

Deep Learning Methods for Colloidal Silver Nanoparticle Concentration and Size Distribution Determination from UV-Vis Extinction Spectra

Tomas Klinavičius^{a,*}, Nadzeya Khinevich^a, Asta Tamulevičienė^{a,b}, Loic Vidal^c, Sigitas Tamulevičius^{a,b} and Tomas Tamulevičius^{a,b,*}

^aInstitute of Materials Science of Kaunas University of Technology, K. Baršausko St. 59, LT-51423, Kaunas, Lithuania

^bDepartment of Physics, Kaunas University of Technology, Studentų St. 50, LT-51368, Kaunas, Lithuania

^cInstitut de Science des Matériaux de Mulhouse IS2M UMR 7361, 15 rue Jean Starcky, F 68100 Mulhouse, France

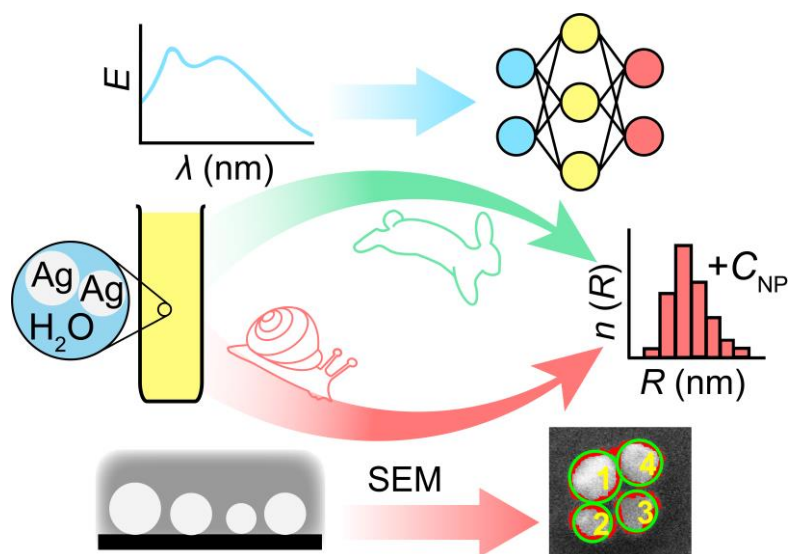
*Corresponding authors:

T. Klinavičius: tomas.klinavicius@ktu.lt

T. Tamulevičius: tomas.tamulevicius@ktu.lt

Abstract

Electron microscopy, while reliable, is an expensive, slow, and inefficient technique for thorough size distribution characterization of both mono- and polydisperse colloidal nanoparticles. If rapid *in-situ* characterization of colloid samples is to be achieved, a different approach, based on fast, widely accessible, and inexpensive optical measurements such as UV-Vis spectroscopy in combination with spectra interpretation related to Mie scattering theory, is needed. In this article, we present a tandem deep neural network (DNN) for the size distribution and concentration prediction of close to spherical silver colloidal nanoparticle batches synthesized via the seeded growth method. The first DNN identified the dipole component of the localized surface plasmon resonance and the second one determined the size distribution from the isolated spectral component. The training data was engineered to be bias-free and generated numerically. High prediction accuracy with typical root mean square percentage error of mean size *ca.* 1.2% was maintained, spanning the entire prediction range up to 150 nm in radius, suggesting the possible extension limits of the effective medium theory used for simulating the spectra. The DNN-predicted nanoparticle concentrations also were very close to the ones expected based on synthesis precursor contents as well as measured by atomic absorption spectroscopy.



The first DNN identified the dipole component of the localized surface plasmon resonance and the second one determined the size distribution from the isolated spectral component. The training data was engineered to be bias-free and generated numerically. High prediction accuracy with typical root mean square percentage error of mean size *ca.* 1.2% was maintained, spanning the entire prediction range up to 150 nm in radius, suggesting the possible extension limits of the effective medium theory used for simulating the spectra. The DNN-predicted nanoparticle concentrations also were very close to the ones expected based on synthesis precursor contents as well as measured by atomic absorption spectroscopy.

Keywords: nanoparticles, plasmonics, silver, size distribution, deep neural networks, effective medium theory, Mie theory

1. Introduction

Virtually any pure solid-state material, such as metals, semiconductors, ceramics, and others¹ or its combination can exist in the nanoparticle (NP) form². NP synthesis methods can be divided into two broad categories – physical and chemical methods³. Physical methods use physical interactions, such as mechanical⁴ or thermal⁵, to produce NPs from the bulk material. Physical synthesis methods also include laser ablation of solids in a liquid⁶ and sputtering⁷. On the other hand, chemical methods use chemical reactions, such as chemical reduction^{8,9} or deposition on a surface from vapor¹⁰. Regardless of the synthesis method, NPs always have a certain size distribution, typically a log-normal one^{11,12}.

Unlike microparticles, NPs exhibit size-dependent properties, which are different from those of their bulk counterparts¹³. Localized Surface Plasmon Resonance (LSPR) is a plasmonic phenomenon¹⁴ that can be optically excited in metallic and semiconductor NPs¹⁵. The LSPR is observed in the extinction ($E(\lambda)$) spectrum of the NPs where multiple peaks, which correspond to different resonance modes, can manifest¹⁶. As a phenomenon, LSPR is highly sensitive to the refractive index of the ambient medium, making it suitable for biomolecule¹⁷ and virus¹⁸ detection applications. This and other unique properties of NPs make them exceptionally useful in biomedicine for the detection of antigens¹⁹, biological warfare agents²⁰ and gene delivery²¹. They also found environmental and energy-related applications for cleaning up pollutants in the water²² and conversion of CO₂ to carbon-based fuels²³, along with water splitting for the generation of green hydrogen²⁴, and enhancing solar energy harvesting^{25,26}. The NP size distribution is one of the most critical parameters for their catalytic efficacy²⁷ as well as other cellular interactions²⁸, especially in cancer detection²⁹, and treatment³⁰ where knowing the NP size distribution is essential. Due to such wide application and the importance of size, simple yet rapid characterization approaches of the NP size distribution information are of special interest in an industrial setting³¹.

Estimating the characteristic size dimension of NPs experimentally is generally a difficult and time-consuming task. One highly popular method is direct imaging of drop-casted NPs by scanning electron microscopy (SEM) or transmission electron microscopy (TEM)³². Producing statistically reliable NP distributions by these methods requires multiple micrographs of the NPs and automatic image processing³³. There are other size measurement techniques, such as dynamic light scattering (DLS)³⁴, atomic force microscopy, and many others^{35–38} each of them with their limitations, such as price and time requirements. UV-Vis $E(\lambda)$ spectra also can be used to characterize the size distribution of liquid-suspended NPs³⁹ but are not straightforward to interpret⁴⁰.

The spherical NP $E(\lambda)$ spectra can be computed for a broad range of sizes either analytically employing Mie theory³⁴ or numerical methods¹⁴. Approaches based on effective media have also been developed⁴¹. Due to the “spectrum to distribution” problem having multiple degrees of freedom, an optimization-based approach⁴² might fail to provide an optimal distribution⁴³. Therefore, a more reliable approach for elucidating the size distributions from $E(\lambda)$ spectra of NPs is required.

NP concentration is another property of interest, especially for those who develop new wet-chemistry colloidal NP synthesis methods^{44,45}. NP seed solutions concentration determination is paramount for computing volumes and chemical precursor concentrations needed to produce specific-size NPs via the seeded growth method⁴⁶. Rapid *in-situ* determination of concentration using spectroscopic data is of particular interest⁴⁷.

In recent years, machine learning (ML) based data analysis tools have exploded in popularity⁴⁸. As a data-driven approach, an ML algorithm can learn the underlying patterns in the data and approximate them effectively, circumventing computational or analytical difficulties^{49,50}. Recently, ML has emerged as a popular method for the characterization of particle size distributions^{51,52}. Some authors use pre-processing of $E(\lambda)$ spectra in order to increase the accuracy of ML algorithms⁵² while others use the raw spectra³⁹. Using computationally derived data to train ML algorithms has also drawn attention^{51,53}. ML tools such as DNNs have been proven to be exceptionally powerful⁵⁴. Due to the wide functionality of DNNs⁵⁵, they have found many scientific applications in biomedicine⁵⁶, materials science⁵⁷, detection of chemical materials⁵⁸, and many more fields, including design of NP colloids⁵⁹, and assisting in their synthesis^{60,61}. Early attempts at using multilayer feedforward DNNs to analyze particle sizes focused on the microparticle size range^{62–65}. In later

years, both DNNs and other ML techniques were used to investigate size distributions of particles in the nanoscale⁶⁶, including more complex shapes such as rod-like NP^{51,67}, but their success was limited at best. Essentially, determining the size distribution from an $E(\lambda)$ spectrum is an inverse modelling problem. A tandem deep neural network (DNN) approach is used quite often for solving inverse problems, for example in photonic design^{68–71}.

In this work, we present a tandem DNN architecture for determining the radius distribution of seeded growth synthesized silver (Ag) colloidal NPs as well as their concentration in the colloid from UV-Vis $E(\lambda)$ spectra. The DNNs were trained using computationally generated data composed of NP distributions, concentration values, and $E(\lambda)$ spectra. The data was engineered to be as bias-free as possible by thoroughly examining the influence of data parameters on the characteristic aspects of the data and selecting said parameters to produce unbiased data instances for the training data. The tandem DNN was tested with experimental wet-chemistry synthesized NP data, providing an excellent match between predictions and empirical results. The first stage of the tandem DNN processed the spectra by first isolating the dipolar peak of the spectrum and the second stage performing the prediction. Using the selected tandem architecture, we have demonstrated that accurate predictions of NP size can be made in a broad range – up to 150 nm in radius, more than other authors – by using only dipole, instead of dipole and quadrupole, features. What is more, our approach allowed to estimate the concentration of both the total mass of Ag and the Ag NPs concentration in the solution.

2. Methods

2.1. Experimental Methods

2.1.1. Synthesis of NPs

Ag NPs were synthesized using the seeded growth method⁸ (**Figure 1a**). The size and size dispersion of Ag NP seeds were controlled by varying the concentrations of trisodium citrate (TC, C_{TC}), tannic acid (TA, C_{TA}), and silver nitrate (SN, C_{SN}) (all from Sigma Aldrich) in the aqueous solution. Ag NPs were grown from seeds (**Figure 1b**) and enlarged in sequential synthesis steps (**Figure 1c**). More details on both seeds and NP synthesis are provided in **Supplementary section S1.1**. Six different NP colloid batches (“A”–“F”) were produced as summarized in **Table S1**. Batches “A” – “E” were synthesized using varying concentrations of TA in the seed solution to induce NP growth, resulting in a greater variation in their size distribution. On the other hand, NPs from batch “F” were synthesized with a reduced TA concentration in the seed solution aiming for a decrease in the concentration of components during the growth process for approaching consistent NP growth and maintaining a constant standard deviation in radii. The production of a wide variation in NP sizes, along with varying size distributions, served as a means to validate the proper functionality of the DNNs.

2.1.2. Characterization of NPs

UV-Vis spectroscopy of Ag colloids

A 1.4 nm spectral resolution optical fiber spectrometer “AvaSpec-2048” (Avantes) and light source “AvaLight-DHc” (Avantes) were used for measuring the $E(\lambda)$ spectra of colloid solutions in 1 mm optical path length quartz cuvette. Initially, colloids were measured with their original particle concentration. All $E(\lambda)$ measurement results were normalized to the $E(\lambda)$ of the same cuvette filled with water.

Determination of NP size distribution

Before imaging, the NP colloids of their original concentration were sonicated in order to break up possible agglomerates. Each NP sample was prepared for scanning electron microscope (SEM) imaging by drop-casting the corresponding colloid solution on a silicon substrate. Field emission gun SEM “Quanta 200 FEG” (FEI) was used. Multiple micrographs of NPs from every sample were made in order to get a sufficient amount of NP images for a reliable effective NP radius statistical analysis. Effective Ag NPs radius distribution was derived from SEM micrographs employing an automated custom micrograph analysis and statistics processing script that was implemented in MATLAB (MathWorks), similar to³³. The principle of SEM micrograph pre- and post-processing is showcased in **Figure 1d** – the red outline shows the actual contour of the NPs while

the green outline is the circle with the resulting effective NP radius. The steps of the automated analysis are depicted in **Figure S1** and described in more detail in **Supplementary section S1.2**. The hydrodynamic particle radius and its distribution were also confirmed by DLS employing a “Zeta Sizer Ultra” particle size analyzer (Malvern Panalytical). Details of the measurement can be found in **Supplementary section S1.3**.

Determination of NP shape and Crystallinity

A high-resolution transmission electron microscope (TEM) “ARM-200F” (JEOL) was used for imaging Ag seeds and NPs. NPs were prepared for imaging by casting them on a TEM wire grid and then drying them at room temperature. Due to the higher magnification of the TEM when compared to SEM, TEM micrographs were used for NP shape and qualitative crystallinity inspection purposes. The crystallinity of several selected samples was also evaluated by X-ray diffraction (XRD) using an X-ray diffractometer D8 Discover” (Bruker AXS GmbH). Measurement details can be found in **Supplementary section S1.4**.

Concentration measurement of colloids

Atomic absorption spectrometer “AAAnalyst 400” (PerkinElmer) was used to measure the mass concentration (mg/L) of Ag in the Ag NP colloids. Colloid samples were specially prepared before measurement – Ag NPs were dissolved in Aqua Regia mixed in a ratio of 2 to 1.

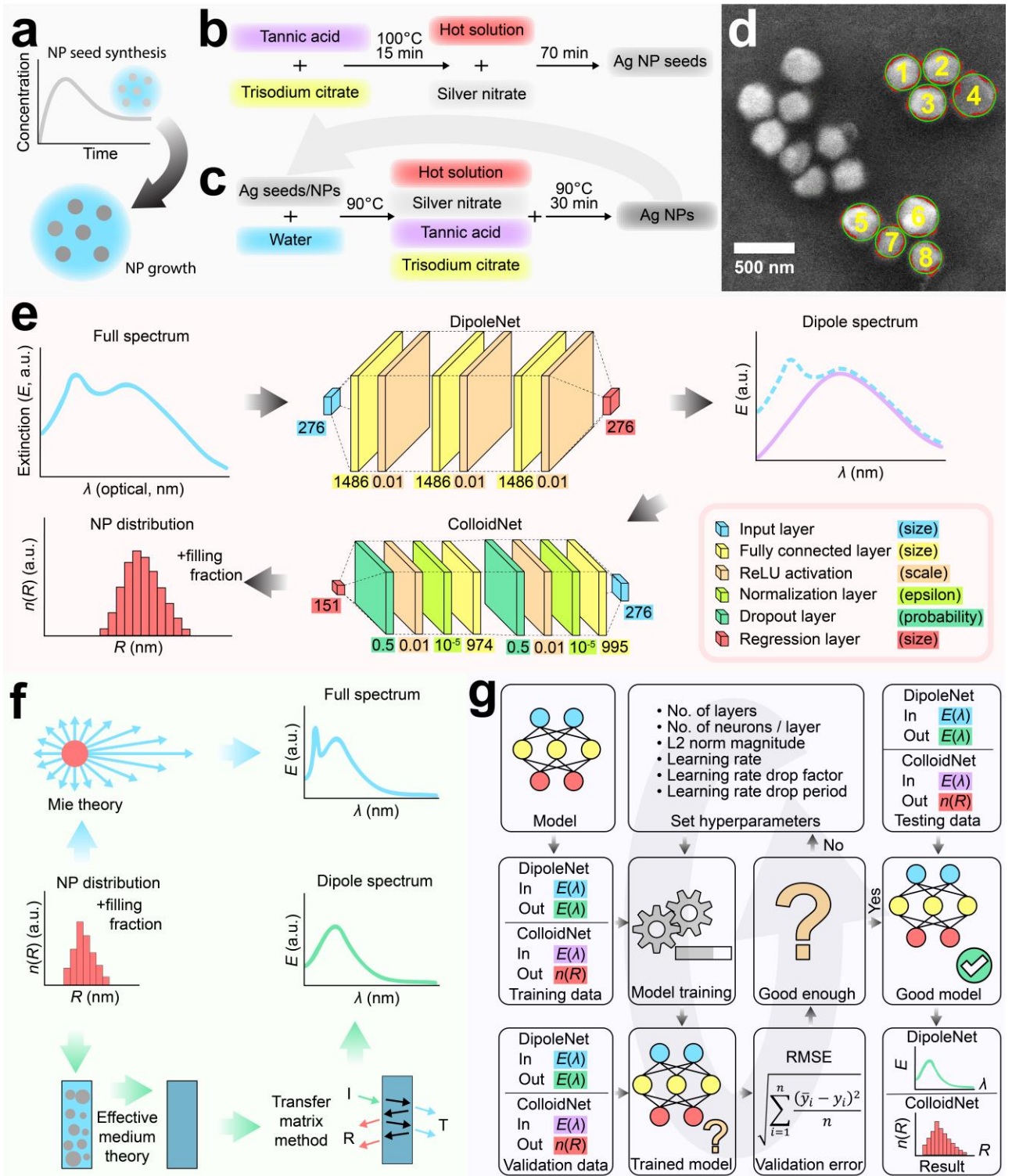


Figure 1. Experimental and computational methods used for Ag NP colloid $E(\lambda)$ spectra analysis for obtaining the effective NP radius size distribution ($n(R)$). (a) Graphical representation of Ag NP seed synthesis and subsequent growth. (b) The chemical reaction for the Ag NP seeds colloid synthesis. (c) The reaction for Ag NPs synthesis. (d) Determination of the empirical NP effective radius distribution from SEM micrographs, and overlaid graphics indicate the automated area analysis results used for extraction of the effective radius distribution of numbered NPs. (e) Working principle of the proposed tandem deep neural network system for determining NP concentration and radius distribution in a colloidal solution from its $E(\lambda)$ spectra, the legend explains the layers and their parameters of the DNNs used where “DipoleNet” is responsible for subtracting the dipole peak and the “ColloidNet” for finding the radius distribution from the pretreated $E(\lambda)$ spectra. (f) Generation of training data for both DNNs using Mie theory and effective medium theory with TMM. (g) Workflow for hyperparameter optimization and network testing for both networks.

2.2. Numerical Methods

2.2.1. Computing Hardware and Software

A desktop computer with a processor “Intel(R) Core (TM) i5-9500 CPU @ 3.00GHz” and 16 GB of RAM running “MATLAB 2021a” (MathWorks) on “Windows 10 Enterprise” (Microsoft) was used to perform all computations. Training of the DNNs as well as their hyperparameter optimization was performed on the same computer, using a graphical processing unit (GPU) “GeForce GTX 1050 Ti” (NVIDIA).

2.2.2. Mie Theory for Calculation of Full Extinction Spectra

Mie theory, as described in³⁴, was used to compute the $E(\lambda)$ of spherical NPs. The spectra $E(\lambda)$ contained peaks corresponding to all resonances depending on NP size, not just the dipolar ones. Such spectra in this work are referred to as *full* spectra (**Figure 1f** blue spectra). The NP distributions used to originate *full* spectra were the same as those used for *dipole* spectra. In Mie theory, the amplitude of $E(\lambda)$ is defined by the number of NPs (N) interacting with the light wave and the thickness of the interaction region (constant in our case). The Palik permittivity was used to define the Ag NPs⁷² while the permittivity of water was taken from⁷³.

2.2.3. Transfer Matrix Method Using Effective Medium Theory for Calculation of Dipolar Extinction Spectra

The Transfer Matrix Method (TMM)⁷⁴ based on scattering matrices⁷⁵ was used to compute $E(\lambda)$ containing only dipolar resonances corresponding to $E(\lambda)$ peaks of spherical NPs. In this work, such spectra are referred to as *dipole* spectra (**Figure 1f** green spectra). Since TMM is only able to perform computations on homogeneous slabs of material, a modified Maxwell – Garnett – Mie (MMGM) effective medium theory (EMT)⁴¹ was used to homogenize the colloidal solution of randomly distributed subwavelength-sized NPs into a homogeneous medium (**Figure 1f**). The same permittivity dispersions were used as in Section 2.2.2. Using an EMT allows to avoid a generally computationally intensive and time-consuming 3D numerical simulation. This was extremely useful for generating training and validation data in vast quantities. The effective medium approximation is valid as long as the NPs are of sufficiently subwavelength size⁷⁶. The effective medium model and its validity for larger NPs are described in more detail in **Supplementary section S1.5**.

In MMGM, concentration is accounted for as a volume-filling fraction (F). The concentration of NPs in the colloid directly impacts the amplitude of the modelled spectra, with a higher concentration corresponding to a higher amplitude as explained in detail in **Supplementary section S2.5**.

2.2.4. Engineering of Training and Validation Data

Using Mie theory for *full* spectra and TMM with MMGM medium for *dipole* spectra, NP distributions, and corresponding $E(\lambda)$ spectra were synthetically generated for both training and validation of DNNs. Assuming the log-normal shape of NP size distributions, the mode (M_L) and full width at half maximum (FWHM_L) were selected as the parameters defining the distribution. Due to lack of a parametrization of log-normal distributions with these parameters, they were subsequently transformed into the canonical parameters, mean μ_L and standard deviation σ_L , which were used to generate the distributions. The derived equations of the transformation (S4) – (S5) are provided in **Supplementary section S1.6** while the ranges of all data parameters are found in **Table S2**.

Two sets of computationally derived training and validation data were generated – one for “DipoleNet” and one for “ColloidNet”. For “DipoleNet”, data consisted of pairs of *full* spectra (inputs, **Figure 1f** blue spectra) and corresponding *dipole* spectra (outputs, **Figure 1f** green spectra). For “ColloidNet”, data pairs were composed of *dipole* spectra and corresponding NP distributions along with their F values (**Figure 1f** red distribution). The size of each training dataset was 626 359 pairs, while the size of each validation dataset was 156 590 pairs – 1/4th the size of the training data.

Testing data for “DipoleNet” used experimentally obtained *full* $E(\lambda)$ spectra for inputs. For the output, the best *dipole* spectrum fits, computed with TMM using MMGM medium, for the *dipolar* resonance peak of *full*

spectra were used. Testing data for “ColloidNet” consisted of *dipole* spectra obtained from “DipoleNet” for the inputs and experimentally obtained distributions of effective NP radius and NP colloid concentration measurements for the outputs. Each testing dataset consisted of 45 corresponding pairs.

2.2.5. The Architecture of Neural Networks, Hyperparameter Optimization and Network Training

DNN “DipoleNet” is constructed as a multilayer feedforward neural network and is depicted in **Figure 1e**. It is composed of an input layer, internal blocks of layers and an output layer. A fully connected layer followed by a Leaky ReLU activation function constitutes a single block. The Leaky ReLU function was chosen to avoid the vanishing gradient problem⁷⁷. Finally, a regression layer was used as the output layer.

Similarly, “ColloidNet”, depicted in **Figure 1e**, was also composed of an input layer, repeating layer blocks and an output layer. This time, a single block was composed of a fully connected layer, a normalization layer, a Leaky ReLU layer, and a dropout layer. Descriptions of all layers are available in⁷⁸.

To determine the best hyperparameters⁷⁹, hyperparameter optimization (HPO) using Bayesian optimization⁸⁰ was performed. The workflow of HPO for both DNNs is depicted in **Figure 1g**. For DNN training during HPO, 1/10th of the training and validation datasets are randomly extracted into reduced training and validation datasets. Validation RMSE was chosen as the cost function to minimize. The reduced datasets were reshuffled once every epoch to ensure maximum data variation during training. It is assumed that hyperparameters which assure the best network performance (lowest validation root-mean-square error, RMSE) will also provide the best network performance when training and validating the DNN with full-sized datasets. It was empirically determined that the root mean square propagation training algorithm⁸¹ works best for training both DNNs. After HPO, optimal hyperparameters were used to train and monitor the DNNs with their respective full-sized training and validation datasets. The same training parameters that were used during HPO were used for the final training process. Ranges of hyperparameters, along with optimal values and other parameters of both DNNs are listed in **Table S3**. The use of training and validation data during training is described in **Supplementary section S1.7** section.

3. Results and Discussion

3.1. Experimental results

Overall, 45 Ag NP samples were synthesized via the seeded-growth method (**Figure 1a, b, c**)⁸. NPs were specifically synthesized to have 2 distinct kinds of distributions – narrow distributions and broad distributions – for various mean radius ranges. This distinction between different batches is clearly visible in the distributions of their statistical parameters, depicted in **Figure S2** and available in **Table S1**. From this empirical data, it was observed that radius distributions with a smaller mean NP radius are narrower, while distributions with a larger mean NP radius are broader. This directly translates into the full width at half maximum (FWHM) of the log-normal fits and is consistent with experimental data provided by other authors⁸. Meanwhile, NP radius distributions (histograms of probability density with a bin width of 1 nm) for all samples (sorted by mean NP radius) are displayed in **Figure 2e**. Mean NP radius values in the study span from 10.5 nm to 107.6 nm. On average, around 1000 (and sometimes up to *ca.* 2000), but never less than 600, NPs are observed for each NP sample, providing high reliability of the collected statistics of NP radius distributions as recommended by the literature⁸². Based on the goodness of fit being $R^2 > 0.87$ for all samples, it was confirmed that the NP radius distributions have a log-normal distribution. Such observation is consistent with the literature^{11,12}.

Because imaging-based measurement techniques (TEM, SEM) determine the size of each NP directly, the quality/noise of the underlying distribution shape and reliability of statistical parameters derived from said distribution are affected by the number of NPs that are analyzed. Indirect measurement methods such as DLS³⁴ produce smooth NP distributions. They confirmed the tendency of the electron microscopy results, but despite that, DLS-determined size distributions did not always absolutely overlap with electron microscopy results (**Figure S3**) and were sub-optimal for DNN validation as explained in more detail in **Supplementary section S2.2**.

Both transmission and scanning electron microcopies confirmed that the NPs seeds were spherical (**Figure 2a**) and became more faceted as NPs grew (**Figure 2b**) but in general preserved symmetrical close to spherical shape, hence suitable to be described by the MMGM effective medium theory⁴¹. This is consistent with the literature describing the synthesis of spherical Ag NPs using various wet chemical methods^{83–86}. In **Figure 2c**, the crystallite planes corresponding to the Ag (111) facet can be attributed to sub-NP-sized crystallites that constitute the polycrystalline NPs and are larger for bigger radius NPs (**Figure S4**) as discussed in more detail in the **Supplementary section S2.3**.

$E(\lambda)$ spectra following the same increasing mean NP radius order are displayed in **Figure 2d**. For narrow NP distributions with a small mean NP radius, the $E(\lambda)$ spectrum consists of a single narrow peak, corresponding to a dipole resonance⁸⁷. As the mean NP radius increases, the peak $E(\lambda)$ wavelength shifts towards longer wavelengths, and the FWHM of the $E(\lambda)$ peak increases due to dynamic effects⁴¹. In general, the $E(\lambda)$ spectrum is composed of a multitude of resonances (**Figure S5 a**), but for small NPs only dipolar resonances manifest. As NP size increases, higher-order LSPR oscillations become supported – a quadrupole $E(\lambda)$ peak emerges and continues to grow in intensity, peak position and FWHM, eventually overtaking the *dipole* peak's intensity (**Figure S6 b**). These findings are consistent with widely known predictions made with Mie theory reported by other authors, *e.g.*,^{34,88}. Once higher-order modes are supported, the quasi-static approximation¹⁶ no longer holds. At that point, scattering is the dominant factor of $E(\lambda)$ (**Figure S6 a**). However, this transition is not clear-cut, as determined in⁷⁶. In our work, it was found that the MMGM effective medium theory is still valid even for NPs with a radius of 150 nm, as long as F is increased to account for the loss of amplitude. For a more in-depth discussion about the validity of the effective medium theory used for generating synthetic $E(\lambda)$ spectra, see **Supplementary section S2.4**.

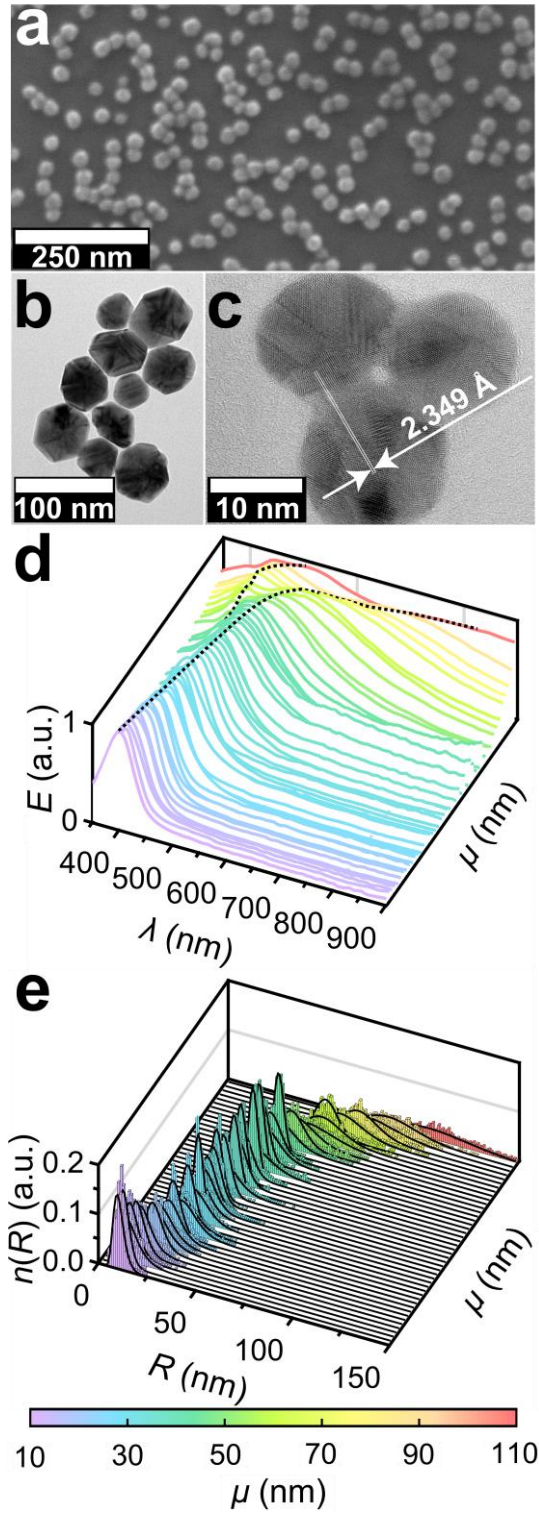


Figure 2. Properties of the synthesized Ag NP colloids. (a) Typical SEM micrograph of Ag NPs from a drop casted colloidal solution on a silicon substrate (sample "C5"). (b) TEM micrographs of NPs after several stages of growth (sample "F15"). (c) High-resolution TEM micrographs of NP seeds for the "F" batch, showing the crystallite boundaries of the NP seeds with indicated interatomic distance corresponding to the Ag (111) plane. (d) $E(\lambda)$ spectra, normalized to one, of all investigated colloidal solutions with increasing mean NP radius, truncated lines indicate *dipole* and quadrupole peak positions. (e) NP radius distributions corresponding to the spectra in (d) extracted from the SEM micrographs, a solid line represents the log-normal fit of experimental distribution.

3.2. Numerical Data Results

Using equations (eq. S4) – (eq. S5) and numerically derived relationships of F (eq. S6) and N (eq. S8), data instances were engineered to be as bias-free as possible by selecting the ranges of the underlying data values to correspond to those of testing data and by equally representing all cases of data values in those ranges (**Table S2**). The relevant data parameter region in the (M_L, FWHM_L) space is depicted in **Figure 3c** (top), the same region post-transformation in the (μ_L, σ_L) parameter space is depicted in **Figure 3c** (bottom). Due to the logarithmic nature of the log-normal distribution parameters (μ_L, σ_L) , the post-transformation region is spaced unevenly compared to the linear (M_L, FWHM_L) space.

The effect of (μ_L, σ_L) on F in MMGM when the spectrum amplitude is maintained is highly non-linear, as depicted in **Figure 3b** and in **Figure 3d** in more detail. An equivalent relation also exists for N in Mie theory. This was taken into account when generating training and validation data. It is observed that for larger NPs F has to increase tremendously in order to preserve the amplitude. This is because the absorbance cross-section decreases in relation to the $E(\lambda)$ cross-section due to the onset of scattering (**Figure S6**). That is why the same value of F produces a smaller amplitude of the spectrum for larger NPs as computed with TMM and MMGM – because MMGM computes absorbance and is based on the 1st (dipolar) electric Mie coefficient. However, this change in amplitude does not have any effect on the shape of the spectrum, as it is an extremely close match when the relevant peak of *full* and *dipole* $E(\lambda)$ spectra is brought to the same amplitude (**Figure 3a**). Due to the effect of F weakening with increasing NP size, it needs to be increased in order to maintain the same spectrum amplitude. Knowing the dependence of F on the $E(\lambda)$ amplitude to be directly proportional⁴¹, this dependence is universal for any spectrum amplitude. An analogous dependence was determined for N , which is a parameter in Mie theory that is equivalent to F in MMGM. Due to scattering being accounted for in Mie theory, the dependence of N was different than that of F . Both of these dependencies were used to generate synthetic spectra with amplitude changing in an exact manner. This allowed to generate training and validation data with a high amount of variation and virtually no bias.

Due to the dependence depicted in **Figure 3d** being derived numerically, it could not be used as a general expression, which is crucial for generating the F value for any given parameters (μ_L, σ_L) . Therefore, it was fitted using a 2 variable 5th order polynomial fit, with goodness of fit parameter R^2 being 0.999. For actual application, the fit of F was used to compute the value of F for log-normal distributions with parameters (μ_L, σ_L) .

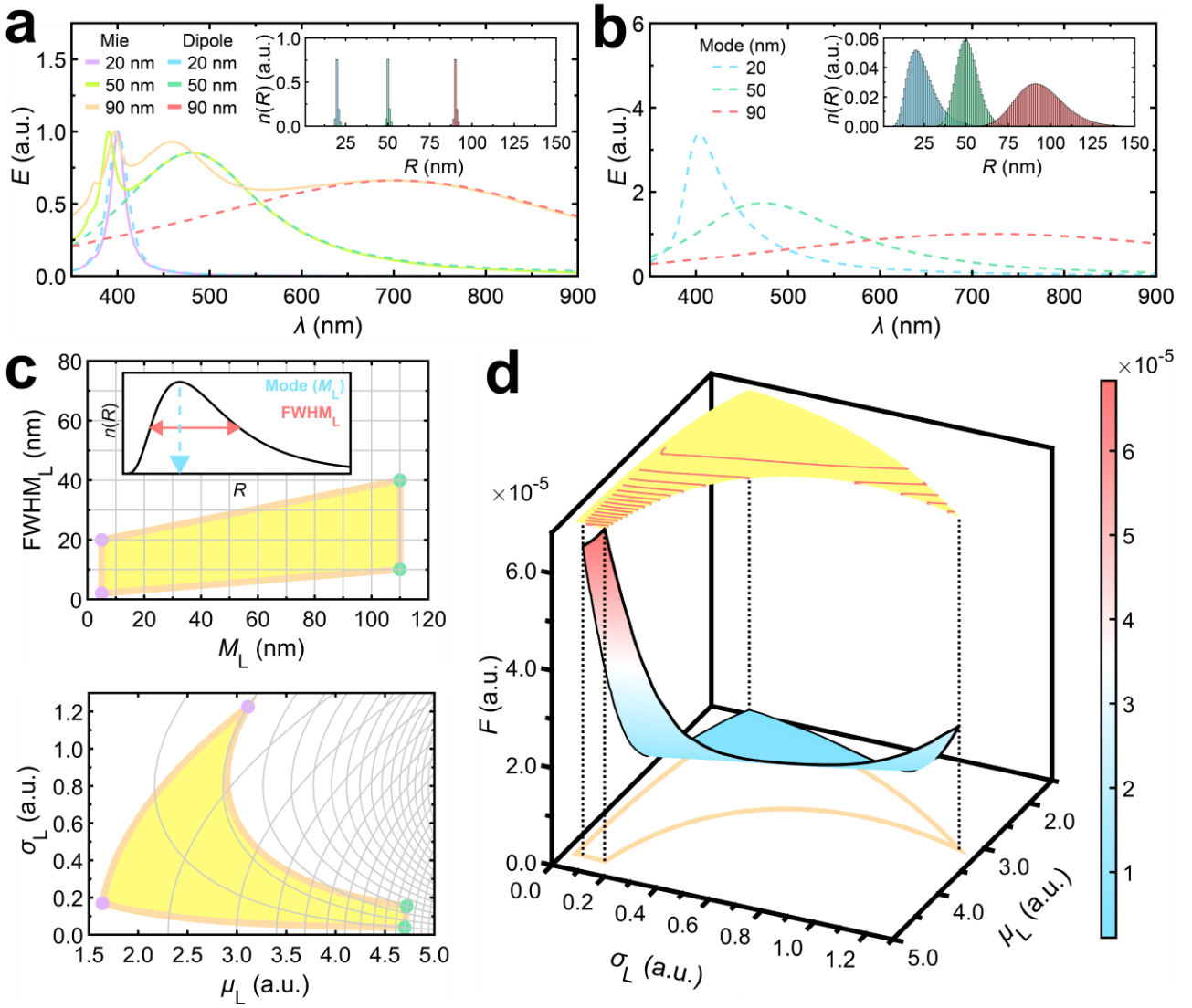


Figure 3. Training data for DNNs. (a) Comparison of $E(\lambda)$ full spectra originated with Mie theory (solid lines) and *dipole* spectra originated with TMM using MMGM medium (dashed lines), the inset depicts narrow radius distributions (M_L 20, 50, and 90 nm, FWHM_L 5 nm) used to compute $E(\lambda)$ spectra employing both origination methods. (b) *Dipole* $E(\lambda)$ spectra for broad radius distribution NPs with M_L identical to those used in (a), depicted in the inset. Spectra were computed with TMM using MMGM medium where the volume filling fraction doubles for each subsequent distribution. (c) Log-normal distribution parameter (M_L and FWHM_L) values in the training data before and after transformation to canonical distribution parameters μ_L and σ_L . (d) Values of filling fraction F depending on canonical distribution parameters required for achieving the amplitude of *dipole* $E(\lambda)$ spectrum equal to one.

3.3. “DipoleNet” DNN Results

The results of “DipoleNet” tests are summarized in **Figure 4**. It is important to mention that the output of “DipoleNet” is purely fictitious as there are no real *dipole* $E(\lambda)$ spectra⁸⁹ and is only used as a stepping stone for achieving the end result – the output of “ColloidNet”. During operation of the tandem DNN, this output is further used as input for “ColloidNet”.

As seen in **Figure 4a**, the $E(\lambda)$ peak which corresponds to higher-order charge oscillations is becoming more prominent with increasing NP size. Results of using “DipoleNet” on its testing data can be seen in **Figure 4b**. The residual between the *dipole* fit spectrum and the prediction of “DipoleNet” at most is about 2 orders of magnitude less than the $E(\lambda)$ amplitude, which corresponds to an absolute error of a few percent at most. This indicates that the “DipoleNet” has correctly learned to predict the characteristics of the $E(\lambda)$ peak which corresponds to the *dipole* component of the $E(\lambda)$. Furthermore, in **Figure 4c** it is seen that the root mean square percentage error (RMSPE,⁹⁰) between the *dipole* fit and “DipoleNet” prediction is also consistently small

throughout a wide range of NP radii. Not only does this further support the conclusion that “DipoleNet” has correctly learned to identify the dipole feature, but it also shows the network’s reliability for NPs of various size. “DipoleNet” conducted each individual prediction in about 0.0031 s.

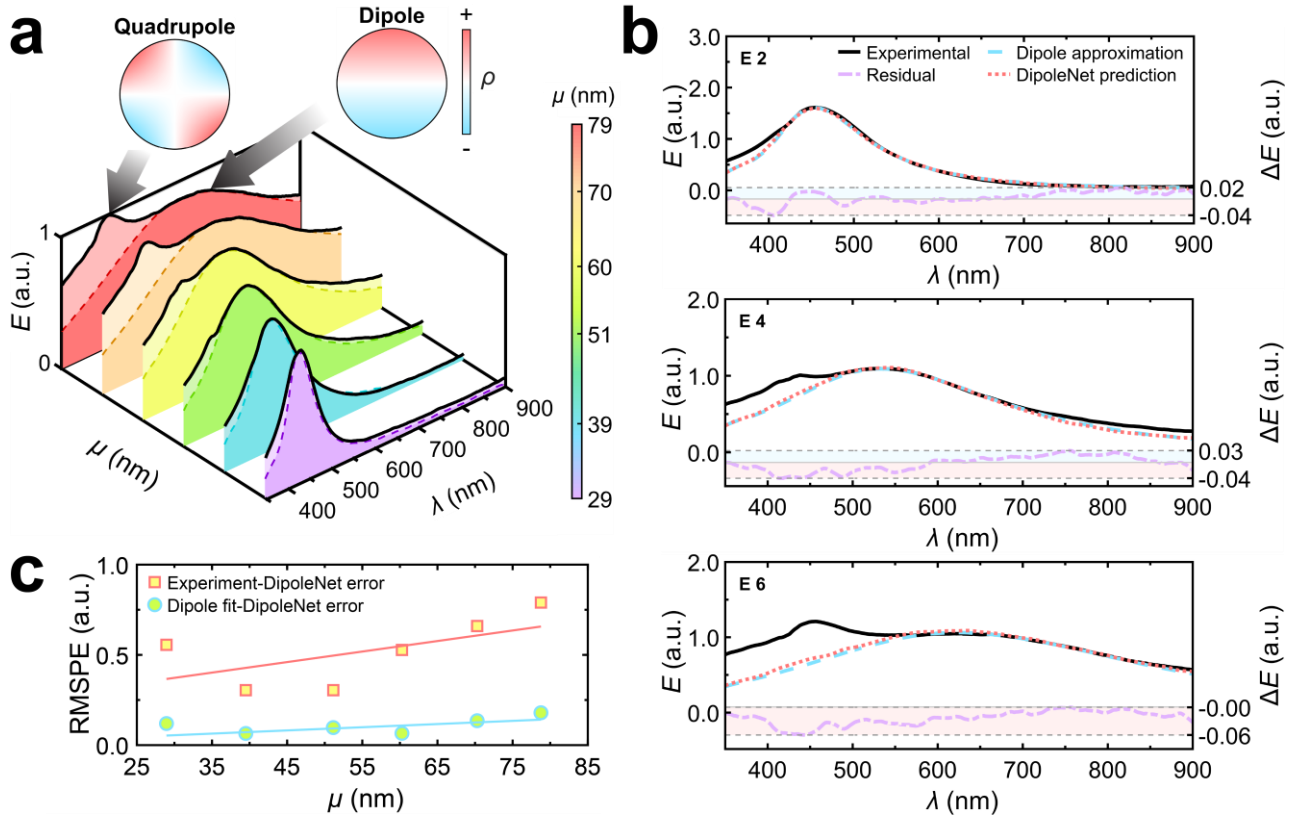


Figure 4. Extraction of a *dipole* component from *full* $E(\lambda)$ spectrum. (a) Examples of *full* spectra (light shade) and *dipole* spectral components (dark shade with dashed line) for the “E” batch of colloid solutions increasing in mean NP radius extracted with “DipoleNet”, along with surface charge density distributions corresponding to distinct resonances. (b) The 2nd, 4th, and 6th spectra from (a) compared with *dipole* spectra predicted by “DipoleNet” and *dipole* spectra computed with effective medium theory and TMM for the best fit of the *dipole* spectral peak on the primary axis, and the difference between *dipole* spectrum and predicted spectrum of “DipoleNet” on the secondary axis. (c) RMSPE between *full* $E(\lambda)$ spectra and “DipoleNet” predicted $E(\lambda)$ spectra compared to the error between “DipoleNet” spectra and *dipole* spectra for various mean NP sizes.

3.4. “ColloidNet” DNN Results

3.4.1. Predictions of NP size

The “DipoleNet”-processed experimental $E(\lambda)$ spectra were fed as inputs to “ColloidNet” and resulted in the NP size predictions depicted in **Figure 5**. Similarly to “DipoleNet”, “ColloidNet” originated its predictions in *ca.* 0.0082 s per spectra. As seen in **Figure 5a**, all distributions have a clearly defined log-normal nature¹¹. By mere visual inspection, it can be seen that the shape of the predicted distributions is close to empirically determined NP radius distribution indicating that “ColloidNet” predictions are correct. The main source of mismatch between the ground truth (experimental data) and network predictions seems to be noise due to having a limited number of samples as *ca.* 1000 NP radius data points might still be not enough to produce a smooth histogram and/or possible analysis errors of the automated SEM micrograph processing algorithm. There are two possible error types in the micrograph processing algorithm – (i) over-segmentation of NPs when one NP is misunderstood as several smaller ones and (ii) non-separation where several NPs clumped together are misunderstood as one large NP. Both can arise due to using the watershed transformation⁹¹ to separate the NPs clumped together in the micrographs.

The true – predicted plots corresponding to the three analyzed histograms in **Figure 5a** are depicted in **Figure 5b**. They confirm a good correlation between the predicted values and the ground truth. Most points follow the trend of this line while overestimates of the histogram bins are above the ground truth line and underestimates are below. These errors mostly represent the smallest values of $n(R)$ (see **Figure 5b**) that are less significant as they have little contribution to $E(\lambda)$. It can be further seen that for distributions with increasing mean NP size the RMSPE is consistently under 40%, indicating the “ColloidNet” has a similar level of prediction accuracy for NPs over a broad size range.

The quantitative statistical distribution parameters – the mean (μ), the mode (M), and the standard deviation (σ) – for both experimental and predicted distributions are provided in **Figure 5a** legends for selected cases and also analyzed for the entire batch “E” in **Figures 5c, d, and e**, respectively, indicate excellent agreement between predicted and experimental values. The mean μ was predicted with the highest level of accuracy while M was slightly underestimated. Haiss *et al.*⁹², proposed a highly simplified method for determining mean NP sizes from their *dipolar* $E(\lambda)$ peak spectral location, but it makes no effort to account for the shape and especially the size distribution width. Meanwhile σ had the largest deviation from the ground truth and was always overestimated. Overestimation of σ values can be explained by the noise in network predictions. Although not visible in the histograms due to its small size, predictions made with “ColloidNet” contain low amplitude noise, which could have influenced the estimate of σ . Another possibility is that NPs much larger or smaller than the mean values were simply not observed due to an insufficient number of NP radius samples, as it would require using tens or hundreds of thousands of samples to observe them because they are just that rare. Still, value prediction is consistently accurate over a large interval, and in most cases, the parameters are within 10% RMSPE bounds or even 1.2% as for μ . This shows the versatility of the proposed tandem DNN system. K. Shiratori *et al.*⁵¹ reported a similar prediction accuracy level for the nanorod sizes. This work has the advantage over⁶⁶ because here size characterization is achieved only from UV-Vis $E(\lambda)$ spectra, without requiring DLS data as input. Moreover, to the best of our knowledge, current work exceeds the maximum plasmonic NP radius prediction range of 110 nm previously reported by E. X. Tan *et al.*⁵² by at least 40 nm. Higher prediction accuracy without sacrificing prediction range could potentially be achieved by using empirical $E(\lambda)$ spectra as training data, but this requires the synthesis of tens if not hundreds of colloids with precision NP size tuning and concentration increment⁵². Moreover, the current study was limited by the spectral range of the silicon-based detector spectrometer and deuterium-tungsten lights source while the LSPR range together with mean sizes could be potentially extended using a multiple detector UV-Vis-NIR spectrometer. Using both dipolar and quadrupolar features/spectra could potentially yield an even wider window of size characterization⁵², but it would require discarding the use of the effective medium theory which brought computational benefits met in this work.

The insets of **Figure 5a** depict the *full* experimental UV-Vis $E(\lambda)$ spectra along with the DNN predicted *dipole* spectra while it is a direct computation result of the “ColloidNet” predicted parameters – NP radii distributions and their respective volume filling fractions. The shape of the *dipole* peak is mainly affected by the NP radius distribution shape while the amplitude of the peak is affected by the filling fraction, which is equivalent to the NP concentration. A good fit of the *dipole* peak is one more quality estimation metric indicator that the prediction is correct or close to it, especially in terms of concentration.

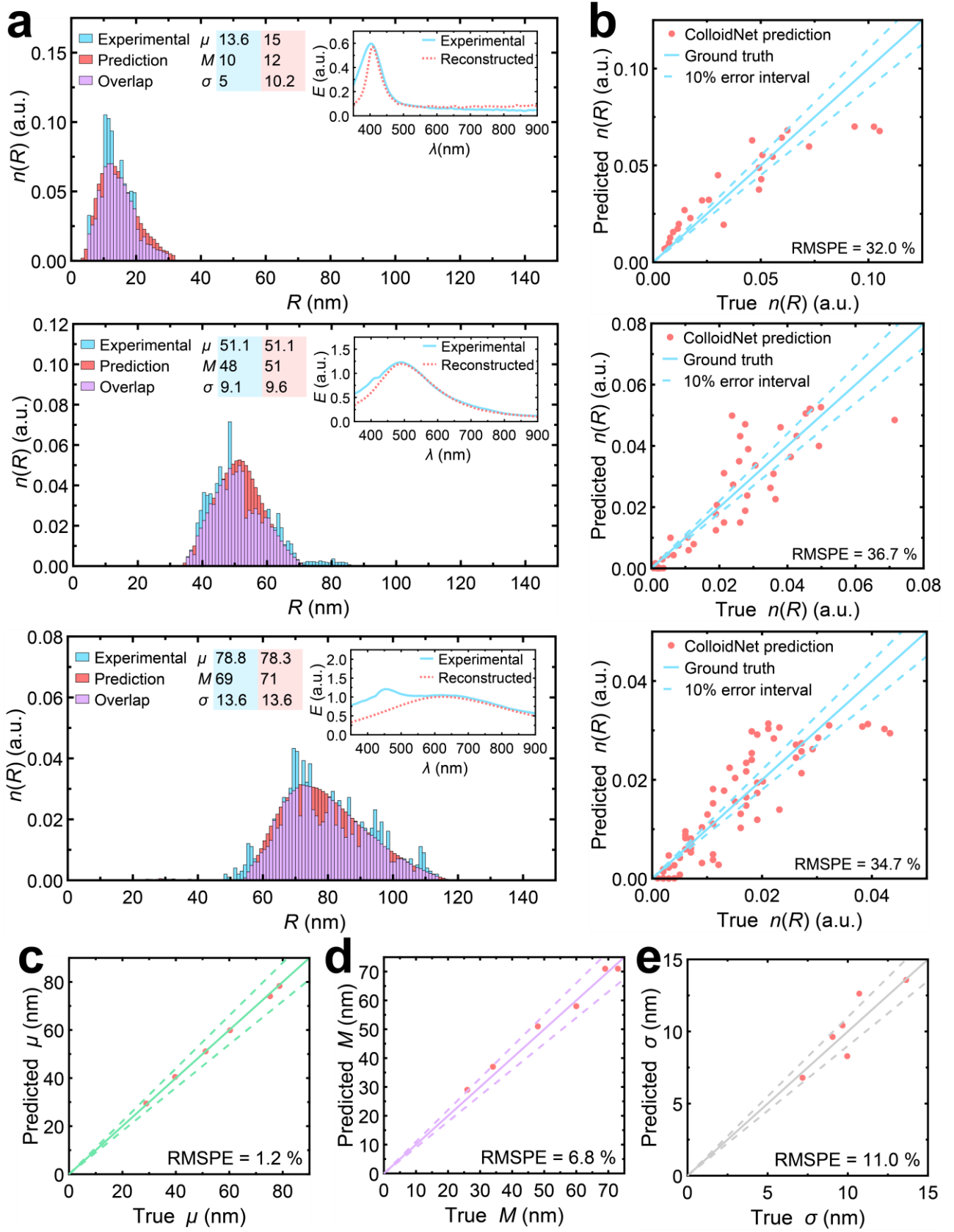


Figure 5. Comparison of NP size distribution characterization by SEM and “ColloidNet”. (a) Experimental and predicted NP size distributions and their statistical distribution parameters for samples “F2”, “E3”, and “E6”, the inset depicts experimental $E(\lambda)$ spectra and *dipole* spectra reconstructed from the predicted NP distributions and their volume filling fractions. (b) True-predicted value plots for distributions of (a) with the estimated RMSPE values and 10% absolute error intervals. True-predicted value plots for mean NP sizes (μ), modes (M), and standard deviations (σ) for distributions of colloidal series “E” are depicted in (c), (d), and (e),

respectively, with solid lines corresponding to the true value while dashed lines correspond to a 10% error interval.

3.4.2. Predictions of NP concentration

Since “ColloidNet” predicts only the volume filling fraction F , additional post-processing is needed to convert it into useful parameters such as Ag mass (C_M) and NP (C_{NP}) concentrations as explained in **Supplementary section S2.5**. Predictions of “F” batch NP colloid Ag concentration acquired from “ColloidNet” predictions are depicted in **Figure 6**. For “ColloidNet”, the LSPR $E(\lambda)$ peak magnitude is the key factor for predicting both C_M and C_{NP} , due to it being directly influenced by F (eq. S7, S9), according to⁴¹. However, the influence of the size distribution $n(R)$ becomes significant when computing C_{NP} , as it is directly impacted by NP size (eq. S9).

The literature described a relation between the intensity of $E(\lambda)$ at the interband wavelength which is 250 nm for Ag and the overall concentration of the material in the colloid^{93,94}. Therefore, the dependence of $E(\lambda)$ magnitude at the interband together with the maximum position of LSPR on the mean NP radius is depicted in **Figure 6a**. The slightly increasing interband trend (**Figure 6a**) can be directly addressed to the stepwise Ag precursor addition at each NP synthesis. The interband $E(\lambda)$ is indicative of the overall Ag present in the solution (C_M), not just the Ag which is in NP form and follows the expected AgNO_3 content trend (**Figure 6b**) computed based on synthesis conditions tabulated data in **Table S1**. Moreover, the predicted C_M is in the very same absolute mg/L level as expected based on the used chemical synthesis recipe and confirmed by the independent atomic mass spectroscopy measurement.

Finally, C_{NP} can be computed from C_M (**Figure 6c**) and compared with an expected value, which is computed from the synthesis conditions assuming the silver precursor is uniformly distributed among all NPs⁴⁶. This is described in more detail in **Supplementary section S2.5**. It is important to note that expected C_{NP} values are computed assuming the NPs are monodisperse, while predicted C_{NP} values have the distribution predicted by “ColloidNet”. It is evident that dilutions of the colloids in each growth step play a key role in the decreasing trend of both predicted and expected C_{NP} and are the ultimate cause of the decrease of $E(\lambda)$ relating to the LSPR peak as depicted in **Figure 6a** and **Figure S7a**. This is mainly because the decrease in the number of NPs cannot be compensated by the increase of the scattering component (**Figure S6a**), causing a net decrease in $E(\lambda)$. A very similar LSPR peak trend (**Figure S7b**) was observed by others as well⁸. While the decreasing C_{NP} may seem like a quirk of the seeded growth synthesis method, dilution of the colloid used as a seed solution for the next growth step is paramount in order to ensure adequate growth of NPs – otherwise, the NPs are observed to hardly grow at all, even after multiple growth steps⁴⁶. The NP size estimations in the seeded-growth synthesis often assume 100% consumption of the Ag precursor⁴⁶ while in our case around 40% provided close result in the semilogarithmic plot depicted in **Figure 6c**. A closer to unit consumption might be expected if the NP distribution was taken into account while here size distribution was not estimated at all. But the expected and experimentally obtained tendencies holds and it for the last time here confirms that the DNN predicted Ag concentrations are trustworthy.

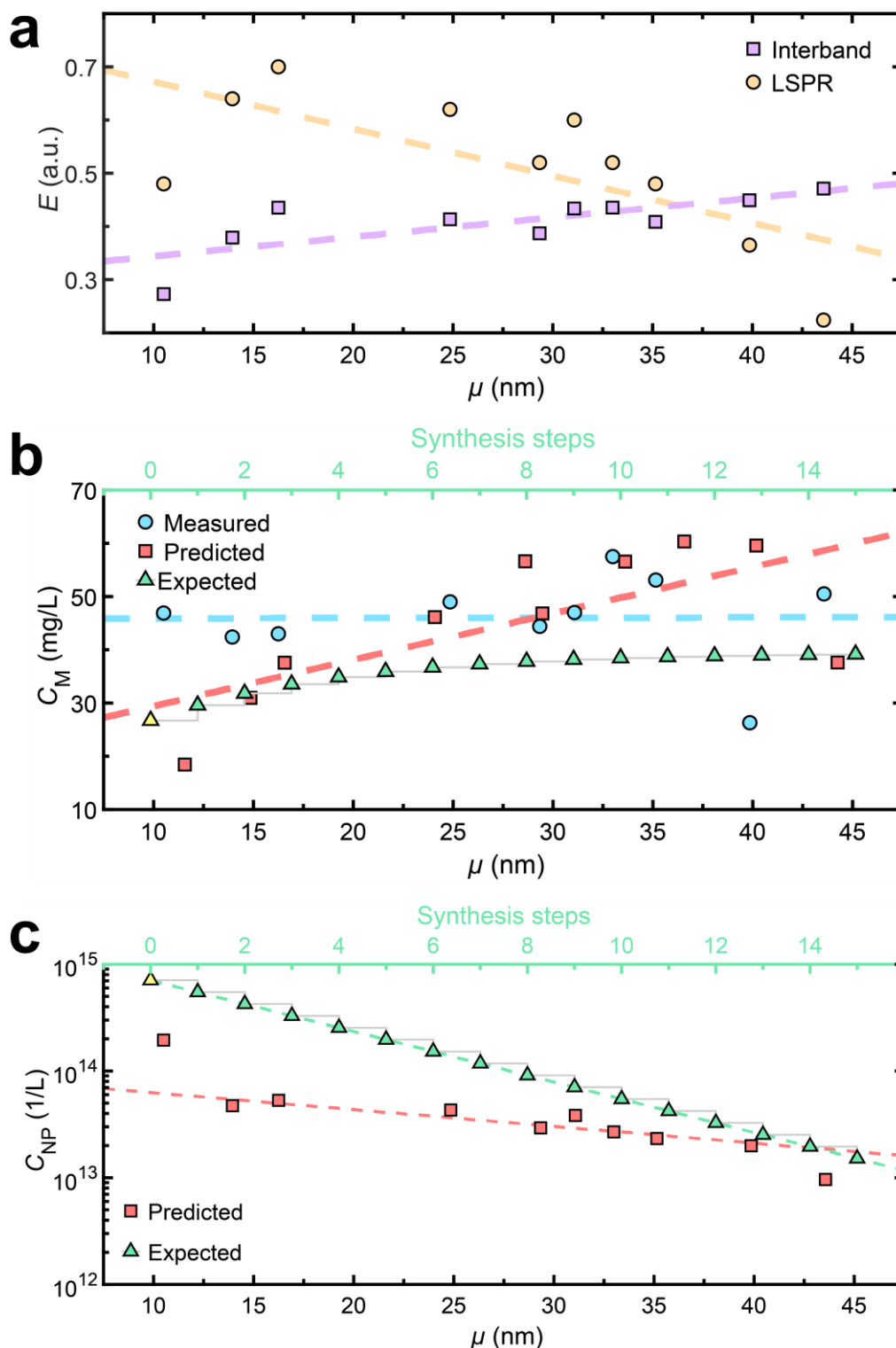


Figure 6. Concentration estimates for selected colloids of the “F” batch. (a) Trends of experimental $E(\lambda)$ of the interband transition (at 250 nm) and the LSPR peak found in **Figure S7a**. (b) Experimental vs. predicted vs. expected mass concentration (C_M). (c) “ColloidNet” predicted NP concentration (C_{NP}) vs. expected NP concentration assuming monodisperse NPs, 8 nm seeds, and 40% reaction efficiency with seed concentration $7.088 \cdot 10^{14}$ and Ag density 10.49 g/cm^3 . The first “expected” values in (b) and (c) indicated in yellow are NP seeds. Lines are guides for the reader's eye.

Conclusions

In this work, we demonstrated the capability of a deep tandem neural network system to perform accurate silver nanoparticle size distribution and concentration predictions using UV-Vis $E(\lambda)$ spectra as input.

Automated identification of the dipole peak contribution with the first DNN helped to accurately predict the NP size distribution with the second DNN in the wide range up to 150 nm retaining down to *ca.* 1.2% root mean square percentage error predicting mean nanoparticle size.

The tandem DNN predictions allowed to attain accurate silver mass and nanoparticle concentration tendencies, and their estimates throughout the seeded-growth synthesis steps as were confirmed by the growing $E(\lambda)$ at the interband and alternative atomic absorption measurement.

The $E(\lambda)$ spectra computed according to predicted nanoparticle size distributions and volume-filling fractions matched the dipolar component of the experimental $E(\lambda)$ spectra extremely well, further indicating an accurate prediction.

It was demonstrated that high accuracy and 11.3 ms per sample prediction rate can be achieved without any spectral data processing before providing it to the tandem DNN, making the proposed approach a good candidate for real-time nanoparticle colloid characterization under both laboratory and industrial conditions.

It was verified that the plasmonic NP size limit validity for the modified Maxwell-Garnett-Mie effective medium theory is much larger than previously thought and that the theory can be reliably used to estimate the size distribution of NPs based on their optical $E(\lambda)$ dipole peak up to 150 nm in radius.

Author Contributions

Conceptualization and methodology – T.K., T.T.; Investigation – T.K., N.K., A.T., L.V.; Data analysis – T.K., N.K., A.T., T.T.; Writing – original draft T.K., T.T., Writing – review & editing; T.K., N.K., A.T., S.T., T.T., Supervision – T.T.; Funding acquisition – T.T.

Funding

This project has received funding from the Research Council of Lithuania (LMTLT), agreement no S-MIP-23-93.

Notes

The authors declare no competing financial interest.

Acknowledgements

A special thanks go to Dr. Karine Mougin from the Institut de Science des Matériaux de Mulhouse IS2M UMR for the technical assistance with the DLS measurements.

Associated Content

The supporting information is available free of charge on the ACS Publications website at DOI:

xx.xxxx/xxxxx.xxxxxxx.

Details of Ag NP synthesis; Volumes and concentrations of Ag NP precursors (Table S1); description of SEM micrograph analysis; example of SEM analysis (Figure S1); description of DLS measurements; description of XRD measurements; description of MMGM effective medium; parametrization of training and validation

data; ranges of data parameters (Table S2); strategy of hyperparameter optimization; list of hyperparameters (Table S3); discussion about statistics of experimental data; statistical parameters of experimental data (Figure S2); discussion about DLS results; DLS results (Figure S3); discussion about XRD results; XRD results (Figure S4); XRD diffractogram parameters (Table S4); XRD crystallite size analysis (Table S5); discussion about the MMGM medium; comparison of MMGM and Mie theory results (Figure S5); comparison of scattering and absorbance (Figure S6); discussion about prediction of concentration; LSPR amplitude trends (PDF).

References

- (1) Khan, I.; Saeed, K.; Khan, I. Nanoparticles: Properties, Applications and Toxicities. *Arabian J. Chem.* **2019**, *12* (7), 908–931. <https://doi.org/10.1016/j.arabjc.2017.05.011>.
- (2) Vert, M.; Doi, Y.; Hellwich, K.-H.; Hess, M.; Hodge, P.; Kubisa, P.; Rinaudo, M.; Schué, F. Terminology for Biorelated Polymers and Applications (IUPAC Recommendations 2012). *Pure Appl. Chem.* **2012**, *84* (2), 377–410. <https://doi.org/10.1351/PAC-REC-10-12-04>.
- (3) Ijaz, I.; Gilani, E.; Nazir, A.; Bukhari, A. Detail Review on Chemical, Physical and Green Synthesis, Classification, Characterizations and Applications of Nanoparticles. *Green Chem. Lett. Rev.* **2020**, *13* (3), 223–245. <https://doi.org/10.1080/17518253.2020.1802517>.
- (4) Xing, T.; Sunarso, J.; Yang, W.; Yin, Y.; Glushenkov, A. M.; Li, L. H.; Howlett, P. C.; Chen, Y. Ball Milling: A Green Mechanochemical Approach for Synthesis of Nitrogen Doped Carbon Nanoparticles. *Nanoscale* **2013**, *5* (17), 7970. <https://doi.org/10.1039/c3nr02328a>.
- (5) Salavati-Niasari, M.; Davar, F.; Mir, N. Synthesis and Characterization of Metallic Copper Nanoparticles via Thermal Decomposition. *Polyhedron* **2008**, *27* (17), 3514–3518. <https://doi.org/10.1016/j.poly.2008.08.020>.
- (6) Amendola, V.; Meneghetti, M. Laser Ablation Synthesis in Solution and Size Manipulation of Noble Metal Nanoparticles. *Phys. Chem. Chem. Phys.* **2009**, *11* (20), 3805. <https://doi.org/10.1039/b900654k>.
- (7) Orozco-Montes, V.; Caillard, A.; Brault, P.; Chamorro-Coral, W.; Bigarre, J.; Saudubois, A.; Andreatza, P.; Cuynet, S.; Baranton, S.; Coutanceau, C. Synthesis of Platinum Nanoparticles by Plasma Sputtering onto Glycerol: Effect of Argon Pressure on Their Physicochemical Properties. *J. Phys. Chem. C* **2021**, *125* (5), 3169–3179. <https://doi.org/10.1021/acs.jpcc.0c09746>.
- (8) Bastús, N. G.; Merkoçi, F.; Piella, J.; Puentes, V. Synthesis of Highly Monodisperse Citrate-Stabilized Silver Nanoparticles of up to 200 Nm: Kinetic Control and Catalytic Properties. *Chem. Mater.* **2014**, *26* (9), 2836–2846. <https://doi.org/10.1021/cm500316k>.
- (9) Hussain, M. H.; Abu Bakar, N. F.; Mustapa, A. N.; Low, K.-F.; Othman, N. H.; Adam, F. Synthesis of Various Size Gold Nanoparticles by Chemical Reduction Method with Different Solvent Polarity. *Nanoscale Res. Lett.* **2020**, *15* (1), 140. <https://doi.org/10.1186/s11671-020-03370-5>.
- (10) Yang, M.; Zhu, H.; Zheng, Y.; Zhang, C.; Luo, G.; Xu, Q.; Li, Q.; Zhang, S.; Goto, T.; Tu, R. One-Step Chemical Vapor Deposition Fabrication of Ni@NiO@graphite Nanoparticles for the Oxygen Evolution Reaction of Water Splitting. *RSC Adv.* **2022**, *12* (17), 10496–10503. <https://doi.org/10.1039/D2RA00947A>.
- (11) Granqvist, C. G. SIZE DISTRIBUTIONS FOR ULTRAFINE METAL PARTICLES. *J. Phys. Colloq.* **1977**, *38* (C2), C2-147-C2-150. <https://doi.org/10.1051/jphyscol:1977230>.
- (12) Kiss, L. B.; Söderlund, J.; Niklasson, G. A.; Granqvist, C. G. New Approach to the Origin of Lognormal Size Distributions of Nanoparticles. *Nanotechnology* **1999**, *10* (1), 25–28. <https://doi.org/10.1088/0957-4484/10/1/006>.
- (13) Amendola, V.; Pilot, R.; Frascioni, M.; Maragò, O. M.; Iatì, M. A. Surface Plasmon Resonance in Gold Nanoparticles: A Review. *J. Phys.: Condens. Matter* **2017**, *29* (20), 203002. <https://doi.org/10.1088/1361-648X/aa60f3>.
- (14) Kelly, K. L.; Coronado, E.; Zhao, L. L.; Schatz, G. C. The Optical Properties of Metal Nanoparticles: The Influence of Size, Shape, and Dielectric Environment. *J. Phys. Chem. B* **2003**, *107* (3), 668–677. <https://doi.org/10.1021/jp026731y>.

- (15) Agrawal, A.; Cho, S. H.; Zandi, O.; Ghosh, S.; Johns, R. W.; Milliron, D. J. Localized Surface Plasmon Resonance in Semiconductor Nanocrystals. *Chem. Rev.* **2018**, *118* (6), 3121–3207. <https://doi.org/10.1021/acs.chemrev.7b00613>.
- (16) Maier, S. A. *Plasmonics: Fundamentals and Applications*, Repr.; Springer: New York, NY, 2007.
- (17) Barbir, R.; Pem, B.; Kalčec, N.; Kastner, S.; Podlesnaia, K.; Csáki, A.; Fritzsche, W.; Vinković Vrček, I. Application of Localized Surface Plasmon Resonance Spectroscopy to Investigate a Nano–Bio Interface. *Langmuir* **2021**, *37* (5), 1991–2000. <https://doi.org/10.1021/acs.langmuir.0c03569>.
- (18) Koushki, E.; Mowlavi, A. A.; Hoseini, S. T. Application of Localized Surface Plasmon Resonance of Conjugated Gold Nanoparticles in Spectral Diagnosis of SARS-CoV-2: A Numerical Study. *Plasmonics* **2023**, *18* (5), 1847–1855. <https://doi.org/10.1007/s11468-023-01901-1>.
- (19) Wang, S.; Mamedova, N.; Kotov, N. A.; Chen, W.; Studer, J. Antigen/Antibody Immunocomplex from CdTe Nanoparticle Bioconjugates. *Nano Lett.* **2002**, *2* (8), 817–822. <https://doi.org/10.1021/nl0255193>.
- (20) Edelstein, R. The BARC Biosensor Applied to the Detection of Biological Warfare Agents. *Biosens. Bioelectron.* **2000**, *14* (10–11), 805–813. [https://doi.org/10.1016/S0956-5663\(99\)00054-8](https://doi.org/10.1016/S0956-5663(99)00054-8).
- (21) Pantarotto, D.; Partidos, C. D.; Hoebeke, J.; Brown, F.; Kramer, E.; Briand, J.-P.; Muller, S.; Prato, M.; Bianco, A. Immunization with Peptide-Functionalized Carbon Nanotubes Enhances Virus-Specific Neutralizing Antibody Responses. *Chem. Biol.* **2003**, *10* (10), 961–966. <https://doi.org/10.1016/j.chembiol.2003.09.011>.
- (22) Jamshed, A.; Iqbal, A.; Ali, S.; Ali, S.; . M. A Quick Review on the Applications of Nanomaterials as Adsorbents. *MOJES* **2023**, *8* (3), 86–89. <https://doi.org/10.15406/moj.2023.08.00278>.
- (23) Xiao, X.; Xu, Y.; Lv, X.; Xie, J.; Liu, J.; Yu, C. Electrochemical CO₂ Reduction on Copper Nanoparticles-Dispersed Carbon Aerogels. *J. Colloid Interface Sci.* **2019**, *545*, 1–7. <https://doi.org/10.1016/j.jcis.2019.03.005>.
- (24) Herbaut, M.; Sij, M.; Claverie, J. P. Nanomaterials-Based Water Splitting: How Far Are We from a Sustainable Solution? *ACS Appl. Nano Mater.* **2021**, *4* (2), 907–910. <https://doi.org/10.1021/acsanm.1c00246>.
- (25) Mirnaziry, S. R.; Shameli, M. A.; Yousefi, L. Design and Analysis of Multi-Layer Silicon Nanoparticle Solar Cells. *Sci. Rep.* **2022**, *12* (1), 13259. <https://doi.org/10.1038/s41598-022-17677-z>.
- (26) Shilpa, G.; Kumar, P. M.; Kumar, D. K.; Deepthi, P. R.; Sadhu, V.; Sukhdev, A.; Kakarla, R. R. Recent Advances in the Development of High Efficiency Quantum Dot Sensitized Solar Cells (QDSSCs): A Review. *Mater. Sci. Energy Technol.* **2023**, *6*, 533–546. <https://doi.org/10.1016/j.mset.2023.05.001>.
- (27) Ishida, T.; Murayama, T.; Taketoshi, A.; Haruta, M. Importance of Size and Contact Structure of Gold Nanoparticles for the Genesis of Unique Catalytic Processes. *Chem. Rev.* **2020**, *120* (2), 464–525. <https://doi.org/10.1021/acs.chemrev.9b00551>.
- (28) Hoshyar, N.; Gray, S.; Han, H.; Bao, G. The Effect of Nanoparticle Size on *in Vivo* Pharmacokinetics and Cellular Interaction. *Nanomedicine* **2016**, *11* (6), 673–692. <https://doi.org/10.2217/nnm.16.5>.
- (29) Shinkai, M.; Yanase, M.; Honda, H.; Wakabayashi, T.; Yoshida, J.; Kobayashi, T. Intracellular Hyperthermia for Cancer Using Magnetite Cationic Liposomes: *In Vitro* Study. *Jpn. J. Cancer Res.* **1996**, *87* (11), 1179–1183. <https://doi.org/10.1111/j.1349-7006.1996.tb03129.x>.
- (30) Chen, F.; Ehlerding, E. B.; Cai, W. Theranostic Nanoparticles. *J. Nucl. Med.* **2014**, *55* (12), 1919–1922. <https://doi.org/10.2967/jnumed.114.146019>.
- (31) Mäki-Lohiluoma, E.; Säkkinen, N.; Palomäki, M.; Winberg, O.; Ta, H. X.; Heikkinen, T.; Kiljunen, E.; Kauppinen, A. Use of Machine Learning in Prediction of Granule Particle Size Distribution and Tablet Tensile Strength in Commercial Pharmaceutical Manufacturing. *Int. J. Pharm.* **2021**, *609*, 121146. <https://doi.org/10.1016/j.ijpharm.2021.121146>.
- (32) Shin, C.; Choi, J.; Kwak, D.; Kim, J.; Yang, J.; Chae, S.; Kim, T. Evaluation of Size Distribution Measurement Methods for Sub-100 Nm Colloidal Silica Nanoparticles and Its Application to CMP Slurry. *ECS J. Solid State Sci. Technol.* **2019**, *8* (5), P3195–P3200. <https://doi.org/10.1149/2.0261905jss>.
- (33) Emil Kaya, E.; Kaya, O.; Alkan, G.; Gürmen, S.; Stopic, S.; Friedrich, B. New Proposal for Size and Size-Distribution Evaluation of Nanoparticles Synthesized via Ultrasonic Spray Pyrolysis Using Search Algorithm Based on Image-Processing Technique. *Materials* **2019**, *13* (1), 38. <https://doi.org/10.3390/ma13010038>.

- (34) Doak, J.; Gupta, R. K.; Manivannan, K.; Ghosh, K.; Kahol, P. K. Effect of Particle Size Distributions on Absorbance Spectra of Gold Nanoparticles. *Phys. E* **2010**, *42* (5), 1605–1609. <https://doi.org/10.1016/j.physe.2010.01.004>.
- (35) Kestens, V.; Roebben, G.; Herrmann, J.; Jämting, Å.; Coleman, V.; Minelli, C.; Clifford, C.; De Temmerman, P.-J.; Mast, J.; Junjie, L.; Babick, F.; Cölfen, H.; Emons, H. Challenges in the Size Analysis of a Silica Nanoparticle Mixture as Candidate Certified Reference Material. *J. Nanopart. Res.* **2016**, *18* (6), 171. <https://doi.org/10.1007/s11051-016-3474-2>.
- (36) Lee, H.; Kwak, D.-B.; Kim, S. C.; Pui, D. Y. H. Characterization of Colloidal Nanoparticles in Mixtures with Polydisperse and Multimodal Size Distributions Using a Particle Tracking Analysis and Electrospray-Scanning Mobility Particle Sizer. *Powder Technol.* **2019**, *355*, 18–25. <https://doi.org/10.1016/j.powtec.2019.07.029>.
- (37) Abbas, Z.; Holmberg, J. P.; Hellström, A. K.; Hagström, M.; Bergenholtz, J.; Hassellöv, M.; Ahlberg, E. Synthesis, Characterization and Particle Size Distribution of TiO₂ Colloidal Nanoparticles. *Colloids Surf., A* **2011**, *384* (1–3), 254–261. <https://doi.org/10.1016/j.colsurfa.2011.03.064>.
- (38) Varenne, F.; Devoille, L.; Makky, A.; Feltin, N.; Violleau, F.; Barratt, G.; Vauthier, C. Evaluation of the Size Distribution of a Multimodal Dispersion of Polymer Nanoparticles by Microscopy after Different Methods of Deposition. *J. Drug Delivery Sci. Technol.* **2020**, *60*, 102047. <https://doi.org/10.1016/j.jddst.2020.102047>.
- (39) Mansour, Y.; Battie, Y.; En Naciri, A.; Chaoui, N. Determination of the Size Distribution of Metallic Colloids from Extinction Spectroscopy. *Nanomaterials* **2021**, *11* (11), 2872. <https://doi.org/10.3390/nano11112872>.
- (40) Grand, J.; Auguié, B.; Le Ru, E. C. Combined Extinction and Absorption UV–Visible Spectroscopy as a Method for Revealing Shape Imperfections of Metallic Nanoparticles. *Anal. Chem.* **2019**, *91* (22), 14639–14648. <https://doi.org/10.1021/acs.analchem.9b03798>.
- (41) Battie, Y.; Resano-Garcia, A.; Chaoui, N.; Zhang, Y.; En Naciri, A. Extended Maxwell-Garnett-Mie Formulation Applied to Size Dispersion of Metallic Nanoparticles Embedded in Host Liquid Matrix. *J. Chem. Phys.* **2014**, *140* (4), 044705. <https://doi.org/10.1063/1.4862995>.
- (42) Mo, J.; Zhang, Q.; Xia, P.; Chen, H.; Lu, Z.; Yang, H.; Chen, Y.; Xiao, W.; Wang, Y.; Liu, M. Another Method to Calculate Grain Size Distribution Parameters: Using Ising Model and Particle Swarm Optimization Algorithm. *Solid State Commun.* **2022**, *356*, 114961. <https://doi.org/10.1016/j.ssc.2022.114961>.
- (43) Faílde, D.; Viqueira, J. D.; Mussa Juane, M.; Gómez, A. Using Differential Evolution to Avoid Local Minima in Variational Quantum Algorithms. *Sci. Rep.* **2023**, *13* (1), 16230. <https://doi.org/10.1038/s41598-023-43404-3>.
- (44) Tuersun, P.; Zhu, C.; Han, X.; Fang Ren, K.; Yin, Y. Light Extinction Spectrometry for Determining the Size Distribution and Concentration of Polydisperse Gold Nanospheres. *Optik* **2020**, *204*, 163676. <https://doi.org/10.1016/j.ijleo.2019.163676>.
- (45) Amendola, V.; Meneghetti, M. Size Evaluation of Gold Nanoparticles by UV–vis Spectroscopy. *J. Phys. Chem. C* **2009**, *113* (11), 4277–4285. <https://doi.org/10.1021/jp8082425>.
- (46) Bastús, N. G.; Comenge, J.; Puntès, V. Kinetically Controlled Seeded Growth Synthesis of Citrate-Stabilized Gold Nanoparticles of up to 200 Nm: Size Focusing versus Ostwald Ripening. *Langmuir* **2011**, *27* (17), 11098–11105. <https://doi.org/10.1021/la201938u>.
- (47) Hendel, T.; Wuthschick, M.; Kettemann, F.; Birnbaum, A.; Rademann, K.; Polte, J. In Situ Determination of Colloidal Gold Concentrations with UV–Vis Spectroscopy: Limitations and Perspectives. *Anal. Chem.* **2014**, *86* (22), 11115–11124. <https://doi.org/10.1021/ac502053s>.
- (48) Pugliese, R.; Regondi, S.; Marini, R. Machine Learning-Based Approach: Global Trends, Research Directions, and Regulatory Standpoints. *DSM.* **2021**, *4*, 19–29. <https://doi.org/10.1016/j.dsm.2021.12.002>.
- (49) Janiesch, C.; Zschech, P.; Heinrich, K. Machine Learning and Deep Learning. *Electron. Mark.* **2021**, *31* (3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>.
- (50) Sarker, I. H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput. Sci.* **2021**, *2* (3), 160. <https://doi.org/10.1007/s42979-021-00592-x>.

- (51) Shiratori, K.; Bishop, L. D. C.; Ostovar, B.; Baiyasi, R.; Cai, Y.-Y.; Rossky, P. J.; Landes, C. F.; Link, S. Machine-Learned Decision Trees for Predicting Gold Nanorod Sizes from Spectra. *J. Phys. Chem. C* **2021**, *125* (35), 19353–19361. <https://doi.org/10.1021/acs.jpcc.1c03937>.
- (52) Tan, E. X.; Chen, Y.; Lee, Y. H.; Leong, Y. X.; Leong, S. X.; Stanley, C. V.; Pun, C. S.; Ling, X. Y. Incorporating Plasmonic Featurization with Machine Learning to Achieve Accurate and Bidirectional Prediction of Nanoparticle Size and Size Distribution. *Nanoscale Horiz.* **2022**, *7* (6), 626–633. <https://doi.org/10.1039/D2NH00146B>.
- (53) Arzola-Flores, J. A.; González, A. L. Machine Learning for Predicting the Surface Plasmon Resonance of Perfect and Concave Gold Nanocubes. *J. Phys. Chem. C* **2020**, *124* (46), 25447–25454. <https://doi.org/10.1021/acs.jpcc.0c05995>.
- (54) Pinkus, A. Approximation Theory of the MLP Model in Neural Networks. *Acta Numer.* **1999**, *8*, 143–195. <https://doi.org/10.1017/S0962492900002919>.
- (55) Emmert-Streib, F.; Yang, Z.; Feng, H.; Tripathi, S.; Dehmer, M. An Introductory Review of Deep Learning for Prediction Models With Big Data. *Front. Artif. Intell.* **2020**, *3*, 4. <https://doi.org/10.3389/frai.2020.00004>.
- (56) Moore, J. A.; Chow, J. C. L. Recent Progress and Applications of Gold Nanotechnology in Medical Biophysics Using Artificial Intelligence and Mathematical Modeling. *Nano Ex.* **2021**, *2* (2), 022001. <https://doi.org/10.1088/2632-959X/abddd3>.
- (57) Wei, J.; Chu, X.; Sun, X.; Xu, K.; Deng, H.; Chen, J.; Wei, Z.; Lei, M. Machine Learning in Materials Science. *InfoMat* **2019**, *1* (3), 338–358. <https://doi.org/10.1002/inf2.12028>.
- (58) Revignas, D.; Amendola, V. Artificial Neural Networks Applied to Colorimetric Nanosensors: An Undergraduate Experience Tailorable from Gold Nanoparticles Synthesis to Optical Spectroscopy and Machine Learning. *J. Chem. Educ.* **2022**, *99* (5), 2112–2120. <https://doi.org/10.1021/acs.jchemed.1c01288>.
- (59) Abbasi Moud, A. Recent Advances in Utility of Artificial Intelligence towards Multiscale Colloidal Based Materials Design. *Colloid Interface Sci. Commun.* **2022**, *47*, 100595. <https://doi.org/10.1016/j.colcom.2022.100595>.
- (60) Tao, H.; Wu, T.; Aldeghi, M.; Wu, T. C.; Aspuru-Guzik, A.; Kumacheva, E. Nanoparticle Synthesis Assisted by Machine Learning. *Nat. Rev. Mater.* **2021**, *6* (8), 701–716. <https://doi.org/10.1038/s41578-021-00337-5>.
- (61) Schletz, D.; Breidung, M.; Fery, A. Validating and Utilizing Machine Learning Methods to Investigate the Impacts of Synthesis Parameters in Gold Nanoparticle Synthesis. *J. Phys. Chem. C* **2023**, *127* (2), 1117–1125. <https://doi.org/10.1021/acs.jpcc.2c07578>.
- (62) Ishimaru, A.; Kitamura, S.; Marks, R. J.; Tsang, L.; Lam, C. M.; Park, D. C. Particle-Size Distribution Determination Using Optical Sensing and Neural Networks. *Opt. Lett.* **1990**, *15* (21), 1221. <https://doi.org/10.1364/OL.15.001221>.
- (63) Nascimento, C. A. O.; Guardani, R.; Giulietti, M. Use of Neural Networks in the Analysis of Particle Size Distributions by Laser Diffraction. *Powder Technol.* **1997**, *90* (1), 89–94. [https://doi.org/10.1016/S0032-5910\(96\)03192-0](https://doi.org/10.1016/S0032-5910(96)03192-0).
- (64) Guardani, R.; Nascimento, C. A. O.; Onimaru, R. S. Use of Neural Networks in the Analysis of Particle Size Distribution by Laser Diffraction: Tests with Different Particle Systems. *Powder Technol.* **2002**, *126* (1), 42–50. [https://doi.org/10.1016/S0032-5910\(02\)00036-0](https://doi.org/10.1016/S0032-5910(02)00036-0).
- (65) Deriemaeker, L.; Finsy, R. Shape and Size Determination by Laser Diffraction: Average Aspect Ratio and Size Distributions by Volume; Feasibility of Data Analysis by Neural Networks. *Part. Part. Syst. Charact.* **2005**, *22* (1), 5–13. <https://doi.org/10.1002/ppsc.200400960>.
- (66) Glaubitz, C.; Bazzoni, A.; Ackermann-Hirschi, L.; Baraldi, L.; Haeffner, M.; Fortunatus, R.; Rothen-Rutishauser, B.; Balog, S.; Petri-Fink, A. Leveraging Machine Learning for Size and Shape Analysis of Nanoparticles: A Shortcut to Electron Microscopy. *J. Phys. Chem. C* **2024**, *128* (1), 421–427. <https://doi.org/10.1021/acs.jpcc.3c05938>.
- (67) He, J.; He, C.; Zheng, C.; Wang, Q.; Ye, J. Plasmonic Nanoparticle Simulations and Inverse Design Using Machine Learning. *Nanoscale* **2019**, *11* (37), 17444–17459. <https://doi.org/10.1039/C9NR03450A>.
- (68) Long, Y.; Ren, J.; Li, Y.; Chen, H. Inverse Design of Photonic Topological State via Machine Learning. *Appl. Phys. Lett.* **2019**, *114* (18), 181105. <https://doi.org/10.1063/1.5094838>.

- (69) Liu, D.; Tan, Y.; Khoram, E.; Yu, Z. Training Deep Neural Networks for the Inverse Design of Nanophotonic Structures. *ACS Photonics* **2018**, *5* (4), 1365–1369. <https://doi.org/10.1021/acsp Photonics.7b01377>.
- (70) Xu, X.; Sun, C.; Li, Y.; Zhao, J.; Han, J.; Huang, W. An Improved Tandem Neural Network for the Inverse Design of Nanophotonics Devices. *Opt. Commun.* **2021**, *481*, 126513. <https://doi.org/10.1016/j.optcom.2020.126513>.
- (71) Unni, R.; Yao, K.; Han, X.; Zhou, M.; Zheng, Y. A Mixture-Density-Based Tandem Optimization Network for on-Demand Inverse Design of Thin-Film High Reflectors. *Nanophotonics* **2021**, *10* (16), 4057–4065. <https://doi.org/10.1515/nanoph-2021-0392>.
- (72) *Handbook of Optical Constants of Solids*; Palik, E. D., Ghosh, G., Eds.; Academic Press: San Diego, 1998.
- (73) Daimon, M.; Masumura, A. Measurement of the Refractive Index of Distilled Water from the Near-Infrared Region to the Ultraviolet Region. *Appl. Opt.* **2007**, *46* (18), 3811. <https://doi.org/10.1364/AO.46.003811>.
- (74) Mackay, T. G.; Lakhtakia, A. *The Transfer-Matrix Method in Electromagnetics and Optics*; Synthesis Lectures on Electromagnetics; Springer International Publishing: Cham, 2020. <https://doi.org/10.1007/978-3-031-02022-3>.
- (75) Rumpf, R. C. IMPROVED FORMULATION OF SCATTERING MATRICES FOR SEMI-ANALYTICAL METHODS THAT IS CONSISTENT WITH CONVENTION. *Prog. Electromagn. Res. B* **2011**, *35*, 241–261. <https://doi.org/10.2528/PIERB11083107>.
- (76) Werdehausen, D.; Staude, I.; Burger, S.; Petschulat, J.; Scharf, T.; Pertsch, T.; Decker, M. Design Rules for Customizable Optical Materials Based on Nanocomposites. *Opt. Mater. Express* **2018**, *8* (11), 3456. <https://doi.org/10.1364/OME.8.003456>.
- (77) Basodi, S.; Ji, C.; Zhang, H.; Pan, Y. Gradient Amplification: An Efficient Way to Train Deep Neural Networks. *Big Data Min. Anal.* **2020**, *3* (3), 196–207. <https://doi.org/10.26599/BDMA.2020.9020004>.
- (78) *List of Deep Learning Layers - MATLAB & Simulink - MathWorks Nordic*. <https://se.mathworks.com/help/deeplearning/ug/list-of-deep-learning-layers.html> (accessed 2024-03-27).
- (79) Yang, L.; Shami, A. On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice. *Neurocomputing* **2020**, *415*, 295–316. <https://doi.org/10.1016/j.neucom.2020.07.061>.
- (80) DeCastro-García, N.; Muñoz Castañeda, Á. L.; Escudero García, D.; Carriegos, M. V. Effect of the Sampling of a Dataset in the Hyperparameter Optimization Phase over the Efficiency of a Machine Learning Algorithm. *Complexity* **2019**, *2019*, 1–16. <https://doi.org/10.1155/2019/6278908>.
- (81) *Options for training deep learning neural network - MATLAB trainingOptions - MathWorks Nordic*. <https://se.mathworks.com/help/deeplearning/ref/trainingoptions.html> (accessed 2024-03-27).
- (82) Teulon, J.-M.; Godon, C.; Chantalat, L.; Moriscot, C.; Cambedouzou, J.; Odorico, M.; Ravaux, J.; Podor, R.; Gerdil, A.; Habert, A.; Herlin-Boime, N.; Chen, S.; Pellequer, J.-L. On the Operational Aspects of Measuring Nanoparticle Sizes. *Nanomaterials* **2018**, *9* (1), 18. <https://doi.org/10.3390/nano9010018>.
- (83) Wu, K.-J.; Torrente-Murciano, L. Continuous Synthesis of Tuneable Sized Silver Nanoparticles via a Tandem Seed-Mediated Method in Coiled Flow Inverter Reactors. *React. Chem. Eng.* **2018**, *3* (3), 267–276. <https://doi.org/10.1039/C7RE00194K>.
- (84) Togashi, T.; Tsuchida, K.; Soma, S.; Nozawa, R.; Matsui, J.; Kanaizuka, K.; Kurihara, M. Size-Tunable Continuous-Seed-Mediated Growth of Silver Nanoparticles in Alkylamine Mixture via the Stepwise Thermal Decomposition of Silver Oxalate. *Chem. Mater.* **2020**, *32* (21), 9363–9370. <https://doi.org/10.1021/acs.chemmater.0c03303>.
- (85) Zong, R.; Wang, X.; Shi, S.; Zhu, Y. Kinetically Controlled Seed-Mediated Growth of Narrow Dispersed Silver Nanoparticles up to 120 Nm: Secondary Nucleation, Size Focusing, and Ostwald Ripening. *Phys. Chem. Chem. Phys.* **2014**, *16* (9), 4236. <https://doi.org/10.1039/c3cp54846e>.
- (86) Pyatenko, A.; Yamaguchi, M.; Suzuki, M. Synthesis of Spherical Silver Nanoparticles with Controllable Sizes in Aqueous Solutions. *J. Phys. Chem. C* **2007**, *111* (22), 7910–7917. <https://doi.org/10.1021/jp071080x>.

- (87) Teranishi, T.; Eguchi, M.; Kanehara, M.; Gwo, S. Controlled Localized Surface Plasmon Resonance Wavelength for Conductive Nanoparticles over the Ultraviolet to Near-Infrared Region. *J. Mater. Chem.* **2011**, *21* (28), 10238. <https://doi.org/10.1039/c0jm04545d>.
- (88) Rashidian Vaziri, M. R.; Omidvar, A.; Jaleh, B.; Partovi Shabestari, N. Investigating the Extrinsic Size Effect of Palladium and Gold Spherical Nanoparticles. *Opt. Mater. (Amsterdam, Neth.)* **2017**, *64*, 413–420. <https://doi.org/10.1016/j.optmat.2017.01.014>.
- (89) Raza, S.; Kadkhodazadeh, S.; Christensen, T.; Di Vece, M.; Wubs, M.; Mortensen, N. A.; Stenger, N. Multipole Plasmons and Their Disappearance in Few-Nanometre Silver Nanoparticles. *Nat. Commun.* **2015**, *6* (1), 8788. <https://doi.org/10.1038/ncomms9788>.
- (90) Shcherbakov, M. V.; Brebels, A.; Shcherbakova, N. L.; Tyukov, A. P.; Janovsky, T. A.; Kamaev, V. A. A Survey of Forecast Error Measures. *World Appl. Sci. J.* **2013**, *24*, 171–176. <https://doi.org/10.5829/idosi.wasj.2013.24.itmies.80032>.
- (91) Meyer, F. Topographic Distance and Watershed Lines. *Signal Process.* **1994**, *38* (1), 113–125. [https://doi.org/10.1016/0165-1684\(94\)90060-4](https://doi.org/10.1016/0165-1684(94)90060-4).
- (92) Haiss, W.; Thanh, N. T. K.; Aveyard, J.; Fernig, D. G. Determination of Size and Concentration of Gold Nanoparticles from UV–Vis Spectra. *Anal. Chem.* **2007**, *79* (11), 4215–4221. <https://doi.org/10.1021/ac0702084>.
- (93) Petit, C.; Lixon, P.; Pileni, M. P. In Situ Synthesis of Silver Nanocluster in AOT Reverse Micelles. *J. Phys. Chem.* **1993**, *97* (49), 12974–12983. <https://doi.org/10.1021/j100151a054>.
- (94) Ershov, V.; Tarasova, N.; Ershov, B. Evolution of Electronic State and Properties of Silver Nanoparticles during Their Formation in Aqueous Solution. *Int. J. Mol. Sci.* **2021**, *22* (19), 10673. <https://doi.org/10.3390/ijms221910673>.

Supplementary Information

Deep Learning Methods for Colloidal Silver Nanoparticle Concentration and Size Distribution Determination from UV-Vis Extinction Spectra

Tomas Klinavičius^{a,*}, Nadzeya Khinevich^a, Asta Tamulevičienė^{a,b}, Loic Vidal^c, Sigitas Tamulevičius^{a,b}, and Tomas Tamulevičius^{a,b,*}

^aInstitute of Materials Science of Kaunas University of Technology, K. Baršausko St. 59, LT-51423, Kaunas, Lithuania

^bDepartment of Physics, Kaunas University of Technology, Studentų St. 50, LT-51368, Kaunas, Lithuania

^cInstitut de Science des Matériaux de Mulhouse IS2M UMR 7361, 15 rue Jean Starcky, F 68100 Mulhouse, France

*Corresponding authors

T. Klinavičius: tomas.klinavicius@ktu.lt

T. Tamulevičius: tomas.tamulevicius@ktu.lt

Figure S1 – key steps of the SEM micrograph analysis algorithm.

Figure S2 – statistical size parameters of synthesized colloids.

Figure S3 – dynamic light scattering results of nanoparticles.

Figure S4 – X-ray diffraction results of nanoparticles.

Figure S5 – comparison of Mie theory and effective medium theory for large nanoparticles.

Figure S6 – comparison of scattering and absorbance for different monodisperse nanoparticle sizes.

Figure S7 – extinction magnitude trends of different mean sizes nanoparticle colloids.

Table S1 – synthesis conditions of nanoparticles and statistical size parameters of synthesized colloids.

Table S2 – parameter ranges used for generation of training and validation data for deep neural networks.

Table S3 – ranges and optimal values of hyperparameters for both deep neural networks.

Table S4 – X-ray diffraction results of nanoparticles.

Table S5 – crystallinity of nanoparticles obtained from X-ray diffraction analysis.

S1. Supplementary Methods

S1.1 Detailed Description of Ag NP Synthesis

First, seeds were generated in the solution. The aqueous solution containing defined volumes of trisodium citrate (TC, with concentration C_{TC}) and tannic acid (TA, C_{TA}), was heated to 100°C and maintained in that temperature for 15 minutes under vigorous stirring. After that, a defined volume of silver nitrate (SN, C_{SN}) was injected. Then the NP seeds were left to form.

Second, NPs were synthesized by growing the seeds. NPs were grown by mixing a defined volume of seed solution with 16.5 ml of de-ionized water and heated to 90°C, then defined volumes of C_{TC} , C_{TA} and C_{SN} were sequentially injected, synthesis lasted for 30 minutes. The produced Ag NP colloid was used as a seed colloid for the next synthesis and the growth process was repeated again (for each colloid the growth process was the same). Reagent volumes and concentrations used for the synthesis of both seeds and NPs are summarized in **Table S1** where seeds are denoted as the “0” samples, other samples are sequential according to their order of growth with letters “A”-“F” identifying different batches.

Table S1. Volumes and concentrations of materials used in seed growth (0) and further synthesis steps of different Ag NP batches (“A”-“F”) along with corresponding statistical parameters of pre-produced NP size distributions. TA – tannic acid, TC – trisodium citrate, SN – silver nitrate along with statistical (μ – mean NP radius, σ – standard deviation of NP radius, R^2 – goodness of fit) and ordering characteristics of empirical UV-Vis spectra (**Figure 2d**) as well as NP radius distributions (**Figure 2e**)

Batch	No.	Synthesis components					Size distribution parameters			R ² of fit	Sample order
		Water (ml)	TA (ml, mM)	TC (ml, mM)	SN (ml, mM)	Previous Solution (ml)	μ (nm)	σ (nm)	σ (% of μ)		
A	A0	76	4, 2.5	20, 25	1, 25	-	-	-	-	-	-
	A1	16.5	1.5, 2.5	0.5, 25		19.5	12.8	3.1	24.2	0.90	2
	A2						15.7	5.3	33.7	0.96	5
	A3						24.1	5.9	24.4	0.94	10
	A4						27.5	5.8	21.1	0.93	14
	A5						30.9	3.2	10.4	0.93	19
	A6						35.7	4.9	13.7	0.93	25
B	B0	60	20, 2.5	20, 25	1, 25	-	-	-	-	-	-
	B1	16.5	1.5, 2.5	0.5, 25		19.5	19.5	4.5	23.1	0.94	7
	B2						21.3	6.0	28.2	0.97	8
	B3						26.6	6.7	25.2	0.96	12
	B4						29.6	4.4	14.9	0.98	17
	B5						37.3	5.6	15.0	0.95	27
	B6						41.0	5.5	13.4	0.96	33
C	C0	79	1, 100	20, 25	1, 25	-	-	-	-	-	-
	C1	16.5	1.5, 2.5	0.5, 25		19.5	26.7	3.5	13.1	0.96	13
	C2						31.1	6.7	21.5	0.97	21
	C3						33.4	8.0	24.0	0.96	23
	C4						40.2	7.4	18.4	0.97	32
	C5						49.9	9.9	19.8	0.90	36
	C6						58.8	10.3	17.5	0.91	39
D	D0	78.5	1.5, 100	20, 25	1, 25	-	-	-	-	-	-
	D1	16.5	1.5,	0.5,		19.5	29.8	6.0	20.1	0.98	18

	D2		2.5	25			44.0	9.6	21.8	0.95	35
	D3						54.3	11.3	20.8	0.92	38
	D4						67.2	12.1	18.0	0.95	41
	D5						86.3	13.5	15.6	0.93	44
	D6						107.6	18.1	16.8	0.87	45
E	E0	77.5	2.5, 100	20, 25	1, 25	-	-	-	-	-	-
	E1	16.5	1.5, 2.5	0.5, 25		19.5	28.9	7.2	24.9	0.97	15
	E2						39.7	10.0	25.2	0.91	29
	E3						51.1	9.1	17.8	0.93	37
	E4						60.3	9.7	16.1	0.91	40
	E5						75.3	10.7	14.2	0.90	42
	E6						78.8	13.6	17.3	0.88	43
F	F0	79.5	0.5, 2.5	20, 25	1, 25	-	-	-	-	-	-
	F1	16.5	0.25, 2.5	0.1, 25	0.25, 25	58.5	10.5	3.1	29.5	0.95	1
	F2						13.6	5.0	36.8	0.94	3
	F3						13.9	4.8	34.5	0.94	4
	F4						16.2	5.0	30.9	0.98	6
	F5						21.7	5.7	26.3	0.97	9
	F6						24.8	4.4	17.7	0.95	11
	F7						29.3	2.6	8.97	0.93	16
	F8						28.1	3.5	12.5	0.98	20
	F9						33.0	3.6	10.9	0.98	22
	F10						35.1	2.9	8.3	0.99	24
	F11						35.8	2.8	7.8	0.99	26
	F12						39.3	2.4	6.1	0.99	28
	F13						39.9	3.8	9.5	0.98	31
	F14						39.7	2.7	6.8	0.99	30
	F15						43.6	2.8	6.4	0.99	34

S1.2 Detailed Description of SEM Micrograph Analysis

An adaptive thresholding algorithm¹ was used to account for the inhomogeneity of image brightness. The scale of each micrograph was determined by automatically determining the length and reading the text (numeric value and units) of the scale bar – this allowed to avoid user error in determining the scale of the micrograph. Binarization of the image was performed according to this local threshold value. In order to eliminate possible binarization artefacts due to uneven brightness of spherical-like NPs resulting in donut-shaped image regions with holes they were filled in². The watershed segmentation employing the Fernand Meyer algorithm³ was used to separate NP regions that were stuck together after binarization in order to estimate the size of a single NP instead of several or an entire aggregate, this increased the reliability of the radius data for statistics. In order to determine the effective radius of NPs, their post-segmentation area in the image was equated to the area of a circle with the effective radius. Radii smaller than 1% of the scale bar were discarded from the statistical analysis in each micrograph because their estimates were found to be unreliable due to the SEM resolution limit. Each detected NP is numbered according to its detection order, this allows to potential outlier identification. NPs touching the boundary of the micrographs were removed from the analysis results in order to have only particles of full size. Probability density functions of effective radius distributions and statistical parameters of said distributions were computed from the effective radius data. The key steps of the used image analysis algorithm are depicted in **Figure S1** by using a fragment of a single SEM micrograph.

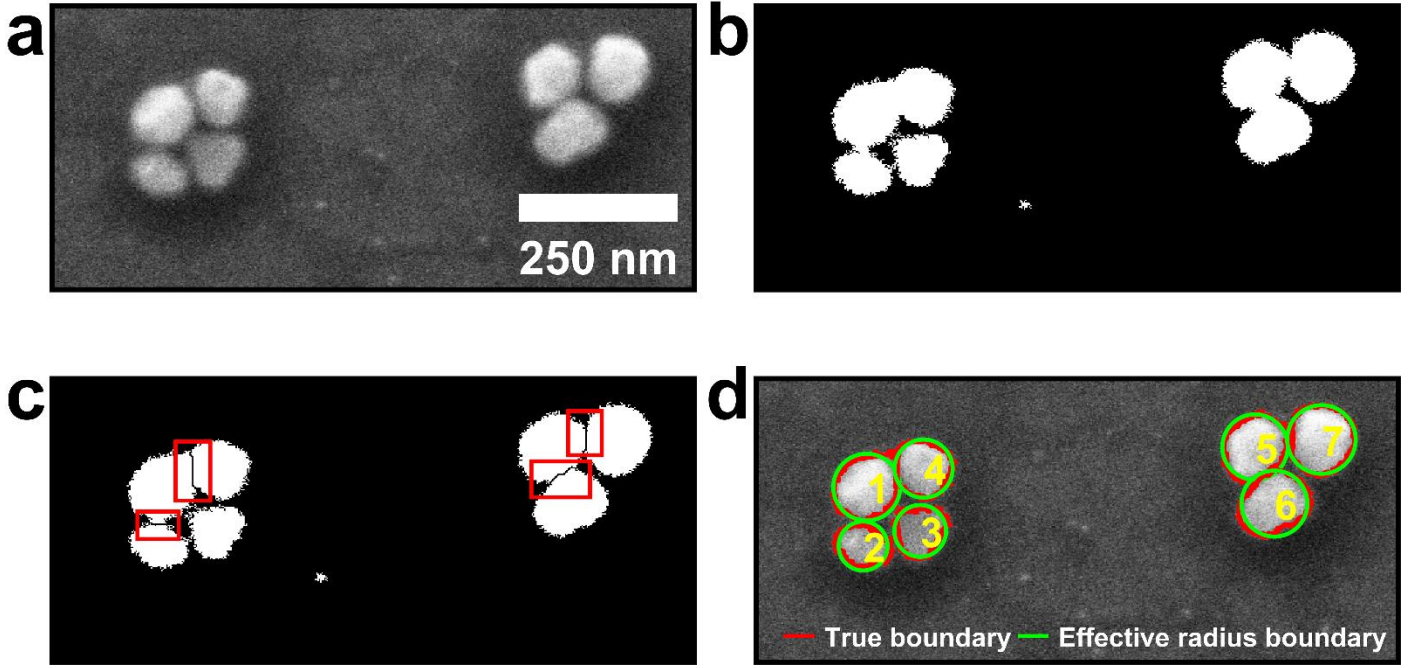


Figure S1. Key steps of the SEM micrograph analysis algorithm. (a) Determination of scale of the input image, performed directly from the scale bar, the scale bar here was added only for reference. (b) Binarization of the image based on adaptive thresholding, in order to separate the NPs from the background. (c) Segmentation of NP clusters into individual NPs using watershed transformation. (d) The fully processed output image with true and effective NP boundaries along with the NP number.

S1.3 Detailed Description of NP Size Measurement by Dynamic Light Scattering

The hydrodynamic particle radius and its distribution were measured using dynamic light scattering (DLS) technique. Measurements were carried out on a “Zeta Sizer Ultra” particle size analyzer (Malvern Panalytical). The scattering angle was fixed at 90°. 2 ml of colloidal solution was used for the measurement in a disposable cuvette. Before the measurement, the colloidal solution was agitated using an ultrasound bath to avoid aggregation of the NPs in the measuring solution.

S1.4 Detailed Description of NP Crystallinity Measurements by X-Ray Diffractometry

X-ray diffractometer “D8 Discover” (Bruker AXS GmbH) was used to perform X-ray diffraction (XRD) measurements and estimate the crystallinity of Ag NPs. Cu K α_1 ($\lambda = 1.5406$ Å) radiation was used as the X-ray source, utilizing a parallel beam-focusing geometry. XRD spectra were measured in the 30°-90° (θ -2 θ) with a step size of 0.012° using a silicon strip LynxEye detector. Data was analyzed using the DIFFRAC.EVA software (Bruker). The crystallite size (C) was computed using the Scherrer equation⁴.

S1.5 Detailed Description of Modified Maxwell-Garnett-Mie Effective Medium Theory

The permittivity of materials composed of weakly interacting sub-wavelength particles disorderly dispersed in a host medium can be described by the Maxwell-Garnett effective medium theory (EMT)⁵. If the particles are spherical or almost spherical, such materials can be more accurately described by Maxwell-Garnett-Mie EMT⁶, because this theory takes into account the polarizability of the particle derived by Mie theory. However, this model also assumes monodisperse inclusions. In general, such materials are composed of particles of various sizes. The modified Maxwell-Garnett-Mie (MMGM)⁷ EMT is able to account for the probability distribution of particles dispersed within a defined size range. For noble metal NPs the permittivity can be computed from:

$$\frac{\epsilon_{eff} - \epsilon_m}{\epsilon_{eff} + 2\epsilon_m} = \frac{3i\lambda^3}{16\pi^3\epsilon_m^{3/2}} \frac{F}{\mu} \int_{R_{min}}^{R_{max}} P(R) a_1(\epsilon_{NP}, R) dR \quad (S1)$$

Here ε_{eff} is the effective permittivity, ε_m is the permittivity of the host medium, λ is the vacuum wavelength of light, F is the volume filling fraction of the particles in the host medium, μ is the mean size of the particles that are distributed according to the probability density law $P(R)$, $a_1(\varepsilon_{np}, R)$ is the first electric Mie coefficient (which accounts for NP permittivity and size) while R_{min} and R_{max} are the lower and upper bounds of the NP distribution range.

Y. Battie *et al.* in⁷ claim that MMGM theory can only be used for noble metal NPs with a clear-cut condition of their radius being <25 nm due to the concept of an effective permittivity having a physical meaning only if scattering does not supersede absorption. However, the formula (S1) does not have any built-in limitation for using it outside the designated size range. Moreover, D. Werdehausen *et al.* in⁸ have shown that the transition from EMT being valid to being invalid is in fact gradual. They have shown the partial validity of EMT for NP radius values at least double that of what is claimed in⁷, making the limit where it is no longer applicable somewhat blurred. Therefore, it cannot be unambiguously stated that EMT is invalid for NPs larger than 25 nm.

S1.6 Log-normal Distributions and Extinction Spectra Parameters for Generation of Training and Validation Data

Both *full* and *dipolar* extinction ($E(\lambda)$) spectra ranged from 350 nm to 900 nm with an increment of 2 nm. Intermediate boundary values of $FWHM_L$ were determined by linearly interpolating between the boundary values of $FWHM_L$ for the minimum and maximum values of M_L using the equations (eq. S2) and (eq. S3) respectively. These equations were derived using the equation for a line going through 2 points.

$$FWHM_{L, lower} = 0.0762M_L + 1.6190 \quad (S2)$$

$$FWHM_{L, upper} = 0.1905M_L + 19.0476 \quad (S3)$$

M_L and $FWHM_L$ are not commonly used to define the log-normal distribution – even an ontology dedicated to distributions offers no parametrization based on M_L and $FWHM_L$ ⁹. Therefore, a relation between (M_L , $FWHM_L$) and (μ_L , σ_L) was derived using the symbolic toolbox of “MATLAB”. The transformation is defined by equations (S4) and (S5):

$$\mu_L = \ln(M_L) + \frac{\left(\ln(M_L) - \ln\left(FWHM_L + \sqrt{4M_L^2 + FWHM_L^2}\right) + \ln(2)\right)^2}{2 \ln(2)} \quad (S4)$$

$$\sigma_L = -\frac{\ln(M_L) - \ln\left(FWHM_L + \sqrt{4M_L^2 + FWHM_L^2}\right) + \ln(2)}{\sqrt{2 \ln(2)}} \quad (S5)$$

These canonical log-normal distribution parameters μ_L and σ_L were then used to originate the log-normal distributions using “MATLAB”. As NPs grow their size distribution tends to become more dispersed¹⁰. To reflect this, for the largest value of M_L the $FWHM_L$ range was selected to be larger. Ranges and iteration steps of both (M_L , $FWHM_L$) and (μ_L , σ_L) are provided in **Table S2**.

Table S2. Parameters for generating training and validation data. R – radius over which the log-normal distributions are defined, M_L – mean of the log-normal distribution, $FWHM_L$ – full width at half maximum of the log-normal distribution

Dataset	Parameter	Lower bound	Upper bound	Step
Log-normal distribution range	R	1 nm	150 nm	1 nm
Training	M_L	5 nm	110 nm	1 nm
	$FWHM_L$ ($M_L = 5$ nm)	2 nm	20 nm	1 nm
	$FWHM_L$ ($M_L = 110$ nm)	10 nm	40 nm	1 nm

	FWHM _L (5 nm < M_L < 110 nm)	Equation (S2)	Equation (S3)	1 nm
	Amplitude of MMGM extinction	0.1	2.5	0.01
Validation	M_L	5 nm	110 nm	Random
	FWHM _L ($M_L = 5$ nm)	2 nm	20 nm	Random
	FWHM _L ($M_L = 110$ nm)	10 nm	40 nm	Random
	FWHM _L (5 nm < M_L < 110 nm)	Equation (S2)	Equation (S3)	Random
	Amplitude of MMGM extinction	0.1	2.5	Random

S1.7 Use of Training and Validation Data to Set Optimal Hyperparameters

For each DNN, training data and validation data consisted of inputs and outputs correspondingly coupled into data pairs. Training was conducted using a minibatch gradient descent process for both DNNs, therefore a number of data pairs were used in each step of both training and validation for both DNNs. During hyperparameter optimization (HPO), and subsequently – the final training using optimal hyperparameters, input members of data pairs were given to the DNN and the DNN provided an output based on those inputs. Then, outputs of the DNN were compared to the output members of the data pairs, and a root-mean-squared error was computed for adjusting the weights and biases of the DNNs using the root mean square propagation algorithm for backpropagation.

Training data was used to train the DNNs, while validation data was used to monitor training performance during HPO. Testing data was used to evaluate performance of the final, post-HPO, model, trained using optimal hyperparameter values and a complete training dataset. HPO parameters along with user-set parameters are provided in **Table S3**.

For “DipoleNet”, training and validation data pairs consisted of numerically generated *full* $E(\lambda)$ spectra for the inputs and numerically generated *dipolar* $E(\lambda)$ spectra for the outputs. Mie theory was used to generate *full* spectra and TMM using MMGM medium was used to generate *dipole* spectra. Testing data inputs consisted of measured $E(\lambda)$ spectra for the inputs and best-fit dipole spectra, computed using TMM with MMGM medium, for the outputs.

Similarly, training and validation data for “ColloidNet” were composed of numerically generated *dipolar* spectra and corresponding numerical NP radius distributions with their F values. TMM using MMGM medium was used to originate the *dipole* spectra, and the size distributions with their F values were numerically generated and used to compute the permittivity of the MMGM medium⁷. For inputs of testing data, outputs of “DipoleNet” were used once it was confirmed that “DipoleNet” has achieved optimal performance. Meanwhile, outputs of testing data were experimentally derived NP distributions which correspond to measured spectra used for testing inputs of “DipoleNet” along with their measured concentrations. Separately using both DNNs one after the other completes the functionality of the “tandem” DNN, allowing the DNN system to take measured spectra as inputs and output predictions of what the NP distribution and its NP concentration should be.

Table S3. Optimizable and user-set hyperparameters for both DNNs and their optimal values. Range bounds of user-set parameters are denoted as “-”.

Network	Hyperparameter	Lower bound	Upper bound	Optimal value
DipoleNet	Size of the input layer	-	-	276
	Number of neurons in each hidden layer	276	1500	1486
	Size of the output layer	-	-	151
	Number of hidden layer blocks	1	5	3
	Leaky ReLU scale	-	-	0.01
	Size of mini-batch	1	100	1
	Initial learning rate	$1 \cdot 10^{-5}$	$1 \cdot 10^{-2}$	$2.5218 \cdot 10^{-5}$
	Learning rate multiplier	0.1	0.9	0.7235

	Learning rate drop period	3	15	11
	Number of training epochs for HPO	-	-	100
	Validation frequency in iterations for HPO	-	-	50
ColloidNet	Size of the input layer	-	-	276
	Number of neurons in each hidden layer	276	1500	995
	Size of the output layer	-	-	151
	Number of hidden layer blocks	1	5	2
	Leaky ReLU scale	-	-	0.01
	Size of mini-batch	1	50	12
	Normalization constant (epsilon)	-	-	10^{-5}
	Dropout probability	-	-	0.5
	Magnitude of L2 regularization	-	-	$1 \cdot 10^{-4}$
	Initial learning rate	$1 \cdot 10^{-5}$	$1 \cdot 10^{-2}$	$6.0203 \cdot 10^{-4}$
	Learning rate multiplier	0.1	0.9	0.5276
	Learning rate drop period	3	15	6
	Number of training epochs for HPO	-	-	100
	Validation frequency in iterations for HPO	-	-	50

S2. Supplementary Results & Discussion

S2.1 Statistical Parameters of Experimental NP Size Distributions

The experimental NP size distribution parameters extracted from the scanning electron microscopy analysis are depicted in **Figure S2a**. The standard deviation (σ) of NP size tends to increase when their mean radius (μ) increases. The exceptions are the most monodisperse batches “A” (red), “B” (Orange), and “F” (blue), where this trend is not observed. In **Figure S2b**, when σ is expressed in units relative to μ , this trend truly becomes noticeable – for each batch with progressively larger μ , the decrease of σ becomes less pronounced, while for largely monodisperse NPs it continues to decrease. Based on this, the monodisperse batches were grouped together and fitted with a least-squares purple line, while the polydisperse batches were grouped together and fitted with a gray line to better illustrate the trend.

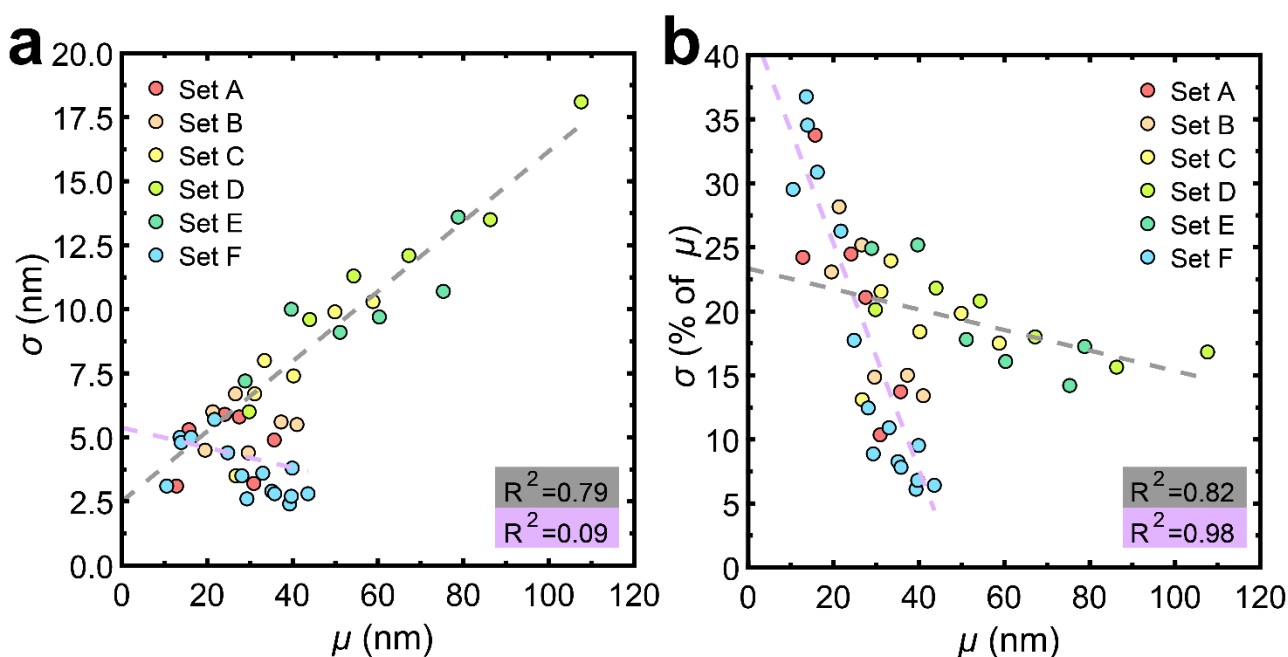


Figure S2. Mean (μ) and standard deviation (σ) of experimental NP size distributions. (a) Dependence of standard deviation on mean size for different batches of NPs. (b) Dependence of standard deviation, expressed as a percentage of the mean, on mean size for different batches of NPs. Corresponding data is available in

Table S1. Truncated lines represent linear regression but in (a) it resembles guide for the reader's eye as R^2 values are low for sets “A” and “F” in (a).

S2.2 Measurement Results of Ag NP Size Distribution by Dynamic Light Scattering

It was observed that for investigated samples DLS tends to significantly underestimate the mode of the NP size distribution while exaggerating the FWHM of the distribution, making it seem like the NPs are more polydisperse than they are according to SEM micrograph analysis, as evident in **Figure S3**. That is because the output radii of DLS are logarithmically spaced due to numerical methods used by the analysis software¹¹. Similar issues when analyzing gold NPs were described by others¹², therefore NP size characterization by SEM was chosen for our work.

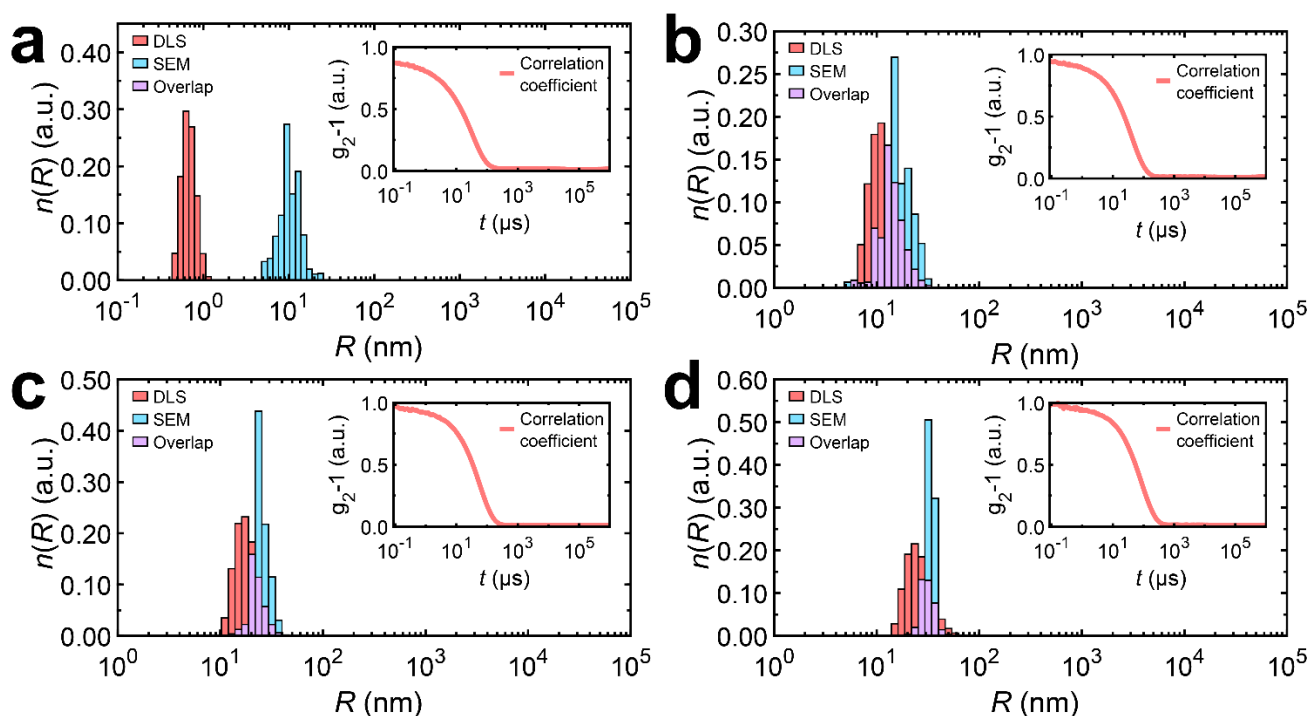


Figure S3. Comparison of Ag NP size distributions obtained by DLS and SEM with corresponding DLS correlogram. (a) Sample “F1”. (b) Sample “F4”. (c) Sample “F6”. (d) Sample “F9”.

S2.3 Results of Crystallinity Measurements of Ag NPs by X-Ray Diffractometry

XRD diffractograms were measured for selected samples of Ag NPs, they are displayed in **Figure S4a**. The diffraction peaks were identified to be the crystal planes (111), (200), (220), (311), and (222) respectively, indicating a face-centered cubic lattice, which is characteristic of pure Ag. The peaks can be seen to correspond fairly well (**Table S4**) to the standard powder diffraction card for Ag of the Joint Committee on Powder Diffraction Standards (JCPDS 04-0783). Estimates of crystal lattice parameters were performed according to¹³. The weak intensity of the peaks shows poor crystallinity of the synthesized Ag NPs, indicating they are likely polycrystalline¹⁴.

Crystallite size with its standard deviation, mean radius of the investigated sample with its standard deviation and also the ratio of crystallite size to the mean radius with the ratio's standard deviation are presented in **Table S5**. The standard deviation of the ratio is computed in accordance with the rules of error propagation¹⁵. Overall, **Figure S4b** shows a clear trend of crystallite size increasing along with the NP radius. Conversely, **Figure S4c** shows a reverse trend, proving that NPs end up becoming more polycrystalline as they grow. This is due to the synthesis method used, as other authors¹⁴ obtained a similar trend.

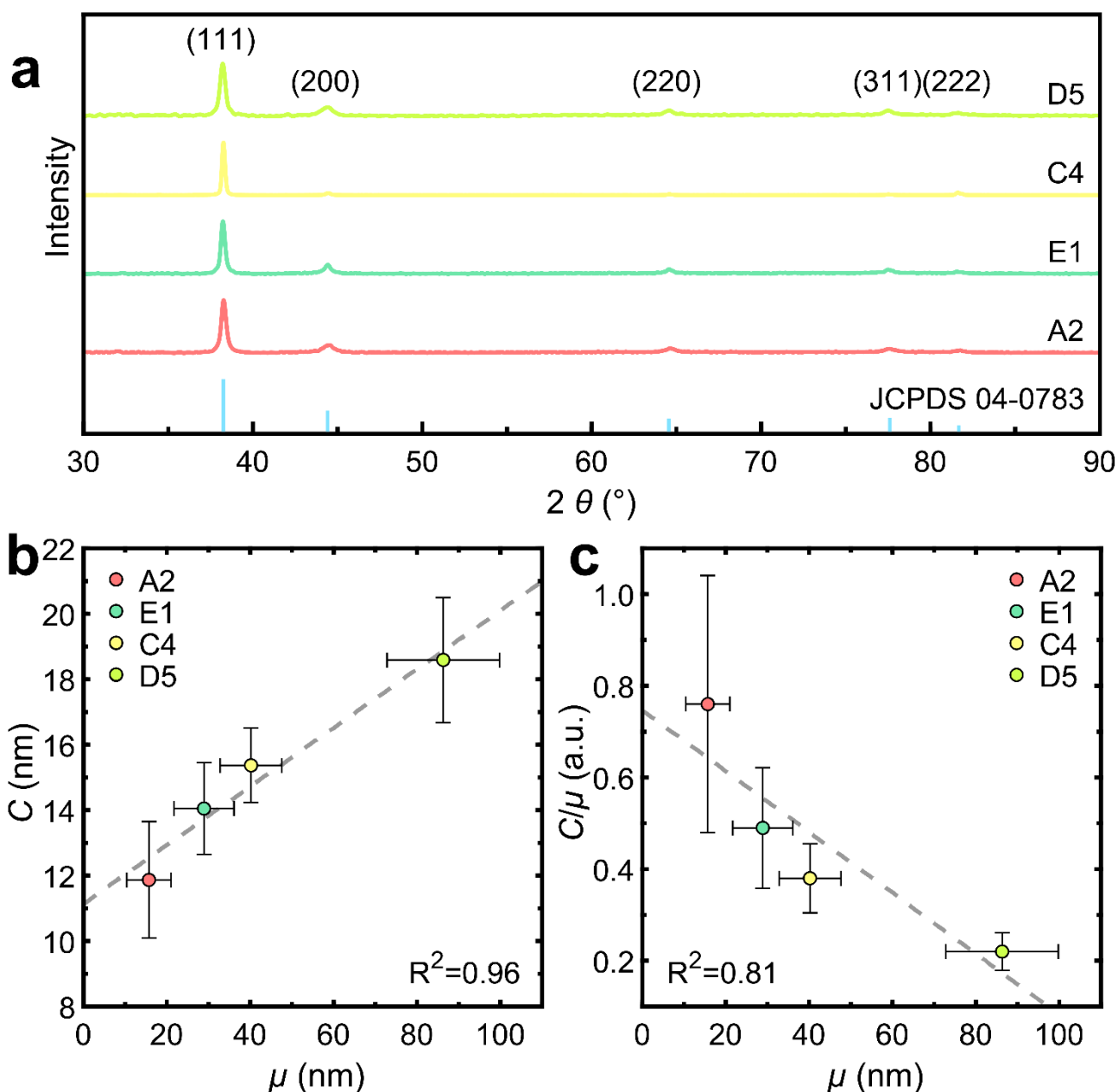


Figure S4. XRD analysis. (a) Several selected Ag NP samples, representing different effective radius values, XRD diffractograms with assigned diffraction peaks and the standard powder diffraction card for Ag (JCPDS 04-0783) (offset for clarity). (b) Dependence of crystallite size (C) on mean NP radius (μ). (c) Dependence of ratio between crystallite size and mean NP radius on mean NP radius. All error bars are 1 standard deviation.

Table S4. XRD results. Miller indices (hkl), peak locations (2θ), full width at half maximum of the peaks (β), estimates of interplanar spacings (d_θ), estimates of lattice parameter (a_θ), mean lattice parameter and standard deviation of lattice parameter estimates (Δa) for select samples

Sample name	(hkl)	2θ (°)	β (°)	d_θ (Å)	a_θ (Å)	a (Å)	Δa (Å)
A2	(111)	38.27	0.38	2.3497	4.0698	4.0734	0.0052
	(200)	44.52	0.48	2.0331	4.0661		
	(220)	64.62	0.49	1.4411	4.0762		
	(311)	77.58	0.39	1.2295	4.0778		
	(222)	81.76	0.49	1.1769	4.0771		
E1	(111)	38.21	0.32	2.3531	4.0756	4.0789	0.0029
	(200)	44.41	0.40	2.0383	4.0766		

	(220)	64.58	0.39	1.4420	4.0786		
	(311)	77.50	0.38	1.2307	4.0819		
	(222)	81.64	0.38	1.1783	4.0818		
C4	(111)	38.22	0.31	2.3531	4.0756	4.0786	0.0023
	(200)	44.38	0.33	2.0396	4.0791		
	(220)	64.58	0.30	1.4419	4.0782		
	(311)	77.49	0.37	1.2307	4.0819		
	(222)	81.74	0.39	1.1772	4.0779		
D5	(111)	38.26	0.24	2.3505	4.0712	4.0775	0.0048
	(200)	44.44	0.30	2.0371	4.0742		
	(220)	64.58	0.25	1.4420	4.0786		
	(311)	81.62	0.33	1.2304	4.0808		
	(222)	77.52	0.29	1.1786	4.0828		
JCPDS 04-0783	(111)	38.12	-	2.3590	4.0859	4.0860	0.0020
	(200)	44.28	-	2.0440	4.0880		
	(220)	64.43	-	1.4450	4.0871		
	(311)	77.47	-	1.2310	4.0828		
	(222)	81.54	-	1.1796	4.0863		

Table S5. Mean NP radius (μ), standard deviation of radius ($\Delta\mu$), mean crystallite radius (C), standard deviation of crystallite radius (ΔC), ratio of mean crystallite radius and mean NP radius (C/μ), and the standard deviation of the ratio ($\Delta(C/\mu)$) for select samples

Sample name	μ (nm)	$\Delta\mu$ (nm)	C (nm)	ΔC (nm)	C/μ	$\Delta(C/\mu)$
A2	15.7	5.3	11.87	1.78	0.76	0.37
E1	28.9	7.2	14.05	1.40	0.49	0.27
C4	40.2	7.4	15.37	1.14	0.38	0.20
D5	86.3	13.5	18.59	1.91	0.22	0.19

S2.4 Validity of the Effective Medium Model

It can be observed that as the mean size of NPs increases, the $E(\lambda)$ spectrum gains more peaks while the previously existing peak redshifts (**Figure 2 d**). This occurs due to multiple modes of oscillation (called multipoles) being allowed to exist simultaneously in large particles due to retardation effects, which are explained by Mie theory¹⁶, while only a single mode exists in particles that are small compared to the wavelength in the host medium (**Figure S5 a**, **Figure S6 b** red curve). What is more, these independent modes exist for both absorption and scattering, which are the constituent parts of $E(\lambda)$ (**Figure S5 a**). For sufficiently small NPs absorbance is larger than scattering by orders of magnitude, but the magnitude of scattering tends to increase more rapidly with size than absorbance (**Figure S6 a**) and eventually overtakes it.

For small NP sizes where absorption is the dominant ingredient of $E(\lambda)$ (*i.e.* **Figure S6 b** red curve), the quasi-static approximation² holds true and a dilute collection of such NPs can be homogenized into an effective medium⁷. The Maxwell-Garnett-Mie (MMGM) effective medium was used in this work. The MMGM medium can only account for the dipolar component of overall absorbance. However, when Ag NP radius reaches ~25 nm absorbance and scattering become equal in amplitude (indicated by an arrow in **Figure S6 a**). From that point forward, scattering begins to dominate in the $E(\lambda)$ spectrum (**Figure S6b** green and blue curves) and the quasi-static approximation is no longer valid. Therefore, the host medium with embedded Ag NPs inclusions can no longer be considered homogeneous for this size range, and said MMGM medium should no longer accurately describe the spectral response of the colloid. However, this is debatable, since the transition from effective medium validity to invalidity is not abrupt. According to⁸, for NPs with radius of

~ 50 nm effective medium theories like MMGM still predict their optical properties with a good deal of accuracy and only NPs larger than 100 nm in radius cannot be homogenized at all. We present a surprising case of the MMGM effective medium being seemingly valid for Ag NPs almost up to 150 nm in radius for predicting the shape of $E(\lambda)$ spectra (**Figure S5 b**). It might seem impossible, but the NP size distribution used to model the colloid clearly matches the experimentally derived NP size distribution (**Figure S5c**), indicating this unlikely claim is empirically true. Moreover, the modelling result obtained using MMGM effective medium also corresponds well with $E(\lambda)$ predictions made using Mie theory (**Figure S5 a**). For NPs with radius <25 nm, where the MMGM theory holds⁷, the filling fraction F corresponds to an actual volume filling fraction (concentration) of NPs in the host medium and can accurately gage their absorbance spectrum. What is more, it directly corresponds to the amplitude of the overall $E(\lambda)$ signal. At that size range $E(\lambda)$ is composed almost exclusively of absorbance, which is what TMM using an MMGM medium actually computes. The amplitudes of $E(\lambda)$ and absorbance match well due to scattering being negligible. The amplitude of the absorbance dipole drops significantly for larger NPs, this is accurately predicted by the MMGM medium. Therefore, in order for the model spectrum amplitude to keep up with the experimental amplitude, F has to be normalized taking into account the NP size-dependent absorbance contributions, essentially making it a free parameter of the MMGM model. Because of this, F no longer directly describes NP concentration, but the part of the spectrum corresponding to the dipole resonance can still be computed and its amplitude can match that of the experimental spectrum, despite the physics not being represented with complete accuracy. Since DNNs need large quantities of training data with the magnitude of said data comparable to testing (“real world”) data and DNNs themselves are not concerned about accurate physics internally, we find this to be a valid technique, especially since the shape of the spectrum is still related to the NP size distribution (**Figure S5c**). Conversion of F into actual NP concentration is described in **Section S6**.

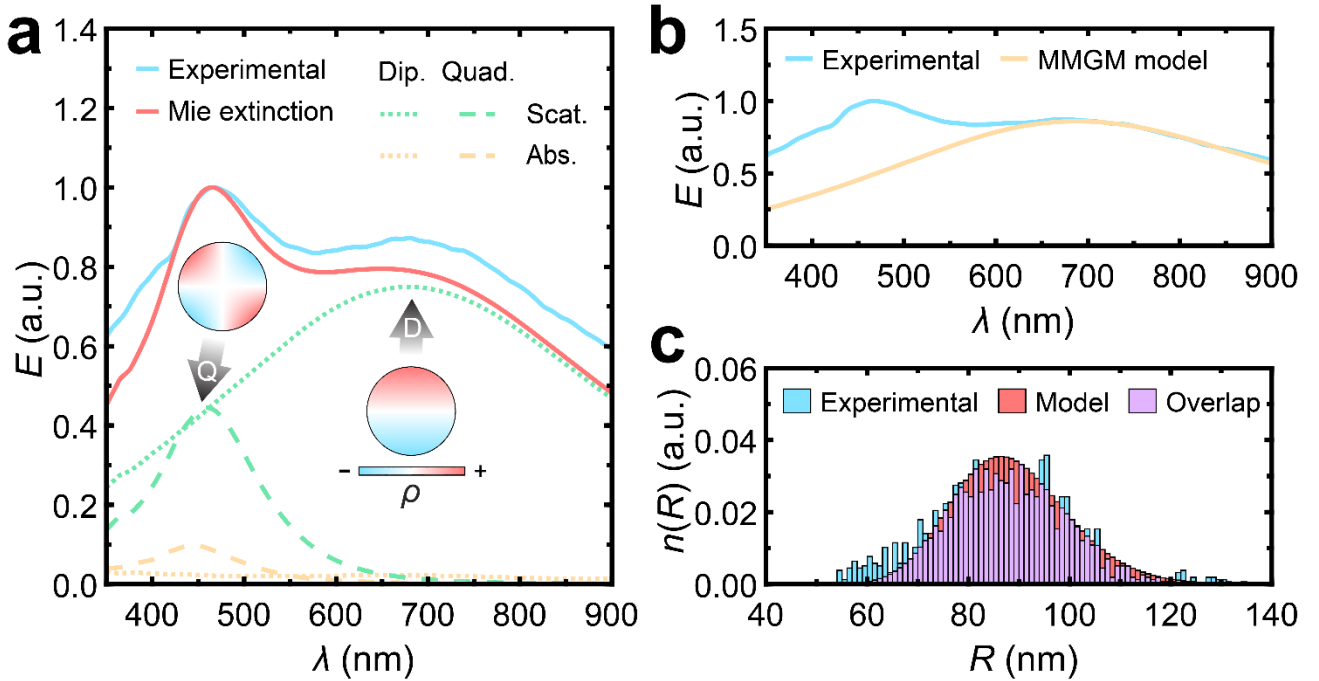


Figure S5. Comparison of Mie theory and MMGM effective medium theory. (a) Decomposition of Ag NP colloid “D5” ($\mu = 86.3$ nm) $E(\lambda)$ spectrum, normalized to one, in terms of multipoles (dipole “D” and quadrupole “Q”) of absorbance and scattering, showing the surface charge distribution (ρ) corresponding to each LSPR oscillation mode. (b) Fitting of the “D5” colloid normalized $E(\lambda)$ dipolar peak amplitude with TMM using the MMGM effective medium permittivity. (c) Experimental NP radius distribution of “D5” compared to the model distribution used for both Mie decomposition and the MMGM effective permittivity computation.

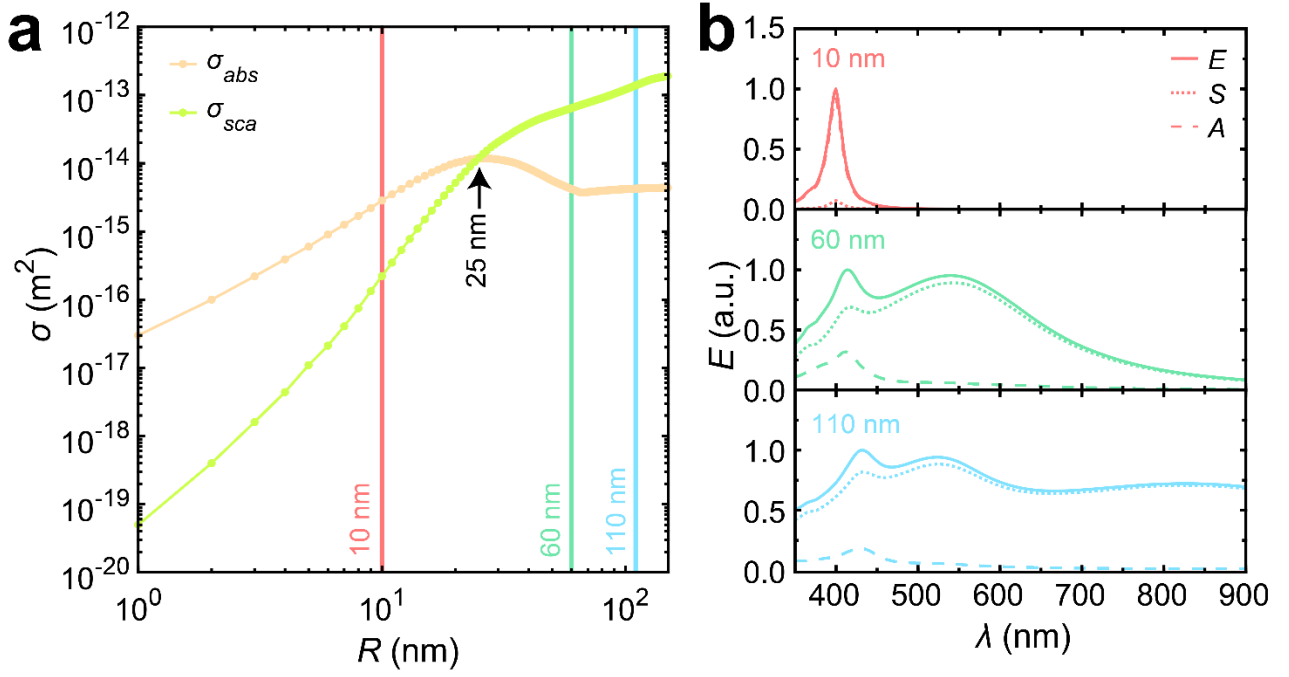


Figure S6. Comparison of absorbance (σ_{abs}) and scattering (σ_{sca}) cross-sections. (a) Logarithmic plot of maximum absorbance and scattering cross-sections for the dipole resonance (selected from the spectral range 350 nm – 900 nm) denoting the radius of equal amplitude indicated by the arrow. (b) $E(\lambda)$, normalized to one, of NPs with distinct radius values, denoted in (a), showing its constituent parts of multipolar scattering (point lines) and absorbance (dashed lines).

S2.5 Ag Mass and NPs Concentration Estimation Using Predicted Filling Fraction

The filling fraction (F), despite being intuitive, is not a parameter that is directly comparable to measurement results like those of mass concentration (mg/L). The concentration of NPs (1/L) is another parameter that may be of use. Therefore, a relation between F and these concentrations must be developed.

Through direct numerical experimentation, it was determined that the relation between F and the scaled absorbance A_F is linear. If F is too large, loss of model accuracy may ensue due to the onset of inter-particle interactions. Given that for every relevant pair of log-normal distribution parameters μ_L and σ_L a relation between a unit amplitude of A_F , denoted as A_1 , and its corresponding F , denoted as $F_1(\mu_L, \sigma_L)$, is known (**Figure 2d** in the main text), the formula for $F(\mu_L, \sigma_L)$ with any given A_F , can be expressed in equation (S1):

$$F(\mu_L, \sigma_L) = \frac{A_F}{A_1} F_1 = A_F F_1(\mu_L, \sigma_L) \quad (S6)$$

Equation (S6) enables the use of E spectra with an arbitrary magnitude, not just that which has a unit amplitude. “ColloidNet” returns the value of F with the probability density of radius, hence equation (S6) is more conceptual rather than practical. The relation between F and concentration of mass (C_M) is simply

$$C_M = F\rho \quad (S7)$$

where ρ_{Ag} is the density of the NPs, in this work taken as the bulk density of Ag (10.49 g/cm³). Keeping in mind that NPs in the colloid are distributed with a relative abundance n (probability density) based on their radius R , the total number of NPs in the volume of optical measurement can be computed:

$$N = \frac{FLr^2}{\frac{4}{3}\sum_i n_i R_i^3} \quad (S8)$$

where L is the thickness of the colloid (the interior thickness of the cuvette, in our case 1 mm) and r is the radius (in our case 2.5 mm) of the cylindrical “light volume” with thickness L . With this, the concentration of NPs can be computed:

$$C_{NP} = \frac{F}{\frac{4}{3}\pi\sum_i n_i R_i^3} \quad (S9)$$

According to **Figure S6 a** (the cross – section is of a single NP), C_{NP} directly impacts the LSPR amplitude. On the other hand, according to^{17,18}, the interband $E(\lambda)$ is related to the total amount of Ag. This is revealed in **Section S7a** – the LSPR $E(\lambda)$ magnitude is decreasing with increasing mean NP radius, while the interband $E(\lambda)$ magnitude is increasing. The decrease of C_{NP} for each subsequent growth step occurs due to dilution of the seed solution. It is an inherent part of the seeded growth method and is necessary for the growth of NPs, as shown in¹⁹. This trend is also observed in the literature describing the similar seeded-growth methodology as used to originate the colloids in this work (**Section S7b**).

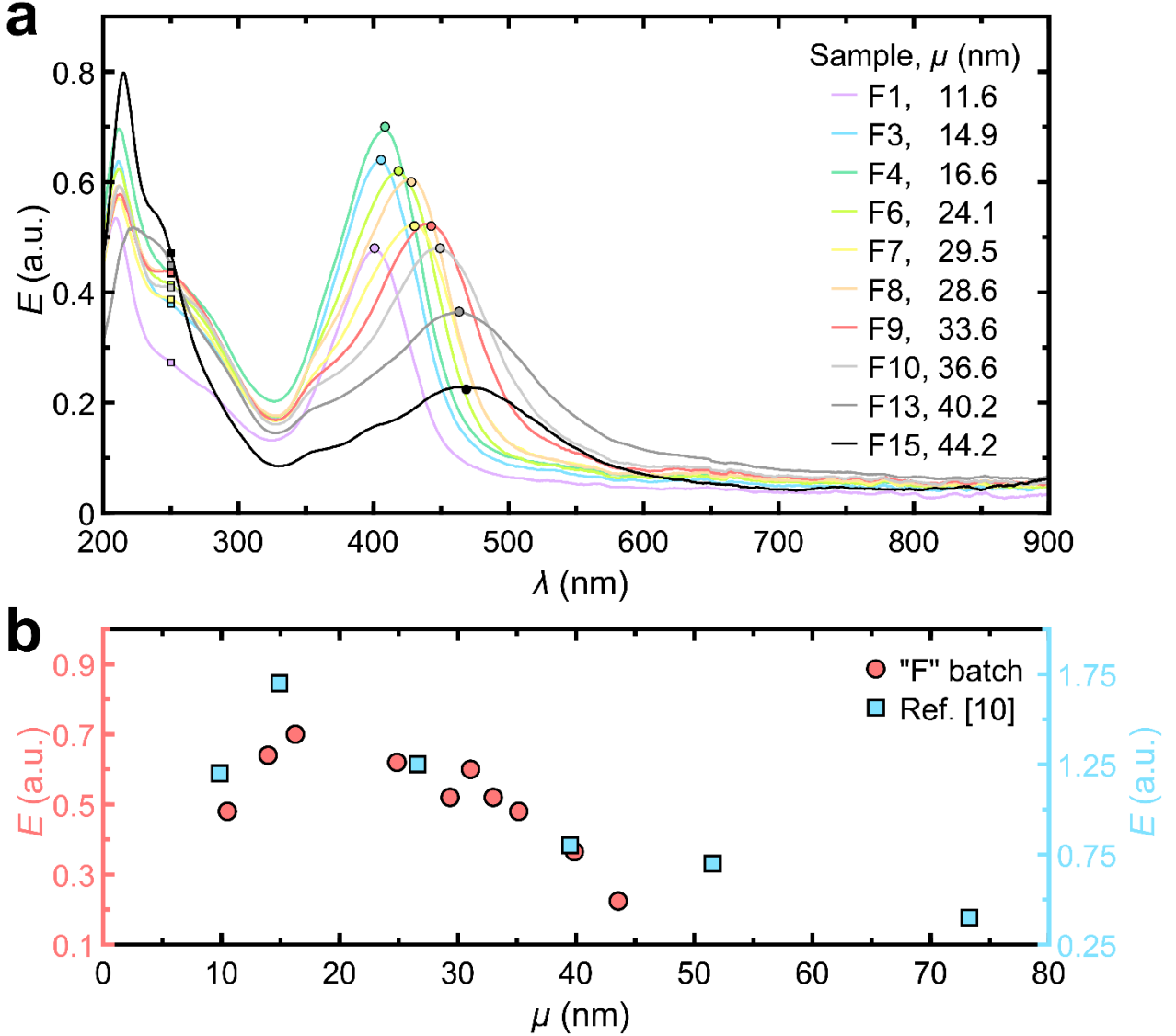


Figure S7. $E(\lambda)$ trends for selected colloids of the “F” batch. (a) experimental $E(\lambda)$ spectra of Ag NPs indicating the interband transition at 250 nm and the LSPR peak. (b) The trend of LSPR peak $E(\lambda)$ of the same colloids as in (a, circles) and reported by N. G. Bastús *et al.* in¹⁰ (squares) describing the similar NP synthesis approach.

For monodisperse NPs, it is possible to compute the expected size knowing the concentration of precursor particles (seeds), the mass of NP material added and its density. According to¹⁹, the radius of NPs after a single growth step will be:

$$R_{NP}^3 = R_{seed}^3 + \frac{3}{4\pi\rho_{Ag}} \frac{m_{Ag}}{C_{seed}} \quad (S10)$$

where R_{seed} is the radius of the seeds, R_{NP} is the NP radius to which the seeds have grown, m_{Ag} is the added mass of Ag, and C_{seed} is the concentration of seeds. Eq. S10 assumes no secondary nucleation, making the concentration of NPs equal to that of seeds, and total use of the added material. In order to model NP growth

with reduced efficiency of material use, the added mass can be multiplied by a coefficient with a value in the range from 0 to 1.

References

- (1) Bradley, D.; Roth, G. Adaptive Thresholding Using the Integral Image. *J. Graph. Tools* **2007**, *12* (2), 13–21. <https://doi.org/10.1080/2151237X.2007.10129236>.
- (2) Maier, S. A. *Plasmonics: Fundamentals and Applications*, Repr.; Springer: New York, NY, 2007.
- (3) Meyer, F. Topographic Distance and Watershed Lines. *Signal Process.* **1994**, *38* (1), 113–125. [https://doi.org/10.1016/0165-1684\(94\)90060-4](https://doi.org/10.1016/0165-1684(94)90060-4).
- (4) Patterson, A. L. The Scherrer Formula for X-Ray Particle Size Determination. *Phys. Rev.* **1939**, *56* (10), 978–982. <https://doi.org/10.1103/PhysRev.56.978>.
- (5) Niklasson, G. A.; Granqvist, C. G.; Hunderi, O. Effective Medium Models for the Optical Properties of Inhomogeneous Materials. *Appl. Opt.* **1981**, *20* (1), 26. <https://doi.org/10.1364/AO.20.000026>.
- (6) Doyle, W. T. Optical Properties of a Suspension of Metal Spheres. *Phys. Rev. B* **1989**, *39* (14), 9852–9858. <https://doi.org/10.1103/PhysRevB.39.9852>.
- (7) Battie, Y.; Resano-Garcia, A.; Chaoui, N.; Zhang, Y.; En Naciri, A. Extended Maxwell-Garnett-Mie Formulation Applied to Size Dispersion of Metallic Nanoparticles Embedded in Host Liquid Matrix. *J. Chem. Phys.* **2014**, *140* (4), 044705. <https://doi.org/10.1063/1.4862995>.
- (8) Werdehausen, D.; Staude, I.; Burger, S.; Petschulat, J.; Scharf, T.; Pertsch, T.; Decker, M. Design Rules for Customizable Optical Materials Based on Nanocomposites. *Opt. Mater. Express* **2018**, *8* (11), 3456. <https://doi.org/10.1364/OME.8.003456>.
- (9) Swat, M. J.; Grenon, P.; Wimalaratne, S. ProbOnto: Ontology and Knowledge Base of Probability Distributions. *Bioinformatics* **2016**, *32* (17), 2719–2721. <https://doi.org/10.1093/bioinformatics/btw170>.
- (10) Bastús, N. G.; Merkoçi, F.; Piella, J.; Puentes, V. Synthesis of Highly Monodisperse Citrate-Stabilized Silver Nanoparticles of up to 200 Nm: Kinetic Control and Catalytic Properties. *Chem. Mater.* **2014**, *26* (9), 2836–2846. <https://doi.org/10.1021/cm500316k>.
- (11) Babick, F. Dynamic Light Scattering (DLS). In *Characterization of Nanoparticles*; Elsevier, 2020; pp 137–172. <https://doi.org/10.1016/B978-0-12-814182-3.00010-9>.
- (12) Farkas, N.; Kramar, J. A. Dynamic Light Scattering Distributions by Any Means. *J. Nanopart. Res.* **2021**, *23* (5), 120. <https://doi.org/10.1007/s11051-021-05220-6>.
- (13) Ali, Md. H.; Azad, Md. A. K.; Khan, K. A.; Rahman, Md. O.; Chakma, U.; Kumer, A. Analysis of Crystallographic Structures and Properties of Silver Nanoparticles Synthesized Using PKL Extract and Nanoscale Characterization Techniques. *ACS Omega* **2023**, *8* (31), 28133–28142. <https://doi.org/10.1021/acsomega.3c01261>.
- (14) Khinevich, N.; Peckus, D.; Tamulevičienė, A.; Klimaitė, G.; Henzie, J.; Tamulevičius, T.; Tamulevičius, S. Size and Crystallinity Effect on the Ultrafast Optical Response of Chemically Synthesized Silver Nanoparticles. *J. Materiomics* **2023**, S2352847823001570. <https://doi.org/10.1016/j.jmat.2023.08.009>.
- (15) Singh, A.; Chaturvedi, P. Error Propagation. *Reson* **2021**, *26* (6), 853–861. <https://doi.org/10.1007/s12045-021-1185-1>.
- (16) Mie, G. Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Ann. Phys.* **1908**, *330* (3), 377–445. <https://doi.org/10.1002/andp.19083300302>.
- (17) Petit, C.; Lixon, P.; Pileni, M. P. In Situ Synthesis of Silver Nanocluster in AOT Reverse Micelles. *J. Phys. Chem.* **1993**, *97* (49), 12974–12983. <https://doi.org/10.1021/j100151a054>.
- (18) Ershov, V.; Tarasova, N.; Ershov, B. Evolution of Electronic State and Properties of Silver Nanoparticles during Their Formation in Aqueous Solution. *Int. J. Mol. Sci.* **2021**, *22* (19), 10673. <https://doi.org/10.3390/ijms221910673>.
- (19) Bastús, N. G.; Comenge, J.; Puentes, V. Kinetically Controlled Seeded Growth Synthesis of Citrate-Stabilized Gold Nanoparticles of up to 200 Nm: Size Focusing versus Ostwald Ripening. *Langmuir* **2011**, *27* (17), 11098–11105. <https://doi.org/10.1021/la201938u>.

