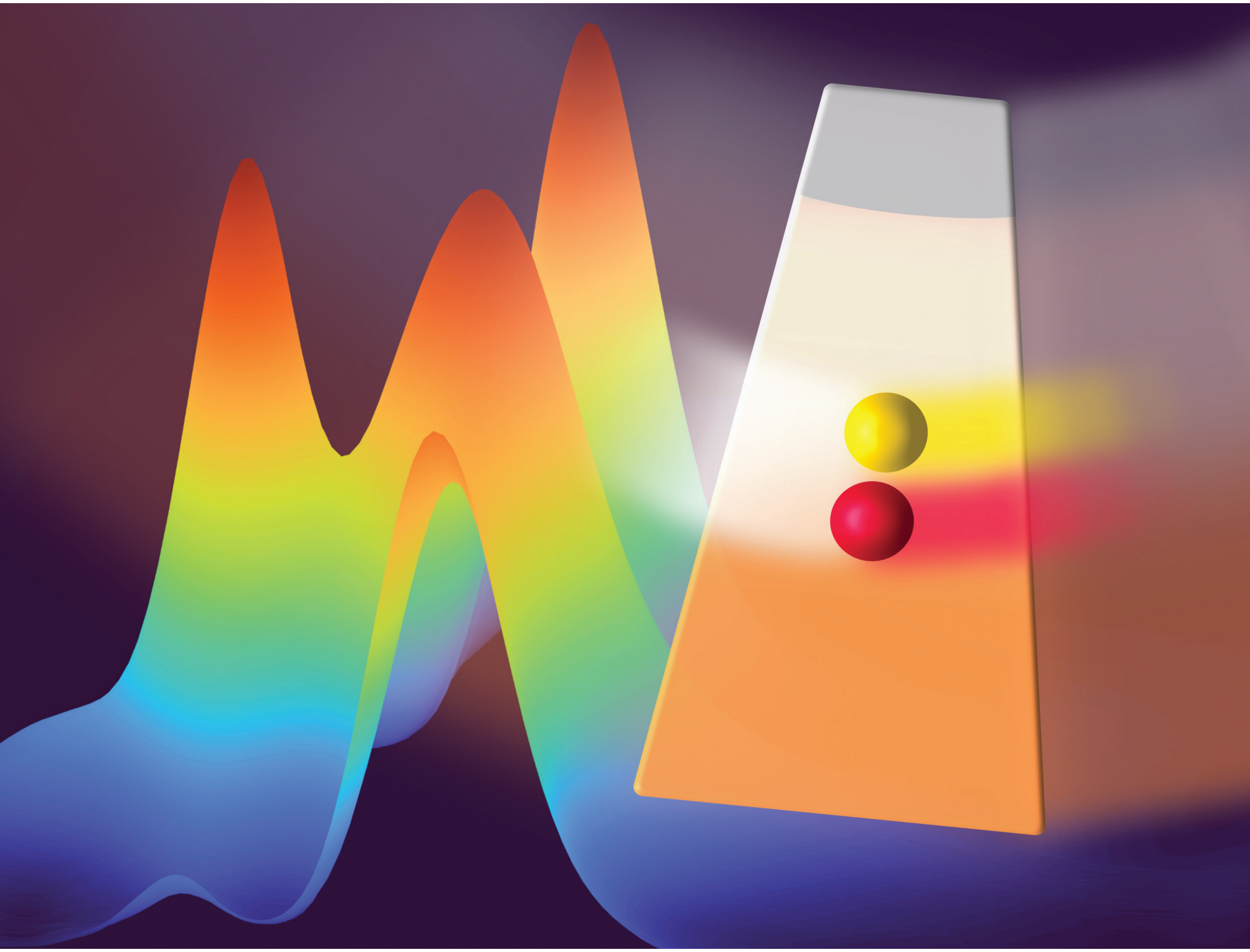
Volume 150
Number 23
7 December 2025
Pages 5115-5364

Analyst

rsc.li/analyst



ISSN 0003-2654

ROYAL SOCIETY
OF **CHEMISTRY**

PAPER
Johannes Walter *et al.*
Global analysis of sedimentation velocity data sets from multiwavelength analytical ultracentrifugation experiments using enhanced regularisation techniques

150 YEARS
ANNIVERSARY

Analyst

ROYAL SOCIETY OF CHEMISTRY

PAPER

Cite this: *Analyst*, 2025, **150**, 5147

Global analysis of sedimentation velocity data sets from multiwavelength analytical ultracentrifugation experiments using enhanced regularisation techniques

Christina Spruck,[a] Lukas Pflug[b] and Johannes Walter*[a,c]

Multiwavelength sedimentation velocity analytical ultracentrifugation experiments are a powerful technique for the simultaneous analysis of the hydrodynamic and optical properties of macromolecules and particles in solution. A fast and accurate analysis of these data sets is an important prerequisite for further investigation of the disperse and spectral properties of the sample, such as the examination of the structure–property function of nanoparticles or the spectral properties of mixtures with multiple species having different extinction or emission values. Various strategies exist for the analysis of these data sets. However, the techniques and tools presented here significantly extend the current analysis methods, since they enable a global analysis of multiwavelength data sets based on a direct boundary modelling approach that can be performed on a standard desktop computer. The distributions of sedimentation coefficients can be regularised in the sedimentation coefficient dimension, whereby the dependency of the regularisation parameter on the wavelength is taken into account. Furthermore, the method enables regularisation of the determined sedimentation coefficient distributions in the wavelength dimension, which avoids the distribution broadening due to regularisation in the sedimentation coefficient dimension. Consequently, the distribution can be determined with higher precision, particularly for narrowly distributed samples. Additionally, the frictional ratio or the partial specific volume can be determined based on a global fit that considers not just one wavelength but rather incorporates a selected range of wavelengths, thereby providing an increased accuracy of the determined parameter, especially for samples with multiple species having different wavelength-dependent extinction coefficients. Our tool and the algorithms implemented were tested and validated using synthetic data sets with known input parameters. Finally, the possibilities arising from the global multidimensional characterisation of dispersed systems are demonstrated for experimental data for several proteins as well as silver and gold nanoparticles. In addition, comparisons are made to state-of-the-art AUC data analysis software.

Received 28th July 2025,
Accepted 7th October 2025

DOI: 10.1039/d5an00793c

rsc.li/analyst

1 Introduction

Analytical ultracentrifugation (AUC) is a versatile analytical technique capable of investigating a wide range of analytes. These include biological materials such as proteins or DNA[1–3] as well as synthetic polymers and inorganic nanoparticles[4–7] with a size range extending from subnanometer to several micrometers, depending on the density of the material. AUC allows for the analysis of size, shape, mass, density, and composition, as well as the underlying distributions of the sample.[8–10] Compared to well-established analytical methods such as electron microscopy or scattering techniques, AUC has several significant advantages. Firstly, an AUC analysis provides excellent statistics, the lack of which is a common drawback of microscopy techniques, and highly accurate sedimentation coefficients (and hence the size distribution of the sample) can be obtained.[11] Secondly, the analysis of broadly distributed samples is possible using AUC, since this method couples the separation in a centrifugal field with a spectroscopic detection of the analyte.[12] A further significant advantage of AUC, particularly in the context of biological samples, is its capacity to examine the analytes in solution. As a consequence, the method allows for the investigation of analytes in their natural environment, thereby preventing artefacts from

---

[a] *Institute of Interfaces and Particle Technology (IPT), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Cauerstr. 4, 91058 Erlangen, Germany. E-mail: johannes.walter@fau.de*

[b] *Competence Center Scientific Computing (CSC), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Martensstr. 5a, 91058 Erlangen, Germany*

[c] *Interdisciplinary Center for Functional Particle Systems (FPS), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Haberstr. 9a, 91058 Erlangen, Germany*

interfering with the analysis due to further sample processing.[8] In addition, AUC is a first-principles technique,[13] and thus the measurement of the characteristics of the analyte is based directly on fundamental physics, without further assumptions required. In this, AUC can measure the absolute molar mass and size of molecules or particles in solution.[14] Thus, AUC is a powerful and accurate method for investigating biomolecules such as proteins and DNA. Although its application is less widespread, it is also an excellent technique for analysing nanoparticles and colloids.[15,16]

A further significant development in this framework is multiwavelength AUC (MW-AUC), in which the optical extinction or emission properties of the analyte can be investigated simultaneously with the hydrodynamic properties.[7,17–20] As a result, it is possible to obtain the optical properties of each species individually.[17,21] Furthermore, with MW-AUC it is also possible to discriminate between species on the basis of their optical properties, even if their hydrodynamic properties overlap.[22] Another advantage is that the signal-to-noise ratio (SNR) increases with the square root of the number of observations. It should be noted that this is only true for wavelengths with similar SNR. So, only wavelengths with significant extinction of the analyte(s) should be considered to enhance the SNR.[17,19,21]

MW-AUC is a powerful tool for the multidimensional characterisation of nanoparticle ensembles. In particle technology, it is well-known that the structure of nanoparticles, described by physical and dispersed parameters of the material, influences their final properties.[23,24] These parameters can, for example, be the size, shape, surface, or composition of the nanoparticles.[25] Since MW-AUC permits the analysis of the optical and hydrodynamic properties at the same time, it is able to connect the structure of the nanoparticles in a sample to their optical properties. For example, the two-dimensional (2D) size distributions of gold nanorods[26] or gold bipyramids[27] can be determined by MW-AUC, since the sedimentation and intrinsic molar extinction properties are both dependent on the shape of the nanoparticles. For silver patches on colloidal silica particles, the coverage and thickness of the patches can be measured.[28] The composition and size of gold–silver alloy nanoparticles can be retrieved simultaneously using MW-AUC.[29]

However, the ability to analyse MW data sets is still limited, since MW-AUC is a rather new technique, and the large amount of data collected in such experiments results in computationally intensive tasks.[21] Currently, only two analytical tools for data sets from MW-AUC sedimentation velocity (SV) experiments are available. First, the apparent distributions of sedimentation coefficients can be determined using the model-free time derivative (DCDT) approach in SEDANAL.[17] Second, UltraScan3 provides analytical approaches such as the van Holde-Weischet (vHW) method or the 2-Dimensional Spectrum Analysis (2DSA), Parametrically Constrained Spectrum Analysis (PCSA) or Genetic Algorithm (GA) models to determine the diffusion-corrected distributions of sedimentation coefficients.[8,21] The individual advantages and disadvantages of these models are discussed in more detail in the following section. To extend the evaluation methods for MW-AUC data sets, we developed a tool that permits a fast and highly efficient global analysis using a standard desktop computer. The main disadvantages of the MW-AUC, the time and cost-intensive analysis, are thereby eliminated, thus offering new possibilities for multidimensional characterisation of complex colloids.

2 The fundamentals of analytical ultracentrifugation

In principle, all AUC experiments are based on the spectroscopic detection of the movement of the analytes in a centrifugal field. In SV experiments, a sample is subjected to a stationary centrifugal field created by the constant angular velocity of the centrifuge rotor over the time of the experiment. However, it is also possible to conduct an SV-AUC experiment with different rotor speeds during the various stages of a single experiment. Thus, the centrifugal field may vary over the course of the experiment for this specific type.[30–32] The centrifugal force acting on the analytes can be up to 250 000 times the value of Earth's gravity, causing the molecules or particles to sediment or float according to their size, shape, and density. The sedimentation pattern is also influenced by the diffusion properties of the analyte. This can be measured as sedimentation boundaries, which are defined as the front between the sedimenting analytes and the remaining, analyte-free, buffer. Consequently, the observed sedimentation boundaries, more precisely their slope and movement over time, describe the diffusional and sedimentation transport of the analyte. Therefore, an SV experiment enables the determination of size and shape as well as the underlying distribution of a sample.[8,12,33]

The hydrodynamic properties of the investigated sample are described by the sedimentation coefficient s, which normalises the sedimentation velocity u of the analyte by the centrifugal acceleration, calculated from the angular velocity ω and the radial position r:

$$s = \frac{u}{\omega^2 r} \tag{1}$$

Since the hydrodynamic properties of the analyte are also influenced by its diffusion, the sedimentation coefficient can be related to the diffusion coefficient D by the Svedberg equation, where M is the molar mass of the analyte, ρ_s is the density of the solvent, $\bar{\nu}_a$ is the partial specific volume of the analyte, R is the universal gas constant, and T is the temperature during the AUC measurement:

$$\frac{s}{D} = \frac{M(1 - \rho_s \bar{\nu}_a)}{RT} \tag{2}$$

The diffusion coefficient itself can be calculated using the Einstein equation based on the universal gas constant, the temperature, the Avogadro number N_A, and the friction coefficient f of the investigated molecule or particle:

$$D = \frac{RT}{N_A f} \tag{3}$$

For creeping flow, which is ensured during an AUC measurement, the friction coefficient is described by the Stokes equation, where η is the dynamic viscosity of the solvent and x_H is the hydrodynamic diameter of the analyte:

$$f = 3\pi\eta x_H \tag{4}$$

Hence, the hydrodynamic diameter is the equivalent diameter of a sphere with the same diffusion properties as that of the analyte. The influence of shape can then be described by the frictional ratio, which is a function of shape as it relates the hydrodynamic diameter to the diameter of a volume-equivalent sphere x_V:

$$\frac{f}{f_0} = \frac{x_H}{x_V} \tag{5}$$

By combining the balance of forces on a particle in a centrifugal field and eqn (1), (3), (4), and (5), the following equation is obtained that allows the diffusion coefficient of an analyte to be calculated based on its sedimentation coefficient, frictional ratio and partial specific volume:

$$D\left(s, \frac{f}{f_0}\right) = RT\left[2\left(\bar{\nu}_a{}^{-1} - \rho_s\right)\right]^{\frac{1}{2}} \left[18 s^{\frac{1}{2}} N_A \pi \left(\frac{f}{f_0}\eta\right)^{\frac{3}{2}}\right]^{-1} \tag{6}$$

During an SV experiment in a sector-shaped cell, the change in concentration C over time t can be described by Lamm's partial differential equation, which takes into account both the flux due to sedimentation as well as the flux due to diffusion:[34]

$$\frac{\partial C}{\partial t} = -\frac{\partial}{r\partial r}\left[C\omega^2 s r^2 - Dr\frac{\partial C}{\partial r}\right] \tag{7}$$

For spherical particles, the diameter x can be calculated based on the sedimentation coefficient, the viscosity and density of the solvent, and the partial specific volume of the particle:

$$x = \left[\frac{18\eta s}{\bar{\nu}_a{}^{-1} - \rho_s}\right]^{\frac{1}{2}} \tag{8}$$

AUC measurements have several sources of error, resulting in noise in the obtained data. The noise can be divided into systematic and random noises. Systematic noise refers to time-invariant (TI) and radial-invariant (RI) noise.[35,36] An example source of TI noise would be scratches on the windows of the measurement cell. The contribution of the TI noise is constant over time but differs for each radial position. RI noise, in contrast, is constant for each radial position of the scan but varies over the duration of the measurement. A possible source of this type of noise would be fluctuations in the intensity of the lamp over the course of the experiment. A third factor is the random noise resulting from the light source and detection electronics used in AUC. This contribution is most often Gaussian distributed and differs for each radial position and time point.[37] The occurrence of these noise contributions must be taken into account during the evaluation of an AUC data set to circumvent influences on the results.

There are a number of different approaches to the evaluation of SV experiments. The methods can be divided into model-free and model-based approaches. The advantage of model-free approaches such as the vHW method and the DCDT method is that these are fast and not computationally intensive, and no prior knowledge about the sample is needed. The vHW method has the advantage that it considers the diffusion of the analyte and therefore determines the distribution of the diffusion-corrected cumulative sedimentation coefficient. However, this method does not consider the TI or RI noise, factors that will affect the calculated distribution.[38] The DCDT method, in contrast, has the advantage that it accounts for the TI noise, but it has the disadvantage that it does not consider the diffusion of the analysed sample and thus determines the apparent sedimentation coefficient distribution.[39] The second group, comprising the model-based approaches, considers both TI and RI noise, leaving only random noise for the calculation of the sedimentation coefficient distribution,[40] thereby circumventing the influence of systematic errors due to noise. These methods are based on a direct boundary modelling (DBM) approach of SV data sets from AUC experiments. Since the calculation of the sedimentation boundaries depends on the meniscus (and bottom) position, depending on the model, these positions can also be fitted to achieve the best possible solution. Within this group there are two different approaches. The ls-$g^*(s)$ method does not consider the diffusion of the sample, and therefore does not require specific prior knowledge of the partial specific volume or the shape of the sample or properties of the solvent.[41] The $c(s)$ method implemented in Sedfit[42] or the 2DSA,[43] PCSA[44] or GA[45] methods implemented in UltraScan3 determine the diffusion-corrected sedimentation coefficient distributions based on the solution of the Lamm equation. These models require prior knowledge of the sample and solvent and assume ideal sedimentation in most cases. The ls-$g^*(s)$ and $c(s)$ models will be explained in more detail later, as they form the basis for the DBM algorithms implemented in this work.

The global analysis of SV data sets from MW-AUC experiments is computationally intensive, as the data sets to be analysed are very large. A data point from a typical single-wavelength (SW) SV-AUC experiment is already in the form of a three-dimensional (3D) vector of time, radial position, and associated extinction, with several tens of thousands to hundreds of thousands of data points in a typical data set. MW-AUC adds a fourth dimension to these data sets, as all data points from a SW SV-AUC experiment are now collected for hundreds of wavelengths, resulting in millions of data points to be analysed.[21]

Currently, only a few analysis methods are available for SV data sets from MW-AUC experiments,[17,21,46] and each has its own advantages and drawbacks. In UltraScan3, the analysis of MW SV-AUC experiments is implemented using different model-based approaches, thus providing accurate distributions of the sedimentation coefficients in many cases with high resolution and advanced analysis tools. However, such implementations require supercomputing resources[8,21,46,47] and in-depth knowledge by the operator. The DCDT method as a model-free approach for MW SV-AUC data sets is implemented in SEDANAL.[17] This program offers the advantages of model-free analysis methods, including greater simplicity and fast analysis that can be performed on a standard desktop computer. However, the DCDT method cannot handle RI noise, and it can only calculate the apparent sedimentation coefficient distribution as described above.[17,48] For many macromolecules and nanoparticles, this results in diffusional broadening of the derived distributions, and determination of frictional ratios or partial specific volumes is not possible.

To provide a tool for the easy and rapid model-based analysis of MW SV-AUC experiments on a standard desktop computer, we developed the stand-alone MATLAB-App HDR-SVFIT (**H**igh **D**ynamic **R**ange – **S**edimentation **V**elocity Data **FIT**ing). The program allows for the global and simultaneous analysis of SV data sets from MW-AUC experiments for all measured wavelengths, and the analysis can be run within a short period of time on standard equipment. The details of this method are described in more detail in the following section.

3 Methods

3.1 Direct boundary modelling

In the tool presented herein, the SV data sets are evaluated using model-based methods, with a choice between the ls-$g^*(s)$ and $c(s)$ approaches. A significant advantage of these methods is that systematic noise can be taken into account; the noise vectors can be determined, and thus excluded from the calculation of the sedimentation coefficient distribution as outlined in.[36] In principle, both methods try to fit the measured sedimentation boundaries with calculated sedimentation boundaries based on various approaches. Both methods are based on solving a Fredholm integral equation of the first kind:

$$g(y) = \int f(x)\, k(x,y)\, \mathrm{d}x \tag{9}$$

The kernel $k(x, y)$ depends on the evaluation method used, as described in more detail in the following sections. For the analysis of SV data sets, the term $g(y)$ corresponds to the signals measured during the AUC experiment, and the unknown function $f(x)$ represents the sedimentation coefficient distribution. The solution of these Fredholm integral equations is based on an inversion of the integral equation followed by solving the resulting system of equations by matrix algebra, since the continuous integral in the algorithm is discretised *via* a Riemann sum.

Since the tool establishes a global analysis method for SV data sets of MW-AUC experiments, the integral equation is solved simultaneously for all measured wavelengths using matrix algebra within several seconds on a standard desktop computer, thus significantly reducing the time required for a global analysis.

3.1.1 The ls-$g^*(s)$ model. Schuck *et al.* developed a method to determine the apparent non-diffusion corrected sedimentation coefficient distribution based on a least-squares modelling of the sedimentation boundaries (ls-$g^*(s)$).[41] This approach is based on the DCDT method developed by Stafford *et al.*,[39] which handles intrinsically produced TI noise in SV data sets. However, the ls-$g^*(s)$ method has the advantage that it also considers RI noise, as outlined in ref. 36, and that it can consider all recorded sedimentation profiles for data analysis, which is not the case for the DCDT method, which should only be performed on a selected subset of scans.

The apparent sedimentation coefficient distribution $g^*(s, \lambda)$ can be determined as follows:

$$e(r,t,\lambda) \cong \int g^*(s,\lambda)\, U(s,r,t)\, \mathrm{d}s \tag{10}$$

where $e(r, t, \lambda)$ denotes the optical extinction or emission signal measured at the radial position r, time t, and wavelength λ. $U(s, r, t)$ is the kernel of the integral equation and describes the temporal and spatial sedimentation behaviour of a non-diffusing species as a step function[41] with a sedimentation coefficient s in the centrifugal field created by an angular velocity ω of the rotor, considering the meniscus position r_m and the radial dilution due to the sector shaped cells used during an SV experiment:

$$U(s,r,t) = e^{-2\omega^2 st} \times \begin{cases} 0 & \text{for } r < r_\mathrm{m} e^{\omega^2 st} \\ 1 & \text{else} \end{cases} \tag{11}$$

Eqn (10) can be solved as an optimisation problem by an inversion of the integral equation, whereby it is transformed into a linear least-squares problem, which can then be solved by matrix algebra.

As previously mentioned, the major advantage of the ls-$g^*(s)$ method is that it does not require prior knowledge about the analysed sample, and it is fast, since only a system of linear equations needs to be solved. However, it is important to note that this method has certain limitations. As it does not take into account the diffusion of the analyte, the calculations return the apparent sedimentation coefficient distribution. This distribution can be artificially broadened due to diffusion, which can introduce inaccuracy into the particle size analysis, especially for small analytes. As a model-based approach, the method assumes ideal sedimentation. Hence, non-ideality effects and noise contributions, which cannot be properly treated by the algorithm, can result in erroneous results. However, especially for large analytes with high sedimentation coefficients where diffusion can be neglected, this is a very fast and suitable analysis approach to determine particle size distributions.

3.1.2 The $c(s)$ model. The $c(s)$ method can be used to calculate the diffusion-corrected distribution of the sedimentation coefficient of the sample.[42] This method considers the diffusion of the analysed sample, as it is based on the solution of the Lamm equation, $L(s, r, t)$, as the kernel for the integral equation:

$$e(r,t,\lambda) \cong \int c(s,\lambda)\, L(s,r,t)\, \mathrm{d}s \tag{12}$$

where $c(s, \lambda)$ is the diffusion corrected sedimentation coefficient distribution. This avoids artificial broadening of the sedimentation coefficient distribution due to the diffusion of the analytes, as is the case for the ls-$g^*(s, \lambda)$ model.

However, the solution of the Lamm equation requires a number of input parameters to be solved, and thus prior knowledge of the sample analysed and the solvent is necessary to perform a $c(s)$ analysis. Since the diffusion coefficient is calculated based on eqn (6), for the solution of the Lamm equation, the density and viscosity of the solvent, the partial specific volume of the analyte, and the frictional ratio of the analyte are required.

Additionally, the calculation of the solution of the Lamm equation, the kernel of the integral equation, is computationally more challenging than the calculation of the kernel for the ls-$g^*(s, \lambda)$ method. The reason for this is that it must be calculated *via* finite element methods (FEM), since the Lamm equation, as a partial differential equation, can only be solved numerically. Claverie *et al.*[49] used hat functions on a radially and temporally equidistantly discretised grid for the FEM modelling of the solution of the Lamm equation. Further developments have included using adaptive non-equidistant grids in the radial dimension to gain a highly accurate solution in a short period of time.[50–53] Since a high $\frac{s}{D}$ ratio is numerically more difficult to simulate, the radial steps are calculated based on the spread of the steepest boundary and increase over the length of the column.[53] This approach is implemented in the developed tool. In the temporal dimension, an equidistant grid with small time steps is chosen for now.

If only an estimate exists for the partial specific volume or the frictional ratio of the analyte, it is possible to fit one of these parameters. This is realised by minimising the root mean squared deviation (RMSD) between the calculated solution and the measured optical signal. Since the solution of the Lamm equation is also dependent on the meniscus and bottom positions, it is also possible to fit these, again by minimising the RMSD.

If an MW SV-AUC data set is evaluated, there are two ways to fit the parameters of the analytes. First, the fitting can be done for one specific wavelength with the maximum measured signal in the linear range of the detector. For extinction data, wavelengths with a signal above one are excluded, since the SNR decreases significantly due to the decreasing light intensity measured by the spectrometer, which can lead to false values of the fitted parameters. Second, it is possible to conduct a global fitting of the parameters for all selected wavelengths in the analysis. In this case, the RMSD between the measured signal for the selected wavelengths and the calculated solutions based on the Lamm equation will be minimised.

Especially for systems with multiple species having different wavelength-dependent extinction coefficients, conducting a global fit can lead to improved accuracy in the estimated parameters, since for all species data with a sufficient amount of extinction can be considered. Since the solution of the Lamm equation depends on the frictional ratio and the partial specific volume of the analyte, for a sample with two or more species with different properties, the difference between these parameters must be considered to avoid artefacts during the data analysis. Up to six species with different frictional ratios and/or partial specific volumes can be considered for pre-defined ranges of the sedimentation coefficients. These ranges are then combined into a concatenated sedimentation coefficient dimension for the analysis. However, it should be noted that an increasing number of species can lead to larger inaccuracies in the fitted parameters. Moreover, conducting a global fit will require more computational effort, since the sedimentation coefficient distribution has to be calculated for each iteration (with potentially multiple parameters to be fitted) and for all selected wavelengths, which needs more computing time.

3.2 Regularisation of the sedimentation coefficient distribution

Although the TI and RI noise can be determined and thus excluded by using the DBM method to analyse the SV data sets, the random noise remains in the measured extinction or emission data. Therefore, regularisation is required to enhance the stability of the calculated solutions for the sedimentation coefficient distributions. This is necessary, as the calculation of the distributions is mathematically an ill-posed inverse problem, since the difference between the calculated solution and the measured signal can never be zero due to the noise in the measured signal.[54–56] Otherwise, without regularisation, the noise can lead to artificial fluctuations in the calculated sedimentation coefficient distribution.

There are several different methods that can be employed to regularise the calculated solution of the sedimentation coefficient distribution. The most common methods are the regularisation in the sedimentation coefficient dimension by the second derivative (Tikhonov–Phillips)[57,58] or by maximising the entropy.[59–61] For SV data sets of MW-AUC, an additional approach featuring a regularisation of the solution in the wavelength dimension is derived and demonstrated here for the first time.

3.2.1 Regularisation in the sedimentation coefficient dimension. Two different implementations are available for the regularisation in the sedimentation coefficient dimension, the Tikhonov–Phillips regularisation and the maximum entropy regularisation.

The Tikhonov–Phillips regularisation method involves the regularisation of the solution by the second derivative of the distribution with respect to the sedimentation coefficient, thereby minimising the curvature of the sedimentation coefficient distribution.[40] This results in a minimisation problem depending on the model used to analyse the data set (see Section 3.1) for $\gamma_1 \geq 0$:

$$\min_{g^*(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int g^*(s,\lambda)\,U(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_1(\lambda)\left[\frac{\partial^2 g^*(s,\lambda)}{\partial s^2}\right]^2\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{13}$$

$$\min_{c(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int c(s,\lambda)\,L(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_1(\lambda)\left[\frac{\partial^2 c(s,\lambda)}{\partial s^2}\right]^2\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{14}$$

Since the Tikhonov–Phillips regularisation is based on the second derivative with respect to the sedimentation coefficient, it enforces a smooth distribution; this leads to an enhanced broadening of the calculated distribution, in particular for monodisperse or paucidisperse samples. In contrast, the maximum entropy regularisation does not assume a smooth distribution, as it is based on the entropy of the distribution derived from information theory.[61] The sedimentation coefficient distribution is subsequently calculated depending on the selected model used to analyse the data set (see Section 3.1) for $\gamma_1 \geq 0$:

$$\min_{g^*(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int g^*(s,\lambda)\,U(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_1(\lambda)\,g^*(s,\lambda)\ln\left[g^*(s,\lambda)\right]\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{15}$$

$$\min_{c(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int c(s,\lambda)\,L(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_1(\lambda)\,c(s,\lambda)\ln\left[c(s,\lambda)\right]\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{16}$$

The maximum entropy regularisation approach is based on the Shannon entropy, which assumes an equal probability for each sedimentation coefficient.[59] A notable disadvantage of this method is that the regularisation term is highly nonlinear, resulting in the calculation of the sedimentation coefficient distribution being quite computationally intensive and consequently requiring significantly more time than the Tikhonov–Phillips regularisation, in particular for MW SV-AUC data sets.

Naturally, both regularisation approaches will result in a higher RMSD of the calculated solution and the measured signal compared to a non-regularised solution.[55] The amount of increase depends on the chosen value of the regularisation parameter γ_1. The increase in RMSD can be related to a probability value in the *F*-statistics, and thus the regularisation parameter is iteratively determined for a measured signal of a specific wavelength *via* the *F*-statistics and a chosen confidence level to obtain a statistically reliable fit.[62]

3.2.2 Regularisation in the wavelength dimension. An inherent drawback of the regularisation in the sedimentation coefficient dimension is that it leads to a broadening of the sedimentation coefficient distribution. However, for MW-AUC data sets this can be avoided by regularising the 2D sedimentation coefficient – wavelength distribution in the wavelength dimension. Since an emission or extinction spectrum is always continuous, a Tikhonov–Phillips regularisation of the optical spectrum is an appropriate choice. In this case, the calculated solution is regularised by the second derivative of the distribution with respect to the wavelength.

Therefore, the sedimentation coefficient distribution can be determined depending on the model chosen for evaluation (see Section 3.1) for $\gamma_2 \geq 0$:

$$\min_{g^*(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int g^*(s,\lambda)\,U(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_2\left[\frac{\partial^2 g^*(s,\lambda)}{\partial\lambda^2}\right]^2\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{17}$$

$$\min_{c(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int c(s,\lambda)\,L(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_2\left[\frac{\partial^2 c(s,\lambda)}{\partial\lambda^2}\right]^2\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{18}$$

As the impact of the regularisation parameter γ_2 scales with the spectral resolution of the analysed MW SV-AUC data set, the regularisation parameter is automatically determined based on the spectral resolution of the evaluated data set.

3.2.3 Simultaneous regularisation in the sedimentation coefficient and wavelength dimension. Both regularisation approaches can be combined to analyse the data set depending on the chosen model (see Section 3.1) and the chosen regularisation approach in the sedimentation coefficient dimension (see Section 3.2.1) for $\gamma_1, \gamma_2 \geq 0$:

$$\min_{g^*(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int g^*(s,\lambda)\,U(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_1(\lambda)\left[\frac{\partial^2 g^*(s,\lambda)}{\partial s^2}\right]^2+\gamma_2\left[\frac{\partial^2 g^*(s,\lambda)}{\partial\lambda^2}\right]^2\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{19}$$

$$\min_{c(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int c(s,\lambda)\,L(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_1(\lambda)\left[\frac{\partial^2 c(s,\lambda)}{\partial s^2}\right]^2+\gamma_2\left[\frac{\partial^2 c(s,\lambda)}{\partial\lambda^2}\right]^2\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{20}$$

$$\min_{g^*(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int g^*(s,\lambda)\,U(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_1(\lambda)\,g^*(s,\lambda)\ln\left[g^*(s,\lambda)\right]+\gamma_2\left[\frac{\partial^2 g^*(s,\lambda)}{\partial\lambda^2}\right]^2\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{21}$$

$$\min_{c(s,\lambda)}\left\{\iiint\left[e(r,t,\lambda)-\int c(s,\lambda)\,L(s,r,t)\,\mathrm{d}s\right]^2\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\lambda + \iint\gamma_1(\lambda)\,c(s,\lambda)\ln\left[c(s,\lambda)\right]+\gamma_2\left[\frac{\partial^2 c(s,\lambda)}{\partial\lambda^2}\right]^2\mathrm{d}s\,\mathrm{d}\lambda\right\} \tag{22}$$

Thus, the calculated distribution is simultaneously regularised in the sedimentation coefficient and wavelength dimension. However, note that at the moment only a simultaneous regularisation in the sedimentation coefficient dimension and in the wavelength dimension is possible in our tool for a Tikhonov–Phillips approach (eqn (19) and (20)) due to the limited number of solvers capable of handling the highly nonlinear systems of equations for the maximum entropy regularisation.

The regularisation parameter in the sedimentation coefficient dimension γ_1 is determined as previously described. The increase in RMSD due to the regularisation in the sedimentation coefficient dimension with a certain value of γ_1 compared to the non-regularised solution is related to a probability value in the *F*-statistics for a given confidence level. It is important to note that the value of the regularisation parameter γ_1 is determined without regularising the distribution in the wavelength dimension, since this would also influence the RMSD. The regularisation parameter for the wavelength dimension γ_2, as described above, must be determined based on the spectral resolution of the evaluated data set.

By combining the regularisation in the sedimentation coefficient and wavelength dimensions, a lower confidence level can be chosen for the sedimentation coefficient dimension if a regularisation in the wavelength dimension alone is not sufficient to suppress artefacts in the obtained distribution. Thus, the artificial broadening of the sedimentation coefficient distribution can be reduced as much as possible.

3.3 Generation of synthetic data sets

Synthetic data sets were used to validate the developed tool and the implemented algorithms. The data sets were generated using a particular workflow to account for the contributions of various factors to noise and species-specific extinction coefficients (Fig. 1).

All SV data sets were simulated using SViMULATE (version 1.1.0),[63] a program that computes a solution to the Lamm equation based on the FEM algorithm of Claverie *et al.*[49] with modifications as outlined by Todd and Haschemeyer[64] and Schuck *et al.*[51] when the numerical algorithm is chosen for the simulation. In this work, the numerical algorithm was used for all synthetic SV data sets, with 1000 radial points and constant time steps of 1 s. Standard deviations of 0.1 and 0.01 were chosen to generate a data set with TI and RI noise, respectively. Random noise was not added. The dimensions of the grid were set to 10 μm; the position of the meniscus was 5.85 cm, and the position of the bottom was set to 7.2 cm.

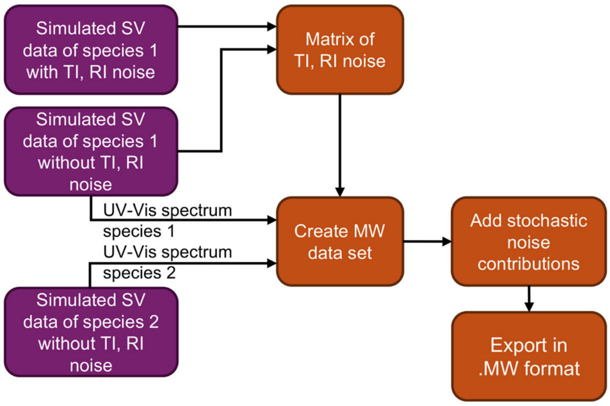


Fig. 1 Workflow for the generation of synthetic MW SV-AUC data sets, which contain time and radial invariant noise as well as stochastic noise contributions. The steps in the purple boxes were done with SViMULATE,[63] all further modifications, shown in the orange boxes, were done with MATLAB.

The hydrodynamic properties of proteins were calculated by HullRad,[65] based on protein structure files in PDB format. For this, the HullRad implementation in SViMULATE was used to obtain the hydrodynamical properties of the substance directly in the software based on the protein structure files. For bovine serum albumin (BSA) and myoglobin, the respective PDB files corresponding to "UniProt-ID P02769"[66] and "RCSB PDB-ID 1WLA"[67] were used.

All subsequent modifications to the simulated SV data sets were made with MATLAB (version R2024b). To generate the MW SV-AUC data sets, the synthetic SV data sets and information concerning the optical properties of the simulated substances were used as input. The extinction coefficients of the proteins were determined by UV/Vis spectroscopy. The concentration data of the synthetic SV data sets were multiplied by the corresponding extinction coefficients for each wavelength and subsequently converted to intensity data. If a bimodal data set was generated, the MW data sets were calculated for both substances individually and then summed for each wavelength. Random noise generated from both the lamp and the detector were then added to the signal at each wavelength. Finally, the resulting data set was exported as intensity data in a binary format for subsequent evaluation in HDR-SVFIT, where these were converted to pseudo-extinction data.

To avoid the scaling of the TI and RI noise over the wavelength by multiplying the SV data sets with the extinction coefficients, the noises were extracted from the synthetic data by simulating the same system twice with the same settings, once with and once without TI and RI noise, and subsequently subtracting both data sets to obtain a systematic noise matrix. This noise matrix was then incorporated into the calculation of the MW SV-AUC data set for each wavelength prior to the conversion of the absorption/extinction signals to intensity data.

3.4 Sample preparation for MW SV-AUC measurements

All proteins were purchased from Sigma-Aldrich and were used without further purification. BSA (product no. A7030, LOT: SLCG9886) in a phosphate buffered saline (PBS) solution, myoglobin (product no. M0630, LOT: 0000249301) in PBS, lysozyme (product no. L6876, LOT: 061M1329V) in PBS, and a mixture of myoglobin and lysozyme in PBS were measured and the experimental MW SV-AUC data sets were analysed using HDR-SVFIT. Additionally, we analysed a sample containing a mixture of silver (SKU: AGCN10-50M, LOT: TJC0256) and gold (SKU: AUCN10-50M, LOT: IAD0033) nanoparticles from nanoComposix both with a mean diameter of 10 nm.

Solutions were prepared in PBS (containing 137 mM NaCl, 10 mM Na_2HPO_4 and 2.7 mM KCl with a pH of 7.4) for all proteins. The concentration of BSA was ∼1 mg ml^{-1}; the concentration of lysozyme ∼0.25 mg ml^{-1}, and the concentration of myoglobin was ∼0.08 mg ml^{-1}. The concentrations were chosen using the criterion of the extinction of the protein solutions being in the linear range of the spectrometer. For the mixture of lysozyme and myoglobin, the pure protein solutions were mixed in a volume ratio of 1 : 1. The protein samples were measured at 20 °C, with a rotor speed of 60 000 rpm, a scan frequency of 60 s, and a radial data resolution of 50 μm. The samples were measured using titanium centrepieces from Nanolytics Instruments with a path length of 12 mm.

The expected sedimentation coefficients, frictional ratios, and partial specific volumes were calculated for the measured density and viscosity of the PBS buffer based on the corresponding PDB files using HullRad.[65] Density and viscosity were measured using a DMA 5000M density meter with an integrated Lovis 2000 viscometer from Anton Paar. The parameters were 4.32 S, 1.33, and 0.735 cm^3 g^{-1} for BSA (UniProt-ID P02769),[66] 1.85 S, 1.16 and 0.747 cm^3 g^{-1} for myoglobin (RCSB PDB-ID 2IN4),[68] and 1.87 S, 1.16 and 0.718 cm^3 g^{-1} for Lysozyme (RCSB PDB-ID 1AZF).[69]

For the mixed nanoparticle sample, the dispersions stabilised with citrate had a silver nanoparticle concentration of 0.02 mg ml^{-1} and a gold nanoparticle concentration of 0.05 mg ml^{-1}. For the MW SV-AUC measurement, both nanoparticle dispersions were mixed in a volume ratio of 1 : 1 and subsequently ultra-sonicated for 30 min. The measurement was carried out at 20 °C with a rotor speed of 9 000 rpm. The scan frequency was set to 50 s and the radial data resolution to 50 μm. The sample was measured in a 3D printed centrepiece[70] with a path length of 12 mm.

All measurements used a modified preparative ultracentrifuge, type Optima L-90K from Beckman Coulter equipped with a mirror-based MW UV/Vis/NIR detector from Nanolytics Instruments.[7,19] For all samples, the sedimentation data were recorded with a spectral resolution of 0.4 nm. The $c(s, \lambda)$ model was used for the analysis of all proteins. A global fit was conducted for the frictional ratio in HDR-SVFIT. For validation with Sedfit, fitting with the simplex algorithm was performed initially, followed by fitting with the Marquardt–Levenberg algorithm to ensure a reproducible result. The ls-$g^*(s)$ model was chosen for the sample with a mixture of gold and silver nanoparticles, since the spherical nanoparticles had a high density, large sedimentation coefficients, and minimal diffusional broadening of the sedimentation boundaries. The determined sedimentation coefficient distributions were regularised by the Tikhonov–Phillips method using a confidence level of 0.9. For the HDR-SVFIT, an additional analysis with regularisation in the wavelength dimension was performed for all samples, where the unscaled, automatically computed value for the regularisation parameter in the wavelength dimension was used (see Section 4.2).

For HDR-SVFIT, sedimentation data were used as intensity data and converted to pseudo-extinction data sets in the tool itself. For the evaluation in Sedfit, data sets for a specific wavelength were exported as a pseudo-extinction data set in ASCII ".RA" format.

A global analysis was performed for all evaluations with HDR-SVFIT. For comparison with Sedfit, 1D sedimentation coefficient distributions were extracted for the respective wavelengths from the determined 2D sedimentation coefficient – wavelength distributions.

3.5 UV/Vis spectroscopy

All UV/Vis spectra were recorded with an Analytik Jena Specord 210 Plus photo spectrometer in the range of 200 nm to 800 nm with a spectral resolution of 1 nm.

4 Results and discussion

4.1 Wavelength-dependent changes in the regularisation parameter in the sedimentation coefficient dimension

As mentioned in Section 3.2.1, the calculation of the regularisation parameter for the distribution of the sedimentation coefficient depends on the signal measured at a specific wavelength. Different considerations are therefore possible for an MW analysis. As the determination of the regularisation parameter is quite computationally intensive and therefore time consuming, the calculation for each wavelength separately is not suited to the initial screening (a typical MW experiment consists of about 2000 wavelengths). Instead, there are two additional options to choose from. One is to neglect the change in the regularisation parameter over the wavelength, in which case the regularisation parameter is determined for the wavelength with the highest measured extinction in the linear range, and this value is used for each wavelength. This can result in the determined regularisation parameter being either too high or too low for the remaining wavelengths, so that the regularisation can be either too strong or too weak for the selected confidence level. Thus, the sedimentation coefficient distributions obtained will be broader or narrower than those expected for the specific confidence level. In contrast, the variation in the regularisation parameter over the wavelengths can be considered. In this case, the regularisation parameter is determined for a specific number of wavelengths with the highest changes in the measured signal and is then interpolated for the remaining wavelengths. Under this option, the user can define the number of supporting points, which is the number of wavelengths for which the regularisation parameter is calculated.

To determine the required number of supporting points, the regularisation parameter was determined for a Tikhonov–Phillips regularisation in the sedimentation coefficient dimension with a confidence level of 0.9 and an increasing number of supporting points, based on the analysis of a synthetic data set. No regularisation in the wavelength dimension was used. A bimodal mixture of BSA and myoglobin with a spectral resolution of 1 nm from 240 nm to 600 nm was used as the data set to ensure different extinction values over a wide range of the optical spectrum. The calculated regularisation parameters over the wavelengths for different numbers of supporting points are shown in Fig. 2.

As the number of supporting points increases, the interpolated regularisation parameters converge to the full solution where the parameters are calculated for each wavelength individually. With 10 supporting points, the key features of the regularisation parameter could not be reproduced. With 25 supporting points, the bimodality of the regularisation parameter could already be observed, and the determined regularisation parameters were quite close to the full solution in the regions with significant absorption of either BSA or myoglobin. With 50 supporting points, the interpolated regularisation parameters closely approximated the correct values, but the time required to calculate the regularisation parameters also increased significantly, by a factor of ~2.2, from 76.8 s to 169 s (Apple MacBook Pro with an M3 Pro and 18 GB RAM). Because of this, for 25 supporting points, the regularisation parameters were considered close enough to the correct values, especially for the wavelengths with a significant absorption of the myoglobin BSA mixture. This appears to be a good compromise between the amount of time required and the accuracy of the calculated values for a screening analysis.

In Fig. 3, the sedimentation coefficient distributions are plotted for three selected wavelengths (298 nm, 320 nm, and 426 nm), illustrating the difference in the regularisation parameters determined for various numbers of supporting points.

The differences between the sedimentation coefficient distributions calculated by having the regularisation parameters determined individually for each wavelength *versus* based on a certain number of supporting points decrease with the number of supporting points. Especially for 298 nm and 426 nm, the differences between the solutions using 25 and 50 supporting points and the full solution with the regularisation parameter calculated individually for each wavelength were almost negligible. At 320 nm there was a recognisable difference between the solutions. However, we note that the signal intensity of the data set at 320 nm was also quite low at around 0.1 absorbance units (AU), since this wavelength is not in a range with significant absorption of BSA or myoglobin and thus is considerably lower in absorbance (~2-fold) than at 298 nm, and 426 nm. Hence, a difference in the value of the regularisation parameter has a higher impact on the resulting sedimentation coefficient distribution. However, since the difference is still quite small, 25 supporting points would be an appropriate choice for the studied bimodal system and can therefore be taken as a blueprint for future analysis of systems spanning a broad spectral range.

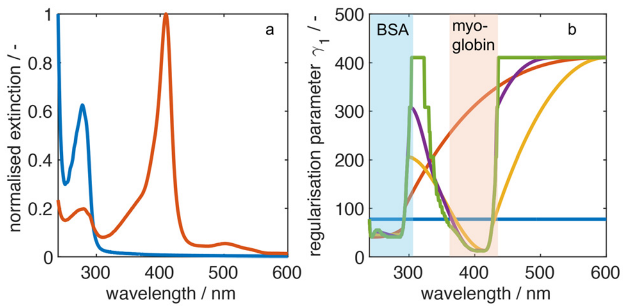


Fig. 2 Normalised extinction spectra of BSA (blue) and myoglobin (orange) (a). The dependency of the regularisation parameter on the wavelength for a Tikhonov–Phillips regularisation in the sedimentation coefficient dimension with a confidence level of 0.9 for different numbers of supporting points (1 = blue, 10 = orange, 25 = yellow, 50 = purple, full solution = green) for a synthetic data set comprising a bimodal mixture of BSA and myoglobin (b). In blue, the range of wavelengths with significant absorption of BSA is highlighted; in orange, the range of wavelengths with significant absorption of myoglobin is highlighted.

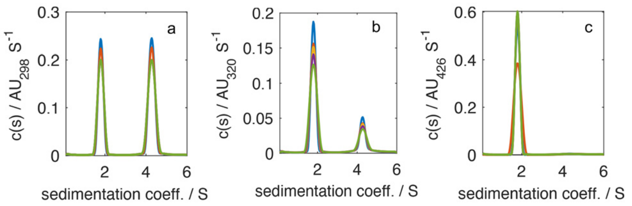


Fig. 3 Selected sedimentation coefficient distributions obtained by HDR-SVFIT at 298 nm (a), 320 nm (b), and 426 nm (c) regularised in the sedimentation coefficient dimension by a Tikhonov–Phillips regularisation with a confidence level of 0.9 and a sedimentation coefficient resolution of 200 for a synthetic data set comprising a bimodal mixture of BSA and myoglobin. Different numbers of supporting points (1 = blue, 10 = orange, 25 = yellow, 50 = purple, full solution = green) were used to account for the wavelength dependency of the regularisation parameter.

The number of supporting points can be selected by the user in HDR-SVFIT for the data sets to be analysed. Furthermore, the full solution with the regularisation parameter calculated for each wavelength individually can be considered as a final step in the analysis process to gain an as accurate and statistically reliable result as possible.

4.2 Scaling of the regularisation parameter in the wavelength dimension with the spectral resolution

As mentioned in Section 3.2.2, the regularisation parameter for the wavelength dimension needs to be scaled with the spectral resolution of the analysed data set; this can also be seen in Fig. 4. Although the regularisation parameter was the same for each spectral resolution of 1 nm, 2 nm, and 3 nm, the broadening of the distribution increased with decreasing spectral resolution. The reason for this is that the level of discretisation of the data points as well as the regularisation parameter influence the solution. The regularisation parameter must therefore be determined as a function of the discretisation of the data set, in this case the spectral resolution of the analysed data set.[71] To this end, synthetic MW SV-AUC data sets comprising a bimodal mixture of BSA and myoglobin with different spectral resolutions were analysed and regularised in the wavelength dimension. No regularisation was used in the sedimentation coefficient dimension.

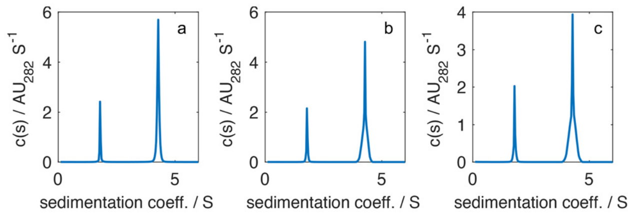


Fig. 4 Selected sedimentation coefficient distributions obtained by HDR-SVFIT at 282 nm for different spectral resolutions (1 nm (a), 2 nm (b), and 3 nm (c)) and a constant regularisation parameter for the regularisation in the wavelength dimension (γ_2 = 40) for a synthetic data set consisting of a bimodal mixture of BSA and myoglobin for a sedimentation coefficient resolution of 200.

The data analyses demonstrate that for wavelengths with a high measured absorbance, artefacts became present in the sedimentation coefficient distribution when the regularisation parameter was set at too high a value. In this case, the shape of the distribution obtained deviated strongly from a Gaussian distribution and became discontinuous. This effect is exemplified in Fig. S1 for a high value of the regularisation parameter and varying spectral resolution.

To avoid such effects, the value for the regularisation parameter in the wavelength dimension should be scaled with the spectral resolution of the analysed data set so that the distribution does not become discontinuous. This is particularly important at high extinction values, so that the broadening of the distribution depending on the spectral resolution can be circumvented. Due to a lack of robust systematic measures, this can be done by visual inspection of the sedimentation coefficient distributions for different spectral resolutions (1 nm, 1.5 nm, 2 nm, 2.5 nm, 3 nm, 3.5 nm, 4 nm and 4.5 nm) and different values for the regularisation parameter in the wavelength dimension. The critical regularisation parameter for a specific spectral resolution can then be chosen as the value before the distribution becomes notably discontinuous. As shown in Fig. 5, the dependency of the regularisation parameter can be approximated to a high degree (R^2 = 0.9938) with a linear fitting of the logarithm of the regularisation parameter over the spectral resolution of the data set.

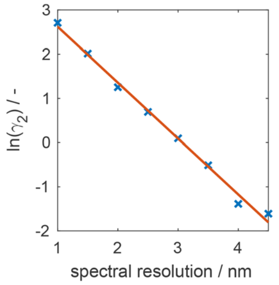


Fig. 5 The dependency of the regularisation parameter in the wavelength dimension over the spectral resolution of the analysed MW data set. The determined regularisation parameters for different spectral resolutions based on a visual inspection are shown in blue, and the fit of the logarithm of the regularisation parameter is shown in orange.

This results in a correlation that allows users to automatically select the value of the regularisation parameter γ_2 based on the spectral resolution *res* of the analysed data set:

$$\ln(\gamma_2) = -1.2647\ \text{nm}^{-1} \cdot res + 3.8851 \tag{23}$$

In HDR-SVFIT, the regularisation parameter is set based on this correlation, but it can also be additionally scaled by the user. The RMSD increase due to the regularisation with the automatically computed regularisation parameter γ_2 was significantly lower than a statistically significant increase would be. This can be seen in Fig. S2 in the SI.

4.3 Comparison of the different regularisation schemes

To assess the advantages of regularising the distributions in the wavelength dimension instead of the sedimentation coefficient dimension, we compared the sedimentation coefficient distributions obtained from analyses of synthetic data sets of BSA and myoglobin with different spectral resolutions carried out with the regularisation done either in the sedimentation coefficient dimension or in the wavelength dimension.

The results depicted in Fig. 6 confirm that the broadening of the distribution due to regularisation is significantly less in the case of regularisation in the wavelength dimension than in the sedimentation coefficient dimension. This can be seen furthermore by the comparison of the full width at half maximum (FWHM) for the respective distributions in Fig. S3 in the SI. Therefore, a significant disadvantage of the regularisation, which is nevertheless needed for a mathematically stable result, can be avoided by regularising the solution in the wavelength dimension. However, this approach is only suitable for a quite high spectral resolution of the MW SV-AUC data set. The reason for this is that for an increasingly coarser spectral resolution, the regularisation parameter decreases to a very low value, and thus the effect of the regularisation becomes negligible and possibly too low to effectively suppress artefacts.

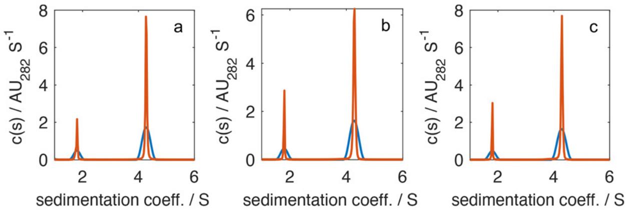


Fig. 6 A comparison of the sedimentation coefficient distributions at 282 nm for different spectral resolutions (1 nm (a), 2 nm (b), and 3 nm (c)), regularised either in the sedimentation coefficient dimension with a Tikhonov–Phillips regularisation and a confidence level of 0.9 (blue) or regularised in the wavelength dimension with the unscaled, automatically computed regularisation parameter (orange) for a synthetic data set consisting of a bimodal mixture of BSA and myoglobin for a sedimentation coefficient resolution of 200.

4.4 Validation by the evaluation of synthetic data sets

After these specific considerations for the regularisation of MW data sets, validation of our implemented algorithms in HDR-SVFIT was conducted for two selected synthetic data sets, a BSA monomer and a mixture of a BSA and a myoglobin monomer. The data sets were evaluated using both the developed tool HDR-SVFIT as well as Sedfit (version 16.1c) as a state-of-the-art benchmark, and the resulting distributions were compared with each other and with the simulation input parameters. For both data sets, a Tikhonov–Phillips regularisation of the sedimentation coefficient dimension with a confidence level of 0.9 was chosen for the analysis in HDR-SVFIT and Sedfit. An analysis with the regularisation in the wavelength dimension was also conducted in HDR-SVFIT, where the automatically selected regularisation parameter was chosen without further scaling. Both data sets were evaluated using the $c(s)$ method to account for the diffusion of the proteins.

4.4.1 Evaluation of synthetic BSA monomer data set. The monomodal data set of the BSA monomer was evaluated initially. The centrifugation process was simulated with a rotor speed of 60 000 rpm and a scanning frequency of 150 s. A total of 110 scans were included in both evaluations. The global analysis with HDR-SVFIT was performed with a spectral resolution of 1 nm from 240 nm to 310 nm. For the analysis in Sedfit, SW data sets were exported at 240 nm, 280 nm, and 295 nm. These wavelengths were chosen due to the absorption of these data sets varying over a wide range from ~1 AU at 240 nm to ~0.15 AU at 295 nm.

Table 1 lists the fitting results for the frictional ratio and the partial specific volume, the mean sedimentation coefficient obtained, and the deviation from the original input parameter of the simulation for both tools. The frictional ratio and the partial specific volume were fitted in two separate evaluations in which the other parameters was set to the value used for the simulation. A global fit in the wavelength range of 240 nm to 310 nm was performed in HDR-SVFIT. For the HDR-SVFIT analysis, the mean sedimentation coefficients obtained for the regularisation conducted either in the sedimentation coefficient dimension or in the wavelength dimension are shown.

Table 1 Comparison of the values for the frictional ratio and partial specific volume fitted by HDR-SVFIT or Sedfit for a synthetic data set of monomodal BSA. The input parameters of the simulation were 1.329 for the frictional ratio, 0.735 cm^3 g^{-1} for the partial specific volume, and 4.266 S for the sedimentation coefficient

	HDR-SVFIT		Sedfit		
	s-dim	WL-dim	240 nm	280 nm	295 nm
f/f_0/—	1.332		1.329	1.329	1.333
Deviation	0.23%		−0.05%	0.00%	0.27%
Mean s/S	4.267	4.266	4.269	4.268	4.270
Deviation	0.04%	0.01%	0.07%	0.05%	0.10%
$\bar{\nu}_a$/cm^3 g^{-1}	0.736		0.735	0.734	0.779
Deviation	0.18%		−0.01%	−0.16%	6.00%
Mean s/S	4.267	4.266	4.269	4.268	4.270
Deviation	0.04%	0.01%	0.07%	0.05%	0.11%

The deviation of the fitted frictional ratio was slightly smaller for the Sedfit results compared to the HDR-SVFIT results. Especially for 280 nm the Sedfit result matched the original frictional ratio exactly. The deviation of the partial specific volume for the Sedfit analysis for 240 nm and 280 nm was slightly smaller than for the HDR-SVFIT analysis. The deviation of the partial specific volume for the analysis at 295 nm was significantly higher than for all other analyses. This can be explained by the fact that the extinction signal at this wavelength is significantly lower than at all other examined wavelengths. On the one hand, the SNR decreases, which can lead to a less accurate result for the fitted value. On the other hand, the absolute value of the RMSD decreases with a decreasing extinction signal. Therefore, the change in the RMSD due to the change in the fitted partial specific volume becomes less significant, so that the solver is more prone to a preliminary stop. This is an advantage of a global MW SV-AUC analysis, since either a wavelength with a significant extinction is automatically chosen for the fitting, or all selected wavelengths are considered in the case of a global fit. The deviations for the mean sedimentation coefficients were generally very low and slightly smaller for the HDR-SVFIT analysis.

However, in order to gain a reproducible fit for the Sedfit analysis, an initial fitting with the simplex algorithm must be conducted followed by a fitting with the Marquardt–Levenberg algorithm. Otherwise, the results may have strong fluctuations between different fits, even though they are conducted with exactly the same boundary conditions and the same starting estimate for the frictional ratio or partial specific volume. In comparison, HDR-SVFIT uses the simplex search method of Lagarias *et al.*[72] With this algorithm, there is no need to change the fitting algorithm within an evaluation, since a reproducible result could be readily obtained with this algorithm for all investigated cases.

In Fig. 7, the comparison between the sedimentation coefficient distributions obtained either by Sedfit or by HDR-SVFIT and the input parameter for a synthetic data set of BSA are

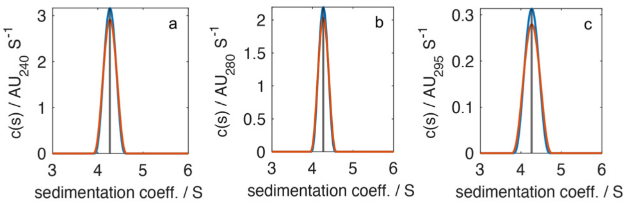


Fig. 7 Sedimentation coefficient distributions of monomodal BSA at 240 nm (a), 280 nm (b), and 295 nm (c) with the fitted frictional ratio and the original partial specific volume with a Tikhonov–Phillips regularisation in the sedimentation coefficient dimension with a confidence level of 0.9 and a sedimentation coefficient resolution of 150. The distributions determined by HDR-SVFIT are shown in blue; those obtained by Sedfit are shown in orange, and the input parameters are in black.

depicted for three different wavelengths. The distributions were derived for the fitted frictional ratio and the simulated partial specific volume.

The sedimentation coefficient distributions determined by Sedfit and by HDR-SVFIT had very minor deviations for all wavelengths. However, the distributions determined by HDR-SVFIT were slightly narrower than those determined by Sedfit, although the values determined for the frictional ratios were quite similar. This hints at slight differences in the determination of the regularisation parameters by the *F*-statistics implementation in HDR-SVFIT and Sedfit.

In Fig. 8, the comparison of the 2D sedimentation coefficient distributions obtained from global analysis with HDR-SVFIT calculated with the fitted frictional ratio and the original partial specific volume is shown. The sedimentation coefficient distributions were either regularised in the sedimentation coefficient dimension or in the wavelength dimension. The comparison of the 2D sedimentation coefficient distributions with the fitted partial specific volume is shown in Fig. S5.

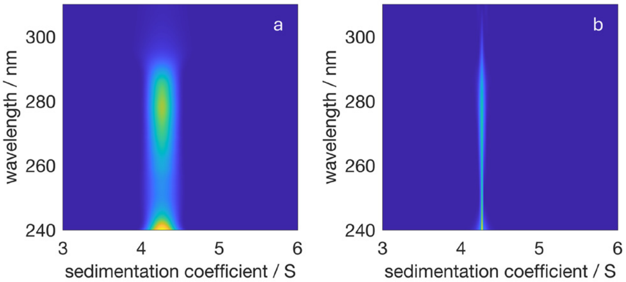


Fig. 8 Comparison of the 2D sedimentation coefficient distributions of synthetic data for monomodal BSA determined by HDR-SVFIT, with the frictional ratio fitted. The resolution for the sedimentation coefficient was 150, and the data were regularised either in the sedimentation coefficient dimension (a) or in the wavelength dimension (b). Colours from blue to yellow indicate higher levels of extinction. The sedimentation coefficients and wavelengths are with linear spacing.

The differences between the distributions with either the frictional ratio or the partial specific volume fitted can be neglected. However, there were pronounced differences between the different regularisation approaches. The distributions regularised in the wavelength dimension are significantly narrower than the distributions regularised in the sedimentation coefficient dimension. This illustrates the advantages of the regularisation in the wavelength dimension for narrowly distributed samples.

4.4.2 Evaluation of synthetic data set consisting of a myoglobin and BSA mixture. Next, a synthetic data set consisting of a bimodal mixture of myoglobin and BSA was evaluated. The simulation used a rotor speed of 60 000 rpm and a scanning frequency of 150 s and comprised 250 scans. A global analysis in HDR-SVFIT was performed with a spectral resolution of 1 nm in the range of 240 nm to 600 nm. For the analysis in Sedfit, only every second scan was included in the analysis, as using Sedfit with every scan was not possible due to memory constraints. The SW data sets were exported at 280 nm, 295 nm, and 410 nm, as the contributions of BSA and myoglobin vary over a wide range of ~0 : 1 at 410 nm to ~4 : 1 at 280 nm.

Table 2 lists the results of the fitted frictional ratios and the mean sedimentation coefficients for the Sedfit analysis as well as for the HDR-SVFIT analysis for BSA and myoglobin. Although myoglobin and BSA are both globular proteins, they differ in their partial specific volumes and frictional ratios. Therefore, a $c(s)$ analysis with a bimodal frictional ratio was chosen to analyse the data set. Since Sedfit does not support bimodality in the partial specific volume for the $c(s)$ model, a mean partial specific volume was considered for evaluation and fitting of the bimodal frictional ratio for both tools. In HDR-SVFIT, a global fit in the range of 240 nm to 600 nm and a fit at 241 nm (the wavelength with the highest extinction in the linear detection range) was performed, and the results for the mean sedimentation coefficients are shown for the regularisation done in the sedimentation coefficient dimension and the wavelength dimension. For HDR-SVFIT an additional global fit in the range of 240 nm to 600 nm was performed,

Table 2 A comparison of the frictional ratios fitted with HDR-SVFIT using a global fit or a fit at 241 nm, and Sedfit for a synthetic data set of a bimodal mixture of BSA and myoglobin. For BSA, the input parameters of the simulation were 1.329, 0.735 cm^3 g^{-1} and 4.266 S for the frictional ratio, partial specific volume, and sedimentation coefficient, respectively. For myoglobin, the respective parameters were 1.171, 0.747 cm^3 g^{-1} and 1.797 S for the frictional ratio, partial specific volume, and sedimentation coefficient

	HDR-SVFIT				Sedfit		
	Global fit		Fit at 241 nm				
	s-dim	WL-dim	s-dim	WL-dim	280 nm	295 nm	410 nm
f/f_0 BSA/—	1.324		1.325		1.320	1.331	*1.670*
Deviation	−0.37%		−0.36%		−0.74%	0.15%	*25.62%*
f/f_0 myoglobin/—	1.175		1.175		1.179	1.180	1.188
Deviation	0.34%		0.35%		0.68%	0.75%	1.46%
Mean s BSA/S	4.268	4.266	4.268	4.266	4.268	4.274	*4.395*
Deviation	0.06%	0.00%	0.06%	0.00%	0.07%	0.20%	*3.04%*
Mean s myoglobin/S	1.789	1.798	1.789	1.798	1.794	1.801	1.799
Deviation	−0.45%	0.07%	−0.45%	0.07%	−0.13%	0.23%	0.13%

where the bimodal partial specific volume was fitted, while the frictional ratios were set to the simulated values. The results can be seen in the SI in Table S1.

For the Sedfit analysis, an initial fit was again obtained using the simplex algorithm, followed by a fit with the Marquardt–Levenberg algorithm to obtain a reproducible result. The deviations from the input parameters were quite similar for the Sedfit and HDR-SVFIT analyses. The deviation of the frictional ratio for the Sedfit analysis at 410 nm was higher than for the analyses at the other wavelengths. The huge deviation for the frictional ratio determined for BSA can be traced to the low absorption of BSA at this wavelength. Thus, the result for the frictional ratio of BSA is not very robust. This highlights the significant advantage of an analysis of a MW SV-AUC data set, since either a wavelength with significant contributions of all species can be easily chosen to fit the frictional ratio, or a global fit can be conducted to gain a reliable fitting result for all species. In this case at 241 nm, there were sufficient signal contributions from both BSA and myoglobin, so that a fit conducted at one specific wavelength also produced satisfactory results. However, the comparison also shows that a global fit results in reliable results; this is especially important for samples containing multiple species, where their extinction coefficients cannot be measured at a common wavelength. The deviations for the fitted bimodal partial specific volumes in HDR-SVFIT are also quite small. Overall, the deviations were larger than for the data set with a single species and a monomodal frictional ratio and partial specific volume distribution. Firstly, as mentioned above, a mean value for the partial specific volume was chosen to evaluate the data set. Thus, the partial specific volume deviates from the original value for both species, and hence has to be equalised by the fit of the frictional ratios, since the diffusional behaviour of the species must be met. Secondly, a fit of two unknown parameters, in this case the frictional ratios in the different sedimentation coefficient ranges, will affect the RMSD, as such a fit is in general more prone to errors than a fit with only one unknown parameter.

Fig. 9 shows a comparison of the sedimentation coefficient distributions obtained from the Sedfit analysis and the HDR-SVFIT analysis with the fitted frictional ratio and the mean partial specific volume together with the input values of the sedimentation coefficient.

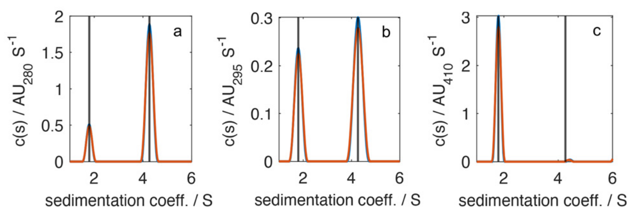


Fig. 9 Sedimentation coefficient distributions for BSA and myoglobin at 280 nm (a), 295 nm (b), and 410 nm (c), with fitted bimodal frictional ratios and a mean partial specific volume. The distributions were regularised by the Tikhonov–Phillips method in the sedimentation coefficient dimension with a confidence level of 0.9 and a sedimentation coefficient resolution of 200. The distributions determined by HDR-SVFIT are shown in blue; those obtained using Sedfit are shown in orange, and the input parameters are in black.

Although there were some minor deviations between the determined values for the frictional ratios, the distributions calculated by Sedfit and HDR-SVFIT were very similar. However, the distributions calculated by HDR-SVFIT were again slightly narrower than the distributions obtained by Sedfit. Only for 410 nm (Fig. 9c) in the Sedfit analysis a small peak is visible with a sedimentation coefficient slightly higher than expected for the BSA; this is not the case for the HDR-SVFIT analysis, since the extinction of the BSA at this wavelength is negligible. Thus, the peak could also be an artefact due to noise in the data set that may not have been sufficiently suppressed by the regularisation.

Fig. 10 shows a plot of the 2D sedimentation coefficient distributions obtained from the global analysis with HDR-SVFIT calculated with the fitted bimodal frictional ratio and the mean partial specific volume, regularised either in the sedimentation coefficient dimension or in the wavelength dimension.

The sedimentation coefficient distribution regularised in the sedimentation coefficient dimension was significantly broader than the distribution regularised in the wavelength dimension. We have thus demonstrated that the regularisation in the wavelength dimension also circumvents the distribution broadening for bimodal and narrowly distributed samples.

4.5 Evaluation of experimental data sets

Finally, five experimental systems were considered to test and validate HDR-SVFIT. These were BSA, myoglobin, lysozyme, a sample containing a mixture of myoglobin and lysozyme, and a sample containing a mixture of gold and silver nanoparticles.

4.5.1 Evaluation of the BSA sample. For the BSA sample, the MW SV-AUC data were evaluated in the range of 242.8 nm to 309.9 nm. We selected a sedimentation coefficient range of

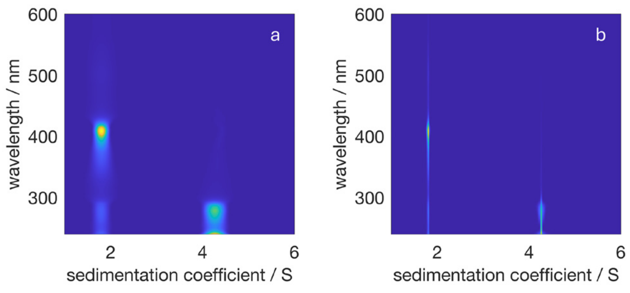


Fig. 10 Comparison of the 2D sedimentation coefficient distributions of synthetic data for a bimodal mixture of BSA and myoglobin determined by HDR-SVFIT, with the fitted bimodal frictional ratio, the mean partial specific volume, a sedimentation coefficient resolution of 200, and regularised in the sedimentation coefficient dimension (a) and the wavelength dimension (b). The colours from blue to yellow indicate higher levels of extinction. The sedimentation coefficients and wavelengths are with linear spacing.

3 S to 15 S with a resolution of 200. For the analysis in Sedfit, data sets at 250.2 nm, 280.2 nm and 294.4 nm were exported. The results for the fitted frictional ratio and the mean sedimentation coefficient of the BSA monomer are displayed in Table 3. For HDR-SVFIT, the mean sedimentation coefficients for the analyses regularised in the sedimentation coefficient dimension and the wavelength dimension are shown.

There was good agreement between the expected values and the fitted frictional ratio and mean sedimentation coefficient for the BSA monomer in all analyses with HDR-SVFIT and Sedfit. Fig. 11a–c shows a comparison of the determined sedimentation coefficient distributions using HDR-SVFIT and Sedfit.

Although the frictional ratio and the mean sedimentation coefficient of the BSA monomer were quite similar for all analyses, there were clear differences between the determined sedimentation coefficient distributions. The sedimentation coefficient distributions determined by HDR-SVFIT were broader than the distributions determined by Sedfit, and the resolution of the oligomers was lower in the distribution determined by HDR-SVFIT than in the distributions determined by Sedfit. A possible reason for this is that the regularisation parameter determined by HDR-SVFIT was greater than that determined by Sedfit, even though the chosen confidence level was the same for both tools. Another aspect that has not been considered so far for the synthetic data is spikes in the raw data; these can occur occasionally for the MW SV-AUC experiments due to timing issues. While Sedfit includes a functionality to filter for such spikes when loading the data, HDR-SVFIT lacks such a feature. Hence, an additional step in the analysis by HDR-SVFIT was applied prior to the analysis. A third-order one-dimensional median filter[73] was applied to the measured extinction signal of the sample to eliminate possible spikes and reduce fluctuation in the measured signal and thereby enhance the SNR. Notably, this had no influence on the fitted frictional ratio or the mean sedimentation coefficient of the BSA monomer, but did affect the determined sedimentation coefficient distributions (Fig. 11).

Table 3 A comparison of the frictional ratio fitted with HDR-SVFIT and Sedfit, with a partial specific volume of 0.735 cm^3 g^{-1}, and of the mean sedimentation coefficient of the monomer of a measured BSA sample. The expected values as determined by HullRad[65] are 4.32 S for the sedimentation coefficient and 1.33 for the frictional ratio

	HDR-SVFIT		Sedfit		
	s-dim	WL-dim	250.2 nm	280.2 nm	294.4 nm
f/f_0/—	1.29		1.30	1.30	1.33
Deviation	−0.04		−0.03	−0.03	−0.00
Mean *s*/S	4.32	4.31	4.29	4.28	4.29
Deviation/S	−0.00	−0.01	−0.03	−0.04	−0.03

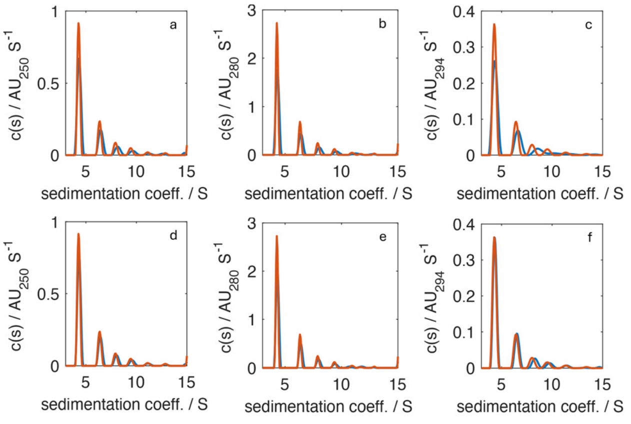


Fig. 11 Sedimentation coefficient distributions of the measured BSA sample at 250.2 nm (a and d), 280.2 nm (b and e), and 294.4 nm (c and f) with a fitted frictional ratio, with a Tikhonov–Phillips regularisation in the sedimentation coefficient dimension with a confidence level of 0.9 and a sedimentation coefficient resolution of 200. For HDR-SVFIT, the results are shown without (a–c) or with (d–f) the application of a third-order one-dimensional median filter to the pseudo-extinction data. Distributions determined by HDR-SVFIT are shown in blue, and those obtained by Sedfit are shown in orange.

The agreement between analyses performed using HDR-SVFIT and Sedfit was better than when the filter was not applied to the pseudo-extinction data in HDR-SVFIT. As the values for the fitted frictional ratio and the mean sedimentation coefficient of the BSA monomer remained constant regardless of whether the filter was applied, these differences were probably due to the determination of the regularisation parameter in HDR-SVFIT and Sedfit. A higher value for the regularisation parameter results in enhanced broadening of the sedimentation coefficient distribution and a decrease in resolution. This can be seen by comparing the values of the regularisation parameters determined with and without the filter. As shown in Fig. S6, the regularisation value for a specific wavelength was always higher for the analysis without pre-filtering than with pre-filtering. This suggests that, in the current implementation, HDR-SVFIT is more rigorous in its regularisation of measured data sets than Sedfit for the same confidence level. However, applying a third-order, one-dimensional median filter to the measured pseudo-extinction signal prior to analysis achieved good agreement between the sedimentation coefficient distributions determined by Sedfit and HDR-SVFIT. Therefore, all subsequent analyses were performed using this filter on the measured pseudo-extinction data.

Fig. 12 shows a comparison of the 2D sedimentation coefficient distribution determined by HDR-SVFIT with the regularisation performed either in the sedimentation coefficient dimension or in the wavelength dimension.

For a measured sample, the broadening of the sedimentation coefficient distribution due to the regularisation in the sedimentation coefficient dimension can also be circumvented by a regularisation of the distribution in the wavelength dimension. Notably, regularisation in the wavelength dimension can suppress artefacts in the determined 2D sedimentation coefficient distribution. We have thus demonstrated that for measured data sets, the sedimentation coefficient dis-

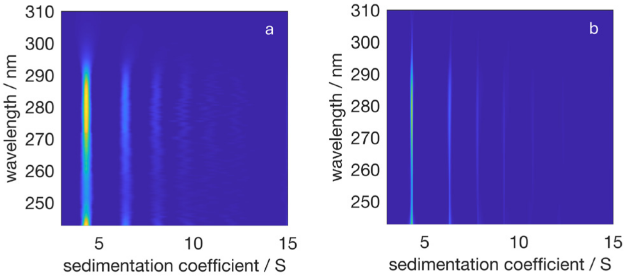


Fig. 12 Comparison of the 2D sedimentation coefficient distributions of the measured BSA sample with a fitted frictional ratio, regularised by a Tikhonov–Phillips regularisation in the sedimentation coefficient dimension (a) with a confidence level of 0.9 or in the wavelength dimension (b) with the unscaled, automatically computed regularisation parameter and a sedimentation coefficient resolution of 200. A third-order one-dimensional median filter was applied to the pseudo-extinction data. Colours from blue to yellow indicate higher levels of extinction. The sedimentation coefficients and wavelengths are with linear spacing.

tribution can be determined with higher resolution for such narrowly distributed samples by regularisation in the wavelength dimension instead of the sedimentation coefficient dimension.

4.5.2 Evaluation of a mixture of myoglobin and lysozyme. For the sample containing a mixture of lysozyme and myoglobin, the MW SV-AUC data were analysed in the range from 240 nm to 429.9 nm, to ensure significant contributions of both proteins to the data set. A sedimentation coefficient range from 1 S to 6 S with a resolution of 150 was selected. For the analysis in Sedfit, data sets at 280.2 nm, 295.3 nm, and 409.9 nm were exported to ensure different protein signal ratios in the data sets. Since myoglobin and lysozyme have different partial specific volumes, the analysis used a mean partial specific volume similar to the synthetic data set studied above. The fitted frictional ratio and the average sedimentation coefficient are listed in Table 4. It was not possible to distinguish both species in the 1D sedimentation coefficient distribution (see Fig. S11 in the SI).

The deviations between the determined frictional ratio and the expected value were low for both tools, although a mean partial specific volume was used for the analyses, since the differences between the partial specific volumes were small. In addition, the determined mean sedimentation coefficients were quite close to the expected values obtained by HullRad for myoglobin and lysozyme and the results of the measurements for the individual proteins (see Tables S2 and S3 in the SI). Fig. S11 shows a comparison of the sedimentation coefficient distributions determined by HDR-SVFIT and Sedfit.

Table 4 A comparison of the frictional ratio fitted with HDR-SVFIT and Sedfit with a mean partial specific volume of 0.732 $cm^3\ g^{-1}$, and of the mean sedimentation coefficient of a measured sample containing a mixture of lysozyme and myoglobin. The expected values determined by HullRad[65] are 1.85 S and 1.87 S for myoglobin and lysozyme, respectively, for the sedimentation coefficient and 1.16 for the frictional ratio

	HDR-SVFIT		Sedfit		
	s-dim	WL-dim	280.2 nm	295.3 nm	409.9 nm
f/f_0/—	1.19		1.15	1.19	1.18
Deviation	+0.03		−0.01	+0.03	+0.02
Mean s/S	1.83	1.83	1.84	1.85	1.85

In general, the agreement between the distributions was quite good. However, for the analyses at 280.2 nm and 295.3 nm artefacts were present at low sedimentation coefficients; these remained when decreasing the lower bound of the fitting range of the sedimentation coefficient. In the HDR-SVFIT analysis, these artefacts were less prominent than in the Sedfit analysis. Since these were not present in the pure protein solutions (see Fig. S7 and S8 in the SI), this could possibly be due to artefacts in the calculated distribution due to assuming a single partial specific volume and frictional ratio for both proteins. The minor contributions at around 3.5 S were also present in the mixture of the protein samples and hint at impurities or small agglomerates (see SI Section S7).

Fig. 13 presents a comparison of the 2D sedimentation coefficient distributions regularised in the sedimentation coefficient dimension and the wavelength dimension.

Also in this case, the distribution broadening due to the regularisation in the sedimentation coefficient dimension can be circumvented by a regularisation in the wavelength dimension, while simultaneously avoiding artefacts in the distribution. Notably, the high-resolution analysis revealed the slightly different sedimentation coefficients of the two species, with myoglobin having a slightly larger value. For the traditional analysis, such minor differences are covered by the broadening of the distributions and are thus barely observable by visual inspection.

4.5.3 Evaluation of a mixture of gold and silver nanoparticles. Finally, HDR-SVFIT was used for a global analysis of a mixture of gold and silver nanoparticles to demonstrate the tool's ability to extract extinction spectra of individual species

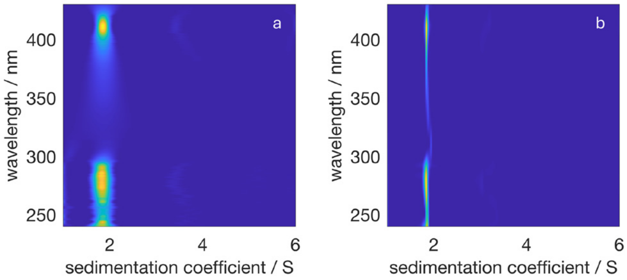


Fig. 13 Comparison of the 2D sedimentation coefficient distributions of a sample containing a mixture of lysozyme and myoglobin with fitted frictional ratio, regularised either by a Tikhonov–Phillips regularisation in the sedimentation coefficient dimension (a) with a confidence level of 0.9 or regularised in the wavelength dimension (b) with the unscaled, automatically computed regularisation parameter and a sedimentation coefficient resolution of 150. A third-order one-dimensional median filter was applied to the pseudo-extinction data. Colours from blue to yellow indicate higher levels of extinction. The sedimentation coefficients and wavelengths are with linear spacing.

in a mixture. The MW SV-AUC data were recorded from 299.9 nm to 800.2 nm, and all wavelengths were included in the global analysis, which considered a sedimentation coefficient range of 100 S to 5000 S with logarithmic bin spacing and a resolution of 150 grid points. This global analysis took only 39 s, when the regularisation parameter was calculated based on 25 supporting points (see Section 4.1). Even when the full solution was considered, so that the regularisation parameter was calculated for all 1177 wavelengths, the time for analysis increased to about 8 min.

Fig. 14 shows the 2D sedimentation coefficient distributions regularised in the sedimentation coefficient dimension and in the wavelength dimension. For the regularisation in the sedimentation coefficient dimension, the Tikhonov–Phillips approach was used with a confidence level of 0.9. For the regularisation in the wavelength dimension, the automatically calculated regularisation parameter was used without further scaling.

The figure shows that two comparatively broadly distributed species with different sedimentation coefficients were present as a result of the silver and gold nanoparticles in the sample. While both nanoparticle species had the same mean diameter, silver has a higher partial specific volume than gold. The expected mean sedimentation coefficient can be calculated using eqn (8) and thus depends on the partial specific volume of the particle species. Consequently, the expected mean sedimentation coefficient of the silver nanoparticles is lower than that of the gold nanoparticles due to their difference in density. The comparatively broad distribution is due to the particle sizes of the nanoparticles in the dispersion being distributed, rather than regularisation in the sedimentation coefficient dimension. This is confirmed by the sedimentation coefficient distribution regularised in the wavelength dimension not being significantly narrower than the distribution regularised in the sedimentation coefficient dimension. However, minor artefacts cannot be sufficiently suppressed by regularisation in the wavelength dimension for such broad distributions. Since a broadening of the distribution due to regularisation in the wavelength dimension is irrelevant for such broad distributions, regularisation in the sedimentation coefficient dimension is a suitable approach. Furthermore, regularisation in the sedimentation coefficient dimension can be combined with regularisation in the wavelength dimension, as outlined in Section 3.2.3. This allows a lower confidence level to be chosen for the sedimentation coefficient dimension regularisation, thereby avoiding the broadening of the distribution as much as possible.

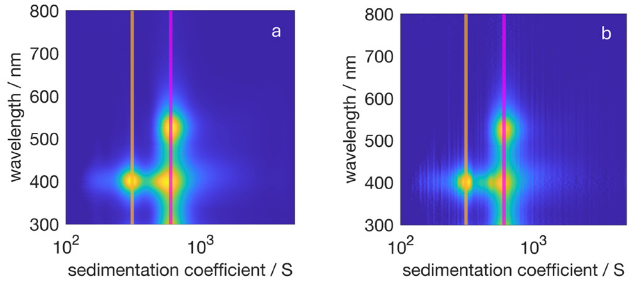


Fig. 14 Comparison of the 2D sedimentation coefficient distributions for the gold silver nanoparticle mixture with a sedimentation coefficient resolution of 150, regularised either in the sedimentation coefficient dimension (a) or in the wavelength dimension (b). Sedimentation coefficients of 310 S and 600 S are indicated by the orange and magenta lines, respectively. Colours from blue to yellow indicate higher levels of extinction. The sedimentation coefficients are with logarithmic spacing and the wavelengths are with linear spacing.

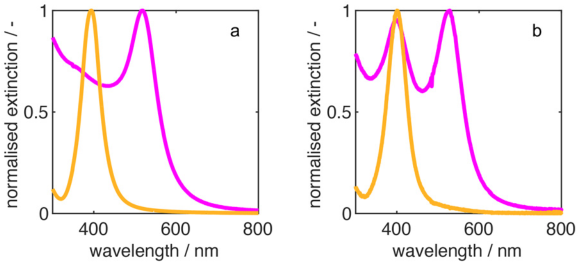


Fig. 15 Measured and normalised extinction spectra of the pure silver (yellow) and gold (magenta) nanoparticle dispersions (a) and at 310 S (yellow) and 600 S (magenta) retrieved from the 2D sedimentation coefficient distribution regularised in the sedimentation coefficient dimension (b).

The extinction spectra can be retrieved for specific sedimentation coefficients from the 2D sedimentation coefficient distribution obtained from the global analysis using HDR-SVFIT. For the gold and silver nanoparticle mixture, the extinction spectra were retrieved for sedimentation coefficients of 310 S and 600 S as indicated in Fig. 14 by the orange and magenta lines, respectively. The gained and normalised extinction spectra are shown in Fig. 15 for the regularisation in the sedimentation coefficient dimension for both sedimentation coefficients. In Fig. S12 in the SI the spectra are shown for the regularisation in the wavelength dimension.

Regardless of the regularisation approach used to analyse the MW SV-AUC dataset, the obtained extinction spectra appeared very similar for both selected sedimentation coefficients. Furthermore, the extinction properties of the species differed depending on the selected sedimentation coefficient. Comparing the extinction spectra for the lower sedimentation coefficient with that of the pure silver nanoparticle dispersion shows that, at a sedimentation coefficient of 310 S, only silver particles were present. In contrast, the extinction spectra for the higher sedimentation coefficient show that both silver and gold nanoparticles were present at 600 S. The reason is that these spectra can be obtained by combining the extinction spectra of pure silver and gold nanoparticle dispersions.

5 Conclusions and outlook

In summary, this work has demonstrated the development of a highly efficient tool for the global analysis of MW SV-AUC data sets on a standard desktop computer. This tool is based on

DBM approaches and features a choice between the ls-$g^*(s, \lambda)$ and $c(s, \lambda)$ methods that intrinsically treat the contributions of systematic noise present in the sedimentation data to avoid incorrectly calculated sedimentation coefficient distributions and to allow determination of fitting boundaries (*i.e.*, the meniscus and bottom positions) as well as analyte parameters (*i.e.*, partial specific volume and frictional ratio) depending on the chosen algorithm.

We demonstrated that the regularisation in the wavelength dimension avoids distribution broadening, a phenomenon that occurs for regularisation in the sedimentation coefficient dimension. As a result, sedimentation coefficient distributions for mono- or paucidisperse samples, such as proteins or other macromolecules and clusters, can be determined with much higher resolution, providing the possibility of resolving minute differences. In addition, the common regularisation approaches in the sedimentation coefficient dimension, the Tikhonov–Phillips and the maximum entropy regularisation, are now available for the first time for the global analysis of MW SV-AUC data sets. In particular, the Tikhonov–Phillips regularisation is still the method of choice for more broadly distributed samples. To obtain a statistically reliable result for such global analysis, the tool considers the wavelength-dependent change in the regularisation parameter in the sedimentation coefficient dimension. Furthermore, the regularisation in the wavelength dimension can be combined with a regularisation in the sedimentation coefficient dimension based on the Tikhonov–Phillips approach to avoid artefacts in the determined 2D sedimentation coefficient distribution. This also successfully prevents broadening of the distribution and is especially useful for data sets with a lower spectral resolution or broadly distributed samples, where regularisation in the wavelength dimension is limited.

Furthermore, the frictional ratio or the partial specific volume can be fitted for an individual wavelength or *via* a global fit, where all wavelengths within a selected range are considered. This approach can significantly enhance the precision of the determined parameters, especially for data sets with a mixture of species that absorb light at different wavelengths, as it allows multiple sedimentation coefficient ranges to be defined to determine frictional ratios or partial specific volumes for multiple species within a single global fit. Further studies are necessary to explore the possibilities offered by such analyses.

The ability to perform a fast and efficient global analysis of MW SV-AUC data sets allows for further examination of the obtained 2D sedimentation coefficient distributions. For example, spectral information about the analytes can be retrieved for specific sedimentation coefficients or ranges. This means that separate optical spectra of different analytes can be obtained without the need for elaborate separation prior to spectroscopy. Furthermore, the 2D sedimentation coefficient distribution can serve as a basis for subsequent analysis.

Future development of the tool could focus on the implemented analysis methods. For example, these could be extended to include a global discrete species analysis, or a global $c(s, f/f_0)$ analysis of MW SV-AUC data in which the size and shape distribution of the analyte could be determined together with their related spectra. The advantage of such a $c(s, f/f_0)$ would be that the solution could be regularised in the wavelength dimension and globally fitted. While fitting an additional dimension such as the frictional ratio would decrease the accuracy, consideration of all available wavelengths would lead to a much more mathematically robust result. Furthermore, regularising the calculated solution in the wavelength dimension can produce more stable results by reducing potential artefacts in the solution. However, due to the significantly larger solution space, memory constraints could be a limiting factor and would need to be investigated in greater detail.

HDR-SVFIT can be used to process data obtained by the custom user-built MW setups with the detection of extinction[7,19,74] and emission[18,20] as well as the commercial preparative ultracentrifuge remodel offered by Nanolytics Instruments.[75] Notably, HDR-SVFIT can also be used to analyse data obtained using Beckman Coulter's XL-A and Optima AUC, even though such data will not really benefit from the MW data analysis capabilities due to the limited number of wavelengths that can be measured in one experiment. In addition, HDR-SVFIT includes all functionalities published previously for the direct boundary analysis of analytical centrifugation experiments.[76,77] Even though only a limited number of groups may benefit from the MW capabilities presented herein, we hope that the development of an easy to use and freely available tool for data analysis will catalyse commercial solutions that provide full MW data to make use of the far-reaching opportunities offered by AUC for multidimensional analysis of complex colloidal systems.

Author contributions

Christina Spruck: data curation, formal analysis, investigation, software, validation, visualisation, writing – original draft; Lukas Pflug: funding acquisition, resources, software, supervision, writing – review & editing; Johannes Walter: conceptualisation, funding acquisition, methodology, project administration, resources, software, supervision, writing – review & editing.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data that support the findings of this study are available in Zenodo at: https://doi.org/10.5281/zenodo.16525558.

A compiled version of the software is freely available at: https://www.ccg.forschung.fau.de/software.

Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d5an00793c.

Acknowledgements

The authors gratefully acknowledge funding of the Collaborative Research Centre 1411 "Design of Particulate Products" (Project-ID 416229255) and the "FPS Core Facility" (Project-ID 539724755), both provided by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). The authors also thank Sandra Wittpahl, Hadi Soltan Moradi, and Sitong Xu for their assistance with the AUC experiments.

References

1 P. Schuck, *Biophys. Rev.*, 2013, **5**, 159–171.

2 J. Lebowitz, M. S. Lewis and P. Schuck, *Protein Sci.*, 2002, **11**, 2067–2079.

3 C. R. Horne, A. Henrickson, B. Demeler and R. C. J. Dobson, *Eur. Biophys. J.*, 2020, **49**, 819–827.

4 H. Cölfen, *Colloid Polym. Sci.*, 2023, **301**, 821–849.

5 G. Cinar, J. I. Solomun, P. Mapfumo, A. Traeger and I. Nischang, *Anal. Chim. Acta*, 2022, **1205**, 339741.

6 H. Lange, *Colloid Polym. Sci.*, 1986, **264**, 488–493.

7 J. Walter, K. Löhr, E. Karabudak, W. Reis, J. Mikhael, W. Peukert, W. Wohlleben and H. Cölfen, *ACS Nano*, 2014, **8**, 8871–8886.

8 Z. Aziz, J. Behlke, G. Bernardi, L. Bourdillon, P. J. G. Butler, N. Carels, O. Clay, H. Colfen, J. J. Correia, M. A. Daugherty, J. G. de la Torre, B. Demeler, C. J. Douady, H. Durchschlag, K. G. Fleming, M. G. Fried, P. B. Furtado, H. E. Gilbert, R. J. C. Gilbert, S. E. Harding, L. A. Holladay, A. Krebs, I. Laidlaw, Y. C. Lee, M. S. Lewis, A. Ortega, S. J. Perkins, M. M. Reily, O. Ristau, A. J. Rowe, H. E. P. Sanchez, P. Schuck, P. J. Sherwood, C. A. Sontag, W. F. Stafford, M. Steinmetz, Z. Sun, C. G. Tate, C. Wandrey, P. R. Wills, D. J. W. Winzor, P. Zipper and J. Cole, in *Analytical Ultracentrifugation*, The Royal Society of Chemistry, 2005.

9 T. Svedberg and J. B. Nichols, *J. Am. Chem. Soc.*, 1923, **45**, 2910–2917.

10 T. Svedberg and H. Rinde, *J. Am. Chem. Soc.*, 1924, **46**, 2677–2693.

11 H. Lange, *Part. Part. Syst. Charact.*, 1995, **12**, 148–157.

12 W. Mächtle and L. Börger, in *Analytical Ultracentrifugation of Polymers and Nanoparticles*, Springer Verlag, Berlin, Heidelberg, 1st edn, 2006, pp. 237.

13 T. M. Laue and W. F. Stafford III, *Annu. Rev. Biophys.*, 1999, **28**, 75–100.

14 H. Zhao, R. Ghirlando, C. Alfonso, F. Arisaka, I. Attali, D. L. Bain, M. M. Bakhtina, D. F. Becker, G. J. Bedwell, A. Bekdemir, T. M. D. Besong, C. Birck, C. A. Brautigam, W. Brennerman, O. Byron, A. Bzowska, J. B. Chaires, C. T. Chaton, H. Cölfen, K. D. Connaghan, K. A. Crowley, U. Curth, T. Daviter, W. L. Dean, A. I. Díez, C. Ebel, D. M. Eckert, L. E. Eisele, E. Eisenstein, P. England, C. Escalante, J. A. Fagan, R. Fairman, R. M. Finn, W. Fischle, J. G. de la Torre, J. Gor, H. Gustafsson, D. Hall, S. E. Harding, J. G. H. Cifre, A. B. Herr, E. E. Howell, R. S. Isaac, S.-C. Jao, D. Jose, S.-J. Kim, B. Kokona, J. A. Kornblatt, D. Kosek, E. Krayukhina, D. Krzizike, E. A. Kusznir, H. Kwon, A. Larson, T. M. Laue, A. L. Roy, A. P. Leech, H. Lilie, K. Luger, J. R. Luque-Ortega, J. Ma, C. A. May, E. L. Maynard, A. Modrak-Wojcik, Y.-F. Mok, N. Mücke, L. Nagel-Steger, G. J. Narlikar, M. Noda, A. Nourse, T. Obsil, C. K. Park, J.-K. Park, P. D. Pawelek, E. E. Perdue, S. J. Perkins, M. A. Perugini, C. L. Peterson, M. G. Peverelli, G. Piszczek, G. Prag, P. E. Prevelige, B. D. E. Raynal, L. Rezabkova, K. Richter, A. E. Ringel, R. Rosenberg, A. J. Rowe, A. C. Rufer, D. J. Scott, J. G. Seravalli, A. S. Solovyova, R. Song, D. Staunton, C. Stoddard, K. Stott, H. M. Strauss, W. W. Streicher, J. P. Sumida, *et al.*, *PLoS One*, 2015, **10**, e0126420.

15 G. B. Edwards, U. M. Muthurajan, S. Bowerman and K. Luger, *Curr. Protoc. Mol. Biol.*, 2020, **133**, e131.

16 K. L. Planken and H. Cölfen, *Nanoscale*, 2010, **2**, 1849–1869.

17 J. Walter, P. J. Sherwood, W. Lin, D. Segets, W. F. Stafford and W. Peukert, *Anal. Chem.*, 2015, **87**, 3396–3403.

18 S. E. Wawra, G. Onishchukov, M. Maranska, S. Eigler, J. Walter and W. Peukert, *Nanoscale Adv.*, 2019, **1**, 4422–4432.

19 J. Pearson, J. Walter, W. Peukert and H. Cölfen, *Anal. Chem.*, 2018, **90**, 1280–1291.

20 V. Lautenbach, G. Onishchukov, S. E. Wawra, U. Frank, L. Hartmann, W. Peukert and J. Walter, *Nanoscale Adv.*, 2024, **6**, 2611–2622.

21 G. E. Gorbet, J. Z. Pearson, A. K. Demeler, H. Cölfen and B. Demeler, in *Chapter Two – Next-Generation AUC: Analysis of Multiwavelength Analytical Ultracentrifugation Data*, ed. J. L. Cole, Academic Press, 2015, vol. 562, pp. 27–47.

22 C. N. Johnson, G. E. Gorbet, H. Ramsower, J. Urquidi, L. Brancaleon and B. Demeler, *Eur. Biophys. J.*, 2018, **47**, 789–797.

23 H. Rumpf, *Staub – Reinhalt. Luft*, 1967, **21**, 3–13.

24 R. Polke, M. Schäfer and N. Scholz, in *Charakterisierung disperser Systeme: Abschnitt 2.1*, ed. H. Schubert, John Wiley and Sons, Hoboken, 2012, pp. 7–45.

25 U. Frank, S. E. Wawra, L. Pflug and W. Peukert, *Part. Part. Syst. Charact.*, 2019, **36**, 1800554.

26 S. E. Wawra, L. Pflug, T. Thajudeen, C. Kryschi, M. Stingl and W. Peukert, *Nat. Commun.*, 2018, **9**, 4898.

27 U. Frank, D. Drobek, A. Sánchez-Iglesias, S. E. Wawra, N. Nees, J. Walter, L. Pflug, B. Apeleo Zubiri, E. Spiecker, L. M. Liz-Marzán and W. Peukert, *ACS Nano*, 2023, **17**, 5785–5798.

28 T. Meincke, J. Walter, L. Pflug, T. Thajudeen, A. Völkl, P. Cardenas Lopez, M. J. Uttinger, M. Stingl, S. Watanabe, W. Peukert and R. N. Klupp Taylor, *J. Colloid Interface Sci.*, 2022, **607**, 698–710.

29 P. Cardenas Lopez, M. J. Uttinger, N. E. Traoré, H. A. Khan, D. Drobek, B. Apeleo Zubiri, E. Spiecker, L. Pflug, W. Peukert and J. Walter, *Nanoscale*, 2022, **14**, 12928–12939.

30 G. E. Gorbet, S. Mohapatra and B. Demeler, *Eur. Biophys. J.*, 2018, **47**, 825–835.

31 J. Ma, H. Zhao, J. Sandmaier, J. A. Liddle and P. Schuck, *Biophys. J.*, 2016, **110**, 103–112.

32 W. F. Stafford and E. H. Braswell, *Biophys. Chem.*, 2004, **108**, 273–279.

33 J. J. Correia and W. F. Stafford, in *Chapter Three – Sedimentation Velocity: A Classical Perspective*, ed. J. L. Cole, Academic Press, 2015, vol. 562, pp. 49–80.

34 O. Lamm, *Ark. Mat., Astron. Fys.*, 1929, **21B**, 1–4.

35 S. Mortezazadeh and B. Demeler, *Eur. Biophys. J.*, 2023, **52**, 203–213.

36 P. Schuck and B. Demeler, *Biophys. J.*, 1999, **76**, 2288–2296.

37 J. Ma, H. Zhao and P. Schuck, *Anal. Biochem.*, 2015, **483**, 1–3.

38 K. E. Van Holde and W. O. Weischet, *Biopolymers*, 1978, **17**, 1387–1403.

39 W. F. Stafford, *Anal. Biochem.*, 1992, **203**, 295–301.

40 P. Schuck, *Sedimentation Velocity Analytical Ultracentrifugation: Discrete Species and Size-Distributions of Macromolecules and Particles*, CRC Press, 1st edn, 2016.

41 P. Schuck and P. Rossmanith, *Biopolymers*, 2000, **54**, 328–341.

42 P. Schuck, *Biophys. J.*, 2000, **78**, 1606–1619.

43 E. Brookes, W. Cao and B. Demeler, *Eur. Biophys. J.*, 2010, **39**, 405–414.

44 G. Gorbet, T. Devlin, B. I. Hernandez Uribe, A. K. Demeler, Z. L. Lindsey, S. Ganji, S. Breton, L. Weise-Cross, E. M. Lafer, E. H. Brookes and B. Demeler, *Biophys. J.*, 2014, **106**, 1741–1750.

45 E. H. Brookes and B. Demeler, *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, New York, NY, USA, 2007, pp. 361–368.

46 B. Demeler and G. E. Gorbet, in *Analytical Ultracentrifugation Data Analysis with UltraScan-III*, ed. S. Uchiyama, F. Arisaka, W. F. Stafford and T. Laue, Springer, Japan, Tokyo, 2016, pp. 119–143.

47 B. Demeler, *Curr. Protoc.*, 2024, **4**, e974.

48 J. L. Cole, J. W. Lary, T. P. Moody and T. M. Laue, *in Analytical Ultracentrifugation: Sedimentation Velocity and Sedimentation Equilibrium*, Academic Press, 2008, vol. 84, pp. 143–179.

49 J.-M. Claverie, H. Dreux and R. Cohen, *Biopolymers*, 1975, **14**, 1685–1700.

50 P. Schuck, *Biophys. J.*, 1998, **75**, 1503–1512.

51 P. Schuck, C. E. MacPhee and G. J. Howlett, *Biophys. J.*, 1998, **74**, 466–474.

52 J. Dam, C. A. Velikovsky, R. A. Mariuzza, C. Urbanke and P. Schuck, *Biophys. J.*, 2005, **89**, 619–634.

53 P. H. Brown and P. Schuck, *Comput. Phys. Commun.*, 2008, **178**, 105–120.

54 M. Richter and S. Schäffler, *Inverse Probleme mit stochastisch modellierten Messdaten: Stochastische und numerische Methoden der Diskretisierung und Optimierung*, Springer Spektrum Berlin, Heidelberg, 2022.

55 M. Richter, *Inverse Probleme: Grundlagen, Theorie und Anwendungsbeispiele*, Springer-Verlag Berlin, Heidelberg, 2015.

56 *Lexikon der Mathematik: Band 4*, ed. G. Walz, Springer Spektrum Berlin, Heidelberg, 2nd edn, 2016.

57 D. L. Phillips, *J. Altern. Complementary Med.*, 1962, **9**, 84–97.

58 F. Haußner and Y. Luchko, *Mathematische Modellierung mit MATLAB® und Octave: Eine praxisorientierte Einführung*, Springer Spektrum Berlin, Heidelberg, 2nd edn, 2019.

59 W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New York, 2nd edn, 1992.

60 C. R. Smith and W. T. Grandy, *Maximum-Entropy and Bayesian Methods in Inverse Problems*, Springer Dordrecht, Dordrecht, 1985.

61 J. Skilling and R. K. Bryan, *Mon. Not. R. Astron. Soc.*, 1984, **211**, 111–124.

62 S. W. Provencher, *Comput. Phys. Commun.*, 1982, **27**, 213–227.

63 C. A. Brautigam, *Eur. Biophys. J.*, 2023, **52**, 293–302.

64 G. P. Todd and R. H. Haschemeyer, *Biophys. Chem.*, 1983, **17**, 321–336.

65 P. J. Fleming and K. G. Fleming, *Biophys. J.*, 2018, **114**, 856–869.

66 J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis, *Nature*, 2021, **596**, 583–589.

67 R. Maurus, C. M. Overall, R. Bogumil, Y. Luo, A. G. Mauk, M. Smith and G. D. Brayer, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1997, **1341**, 1–13.

68 K. E. Wheeler, J. M. Nocek, D. A. Cull, L. A. Yatsunyk, A. C. Rosenzweig and B. M. Hoffman, *J. Am. Chem. Soc.*, 2007, **129**, 3906–3917.

69 K. Lim, A. Nadarajah, E. L. Forsythe and M. L. Pusey, *Acta Crystallogr., Sect. D:Biol. Crystallogr.*, 1998, **54**, 899–904.

70 A. Desai, J. Krynitsky, T. J. Pohida, H. Zhao and P. Schuck, *PLoS One*, 2016, **11**, e0155201.

71 V. Albani, A. D. Cezaro and J. P. Zubelli, *Inverse Probl. Imaging*, 2016, **10**, 1–25.

72 J. C. Lagarias, J. A. Reeds, M. H. Wright and P. E. Wright, *SIAM J. Control*, 1998, **9**, 112–147.

73 W. K. Pratt, *Digital image processing*, Wiley, New York, 2007.

74 *Analytical Ultracentrifugation VIII*, ed. C. Wandrey and H. Cölfen, Springer-Verlag, Berlin/Heidelberg, 2006.

75 J. Z. Pearson, F. Krause, D. Haffke, B. Demeler, K. Schilling and H. Cölfen, in *Chapter One – Next-Generation AUC Adds a Spectral Dimension: Development of Multiwavelength Detectors for the Analytical Ultracentrifuge*, ed. J. L. Cole, Academic Press, 2015, vol. 562, pp. 1–26.

76 J. Walter, T. Thajudeen, S. Süß, D. Segets and W. Peukert, *Nanoscale*, 2015, **7**, 6574–6587.

77 M. J. Uttinger, S. Boldt, S. E. Wawra, T. D. Freiwald, C. Damm, J. Walter, D. Lerche and W. Peukert, *Part. Part. Syst. Charact.*, 2020, **37**, 2000108.