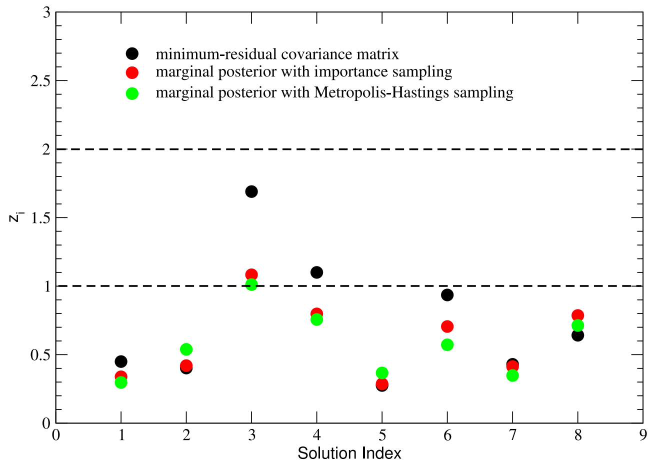

Fig. 9. Absolute normalized errors $|z_i|$ for the three methods. Horizontal dashed lines indicate the levels 1 and 2, corresponding to standard normal thresholds.

CRediT authorship contribution statement

Alexandru Doicu: Conceptualization, Methodology, Theoretical analysis, Software, Writing – review & editing, Writing – original draft. **Dmitry S. Efremenko:** Software, Numerical analysis, Visualization, Numerical analysis, Writing – review & editing, Writing – original draft. **Maxim A. Yurkin:** Methodology, Formal analysis, Validation, Interpretation of results, Writing – review & editing, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

MYu acknowledges support by the Normandy Region (project RADDAERO). DDA simulations were performed using computing resources of CRIANN (Normandy, France).

Appendix A. Scattering by a multilayered spherical particle

In this appendix, we first outline general theoretical considerations regarding the computation of the scattered intensity and the associated cross sections for an arbitrary particle, and then specialize the discussion to the case of a multilayered spherical particle. For this class of particles, we provide a set of recurrence relations for computing the diagonal elements of the transition matrix and, finally, present an algorithm based on the transfer-matrix method for the numerical evaluation of the scattering and absorption cross sections.

General considerations

Consider a particle embedded in a homogeneous medium of wavenumber k. Outside the smallest sphere that circumscribes the particle, the scattered field can be expanded in terms of radiating spherical vector wave functions:

$$\mathbf{E}_{\mathrm{s}}(\mathbf{r})=\sum_{n=1}^{\infty}\sum_{m=-n}^{n}[f_{mn}\mathbf{M}^{3}_{mn}(k\mathbf{r})+g_{mn}\mathbf{N}^{3}_{mn}(k\mathbf{r})], \tag{A.1}$$

whereas the incident field, as an entire solution to the Maxwell equations, admits an expansion in regular spherical vector wave functions:

$$\mathbf{E}_{0}(\mathbf{r})=\sum_{n=1}^{\infty}\sum_{m=-n}^{n}[a_{mn}\mathbf{M}^{1}_{mn}(k\mathbf{r})+b_{mn}\mathbf{N}^{1}_{mn}(k\mathbf{r})]. \tag{A.2}$$

The expansion coefficients of the scattered field f_{mn} and g_{mn} are linearly related to those of the incident field a_{mn} and b_{mn} through the transition matrix T, i.e.,

$$\begin{bmatrix} f_{mn} \\ g_{mn} \end{bmatrix}=\mathsf{T}\begin{bmatrix} a_{mn} \\ b_{mn} \end{bmatrix}. \tag{A.3}$$

In the far-field region, and assuming a time dependence of the form $\exp(-\mathrm{j}\omega t)$, where ω is the angular frequency and t is time, the scattered field takes the asymptotic form

$$\mathbf{E}_{\mathrm{s}}(\mathbf{r})=\frac{\mathrm{e}^{\mathrm{j}kr}}{r}\left[\mathbf{E}_{\mathrm{s}\infty}(\hat{\mathbf{r}})+O\left(\frac{1}{r}\right)\right]\quad r\to\infty, \tag{A.4}$$

where $\hat{\mathbf{r}}=\mathbf{r}/r$ is the radial direction. The leading term $\mathbf{E}_{\mathrm{s}\infty}$, known as the electric far-field pattern, satisfies the transversality condition $\hat{\mathbf{r}}\cdot\mathbf{E}_{\mathrm{s}\infty}=0$. Its expansion in terms of vector spherical harmonics reads

$$\mathbf{E}_{\mathrm{s}\infty}(\hat{\mathbf{r}})=\frac{1}{k}\sum_{n=1}^{\infty}\sum_{m=-n}^{n}(-\mathrm{j})^{n+1}[f_{mn}\mathbf{m}_{mn}(\hat{\mathbf{r}})+\mathrm{j}g_{mn}\mathbf{n}_{mn}(\hat{\mathbf{r}})]. \tag{A.5}$$

The vector spherical harmonics are defined by

$$\mathbf{m}_{mn}(\theta,\varphi)=\frac{1}{\sqrt{2n(n+1)}}[\mathrm{j}m\pi_n^{|m|}(\theta)\hat{\boldsymbol{\theta}}-\tau_n^{|m|}(\theta)\hat{\boldsymbol{\varphi}}]\mathrm{e}^{\mathrm{j}m\varphi}, \tag{A.6}$$

$$\mathbf{n}_{mn}(\theta,\varphi)=\frac{1}{\sqrt{2n(n+1)}}[\tau_n^{|m|}(\theta)\hat{\boldsymbol{\theta}}+\mathrm{j}m\pi_n^{|m|}\hat{\boldsymbol{\varphi}}]\mathrm{e}^{\mathrm{j}m\varphi}, \tag{A.7}$$

where $\hat{\mathbf{r}}=(\theta,\varphi)$ denotes the angular coordinates and $(\hat{\mathbf{r}},\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\varphi}})$ are the spherical unit vectors associated with $\mathbf{r}$. The angular functions π_n^m and τ_n^m are related to the associated Legendre functions P_n^m by

$$\pi_n^m(\theta)=\frac{P_n^m(\cos\theta)}{\sin\theta},\quad \tau_n^m(\theta)=\frac{\mathrm{d}}{\mathrm{d}\theta}P_n^m(\cos\theta). \tag{A.8}$$

For a plane-wave incidence, the amplitude matrix S describes how the transverse components of the incident electric field are transformed into the transverse components of the scattered field in the far-field region. To make this relationship explicit, we first express the scattered electric field in the far-field as

$$\mathbf{E}_{\mathrm{s}\infty}(\hat{\mathbf{r}})=E_{\mathrm{s}\infty,\theta}(\hat{\mathbf{r}})\,\hat{\boldsymbol{\theta}}+E_{\mathrm{s}\infty,\varphi}(\hat{\mathbf{r}})\,\hat{\boldsymbol{\varphi}}. \tag{A.9}$$

Similarly, we represent the incident field as a plane electromagnetic wave propagating along the direction $\hat{\mathbf{s}}$:

$$\mathbf{E}_{0}(\mathbf{r})=(E_{0,\beta}\,\hat{\boldsymbol{\beta}}+E_{0,\alpha}\,\hat{\boldsymbol{\alpha}})\mathrm{e}^{\mathrm{j}k\hat{\mathbf{s}}\cdot\mathbf{r}}, \tag{A.10}$$

where $(\hat{\mathbf{s}},\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\alpha}})$ form a right-handed orthonormal triad associated with the incident direction. The transverse components are parameterized by the polarization angle α_{pol}: $E_{0,\beta}=\cos\alpha_{\mathrm{pol}}$ and $E_{0,\alpha}=\sin\alpha_{\mathrm{pol}}$, where we have assumed a unit incident amplitude ($E_0=1$). The relation between the incident and scattered transverse components is then

$$\begin{bmatrix} E_{\mathrm{s}\infty,\theta}(\hat{\mathbf{r}}) \\ E_{\mathrm{s}\infty,\varphi}(\hat{\mathbf{r}}) \end{bmatrix}=\mathsf{S}(\hat{\mathbf{r}},\hat{\mathbf{s}})\begin{bmatrix} E_{0,\beta} \\ E_{0,\alpha} \end{bmatrix},\quad \mathsf{S}=\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}. \tag{A.11}$$

For the special case of normal incidence ($\hat{\mathbf{s}}=\hat{\mathbf{z}}$), the columns of S are obtained by exciting the system with two orthogonal polarizations:

- For $\alpha_{\mathrm{pol}}=0$ (incident field polarized along $\hat{\boldsymbol{\beta}}=\hat{\mathbf{x}}$):

$$S_{11}(\theta,\varphi)=E_{\mathrm{s}\infty,\theta}(\theta,\varphi),\quad S_{21}(\theta,\varphi)=E_{\mathrm{s}\infty,\varphi}(\theta,\varphi). \tag{A.12}$$

- For $\alpha_{\mathrm{pol}}=\pi/2$ (incident field polarized along $\hat{\boldsymbol{\alpha}}=\hat{\mathbf{y}}$):

$$S_{12}(\theta,\varphi)=E_{\mathrm{s}\infty,\theta}(\theta,\varphi),\quad S_{22}(\theta,\varphi)=E_{\mathrm{s}\infty,\varphi}(\theta,\varphi). \tag{A.13}$$

The measurable quantities that characterize scattering are the angular distribution of the scattered intensity and the wavelength-dependent cross sections:

1. Scattered intensity: For a macroscopically isotropic and mirror-symmetric scattering medium, the scattered intensity — corresponding to the first component of the Stokes vector — is given by

$$\begin{aligned} I(\theta) = \frac{1}{2}(&\left|S_{11}(\theta,\varphi=0)\right|^2 + \left|S_{12}(\theta,\varphi=0)\right|^2 \\ &+ \left|S_{21}(\theta,\varphi=0)\right|^2 + \left|S_{22}(\theta,\varphi=0)\right|^2). \end{aligned} \tag{A.14}$$

2. Extinction and scattering cross sections: These are computed from the expansion coefficients f_{mn}, g_{mn}, a_{mn}, and b_{mn} as

$$C_{\text{ext}} = -\frac{\pi}{k^2}\sum_{n=1}^{\infty}\sum_{m=-n}^{n} \operatorname{Re}\{f_{mn}a_{mn}^* + g_{mn}b_{mn}^*\}. \tag{A.15}$$

and

$$C_{\text{scat}} = \frac{\pi}{k^2}\sum_{n=1}^{\infty}\sum_{m=-n}^{n} (\left|f_{mn}\right|^2 + \left|g_{mn}\right|^2), \tag{A.16}$$

In the case of a multilayered sphere, several simplifications occur:

1. With the incident plane wave propagating along the z-axis, only the azimuthal modes $m = \pm 1$ contribute. Consequently, the infinite sums $\sum_{n=1}^{\infty}\sum_{m=-n}^{n}$ are truncated as $\sum_{n=1}^{N}\sum_{m=\pm 1}$, where N is the maximum expansion order.
2. The expansion coefficients take simple forms:

$$a_{1n} = -a_{-1n} = \mathrm{j}^{n-1}\sqrt{2n+1}, \tag{A.17}$$

$$b_{1n} = b_{-1n} = \mathrm{j}^{n-1}\sqrt{2n+1}, \tag{A.18}$$

for a plane wave polarized along the x-axis, and

$$a_{1n} = a_{-1n} = \mathrm{j}^{n-2}\sqrt{2n+1}, \tag{A.19}$$

$$b_{1n} = -b_{-1n} = \mathrm{j}^{n-2}\sqrt{2n+1}. \tag{A.20}$$

for a plane wave polarized along the y-axis

3. The transition matrix becomes diagonal

$$T^{11}_{mn,m_1n_1} = T^1_n\delta_{mm_1}\delta_{nn_1},\ \ T^{22}_{mn,m_1n_1} = T^2_n\delta_{mm_1}\delta_{nn_1}, \tag{A.21}$$

so that the scattered-field coefficients are simply

$$f_{\pm1,n} = T^1_n a_{\pm1,n},\ \ g_{\pm1,n} = T^1_n b_{\pm1,n}. \tag{A.22}$$

Recurrence relations for T-matrix calculations

Consider a multilayered sphere with L concentric layers of radii $r_1 > r_2 > \cdots > r_L$ (outside to inside) and relative refractive indices $m_1, m_2, \dots, m_L$ with respect to the ambient medium. Let k be the ambient-medium wavenumber and define

$$x_\ell = kr_\ell,\ \ \overline{x}_\ell = \begin{cases} x_1, & \ell = 1, \\ m_{\ell-1}x_\ell, & \ell \geq 2, \end{cases}\ \ \overline{m}_\ell = \begin{cases} m_1, & \ell = 1, \\ \dfrac{m_\ell}{m_{\ell-1}}, & \ell \geq 2, \end{cases} \tag{A.23}$$

and

$$\overline{z}_\ell = \overline{m}_\ell\,\overline{x}_\ell = m_\ell x_\ell. \tag{A.24}$$

Denote the Riccati–Bessel and Riccati–Hankel functions by $\psi_n(z) = zj_n(z)$ and $\xi_n(z) = zh_n^{(1)}(z)$, where $j_n(z)$ and $h_n^{(1)}(z)$ are the spherical Bessel function and the spherical Hankel function of the first kind, respectively. Their logarithmic derivatives are defined by

$$A_n(z) = \frac{\psi_n'(z)}{\psi_n(z)} \text{ and } B_n(z) = \frac{\xi_n'(z)}{\xi_n(z)}.$$

Following Doicu et al. [38], the recurrence for the diagonal elements of the T-matrix is initialized at the core boundary $\ell = L$ as

$$T^{(1)}_{n,L} = -\frac{\left[\overline{m}_L A_n(\overline{z}_L) + \dfrac{n}{\overline{x}_L}\right]\psi_n(\overline{x}_L) - \psi_{n-1}(\overline{x}_L)}{\left[\overline{m}_L A_n(\overline{z}_L) + \dfrac{n}{\overline{x}_L}\right]\xi_n(\overline{x}_L) - \xi_{n-1}(\overline{x}_L)}, \tag{A.25}$$

$$T^{(2)}_{n,L} = -\frac{\left[\dfrac{A_n(\overline{z}_L)}{\overline{m}_L} + \dfrac{n}{\overline{x}_L}\right]\psi_n(\overline{x}_L) - \psi_{n-1}(\overline{x}_L)}{\left[\dfrac{A_n(\overline{z}_L)}{\overline{m}_L} + \dfrac{n}{\overline{x}_L}\right]\xi_n(\overline{x}_L) - \xi_{n-1}(\overline{x}_L)}. \tag{A.26}$$

For the remaining layers, $\ell = L-1, L-2, \dots, 1$, the downward recurrence reads

$$T^{(1)}_{n,\ell} = -\frac{\left[r^{(1)}_{n,\ell}\overline{m}_\ell A_n(\overline{z}_\ell) + p^{(1)}_{n,\ell}\dfrac{n}{\overline{x}_\ell}\right]\psi_n(\overline{x}_\ell) - p^{(1)}_{n,\ell}\psi_{n-1}(\overline{x}_\ell)}{\left[r^{(1)}_{n,\ell}\overline{m}_\ell A_n(\overline{z}_\ell) + p^{(1)}_{n,\ell}\dfrac{n}{\overline{x}_\ell}\right]\xi_n(\overline{x}_\ell) - p^{(1)}_{n,\ell}\xi_{n-1}(\overline{x}_\ell)}, \tag{A.27}$$

$$T^{(2)}_{n,\ell} = -\frac{\left[r^{(2)}_{n,\ell}\dfrac{A_n(\overline{z}_\ell)}{\overline{m}_\ell} + p^{(2)}_{n,\ell}\dfrac{n}{\overline{x}_\ell}\right]\psi_n(\overline{x}_\ell) - p^{(2)}_{n,\ell}\psi_{n-1}(\overline{x}_\ell)}{\left[r^{(2)}_{n,\ell}\dfrac{A_n(\overline{z}_\ell)}{\overline{m}_\ell} + p^{(2)}_{n,\ell}\dfrac{n}{\overline{x}_\ell}\right]\xi_n(\overline{x}_\ell) - p^{(2)}_{n,\ell}\xi_{n-1}(\overline{x}_\ell)}, \tag{A.28}$$

where the layer-coupling coefficients are

$$p^{(1,2)}_{n,l} = 1 + T^{(1,2)}_{n,\ell+1}C_n(\overline{z}_\ell), \tag{A.29}$$

$$r^{(1,2)}_{n,\ell} = 1 + T^{(1,2)}_{n,\ell+1}D_n(\overline{z}_\ell). \tag{A.30}$$

with

$$C_n(z) = \frac{\xi_n(z)}{\psi_n(z)} \text{ and } D_n(z) = \frac{\xi_n'(z)}{\psi_n'(z)} = \frac{B_n(z)}{A_n(z)}C_n(z). \tag{A.31}$$

By construction, $p^{(1,2)}_{n,\ell}$ and $r^{(1,2)}_{n,\ell}$ depend on the current-layer parameters $\overline{\zeta}_\ell = (\overline{m}_\ell, \overline{x}_\ell)$ through $C_n(\overline{z}_\ell)$ and $D_n(\overline{z}_\ell)$, and on all inner layers $\overline{\zeta}_L, \dots, \overline{\zeta}_{\ell+1}$ through $T^{(1,2)}_{n,\ell+1}$. Each $T^{(1,2)}_{n,\ell}$ therefore depends on the parameters of layers $L, L-1, \dots, \ell$, which we write succinctly as

$$T^{(1,2)}_{n,\ell} = T^{(1,2)}_{n,\ell}\left(\overline{\zeta}_L, \overline{\zeta}_{L-1}, \dots, \overline{\zeta}_\ell\right).$$

The total diagonal T-matrix elements are

$$T^{(1,2)}_n = T^{(1,2)}_{n,1}. \tag{A.32}$$

An equivalent recurrence follows from the identities

$$\psi_n'(z) = \psi_{n-1}(z) - \frac{n}{z}\psi_n(z), \tag{A.33}$$

$$\xi_n'(z) = \xi_{n-1}(z) - \frac{n}{z}\xi_n(z), \tag{A.34}$$

which enter D_n. Introducing

$$R_n(z) = \frac{1}{C_n(z)} = \frac{\psi_n(z)}{\xi_n(z)}, \tag{A.35}$$

we obtain

$$T^{(1)}_{n,\ell} = -R_n(\overline{x}_\ell)\frac{\overline{m}_\ell r^{(1)}_{n,\ell} - p^{(1)}_{n,\ell}A_n(\overline{x}_\ell)}{\overline{m}_\ell r^{(1)}_{n,\ell} - p^{(1)}_{n,\ell}B_n(\overline{x}_\ell)}, \tag{A.36}$$

$$T^{(2)}_{n,\ell} = -R_n(\overline{x}_\ell)\frac{r^{(2)}_{n,\ell} - \overline{m}_\ell p^{(1)}_{n,\ell}A_n(\overline{x}_\ell)}{r^{(2)}_{n,\ell} - \overline{m}_\ell p^{(1)}_{n,\ell}B_n(\overline{x}_\ell)}, \tag{A.37}$$

where now

$$p^{(1,2)}_{n,l} = R_n(\overline{z}_\ell) + T^{(1,2)}_{n,\ell+1}, \tag{A.38}$$

$$r^{(1,2)}_{n,\ell} = R_n(\overline{z}_\ell)A_n(\overline{z}_\ell) + T^{(1)}_{n,\ell+1}B_n(\overline{z}_\ell). \tag{A.39}$$

Defining the ratio

$$\mathcal{T}^{(1,2)}_{n,l} = \frac{r^{(1,2)}_{n,\ell}}{p^{(1,2)}_{n,\ell}} = \frac{R_n(\overline{z}_\ell)A_n(\overline{z}_\ell) + B_n(\overline{z}_\ell)T^{(1,2)}_{n,\ell+1}}{R_n(\overline{z}_\ell) + T^{(1,2)}_{n,\ell+1}}, \tag{A.40}$$

the previous expressions compactify to

$$T^{(1)}_{n,\ell} = -R_n(\overline{x}_\ell)\frac{\overline{m}_\ell \mathcal{T}^{(1)}_{n,l} - A_n(\overline{x}_\ell)}{\overline{m}_\ell \mathcal{T}^{(1)}_{n,l} - B_n(\overline{x}_\ell)}, \tag{A.41}$$

$$T^{(2)}_{n,\ell} = -R_n(\overline{x}_\ell)\frac{\mathcal{T}^{(2)}_{n,l} - \overline{m}_\ell A_n(\overline{x}_\ell)}{\mathcal{T}^{(2)}_{n,l} - \overline{m}_\ell B_n(\overline{x}_\ell)}. \tag{A.42}$$

The spherical Bessel, spherical Neumann, and Riccati–Hankel functions are obtained as follows.

1. The spherical Bessel functions $j_n(z)$ are computed using the backward Miller recursion, which proceeds through the following steps:
   (a) Compute
   $$j_0 = \frac{\sin z}{z}, \quad j_1 = \frac{\sin z}{z^2} - \frac{\cos z}{z}.$$
   (b) Choose $M \geq N + 5$ and initialize small nonzero seeds: $\tilde{j}_{M+1} = 0$ and $\tilde{j}_M = 10^{-100}$.
   (c) Sweep backwards using
   $$\tilde{j}_{n-1}(z) = \frac{2n+1}{z}\tilde{j}_n(z) - \tilde{j}_{n+1}(z), \quad n = M, M-1, \ldots, 1.$$
   (d) Normalize the sequence using a known reference:
   $$c = \begin{cases} j_0(z)/\tilde{j}_0(z), & \text{if } |\tilde{j}_0| > |\tilde{j}_1|, \\ j_1(z)/\tilde{j}_1(z), & \text{otherwise.} \end{cases}$$
   (e) Scale all terms: $j_n(z) = c\,\tilde{j}_n(z)$, $n = 0, 1, \ldots, N$.
2. The spherical Neumann functions $y_n(z)$ are computed using stable Wronskian-based recurrences, rather than direct forward recursion, as follows:
   (a) Compute
   $$y_0(z) = -\frac{\cos z}{z}, \quad y_1(z) = -\frac{\cos z}{z^2} - \frac{\sin z}{z}.$$
   (b) For $k \geq 2$, compute
   $$y_k(z) = \begin{cases} \dfrac{j_k(z)\, y_{k-1}(z) - 1/z^2}{j_{k-1}(z)}, & \text{if } |j_{k-1}| \geq |j_{k-2}|, \\[2ex] \dfrac{j_k(z)\, y_{k-2}(z) - (2k-1)/z^3}{j_{k-2}(z)}, & \text{otherwise.} \end{cases}$$
3. The Riccati–Hankel functions $\xi_n(z) = z h^{(1)}_n(z) = z[j_n(z) + \mathrm{j} y_n(z)]$ are computed either by an upward recurrence for ξ_n or by reconstruction from j_n and y_n, as follows:
   (a) Compute
   $$\xi_0(z) = -\mathrm{j}\mathrm{e}^{\mathrm{j}z}, \quad \xi_1(z) = -\mathrm{e}^{\mathrm{j}z}\left(1 + \frac{\mathrm{j}}{z}\right), \quad \psi_0 = \sin z$$
   and
   $$\rho = \frac{|\xi_0(z)|}{|\psi_0|}.$$
   (b) If $\rho < 10^{-6}$, use the three-term recurrence relation
   $$\xi_n(z) = \frac{2n-1}{z}\xi_{n-1}(z) - \xi_{n-2}(z), \quad n = 2, 3, \ldots, N,$$
   with the seeds ξ_0 and ξ_1 from Step 1; otherwise, form the Riccati–Hankel functions as
   $$\xi_n(z) = z[j_n(z) + \mathrm{j} y_n(z)], \quad n = 0, 1, \ldots, N.$$

The coefficients A_n and B_n are computed by continued fractions.

1. For A_n, use the downward recurrence
   $$A_n(z) = \frac{n+1}{z} - \cfrac{1}{A_{n+1}(z) + \cfrac{n+1}{z}}, \quad n = N_{\max} - 1, N_{\max} - 2, \ldots, 1, \tag{A.43}$$
   with $A_{N_{\max}}(z) = 0$ and $N_{\max}$ chosen sufficiently large.
2. For B_n, introduce an auxiliary sequence $\{\rho_n\}$ via the upward recurrence
   $$\rho_{-1} = \mathrm{j}, \quad \rho_n(z) = \cfrac{1}{\cfrac{2n+1}{z} - \rho_{n-1}(z)}, \quad n \geq 0, \tag{A.44}$$
   and then compute
   $$B_n(z) = -\frac{n}{z} + \rho_{n-1}(z), \quad n = 1, 2, \ldots, N. \tag{A.45}$$
   Another algorithm for computing B_n, due to Mackowski et al. [28], performs a simultaneous update of the product $\psi_n\xi_n$ and the logarithmic derivative B_n:
   $$\psi_n(z)\xi_n(z) = \psi_{n-1}(z)\xi_{n-1}(z)\left[\frac{n}{z} - A_{n-1}(z)\right]\left[\frac{n}{z} - B_{n-1}(z)\right],$$
   $$B_n(z) = A_n(z) + \frac{\mathrm{j}}{\psi_n(z)\xi_n(z)}, \quad n \geq 1,$$
   starting from
   $$B_0(z) = \mathrm{j}, \quad \psi_0(z)\xi_0(z) = \tfrac{1}{2}\left(1 - \mathrm{e}^{2\mathrm{j}z}\right)$$
   It is worth noting that, based on the above algorithms for computing the logarithmic derivatives A_n and B_n, Mackowski et al. [28] proposed the following forward scheme for the Riccati functions:
   $$\psi_n(z) = \psi_{n-1}(z)\left[\frac{n}{z} - A_{n-1}(z)\right],$$
   $$\xi_n(z) = \xi_{n-1}(z)\left[\frac{n}{z} - B_{n-1}(z)\right],$$
   valid for $n \geq 1$.
3. The coefficient R_n can be computed either directly as the ratio $\psi_n(z)/\xi_n(z)$ or by means of the upward recurrence:
   $$R_n(z) = R_{n-1}(z)\frac{B_n(z) + \frac{n}{z}}{A_n(z) + \frac{n}{z}}, \quad n \geq 1,$$
   with the starting value
   $$R_0(z) = \frac{\psi_0(z)}{\xi_0(z)} = \tfrac{1}{2}\left(1 - \mathrm{e}^{-2\mathrm{j}z}\right).$$
   Equivalently, taking logarithms yields the update formula
   $$\ln R_n(z) = \ln R_0(z) + \sum_{k=1}^{n} \ln\left[1 + \frac{B_k(z) - A_k(z)}{A_k(z) + \frac{k}{z}}\right].$$
4. Similarly, the coefficient C_n can be computed as the ratio $\xi_n(z)/\psi_n(z)$ or via the upward recurrence:
   $$C_n(z) = C_{n-1}(z)\frac{A_n(z) + \frac{n}{z}}{B_n(z) + \frac{n}{z}}, \quad n \geq 1,$$
   with the starting value
   $$C_0(z) = \frac{\xi_0(z)}{\psi_0(z)} = \frac{2}{1 - \mathrm{e}^{-2\mathrm{j}z}}.$$

The derivatives $\partial T^{(1,2)}_n/\partial\overline{m}_\ell$ and $\partial T^{(1,2)}_n/\partial\overline{x}_\ell$ are computed by a set recurrence relations obtained by taking the derivatives of the Eqs. (A.27)–(A.28), (A.36)–(A.37), and (A.41)–(A.42) with respect to $\overline{m}_\ell$ and $\overline{x}_\ell$, respectively. In particular:

1. The derivatives of $A_n(z)$ and $B_n(z)$ follow from the Riccati-type differential equations

$$A_n'(z) = \frac{n(n+1)}{z^2} - 1 - A_n^2(z), \quad B_n'(z) = \frac{n(n+1)}{z^2} - 1 - B_n^2(z). \tag{A.46}$$

2. The derivatives of $C_n(z)$,$R_n(z)$, and $D_n(z)$ are calculated as

$$C_n'(z) = C_n(z)\,[B_n(z) - A_n(z)], \tag{A.47}$$

$$R_n'(z) = R_n(z)\,[A_n(z) - B_n(z)], \tag{A.48}$$

$$D_n'(z) = \frac{C_n(z)\,B_n'(z) + B_n(z)\,C_n'(z) - D_n(z)\,A_n'(z)}{A_n(z)}. \tag{A.49}$$

The derivatives with respect to the physical parameters m_ℓ and r_ℓ are computed as

$$\frac{\partial T_n^{(1,2)}}{\partial m_1} = \frac{\partial T_n^{(1,2)}}{\partial \overline{m}_1} - \frac{m_2}{m_1^2}\frac{\partial T_n^{(1,2)}}{\partial \overline{m}_2} + kr_2\frac{\partial T_n^{(1,2)}}{\partial \overline{x}_2}, \tag{A.50}$$

$$\frac{\partial T_n^{(1,2)}}{\partial m_\ell} = \frac{1}{m_{\ell-1}}\frac{\partial T_n^{(1,2)}}{\partial \overline{m}_\ell} - \frac{m_{\ell+1}}{m_\ell^2}\frac{\partial T_n^{(1,2)}}{\partial \overline{m}_{\ell+1}} + kr_{\ell+1}\frac{\partial T_n^{(1,2)}}{\partial \overline{x}_{\ell+1}}, \quad \ell = 2,\dots,L-1, \tag{A.51}$$

$$\frac{\partial T_n^{(1,2)}}{\partial m_L} = \frac{1}{m_{L-1}}\frac{\partial T_n^{(1,2)}}{\partial \overline{m}_L}, \tag{A.52}$$

and

$$\frac{\partial T_n^{(1,2)}}{\partial r_1} = k\frac{\partial T_n^{(1,2)}}{\partial \overline{x}_1}, \tag{A.53}$$

$$\frac{\partial T_n^{(1,2)}}{\partial r_\ell} = km_{\ell-1}\frac{\partial T_n^{(1,2)}}{\partial \overline{x}_\ell}, \quad \ell = 2,\dots,L, \tag{A.54}$$

respectively.

When the layer thicknesses d_ℓ are specified, the radii of the layers are computed as

$$r_\ell = \begin{cases} r_{\ell+1} + d_\ell, & \ell = 1,\dots,L-1, \\ d_L, & \ell = L, \end{cases} \tag{A.55}$$

and the derivatives of a quantity F (intensity or cross section) with respect to the layer thicknesses are compute as

$$\frac{\partial F}{\partial d_1} = \frac{\partial F}{\partial r_1}, \quad \frac{\partial F}{\partial d_\ell} = \frac{\partial F}{\partial d_{\ell-1}} + \frac{\partial F}{\partial r_\ell}, \quad \ell = 2,\dots,L. \tag{A.56}$$

Cross-section calculation using the transfer-matrix method

Consider a multilayered sphere with L concentric layers of radii $r_1 < r_2 < \cdots < r_L$ (inside to outside) and relative refractive indices $m_1, m_2, \dots, m_L$ with respect to the ambient medium ($m_{L+1} = 1$). Define

$$z_\ell = km_\ell r_\ell, \quad \overline{z}_\ell = km_{\ell+1}r_\ell = \overline{m}_\ell z_\ell, \quad \overline{m}_\ell = \frac{m_{\ell+1}}{m_\ell}, \tag{A.57}$$

where k is the wavenumber of the ambient medium. Furthermore, let $\psi_n(z)$ and $\chi_n(z)$ denote the Riccati–Bessel functions.

1. For each $n = 1,\dots,N_{\mathrm{rank}}$, build the cumulative 2×2 matrices for TE and TM polarizations as ordered products of interface blocks:

$$\mathsf{M}_n^{\mathrm{TE}} = \mathsf{M}_{n,L}^{\mathrm{TE}}\mathsf{M}_{n,L-1}^{\mathrm{TE}}\cdots\mathsf{M}_{n,1}^{\mathrm{TE}}, \tag{A.58}$$

$$\mathsf{M}_n^{\mathrm{TM}} = \mathsf{M}_{n,L}^{\mathrm{TM}}\mathsf{M}_{n,L-1}^{\mathrm{TM}}\cdots\mathsf{M}_{n,1}^{\mathrm{TM}}. \tag{A.59}$$

Each interface block is

$$\mathsf{M}_{n,\ell}^{\mathrm{TE}} = \mathsf{A}\,\mathsf{B}^{\mathrm{TE}}, \quad \mathsf{M}_{n,\ell}^{\mathrm{TM}} = \mathsf{A}\,\mathsf{B}^{\mathrm{TM}}, \tag{A.60}$$

where

$$\mathsf{A} = \begin{bmatrix} \chi_n'(\overline{z}_\ell) & -\chi_n(\overline{z}_\ell) \\ -\psi_n'(\overline{z}_\ell) & \psi_n(\overline{z}_\ell) \end{bmatrix} \propto \begin{bmatrix} \psi_n(\overline{z}_\ell) & \chi_n(\overline{z}_\ell) \\ \psi_n'(\overline{z}_\ell) & \chi_n'(\overline{z}_\ell) \end{bmatrix}^{-1} \tag{A.61}$$

and

$$\mathsf{B}^{\mathrm{TE}} = \begin{bmatrix} \psi_n(z_\ell) & \chi_n(z_\ell) \\ \psi_n'(z_\ell) & \chi_n'(z_\ell) \end{bmatrix}, \quad \mathsf{B}^{\mathrm{TM}} = \begin{bmatrix} \psi_n(z_\ell) & \chi_n(z_\ell) \\ \overline{m}_\ell^{2}\,\psi_n'(z_\ell) & \overline{m}_\ell^{2}\,\chi_n'(z_\ell) \end{bmatrix}. \tag{A.62}$$

In the above equations, $\chi_n(z) = zy_n(z)$, where $y_n(z)$ denotes the spherical Neumann function.

2. After sweeping all interfaces, form for each n, the effective reflection coefficients:

$$R_n^{\mathrm{TE}} = \frac{[\mathsf{M}_n^{\mathrm{TE}}]_{11} - \mathrm{j}[\mathsf{M}_n^{\mathrm{TE}}]_{21}}{[\mathsf{M}_n^{\mathrm{TE}}]_{11} + \mathrm{j}[\mathsf{M}_n^{\mathrm{TE}}]_{21}}, \tag{A.63}$$

$$R_n^{\mathrm{TE}} = \frac{[\mathsf{M}_n^{\mathrm{TM}}]_{11} - \mathrm{j}[\mathsf{M}_n^{\mathrm{TM}}]_{21}}{[\mathsf{M}_n^{\mathrm{TM}}]_{11} + \mathrm{j}[\mathsf{M}_n^{\mathrm{TM}}]_{21}}, \tag{A.64}$$

and compute the scattering and absorption cross sections as

$$C_{\mathrm{scat}} = \frac{\pi}{2k^2}\sum_{n=1}^{N_{\mathrm{rank}}}(2n+1)(|1 - R_n^{\mathrm{TE}}|^2 + |1 - R_n^{\mathrm{TM}}|^2), \tag{A.65}$$

$$C_{\mathrm{abs}} = \frac{\pi}{2k^2}\sum_{n=1}^{N_{\mathrm{rank}}}(2n+1)(2 - |R_n^{\mathrm{TE}}|^2 - |R_n^{\mathrm{TM}}|^2). \tag{A.66}$$

Note that, in terms of the modal amplitudes

$$f_n = \sqrt{\frac{2n+1}{2}}\,(1 - R_n^{\mathrm{TM}}), \quad g_n = \sqrt{\frac{2n+1}{2}}\,(1 - R_n^{\mathrm{TE}}), \tag{A.67}$$

the scattering cross section takes the familiar form

$$C_{\mathrm{scat}} = \frac{\pi}{k^2}\sum_{n=1}^{N_{\mathrm{rank}}}(|f_n|^2 + |g_n|^2). \tag{A.68}$$

Appendix B. Forward model and data transformations

This appendix summarizes the forward model and the optional data transformations used in the inversion procedure. The purpose of these transformations is either to reduce the data dimension or to remove smooth components, while preserving the Gaussian error structure. In each case, the transformed inverse problem has the form

$$\widehat{\mathbf{F}}(\mathbf{x}) = \widehat{\mathbf{y}}^{\delta}, \qquad \widehat{\mathbf{y}}^{\delta} = \widehat{\mathbf{y}} + \widehat{\boldsymbol{\delta}}, \tag{B.1}$$

where $\widehat{\boldsymbol{\delta}}$ is the transformed data error with covariance matrix $\widehat{\mathsf{C}}_\delta$. After applying a transformation, the notation is reset for simplicity, that is, $\mathbf{F} \leftarrow \widehat{\mathbf{F}}$, $\mathbf{y}^\delta \leftarrow \widehat{\mathbf{y}}^\delta$, $\boldsymbol{\delta} \leftarrow \widehat{\boldsymbol{\delta}}$, and $\mathsf{C}_\delta \leftarrow \widehat{\mathsf{C}}_\delta$.

Forward model for intensities and cross sections

For Mie-scattering inverse problems, the forward model is defined componentwise by

$$F_i(\mathbf{x}) = y_i(\mathbf{x}) = \begin{cases} w(\theta_i)\,I(\theta_i,\mathbf{x}), & \text{intensity-based inversion,} \\ C_{\mathrm{scat}}(\lambda_i,\mathbf{x}), & \text{cross-section-based inversion,} \end{cases} \tag{B.2}$$

where $I(\theta_i,\mathbf{x})$ is the scattering intensity at angle θ_i, $C_{\mathrm{scat}}(\lambda_i,\mathbf{x})$ is the scattering cross section at wavelength λ_i, and $w(\theta_i)$ is a weighting function. Typical choices are

$$w(\theta) = 1, \qquad w(\theta) = \frac{1^\circ}{\theta}\exp\left[-2\ln^2\left(\frac{\theta}{54^\circ}\right)\right], \qquad w(\theta) = \sin^4\theta. \tag{B.3}$$

The Jacobian matrix is given by

$$\frac{\partial F_i}{\partial x_j}(\mathbf{x}) = \begin{cases} w(\theta_i)\,\dfrac{\partial I(\theta_i,\mathbf{x})}{\partial x_j}, & \text{intensity-based inversion,} \\ \dfrac{\partial C_{\mathrm{scat}}(\lambda_i,\mathbf{x})}{\partial x_j}, & \text{cross-section-based inversion.} \end{cases} \tag{B.4}$$

Principal component analysis

Let $\{\mathbf{y}_p\}_{p=1}^{P} \subset \mathbb{R}^m$ be a training data set and $\bar{\mathbf{y}} = (1/P)\sum_{p=1}^{P}\mathbf{y}_p$ its sample mean. In principal component analysis (PCA), the data are approximated in a subspace of dimension $m_{\mathrm{PC}} < m$,

$$\mathbf{y}_p \approx \bar{\mathbf{y}} + \mathsf{A}\widehat{\mathbf{y}}_p, \qquad \widehat{\mathbf{y}}_p = \mathsf{A}^{\dagger}(\mathbf{y}_p - \bar{\mathbf{y}}), \tag{B.5}$$

where $\mathsf{A} \in \mathbb{R}^{m\times m_{\mathrm{PC}}}$ is the basis matrix and $\mathsf{A}^{\dagger}$ is the Moore–Penrose pseudoinverse. Applied to the inverse problem, this gives

$$\widehat{\mathbf{F}}(\mathbf{x}) = \mathsf{A}^{\dagger}\big(\mathbf{F}(\mathbf{x}) - \bar{\mathbf{y}}\big), \qquad \widehat{\mathbf{y}}^{\delta} = \mathsf{A}^{\dagger}\big(\mathbf{y}^{\delta} - \bar{\mathbf{y}}\big), \qquad \widehat{\boldsymbol{\delta}} = \mathsf{A}^{\dagger}\boldsymbol{\delta}. \tag{B.6}$$

If $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \mathsf{C}_\delta)$, then

$$\widehat{\mathsf{C}}_\delta = \mathsf{A}^{\dagger}\mathsf{C}_\delta(\mathsf{A}^{\dagger})^{\top}, \qquad \widehat{\mathscr{C}}_\delta = \mathsf{A}^{\dagger}\mathscr{C}_\delta(\mathsf{A}^{\dagger})^{\top}. \tag{B.7}$$

In PCA, the centered data matrix $\mathsf{Y} = [\mathbf{y}_1 - \bar{\mathbf{y}}, \ldots, \mathbf{y}_P - \bar{\mathbf{y}}]$ defines the sample covariance $\mathsf{C} = (1/P)\mathsf{Y}\mathsf{Y}^{\top}$. Let $\mathsf{C} = \mathsf{U}\boldsymbol{\Sigma}\mathsf{U}^{\top}$ be its spectral decomposition. Choosing $\mathsf{A} = \mathsf{U}_{m_{\mathrm{PC}}}$, formed by the first m_{PC} eigenvectors, yields

$$\mathsf{A}^{\dagger} = \mathsf{U}_{m_{\mathrm{PC}}}^{\top}, \qquad \widehat{\mathscr{C}}_\delta = \mathsf{U}_{m_{\mathrm{PC}}}^{\top}\mathscr{C}_\delta\mathsf{U}_{m_{\mathrm{PC}}}. \tag{B.8}$$

To preserve the trace normalization in the reduced space, one defines

$$\widehat{\mathscr{C}}_\delta^{(\mathrm{norm})} = \frac{m_{\mathrm{PC}}}{\mathrm{trace}(\widehat{\mathscr{C}}_\delta)}\widehat{\mathscr{C}}_\delta, \qquad \mathrm{trace}\Big(\widehat{\mathscr{C}}_\delta^{(\mathrm{norm})}\Big) = m_{\mathrm{PC}}. \tag{B.9}$$

The corresponding squared error level and average variance are

$$\Delta^2_{\mathrm{PCA}} = \bar{\sigma}^2\,\mathrm{trace}(\widehat{\mathscr{C}}_\delta), \qquad \bar{\sigma}^2_{\mathrm{PCA}} = \frac{1}{m_{\mathrm{PC}}}\Delta^2_{\mathrm{PCA}}. \tag{B.10}$$

Note that the number of principal components m_{PC} can be obtained by imposing that the reconstruction error defined by

$$\varepsilon_{\mathrm{PCA}} = \max_{p=1,\ldots,P}\frac{\left\|\mathbf{y}_p - (\bar{\mathbf{y}} + \mathsf{A}\widehat{\mathbf{y}}_p)\right\|}{\left\|\mathbf{y}_p\right\|} \tag{B.11}$$

is below a prescribed tolerance.

Low-pass filtering by the discrete cosine transform

Let $M \geq m$ be the smallest power of two, $\mathsf{P} : \mathbb{R}^m \to \mathbb{R}^M$ the zero-padding operator, and $\mathsf{F}_{\mathrm{DCT}}$ the orthonormal discrete cosine transform matrix. For $\mathbf{z} = \mathsf{P}\mathbf{y}$, the DCT coefficients are $\widehat{\mathbf{z}} = \mathsf{F}_{\mathrm{DCT}}\mathbf{z}$. Given an energy fraction $\rho_{\mathrm{FFT}} \in (0, 1]$, the cutoff index M_{cut} is chosen such that the retained coefficients contain the prescribed fraction of spectral energy. Let M be the diagonal mask that keeps the coefficients $0, \ldots, M_{\mathrm{cut}}$, and let $\mathsf{R} : \mathbb{R}^M \to \mathbb{R}^m$ be the restriction operator. The low-pass filter is

$$\mathsf{S} = \mathsf{R}\mathsf{F}_{\mathrm{DCT}}^{\top}\mathsf{M}\mathsf{F}_{\mathrm{DCT}}\mathsf{P}. \tag{B.12}$$

The transformed problem becomes

$$\widehat{\mathbf{F}}(\mathbf{x}) = \mathsf{S}\mathbf{F}(\mathbf{x}), \qquad \widehat{\mathbf{y}}^{\delta} = \mathsf{S}\mathbf{y}^{\delta}. \tag{B.13}$$

If $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \mathsf{C}_\delta)$, then

$$\widehat{\mathsf{C}}_\delta = \mathsf{S}\mathsf{C}_\delta\mathsf{S}^{\top}, \qquad \widehat{\mathscr{C}}_\delta = \mathsf{S}\mathscr{C}_\delta\mathsf{S}^{\top}. \tag{B.14}$$

With $m_* = \mathrm{rank}(\mathsf{S})$, the normalization and corresponding noise quantities are defined in the same way as in Eqs. (B.9)–(B.10), replacing m_{PC} by m_*.

Differential model with polynomial detrending

Let the data be sampled on a uniform grid $\theta_i = \theta_{\min} + (i-1)\Delta\theta$, $i = 1, \ldots, m$, where θ_i denotes either the scattering angle θ or the wavelength λ. Define the normalized coordinates

$$t_i = \frac{\theta_i - \bar{\theta}}{L}, \qquad \bar{\theta} = \frac{1}{m}\sum_{i=1}^{m}\theta_i, \qquad L = \frac{\theta_m - \theta_1}{2}. \tag{B.15}$$

For a polynomial of degree K, define the design matrix $\mathsf{T} \in \mathbb{R}^{m\times(K+1)}$ by

$$[\mathsf{T}]_{ik} = t_i^{k-1}, \qquad i = 1, \ldots, m, \quad k = 1, \ldots, K+1. \tag{B.16}$$

The least-squares trend coefficients are

$$\mathbf{c}(\mathbf{x}) = (\mathsf{T}^{\top}\mathsf{T})^{-1}\mathsf{T}^{\top}\mathbf{y}(\mathbf{x}), \tag{B.17}$$

and the associated projection matrix is

$$\mathsf{H} = \mathsf{T}(\mathsf{T}^{\top}\mathsf{T})^{-1}\mathsf{T}^{\top}. \tag{B.18}$$

The detrended forward model and data are therefore

$$\widehat{\mathbf{y}}(\mathbf{x}) = (\mathrm{I}_m - \mathsf{H})\mathbf{y}(\mathbf{x}), \qquad \widehat{\mathbf{y}}^{\delta} = (\mathrm{I}_m - \mathsf{H})\mathbf{y}^{\delta}. \tag{B.19}$$

Hence, $\widehat{\boldsymbol{\delta}} = (\mathrm{I}_m - \mathsf{H})\boldsymbol{\delta}$, and, if $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \mathsf{C}_\delta)$,

$$\widehat{\mathsf{C}}_\delta = (\mathrm{I}_m - \mathsf{H})\mathsf{C}_\delta(\mathrm{I}_m - \mathsf{H})^{\top}, \qquad \widehat{\mathscr{C}}_\delta = (\mathrm{I}_m - \mathsf{H})\mathscr{C}_\delta(\mathrm{I}_m - \mathsf{H}). \tag{B.20}$$

Since $\mathrm{I}_m - \mathsf{H}$ is a projection matrix of rank $m_* = m - (K+1)$ the effective data dimension is reduced accordingly. The normalized shape covariance matrix and the associated squared error level and average variance are then defined as in Eqs. (B.9)–(B.10), with m_{PC} replaced by m_*.

In summary, PCA, DCT low-pass filtering, and polynomial detrending all preserve the Gaussian structure of the transformed inverse problem while modifying the covariance matrix and the effective data dimension in a consistent way.

Appendix C. Marginal posterior expectation with effective degrees of freedom

Effective degrees of freedom from correlated residuals

Under local linearization, the whitened residual $\widehat{\mathbf{r}} = \widehat{\mathbf{y}}^{\delta} - \widehat{\mathbf{F}}(\mathbf{x}^{\delta})$ has covariance

$$\mathrm{Cov}(\widehat{\mathbf{r}}) = \sigma^2(\mathrm{I}_m - \widehat{\mathsf{A}}), \tag{C.1}$$

and its squared norm satisfies

$$\mathbb{E}\{\|\widehat{\mathbf{r}}\|^2\} = (m-n)\sigma^2, \qquad \mathrm{Var}(\|\widehat{\mathbf{r}}\|^2) = 2(m-n)\sigma^4. \tag{C.2}$$

The matrix $\mathrm{I}_m - \widehat{\mathsf{A}}$ is a symmetric positive-semidefinite projection matrix and, in general, is neither a correlation matrix nor of full rank.

To estimate an effective number of independent observations from the residual correlations, an empirical surrogate model is introduced:

$$\widehat{\mathbf{r}} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathscr{R}), \tag{C.3}$$

where $\mathscr{R} \in \mathbb{R}^{m\times m}$ is symmetric positive semidefinite and has unit diagonal. This model is used only for estimating the effective degrees of freedom and is not identified with the linearized covariance (C.1). Let $R_\delta = \|\widehat{\mathbf{r}}\|^2$. Under assumption (C.3),

$$\mathbb{E}\{R_\delta\} = m\sigma^2, \qquad \mathrm{Var}(R_\delta) = 2\sigma^4\,\mathrm{trace}(\mathscr{R}^2). \tag{C.4}$$

If $\mathscr{R}$ is Toeplitz with lag correlations ρ_k, then

$$\mathrm{trace}(\mathscr{R}^2) = m + 2\sum_{k=1}^{m-1}(m-k)\rho_k^2, \tag{C.5}$$

and hence

$$\mathrm{Var}(R_\delta) = 2\sigma^4\Big[m + 2\sum_{k=1}^{m-1}(m-k)\rho_k^2\Big]. \tag{C.6}$$

Following a classical moment-matching argument, R_δ is approximated by a scaled chi-square law,

$$R_\delta \approx c\,\sigma^2\chi^2_\kappa, \tag{C.7}$$

where κ is interpreted as the effective number of independent observations. Since

$$\mathbb{E}\{\chi_\kappa^2\} = \kappa, \qquad \mathrm{Var}(\chi_\kappa^2) = 2\kappa, \tag{C.8}$$

matching the moments (C.7) with the moments (C.8) gives

$$c\kappa = m, \qquad c^2\kappa = \mathrm{trace}(\mathscr{R}^2),$$

and therefore,

$$\kappa = \frac{m^2}{\mathrm{trace}(\mathscr{R}^2)}. \tag{C.9}$$

Using Eq. (C.5), we obtain

$$\kappa = \frac{m^2}{m + 2\sum_{k=1}^{m-1}(m-k)\rho_k^2}. \tag{C.10}$$

This quantity is the classical effective number of independent observations.

Effective marginal posterior

To account approximately for correlated residuals, an effective isotropic Gaussian surrogate is introduced. The correlated residual vector is replaced by an effective isotropic residual vector with κ independent components,

$$\widehat{\mathbf{r}}_\kappa(\mathbf{x}) \sim \mathcal{N}(\mathbf{0}, \sigma_\kappa^2 \mathrm{I}), \qquad \|\widehat{\mathbf{r}}_\kappa(\mathbf{x})\|^2 \approx \|\widehat{\mathbf{r}}(\mathbf{x})\|^2, \tag{C.11}$$

where $\widehat{\mathbf{r}}(\mathbf{x}) = \widehat{\mathbf{y}}^\delta - \widehat{\mathbf{F}}(\mathbf{x})$. Matching the expected residual energies at the solution $\mathbf{x}^\delta$ yields

$$\kappa\sigma_\kappa^2 = m\sigma^2, \qquad \sigma_\kappa^2 = \frac{m}{\kappa}\sigma^2.$$

Under this effective isotropic surrogate, the likelihood is approximated by

$$p(\mathbf{y}^\delta \mid \mathbf{x}, \sigma^2) \propto (\sigma^2)^{-\kappa/2} \exp\left(-\frac{\kappa}{2m\sigma^2} R(\mathbf{x})\right), \tag{C.12}$$

where $R(\mathbf{x}) = \|\widehat{\mathbf{r}}(\mathbf{x})\|^2$. Assuming the Jeffreys prior $p(\sigma^2) \propto 1/\sigma^2$, the marginal posterior is

$$p(\mathbf{x} \mid \mathbf{y}^\delta) \propto \int_0^\infty (\sigma^2)^{-(\kappa/2+1)} \exp\left(-\frac{\kappa}{2m\sigma^2} R(\mathbf{x})\right) d\sigma^2. \tag{C.13}$$

Using the result

$$\int_0^\infty t^{-(\nu+1)} \exp\left(-\frac{\beta}{t}\right) dt = \Gamma(\nu)\beta^{-\nu}, \qquad \nu > 0,\ \beta > 0,$$

we obtain

$$p(\mathbf{x} \mid \mathbf{y}^\delta) \propto [R(\mathbf{x})]^{-\kappa/2}. \tag{C.14}$$

Thus, compared with the independent case, the exponent $m/2$ is replaced by the effective exponent $\kappa/2$. The posterior expectation of any quantity $f(\mathbf{x})$ is then approximated by

$$\mathbb{E}\{f(\mathbf{x})\} = \frac{\int f(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y}^\delta)\, d\mathbf{x}}{\int p(\mathbf{x} \mid \mathbf{y}^\delta)\, d\mathbf{x}}. \tag{C.15}$$

In this work, we calculate the expectation value $\mathbb{E}\{\mathbf{x}\} = \langle\mathbf{x}\rangle$ and the posterior covariance matrix $\widehat{\mathrm{C}}_x = \mathbb{E}\{(\mathbf{x} - \langle\mathbf{x}\rangle)(\mathbf{x} - \langle\mathbf{x}\rangle)^\mathrm{T}\}$ using Eq. (C.15) since in most cases the probability distribution is well described by a Gaussian distribution.

Importance sampling

Direct Monte Carlo sampling of Eq. (C.15) is generally inefficient because the posterior mass is concentrated near the minimizers of $R(\mathbf{x})$. Therefore, importance sampling is used. Let $q(\mathbf{x})$ be a proposal density and N_s the number of samples. Then

$$\langle f(\mathbf{x})\rangle \approx \frac{\sum_{k=1}^{N_\mathrm{s}} f(\mathbf{x}_k)\, w_k}{\sum_{k=1}^{N_\mathrm{s}} w_k}, \qquad \mathbf{x}_k \sim q, \qquad w_k = \frac{\widetilde{p}(\mathbf{x}_k)}{q(\mathbf{x}_k)}, \tag{C.16}$$

with unnormalized target density $\widetilde{p}(\mathbf{x}) = [R(\mathbf{x})]^{-\kappa/2}$ [65]. The proposal is chosen as a Gaussian distribution centered at the minimum-residual solution $\mathbf{x}^\delta$,

$$q(\mathbf{x}) = \mathcal{N}\big(\mathbf{x} \mid \mathbf{x}^\delta, \widehat{\mathrm{C}}_{\mathrm{xlin}}\big), \tag{C.17}$$

where $\widehat{\mathrm{C}}_{\mathrm{xlin}} = \widehat{\sigma}^2(\widehat{\mathrm{K}}^\top\widehat{\mathrm{K}})^{-1}$ is the covariance matrix obtained from a linearized model around the solution. The log-density of the proposal is then

$$\ln q(\mathbf{x}) = -\frac{1}{2}\left[n\ln(2\pi) + \ln\det(\widehat{\mathrm{C}}_{\mathrm{xlin}}) + (\mathbf{x} - \mathbf{x}^\delta)^\top(\widehat{\mathrm{C}}_{\mathrm{xlin}})^{-1}(\mathbf{x} - \mathbf{x}^\delta)\right]. \tag{C.18}$$

A useful diagnostic for the quality of the importance sampling approximation is the effective sample size N_eff, defined by [66]

$$N_\mathrm{eff} = \frac{1}{\sum_{k=1}^{N_\mathrm{s}} \widetilde{w}_k^2} = \frac{\left(\sum_{k=1}^{N_\mathrm{s}} w_k\right)^2}{\sum_{k=1}^{N_\mathrm{s}} w_k^2}, \tag{C.19}$$

where

$$\widetilde{w}_k = \frac{w_k}{\sum_{k=1}^{N_\mathrm{s}} w_k}. \tag{C.20}$$

A rigorous interpretation of the effective sample size is as follows. Since the importance sampling estimator is invariant under scaling of the target density, we may assume without loss of generality that $\widetilde{p}$ is normalized, i.e., $\int \widetilde{p}(\mathbf{x})\, d\mathbf{x} = 1$. From Eq. (C.19), we find

$$\frac{N_\mathrm{eff}}{N_\mathrm{s}} = \frac{\left(\frac{1}{N_\mathrm{s}}\sum_{k=1}^{N_\mathrm{s}} w_k\right)^2}{\frac{1}{N_\mathrm{s}}\sum_{k=1}^{N_\mathrm{s}} w_k^2}. \tag{C.21}$$

so that, for $N_\mathrm{s} \to \infty$, the law of large numbers yields

$$\frac{N_\mathrm{eff}}{N_\mathrm{s}} \approx \frac{(\mathbb{E}_q[w])^2}{\mathbb{E}_q[w^2]} = \frac{1}{\mathbb{E}_q[w^2]}, \tag{C.22}$$

where

$$\mathbb{E}_q[w] = \int \widetilde{p}(\mathbf{x})\, d\mathbf{x} = 1, \quad \mathbb{E}_q[w^2] = \int \frac{\widetilde{p}(\mathbf{x})^2}{q(\mathbf{x})}\, d\mathbf{x}. \tag{C.23}$$

The expected value $\mathbb{E}_q[w^2]$ is directly related to the χ^2-divergence:

$$\chi^2(\widetilde{p} \parallel q) = \int \frac{(\widetilde{p}(\mathbf{x}) - q(\mathbf{x}))^2}{q(\mathbf{x})}\, d\mathbf{x} = \mathbb{E}_q[w^2] - 1, \tag{C.24}$$

yielding

$$\frac{N_\mathrm{eff}}{N_\mathrm{s}} \approx \frac{1}{1 + \chi^2(\widetilde{p} \parallel q)}. \tag{C.25}$$

Thus, N_eff measures the discrepancy between the proposal distribution q and the target distribution $\widetilde{p}$: it is large if $q \approx \widetilde{p}$, and small if the two distributions differ significantly. In the extreme case of weight degeneracy, where one weight dominates,

$$\widetilde{w}_k \approx \begin{cases} 1, & k = k^*, \\ 0, & k \neq k^*, \end{cases}$$

we have $\sum_{k=1}^{N_\mathrm{s}} \widetilde{w}_k^2 \approx 1$, and hence $N_\mathrm{eff} \approx 1$. In this case, the proposal distribution q has very poor overlap with the target distribution $\widetilde{p}$, and the importance sampling estimator is unreliable. From a physical point of view, the quantity $R(\mathbf{x})$ plays the role of an energy or misfit. The posterior density behaves as $\widetilde{p}(\mathbf{x}) \sim R(\mathbf{x})^{-\kappa/2}$, and for large κ, this distribution becomes highly concentrated around the minimizers of $R(\mathbf{x})$. Consequently: (i) the posterior mass is localized in a small region of the parameter space, (ii) most samples drawn from q fall in regions where $\widetilde{p}(\mathbf{x})$ is very small, and (iii) a few samples near the minima receive very large weights. This leads to a collapse of the effective sample size. To improve the effective sample size, we construct the covariance matrix as $\alpha\widehat{\mathrm{C}}_{\mathrm{xlin}}$, where α is a scaling parameter determined by a grid search, based on monitoring N_eff using a reduced set of samples, for example, $0.2N_\mathrm{s}$.

Instead of using the full covariance matrix $\widehat{\mathsf{C}}_{\mathrm{xlin}}$, we may use a diagonal approximation $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2)$ with $\sigma_i^2 = [\widehat{\mathsf{C}}_{\mathrm{xlin}}]_{ii}$, $i = 1, \ldots, n$. In this case, the log-density of the proposal is computed as

$$\ln q(\mathbf{x}) = -\frac{1}{2}\left[n \ln(2\pi) + \sum_{i=1}^{n} \ln(\sigma_i^2) + \sum_{i=1}^{n} \frac{(x_i - x_i^{\delta})^2}{\sigma_i^2} \right]. \tag{C.26}$$

Metropolis–Hastings sampling

As an alternative to importance sampling, posterior expectations can be approximated by a Metropolis–Hastings Markov chain Monte Carlo method [67,68]. This approach is particularly useful when, for large κ, the posterior distribution is highly concentrated and importance sampling suffers from severe weight degeneracy. As before, the target density is the unnormalized marginal posterior $\widetilde{p}(\mathbf{x}) = [R(\mathbf{x})]^{-\kappa/2}$, where $R(\mathbf{x}) = \|\widehat{\mathbf{r}}(\mathbf{x})\|^2$, and $\widehat{\mathsf{C}}_{\mathrm{xlin}} = \widehat{\sigma}^2(\widehat{\mathsf{K}}^{\top}\widehat{\mathsf{K}})^{-1}$ is the covariance matrix obtained from the linearized model around the minimum-residual solution $\mathbf{x}^{\delta}$. Starting from an initial state $\mathbf{x}^{(0)}$, the MH method constructs a Markov chain $\{\mathbf{x}^{(k)}\}_{k=0}^{N-1}$ of length N, whose stationary distribution is proportional to $\widetilde{p}(\mathbf{x})$ (here N denotes the total number of MH iterations). The chain is initialized at $\mathbf{x}^{(0)} = \mathbf{x}^{\delta}$, and a Gaussian random-walk proposal is used,

$$q(\mathbf{x}' \mid \mathbf{x}^{(k)}) = \mathcal{N}\big(\mathbf{x}' \mid \mathbf{x}^{(k)}, \alpha\widehat{\mathsf{C}}_{\mathrm{xlin}}\big), \tag{C.27}$$

where $\alpha > 0$ is a scaling parameter controlling the proposal size. Thus, a candidate point is generated as

$$\mathbf{x}' = \mathbf{x}^{(k)} + \boldsymbol{\eta}^{(k)}, \qquad \boldsymbol{\eta}^{(k)} \sim \mathcal{N}\big(\mathbf{0}, \alpha\widehat{\mathsf{C}}_{\mathrm{xlin}}\big). \tag{C.28}$$

If box constraints are imposed, proposals outside the admissible parameter domain are rejected immediately. Since the Gaussian random-walk proposal is symmetric, i.e., $q(\mathbf{x}' \mid \mathbf{x}^{(k)}) = q(\mathbf{x}^{(k)} \mid \mathbf{x}')$, the acceptance probability reduces to

$$a(\mathbf{x}^{(k)}, \mathbf{x}') = \min\left\{1, \frac{\widetilde{p}(\mathbf{x}')}{\widetilde{p}(\mathbf{x}^{(k)})}\right\} = \min\left\{1, \left(\frac{R(\mathbf{x}')}{R(\mathbf{x}^{(k)})}\right)^{-\kappa/2}\right\}. \tag{C.29}$$

The next state of the chain is then defined by

$$\mathbf{x}^{(k+1)} = \begin{cases} \mathbf{x}', & \text{with probability } a(\mathbf{x}^{(k)}, \mathbf{x}'), \\ \mathbf{x}^{(k)}, & \text{with probability } 1 - a(\mathbf{x}^{(k)}, \mathbf{x}'). \end{cases} \tag{C.30}$$

After discarding an initial burn-in period of length N_{burn}, the remaining samples are used to estimate posterior moments. If every Lth sample is retained, the posterior expectation of any quantity $f(\mathbf{x})$ is approximated by

$$\langle f(\mathbf{x})\rangle \approx \frac{1}{N_{\mathrm{keep}}} \sum_{j=1}^{N_{\mathrm{keep}}} f(\mathbf{x}^{(k_j)}), \tag{C.31}$$

where $\{k_j\}$ denotes the retained sample indices and N_{keep} is the number of retained samples. In particular, the posterior mean and covariance matrix are estimated by

$$\langle \mathbf{x}\rangle \approx \frac{1}{N_{\mathrm{keep}}} \sum_{j=1}^{N_{\mathrm{keep}}} \mathbf{x}^{(k_j)}, \tag{C.32}$$

and

$$\widehat{\mathsf{C}}_{\mathrm{x}} \approx \frac{1}{N_{\mathrm{keep}} - 1} \sum_{j=1}^{N_{\mathrm{keep}}} \big(\mathbf{x}^{(k_j)} - \langle \mathbf{x}\rangle\big)\big(\mathbf{x}^{(k_j)} - \langle \mathbf{x}\rangle\big)^{\top}. \tag{C.33}$$

A useful diagnostic for the quality of the MH approximation is the acceptance rate,

$$A = \frac{N_{\mathrm{acc}}}{N}, \tag{C.34}$$

where N_{acc} is the number of accepted proposals [65]. If A is too small, then the proposal variance is too large and the chain moves only rarely; if A is too large, then the proposal variance is too small and the chain explores the posterior only slowly. In practice, the scaling parameter α is tuned so that the acceptance rate remains in a moderate range. Note that the generated samples are correlated, and the quality of the approximation depends on the mixing properties of the Markov chain. Consequently, a burn-in phase must be discarded and, if necessary, thinning is applied to reduce serial correlation.

Instead of using a random-walk proposal, one may employ an independence sampler, in which the proposal density does not depend on the current state:

$$q(\mathbf{x}') = \mathcal{N}\big(\mathbf{x}' \mid \mathbf{x}^{\delta}, \alpha\widehat{\mathsf{C}}_{\mathrm{xlin}}\big). \tag{C.35}$$

A candidate point is drawn independently as $\mathbf{x}' \sim q(\mathbf{x}')$. Since the proposal is no longer symmetric, the acceptance probability is given by ($\widetilde{p}(\mathbf{x}) = [R(\mathbf{x})]^{-\kappa/2}$)

$$a(\mathbf{x}^{(k)}, \mathbf{x}') = \min\left\{1, \frac{\widetilde{p}(\mathbf{x}')\, q(\mathbf{x}^{(k)})}{\widetilde{p}(\mathbf{x}^{(k)})\, q(\mathbf{x}')}\right\} = \min\left\{1, \left(\frac{R(\mathbf{x}')}{R(\mathbf{x}^{(k)})}\right)^{-\kappa/2} \frac{q(\mathbf{x}^{(k)})}{q(\mathbf{x}')}\right\}. \tag{C.36}$$

The chain is updated according to the same rule as above.

Estimation of the autocorrelation function

In practice, the correlations ρ_k are unknown and must be estimated from the whitened residuals. Let

$$\widehat{\mathbf{r}} = [\widehat{r}_1, \ldots, \widehat{r}_m]^{\top}, \qquad \bar{r} = \frac{1}{m}\sum_{i=1}^{m} \widehat{r}_i.$$

The empirical autocovariance function is defined by

$$\widehat{\gamma}(k) = \frac{1}{m-k} \sum_{i=1}^{m-k} (\widehat{r}_i - \bar{r})(\widehat{r}_{i+k} - \bar{r}), \qquad k = 0, \ldots, K_{\max}, \tag{C.37}$$

which is the usual approximately unbiased estimator. The corresponding sample autocorrelation function is

$$\widehat{\rho}(k) = \frac{\widehat{\gamma}(k)}{\widehat{\gamma}(0)}, \qquad k = 0, \ldots, K_{\max}. \tag{C.38}$$

Replacing ρ_k by $\widehat{\rho}(k)$ in Eq. (C.10) gives the empirical estimate

$$\kappa(K) = \frac{m^2}{m + 2\sum_{k=1}^{K}(m-k)\widehat{\rho}(k)^2}, \qquad K = 0, 1, \ldots, K_{\max}. \tag{C.39}$$

To reduce the influence of sampling noise at large lags, a cutoff lag K^* is selected by monitoring the stabilization of $\kappa(K)$. Given a window length $L \geq 1$ and a relative tolerance $\varepsilon_{\kappa} > 0$, define

$$\Delta_{\mathrm{win}}(K) = \max_{j=1,\ldots,L} \frac{|\kappa(K+j) - \kappa(K)|}{\kappa(K)}. \tag{C.40}$$

Then the cutoff lag is chosen as

$$K^* = \min\left\{K \in \{0, \ldots, K_{\max} - L\} : \Delta_{\mathrm{win}}(K) \leq \varepsilon_{\kappa}\right\}. \tag{C.41}$$

Finally, the sample autocorrelation may be tapered to improve stability. With taper weights $w(k) \in [0, 1]$, define

$$\widetilde{\rho}(k) = \begin{cases} w(k)\widehat{\rho}(k), & 0 \leq k \leq K^*, \\ 0, & k > K^*. \end{cases} \tag{C.42}$$

Typical choices are the Bartlett taper

$$w(k) = 1 - \frac{k}{K^*}, \qquad 0 \leq k \leq K^*,$$

and the Tukey taper

$$w(k) = \frac{1}{2}\left(1 + \cos\frac{\pi k}{K^*}\right), \qquad 0 \leq k \leq K^*.$$

The tapered estimate of the effective degrees of freedom is then

$$\widetilde{\kappa} = \frac{m^2}{m + 2\sum_{k=1}^{K^*}(m-k)\widetilde{\rho}(k)^2}, \tag{C.43}$$

which is generally more stable than the untapered estimator.

Data and code availability

The developed optimization tool together with detailed documentation is available at https://github.com/DmitryEfremenko/MLSI. This repository also contains the raw output of the DDA simulations.

References

[1] Romanov AV, Yurkin MA. Single-particle characterization by elastic light scattering. Laser Photon. Rev 2021;15:2000368. http://dx.doi.org/10.1002/lpor.202000368.

[2] Tycko DH, Metz MH, Epstein EA, Grinbaum A. Flow-cytometric light scattering measurement of red blood cell volume and hemoglobin concentration. Appl Opt 1985;24(9):1355–65. http://dx.doi.org/10.1364/AO.24.001355.

[3] Potenza MAC, Cremonesi L. An overview of the optical characterization of free microparticles and their radiative properties. J Quant Spectrosc Radiat Transfer 2023;311:108773. http://dx.doi.org/10.1016/j.jqsrt.2023.108773.

[4] Semyanov KA, Tarasov PA, Zharinov AE, Chernyshev AV, Hoekstra AG, Maltsev VP. Single-particle sizing from light scattering by spectral decomposition. Appl Opt 2004;43(26):5110–5. http://dx.doi.org/10.1364/AO.43.005110.

[5] Romanov AV, Maltsev VP, Yurkin MA. Retrieving refractive index of single spheres using the phase spectrum of light-scattering pattern. Opt Laser Technol 2023;161:109141. http://dx.doi.org/10.1016/j.optlastec.2023.109141.

[6] Bartholomew-Biggs M, Ulanowski Z, Zakovic S. Using global optimization for a microparticle identification problem with noisy data. J Global Optim 2005;32(3):325–47. http://dx.doi.org/10.1007/s10898-004-1943-0.

[7] Dimiduk TG, Manoharan VN. Bayesian approach to analyzing holograms of colloidal particles. Opt Express 2016;24(21):24045–60. http://dx.doi.org/10.1364/OE.24.024045.

[8] Strokotov DI, Yurkin MA, Gilev KV, van Bockstaele DR, Hoekstra AG, Rubtsov NB, Maltsev VP. Is there a difference between T- and B-lymphocyte morphology? J. Biomed Optics 2009;14(6):064036. http://dx.doi.org/10.1117/1.3275471.

[9] Konokhova AI, Chernova DN, Moskalensky AE, Strokotov DI, Yurkin MA, Chernyshev AV, Maltsev VP. Super-resolved calibration-free flow cytometric characterization of platelets and cell-derived microparticles in platelet-rich plasma. Cytom A 2016;89:159–68. http://dx.doi.org/10.1002/cyto.a.22621.

[10] Ulanowski Z, Wang Z, Kaye PH, Ludlow IK. Application of neural networks to the inverse light scattering problem for spheres. Appl. Optics 1998;37(18):4027–33. http://dx.doi.org/10.1364/AO.37.004027.

[11] Berdnik VV, Loiko VA. Retrieval of size and refractive index of spherical particles by multiangle light scattering: neural network method application. Appl Opt 2009;48(32):6178–87. http://dx.doi.org/10.1364/AO.48.006178.

[12] Moskalensky AE, Yurkin MA, Konokhova AI, Strokotov DI, Nekrasov VM, Chernyshev AV, Tsvetovskaya GA, Chikova ED, Maltsev VP. Accurate measurement of volume and shape of resting and activated blood platelets from light scattering. J Biomed Opt 2013;18(1):017001. http://dx.doi.org/10.1117/1.JBO.18.1.017001.

[13] Gilev KV, Yurkin MA, Chernyshova ES, Strokotov DI, Chernyshev AV, Maltsev VP. Mature red blood cells: From optical model to inverse light-scattering problem. Biomed Opt Express 2016;7(4):1305–10. http://dx.doi.org/10.1364/BOE.7.001305.

[14] Wang X, Bi L, Han W, Zhang X. Single-scattering properties of encapsulated fractal black carbon particles computed using the invariant imbedding T-matrix method and deep learning approaches. J Geophys Res Atmos. 2023;128(21). http://dx.doi.org/10.1029/2023JD039568.

[15] Lin Z-YD, Weinberger AJ, Zubko E, Arnold JA, Videen G. Glitterin: Toward replacing the role of Lorenz–Mie theory in astronomy using neural networks trained on light scattering of irregularly shaped grains. Publ Astron Soc Pac 2025;137(12):124502. http://dx.doi.org/10.1088/1538-3873/ae1e2f.

[16] Yurkin MA. Discrete dipole approximation. In: Menguc MP, Francoeur M, editors. Light, plasmonics and particles. Amsterdam: Elsevier; 2023, p. 167–98. http://dx.doi.org/10.1016/B978-0-323-99901-4.00020-2.

[17] Romanov AV, Yurkin MA. Rigorous analysis of the spectral sizing of single particles based on light scattering patterns. Opt Laser Technol 2022;151:108047. http://dx.doi.org/10.1016/j.optlastec.2022.108047.

[18] Peurifoy J, Shen Y, Jing L, Yang Y, Cano-Renteria F, DeLacy B, Tegmark M, Joannopoulos JD, Soljačić M. Nanophotonic particle simulation and inverse design using artificial neural networks. Sci. Advances 2018;4(6):eaar4206. http://dx.doi.org/10.1126/sciadv.aar4206.

[19] So S, Mun J, Rho J. Simultaneous inverse design of materials and structures via deep learning: Demonstration of dipole resonance engineering using core–shell nanoparticles. ACS Appl Mater Interfaces 2019;11(27):24264–8. http://dx.doi.org/10.1021/acsami.9b05857.

[20] Wiecha PR, Arbouet A, Girard C, Muskens OL. Deep learning in nano-photonics: inverse design and beyond. Photon. Res 2021;9(5):B182–200. http://dx.doi.org/10.1364/PRJ.415960.

[21] Aden AL, Kerker M. Scattering of electromagnetic waves from two concentric spheres. J. Appl Phys. 1951;22(10):1242–6. http://dx.doi.org/10.1063/1.1699834.

[22] Bohren CF, Huffman DR. Absorption and Scattering of Light by Small Particles. New York: Wiley; 1983, Wiley Professional Paperback Edition, 1998; Wiley-VCH, Weinheim.

[23] Toon OB, Ackerman TP. Algorithms for the calculation of scattering by stratified spheres. Appl. Optics 1981;20(20):3657–60. http://dx.doi.org/10.1364/AO.20.003657.

[24] Bhandari R. Scattering coefficients for a multilayered sphere: analytic expressions and algorithms. Appl. Optics 1985;24(13):1960–7. http://dx.doi.org/10.1364/AO.24.001960.

[25] Wu ZS, Wang YP. Electromagnetic scattering for multilayered sphere: recursive algorithms. Radio Sci. 1991;26:1393–401. http://dx.doi.org/10.1029/91RS01192.

[26] Wu ZS, Guo LX, Ren KF, Gouesbet G, Gréhan G. Improved algorithm for electromagnetic scattering of plane waves and shaped beams by multilayered spheres. Appl. Optics 1997;36(21):5188–98. http://dx.doi.org/10.1364/AO.36.005188.

[27] Johnson BR. Light scattering by a multilayer sphere. Appl. Optics 1996;35(18):3286–96. http://dx.doi.org/10.1364/AO.35.003286.

[28] Mackowski DW, Altenkirch RA, Mengüç MP. Internal absorption cross sections in a stratified sphere. Appl. Optics 1990;29(10):1551–9. http://dx.doi.org/10.1364/AO.29.001551.

[29] Kai L, Massoli P. Scattering of electromagnetic-plane waves by radially inhomogeneous spheres: a finely stratified sphere model. Appl. Optics 1994;33(3):501–11. http://dx.doi.org/10.1364/AO.33.000501.

[30] Yang W. Improved recursive algorithm for light scattering by a multilayered sphere. Appl. Optics 2003;42(9):1710–20. http://dx.doi.org/10.1364/AO.42.001710.

[31] Lock JA. Scattering of an electromagnetic plane wave by a luneburg lens III finely stratified sphere model. J Opt Soc Amer A 2008;25(12):2991. http://dx.doi.org/10.1364/josaa.25.002991.

[32] Moroz A. A recursive transfer-matrix solution for a dipole radiating inside and outside a stratified sphere. Annals Phys. 2005;315(2):352–418. http://dx.doi.org/10.1016/j.aop.2004.07.002.

[33] Qiu W, Shen Y, DeLacy BG, Joannopoulos JD, Soljačić M, Jing L. Optimization of broadband optical response of multilayer nanospheres. Optics Express 2012;20(16):18494–504. http://dx.doi.org/10.1364/OE.20.018494.

[34] Rasskazov IL, Moroz A, Carney PS. STRATIFY: a comprehensive and versatile MATLAB code for a multilayered sphere. OSA Continuum 2020;3(8):2290–306. http://dx.doi.org/10.1364/OSAC.399979.

[35] Zhang J. A modified recursive transfer matrix algorithm for radiation and scattering computation of multilayer spheres. 2024, http://dx.doi.org/10.48550/arXiv.2409.10877, arXiv:2409.10877.

[36] Grainger RG, Lucas J, Thomas GE, Ewen GBL. Calculation of Mie derivatives. Appl. Optics 2004;43(28):5386–93. http://dx.doi.org/10.1364/AO.43.005386.

[37] Li Y, Bowler N. Computation of Mie derivatives. Appl. Optics 2013;52(20):4997–5006. http://dx.doi.org/10.1364/AO.52.004997.

[38] Doicu A, Wriedt T, Eremin YA. Light Scattering by Systems of Particles: Null-Field Method with Discrete Sources — Theory and Programs. Springer Berlin Heidelberg; 2006, http://dx.doi.org/10.1007/978-3-540-33697-6.

[39] Törn AA. A search clustering approach to global optimization. In: Dixon LCW, Szegö GP, editors. Towards global optimization. Amsterdam: North-Holland; 1978, p. 49–62.

[40] Boender CGE, Rinnooy Kan HG, Timmer GT, Stougie L. A stochastic method for global optimization. Math. Programming 1982;22:125–40. http://dx.doi.org/10.1007/BF01581033.

[41] Rinnooy Kan AHG, Timmer GT. Stochastic global optimization methods, part I: Clustering methods. Math. Programming 1987;39:27–56. http://dx.doi.org/10.1007/BF02592070.

[42] Rinnooy Kan AHG, Timmer GT. Stochastic global optimization methods part II: Multi level methods. Math Program 1987;39(1):57–78. http://dx.doi.org/10.1007/bf02592071.

[43] Tsoulos IG, Lagaris IE. MinFinder: Locating all the local minima of a function. Comput. Phys Commun. 2006;174:166–79. http://dx.doi.org/10.1016/j.cpc.2005.10.001.

[44] Ali MM, Storey C. Topographical multilevel single linkage. J. Glob Optimization 1994;5:349–58. http://dx.doi.org/10.1007/BF01096684.

[45] Doicu A, Efremenko DS, Wirth CL, Wriedt T. An algorithm for solving the inverse problem in total internal reflection microscopy. J Quant Spectrosc Radiat Transfer 2025;345:109534. http://dx.doi.org/10.1016/j.jqsrt.2025.109534.

[46] Son H, Hochstaffl P, Efremenko D. Synthetic study of a surrogate-model-based CO2/CH4 retrieval algorithm from infrared reflectance measurements. Light Eng 2026;34(2):19–27. http://dx.doi.org/10.33383/2025-042.

[47] Marsaglia G, Tsang WW. The ziggurat method for generating random variables. J. Stat Software 2000;5:1–7. http://dx.doi.org/10.18637/jss.v005.i08.

[48] Matsumoto M, Nishimura T. Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. ACM Trans Model Comput Simulation 1998;1:3–30. http://dx.doi.org/10.1145/272991.272995.

[49] Dong N, Wu CH, Ip WH, Chen ZQ, Chan CY, Yung KL. An opposition-based chaotic GA/PSO hybrid algorithm and its application in circle detection. Comput. Math Appl. 2012;64:1886–902. http://dx.doi.org/10.1016/j.camwa.2012.03.040.

[50] Gao Y, Wang YJ. A memetic differential evolutionary algorithm for high dimensional functions optimization. In: Third international conference on natural computation (ICNC 2007). vol. 4, IEEE; 2007, p. 188–92.

[51] Gao WF, Liu SY. A modified artificial bee colony algorithm. Comput. Oper Res. 2012;39:687–97. http://dx.doi.org/10.1016/j.cor.2011.06.007.

[52] McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 1979;21:239–45. http://dx.doi.org/10.2307/1268522.

[53] Rahnamayan S, Tizhoosh HR, Salama MM. Opposition-based differential evolution for optimization of noisy problems. In: 2006 IEEE international conference on evolutionary computation. IEEE; 2006, p. 1865–72.

[54] Rahnamayan S, Tizhoosh HR, Salama MM. Opposition-based differential evolution. IEEE Trans Evol Comput. 2008;12:64–79. http://dx.doi.org/10.1109/TEVC.2007.894200.

[55] Du Q, Gunzburger M, Ju L. Advances in studies and applications of centroidal Voronoi tessellations. Numerical Math: Theory, Methods Appl. 2010;3:119–42. http://dx.doi.org/10.4208/nmtma.2010.32s.1.

[56] Dennis Jr JE, Gay DM, Welsch RE. An adaptive nonlinear least-squares algorithm. ACM Trans Math Software 1981;7:348–68. http://dx.doi.org/10.1145/355958.355965.

[57] Dennis Jr JE, Gay DM, Welsch RE. Algorithm 573. NL2SOL — An adaptive nonlinear least-squares algorithm. ACM Trans Math Software 1981;7:369–83. http://dx.doi.org/10.1145/355958.355966.

[58] Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. SIAM J Sci Comput. 1995;16:1190–208. http://dx.doi.org/10.1137/0916069.

[59] Jones DR, Perttunen CD, Stuckman BE. Lipschitzian optimization without the Lipschitz constant. J. Optim Theory Appl. 1993;79(1):157–81. http://dx.doi.org/10.1007/BF00941892.

[60] Zhang J, Sanderson AC. JADE: Adaptive differential evolution with optional external archive. IEEE Trans Evol Comput. 2009;13(5):945–58. http://dx.doi.org/10.1109/TEVC.2009.2014613.

[61] Tanabe R, Fukunaga A. Improving the search performance of SHADE using linear population size reduction. In: Proc. 2014 IEEE congress on evolutionary computation. 2014, p. 1658–65. http://dx.doi.org/10.1109/CEC.2014.6900380.

[62] Chuprov I, Gao J, Efremenko DS, Buzaev F. Application of global optimization for retrievals from synthetic multi-angle measurements. Light Eng 2024;(03–2024):11–9. http://dx.doi.org/10.33383/2023-090.

[63] Yurkin MA, Hoekstra AG. The discrete-dipole-approximation code ADDA: Capabilities and known limitations. J Quant Spectrosc Radiat Transfer 2011;112:2234–47. http://dx.doi.org/10.1016/j.jqsrt.2011.01.031.

[64] Argentin C, Chaumet PC, Gross M, Yurkin MA. A floating-point-consistent cross-verification framework for reproducible and interoperable DDA solvers with fair benchmarking. Comput Phys Comm 2026;325:110172. http://dx.doi.org/10.1016/j.cpc.2026.110172.

[65] Robert CP, Casella G. Monte Carlo Statistical Methods. Springer New York; 2004, http://dx.doi.org/10.1007/978-1-4757-4145-2.

[66] Liu JS. Monte Carlo Strategies in Scientific Computing. Springer New York; 2004, http://dx.doi.org/10.1007/978-0-387-76371-2.

[67] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. J Chem Phys 1953;21(6):1087–92. http://dx.doi.org/10.1063/1.1699114.

[68] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970;57(1):97–109. http://dx.doi.org/10.1093/biomet/57.1.97.