Eur Biophys J (2010) 39:449–455
DOI 10.1007/s00249-009-0425-1

ORIGINAL PAPER

Enhanced resolution of sedimentation coefficient distribution profiles by extrapolation to infinite time

Joachim Behlke · Otto Ristau

Received: 8 November 2008 / Revised: 3 February 2009 / Accepted: 15 February 2009 / Published online: 6 March 2009

Abstract Sedimentation velocity experiments can be used to identify two or more independent non-interacting macromolecules, which differ in their size by only a few percent. The procedure requires the extrapolation of differential apparent sedimentation coefficient distributions obtained at different running time to $t \rightarrow \infty$ and works because it eliminates or greatly reduces diffusion effects. Here, we present an improved time extrapolation function of sedimentation distribution profiles originally presented by Stafford (In: Harding, Rowe, Horton (eds.) Analytical ultracentrifugation in biochemistry and polymer science, 1992). We describe a computing procedure with the program LAMM to analyze concentration profiles obtained by absorbance or interference optics that utilizes suitable smoothing methods for noisy data sets and present examples which include time invariant noises.

Keywords Computer simulation · Sedimentation coefficient distribution · Recognition of non-interactingmacromolecules

Introduction

Several methods were developed to detect independent components in sedimentation velocity experiments. The procedures success depends on eliminating or reducing the effects of diffusion (Baldwin and Williams 1950, Gralen and Lagermalm 1951, Van Holde and Weischet 1978). These methods are based on extrapolation of concentration profiles or sedimentation coefficient distributions obtained at different running times to $t \rightarrow \infty$. The information from several curves recorded at different time is more informative than that obtained from a single curve. Two newer approaches developed by Schuck (2000) and Schuck and Rossmanith (2000) are based on direct least squares modeling of sedimentation boundary using linear combinations of sedimentation profiles for non-diffusing species or Lamm equation solutions.

Stafford (1992) has presented some (model free) extrapolation procedures using apparent sedimentation coefficient distribution curves g*. Here we present improvements of Stafford's extrapolation equation. Smoothing methods for noisy data sets are also shown.

AUC&HYDRO 2008—Contributions from 17th International Symposium on Analytical Ultracentrifugation and Hydrodynamics, Newcastle, UK, 11–12 September 2008.

J. Behlke (✉) · O. Ristau
Max Delbrück Center for Molecular Medicine,
Robert Rössle Str. 10, 13125 Berlin, Germany
e-mail: behlke@mdc-berlin.de

Methods

Synthetic concentration profiles were calculated with the finite element method of Claverie et al. (1975). The following conditions were specified: rotor speed = 55,000 rpm, meniscus radius = 6.0 cm, bottom radius = 7.2 cm, time interval = 0.1 s, partial specific volume = 0.730 ml/g, frictional ratio = 1.2 for most data sets, summarized relative loading concentration $c/c_0 = 1$ and random noise = 1%.

The chosen column height is of advantage for separating neighboring boundaries as can be demonstrated in the following way. The moving of the boundary is described by Eq. 1

$$r = r_m \exp\left(\omega^2 s\, t\right) \approx r_m\left(1 + \omega^2 s\, t\right). \quad (1)$$

The radial distance between two species is

$$\Delta r = r_m(s_1 - s_2)\omega^2 t. \quad (2)$$

It grows with time. The maximal possible running time ($r = r_b$) is

$$t = \frac{r_b - r_m}{r_m \omega^2 s_1}. \quad (3)$$

The greatest possible distance is

$$\Delta r = (s_1 - s_2)(r_b - r_m)/s_1. \quad (4)$$

Thus the maximal radial distance of two species depends only by the column high. The rotor speed effects on the other hand the slope of the moving boundaries important for the width of the gaussian curve obtained by differentiation. Small bandwidth diminishes or avoids overlapping of neighboring gaussian curves. This effect can be demonstrated by differentiation, the approximate formula for the moving boundary (Fujita 1962, Eq. 2.147).

$$c = \frac{c_0}{2}\exp\left(-2\omega^2 s\, t\right)\left\{1 - erf\left(\frac{r_m\omega^2 s\, t - r_m \ln(r/r_m)}{2\sqrt{Dt}}\right)\right\} \quad (5)$$

$$\frac{\partial c}{\partial r} = \frac{c_0}{2\sqrt{\pi}}\exp\left(-2\omega^2 s\, t\right) \times \exp\left(-\frac{\left(r_m\omega^2 s\, t - r_m \ln(r/r_m)\right)^2}{4Dt}\right)\frac{r_m}{r\sqrt{Dt}} \quad (6)$$

For the midpoint (flexion point) of the boundary results

$$\frac{\partial c}{\partial r} = \frac{c_0}{2\sqrt{\pi}}\exp\left(-2\omega^2 s\, t\right)\frac{r_m}{r\sqrt{Dt}} \quad (7)$$

Together with the maximal possible running time Eq. 3 follows

$$\frac{\partial c}{\partial r} = \frac{c_0}{2\sqrt{\pi}}\exp\left[-2\left(r_b/r_m - 1\right)\right]\frac{r_m\omega}{r_b}\sqrt{\frac{s_1}{D(r_b/r_m - 1)}} \quad (8)$$

Therefore, the slope of the boundary is proportional to the rotor speed given by ω, which holds also to other running times. The bottom radius r_b is then to substitute for r_t the radius reached at running time t.

The best method to abolish time independent noise is the subtraction of a trace recorded at low-rotation speed (e.g., 3,000 rpm). Alternatively, the first data trace stored immediately after reaching the rotation speed can be used. Only the plateau region may be taken without a small meniscus and bottom region to cancel the influence of beginning sedimentation. The first possibility is of advantage when low-sedimentation coefficients ($s < 3$ S) are included because the lost meniscus and bottom regions are important to get complete bell-shaped derivatives.

The smoothing of noisy concentration profiles was done by repetition of a sliding 53-point centered least squares parabolic formula. The smoothed files were then differentiated again using a sliding 53-point centered least squares parabolic formula but now employing the first derivation at the center. Next, the derivatives were again repeatedly smoothed using the 53-point formula. The apparent sedimentation coefficient distribution was calculated with the Bridgman (1942) equation:

$$g^*(s) = \frac{\partial c}{\partial r}\omega^2 t\left(\frac{r}{r_m}\right)^2 \quad (9)$$

The s-values for abscissa were derived by

$$s = \frac{1}{\omega^2 t}\ln\left(\frac{r}{r_m}\right) \quad (10)$$

Common equidistant abscissa values (2001) for all curves at different running times and the corresponding $g^*(s)$-values were obtained by interpolation of all distribution curves using the natural cubic spline procedure SMOOTH (De Boor 2001). This procedure requires the expected standard deviation to be a smooth regulating parameter. This apparent standard deviation was calculated by least squares approximation of parabolic splines (De Boor 2001).

The fitting of the time extrapolation function to the $g^*(s)$-values at each s-value for different running times represents a linear regression with respect to the parameter. The estimation of the exponents of the time extrapolation function requires a non-linear regression. The solution of the normal equations in both cases was carried out using the Cholesky (square root) method.

Results

Stafford (1992) in his pioneering work has probed many time extrapolation functions for $g^*(s)$. To avoid negative estimated results we did not directly use $g^*(s)$ but rather the logarithmic form ln $[g^*(s)]$ was extrapolated. From Stafford's three tested extrapolation functions the following one yielded the best results:

$$\ln(g^*(s)) = A + B \cdot t^{-0.5} + C \cdot t^{-1} \quad (11)$$

Parameter A represents the logarithmic amplitude of $g(s)$ at $t \to \infty$. In order to enhance the resolution for two independent species we have optimized the exponents by non-linear regression. Tests with many synthetic sedimentation distribution profiles $g^*(s)$ showed that an optimal fit and resolution was obtained with the function

$$\ln(g^*(s)) = A + B \cdot t^{-0.304} + C \cdot t^{-0.723} \quad (12)$$

Synthetic distribution profiles were obtained by numeric differentiation of synthetic concentration profiles and use of the Bridgman equation, Eq. 9 (details are

described in the section "Methods"). Figure 1a shows the differentiated sedimentation distribution profiles and the time extrapolation calculated with Eq. 11, while Fig. 1b demonstrates the improvement calculated using Eq. 12. In this example the concentrations of the two components, 5 S and 6 S, were equal and the traces were simulated without random noise. The normal apparent sedimentation distribution plot (Bridgman) $g^*(s)$ also shown did not resolve the two components. However, when using the time extrapolation according to Eq. 11 or 12 the two components were visible and the observed data can be taken as start values for the fitting procedure.

To simulate experimental concentration profiles, 1% random noise was added. This made it necessary to smooth the concentration and the sedimentation distribution profiles. Here the number of repetitions was 200 (Fig. 2).

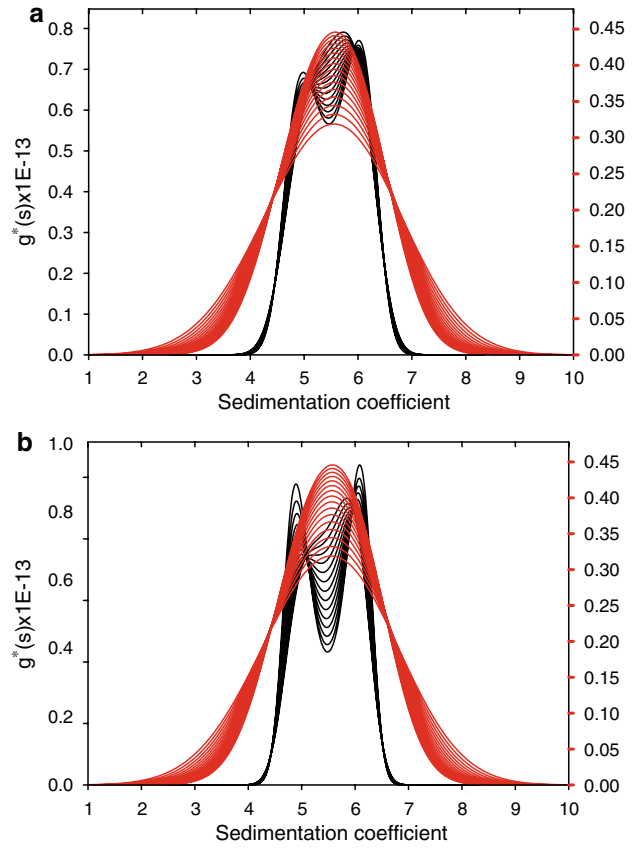

Fig. 1 **a** Sedimentation coefficient distribution plot (g^*, *right ordinate, red curves*) and the resolved time extrapolated presentation (*left ordinate, black curves*) of simulated concentration profiles for two species: with equal concentrations. $s_1 = 5$ S, $D_1 = 6.54\ 10^{-7}$ cm^2 s^{-1} and $s_2 = 6$ S, $D_2 = 5.97\ 10^{-7}$ cm^2 s^{-1}. Time extrapolation according to Eq. 11; **b** Sedimentation coefficient distribution plot (g^*, *right ordinate, red curves*) and the resolved time extrapolated presentation (*left ordinate, black curves*) of simulated concentration profiles for two species: with equal concentrations, $s_1 = 5$ S, $D_1 = 6.54\ 10^{-7}$ cm^2 s^{-1} and $s_2 = 6$ S, $D_2 = 5.97\ 10^{-7}$ cm^2 s^{-1}. Time extrapolation according to Eq. 12

The use of time extrapolation methods for sedimentation distribution profiles requires completely bell-shaped (gaussian) curves optimally produced by the differentiation. This means that the corresponding concentration profile must be depleted in the meniscus region and reach a plateau in the bottom region. Both conditions are necessary to get complete gaussian curves by differentiation, difficult to achieve for substances with low sedimentation coefficients. Such an example of an equal mixture of two components with $s = 2$ S and $s = 3$ S will be presented next. The use of a long path of the moving boundary and high-rotor speed are critical if the described conditions are to be fulfilled. A path of 1.2 cm and a rotor speed of 55,000 rpm are sufficient (Fig. 3).

The third example to be considered is a pair of components with high-sedimentation coefficients, $s = 20$ S and $s = 21.5$ S (Fig. 4). In this case the bottom plateau requirement is readily met. However, the influence of the bottom on the last narrow bell-shaped distribution curve is surprisingly high, stressing the requirement for a sufficiently long distance to the bottom for valid measurements. Although the new exponents were estimated with mixtures of two components, solutions of three or more components can also be analyzed. Figure 5 presents a mixture of equal amounts of three components with $s = 4$ S, $s = 5.5$ S, and $s = 7$ S. The smoothing procedure was repeated 50-fold.

Two further given examples are mixtures with different friction ratios and additionally added time independent noise. The first represents a mixture of two components with $s_1 = 5$ S and $s_2 = 6$ S as used in example 1 but with different frictional ratios 1.2 for the first and 1.5 for the second species (Figs. 6, 7). The elimination of the time independent noise was performed by subtraction of the first

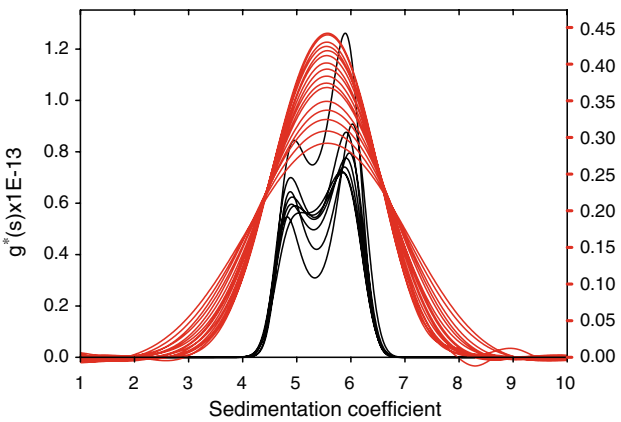

Fig. 2 Sedimentation coefficient distribution plot (g^*, *right ordinate, red curves*) and the resolved time extrapolated presentation (*left ordinate, black curves*) of simulated concentration profiles for two species: with equal concentrations. $s_1 = 5$ S, $D_1 = 6.54\ 10^{-7}$ cm^2 s^{-1} and $s_2 = 6$ S, $D_2 = 5.97\ 10^{-7}$ cm^2 s^{-1}, Time extrapolated by Eq. 12

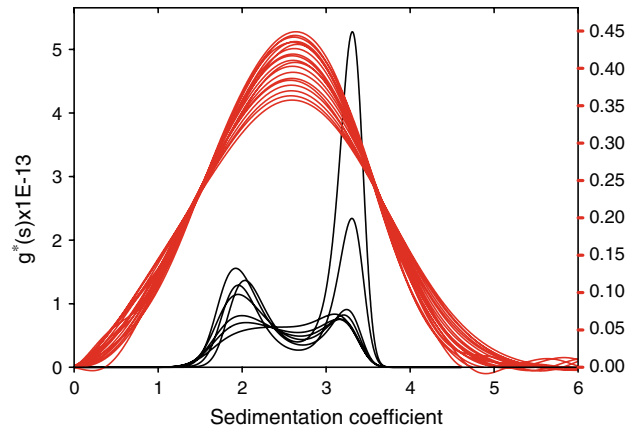

Fig. 3 Concentration profiles used in Fig. 2 for two species: with equal concentrations. $s_1 = 2$ S, $D_1 = 10.3\ 10^{-7}$ cm^2 s^{-1} and $s_2 = 3$ S, $D_2 = 8.44\ 10^{-7}$ cm^2 s^{-1}. Random noise 1%

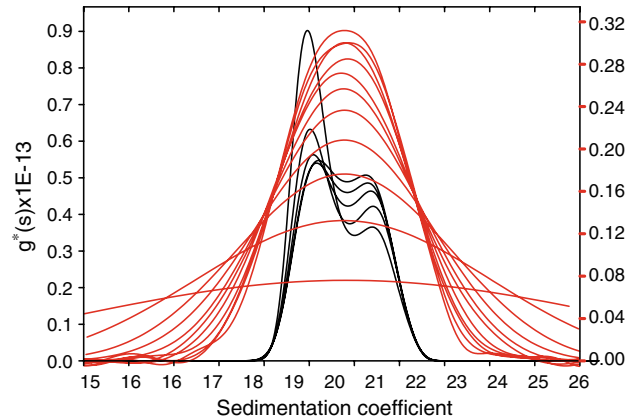

Fig. 4 Sedimentation coefficient distribution plot for two species: with equal concentrations. $s_1 = 20$ S, $D_1 = 3.27\ 10^{-7}$ cm^2 s^{-1} and $s_2 = 21.5$ S, $D_2 = 3.15\ 10^{-7}$ cm^2 s^{-1}. Random noise 1%

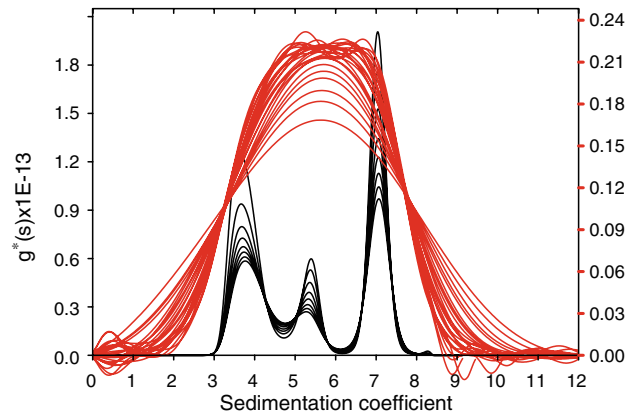

Fig. 5 Sedimentation coefficient distribution plot for three species with equal concentrations. $s_1 = 4$ S, $D_1 = 7.31\ 10^{-7}$ cm^2 s^{-1}; $s_2 = 5.5$ S, $D_2 = 6.23\ 10^{-7}$ cm^2 s^{-1} and $s_3 = 7$ S, $D_3 = 5.53\ 10^{-7}$ cm^2 s^{-1}. Random noise 1%

curve recorded at 2 min. The normal chosen time difference was 4 min. The smoothing procedure was repeated 10-fold.

The next example presented first by Brookes and Demeler (2008) is a mixture of four species with friction ratios 1.2, 1.6, 2.0, and 2.4 and molecular weights of 25, 75, 150, and 300 kDa (Figs. 8, 9). Here the TI-noise was eliminated by subtraction of a curve calculated at low-rotation speed (3,000 rpm). The base line is modified and must be estimated by the fitting procedure. The smoothing procedure was repeated 10-fold. The example in Fig. 8 shows the excellent resolution power (convergence) of our program LAMM. The time resolution method offers four poor s-values to start the fit procedure. $s_1 = 1.88$ S, $s_2 = 3.54$ S, $s_3 = 4.82$ S, and $s_4 = 6.64$ S. The results are (true values in brackets): $s_1 = 2.55$ (2.54) S, $s_2 = 3.94$ (3.96) S, $s_3 = 5.02$ (5.03), and $s_4 = 6.66$ (6.65) S.

Moreover the time extrapolation approach can separate mixtures of equal molecular weights but with different friction ratios and of course something different sedimentation coefficients. This possibility is demonstrated in the last example (Figs. 10, 11), a mixture of two species with the same molecular weight of 75 kDa but with frictional ratios of 1.2 and 1.35. The smoothing procedure was repeated 40-fold.

Discussion

Centrifugation is a powerful and generally applicable method for separating macromolecules based on their different mass or shape. The success to get correct results depends on (1) the quality of row data, (2) the start values, and (3) the chosen fitting procedure used in the software. Here, in order to optimize the start values, we have

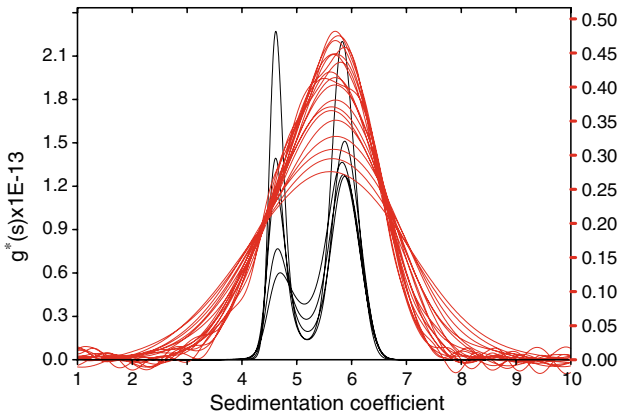

Fig. 6 Sedimentation coefficient distribution plot for two species with equal concentrations but different friction ratios $s_1 = 5$ S, $D_1 = 6.54\ 10^{-7}$ cm^2 s^1, $f/f_0 = 1.2$, $s_2 = 6$ S, $D_2 = 4.27\ 10^{-7}$ cm^2 s^{-1}, $f/f_0 = 1.5$

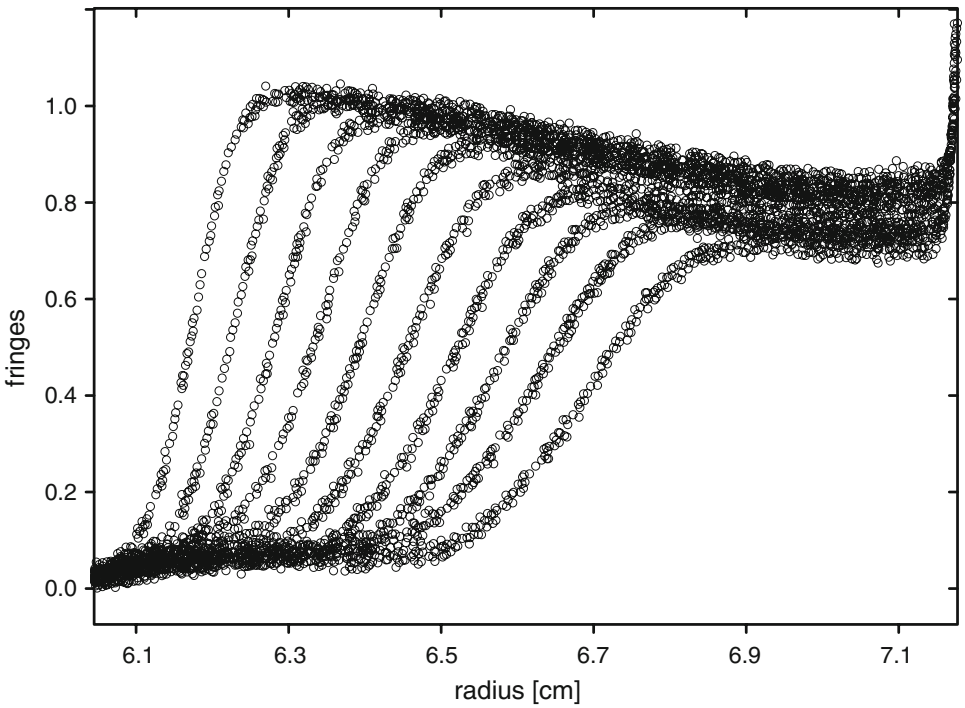

Fig. 7 The original concentration profiles used in Fig. 6. Only every second trace is shown. Random noise is 1%

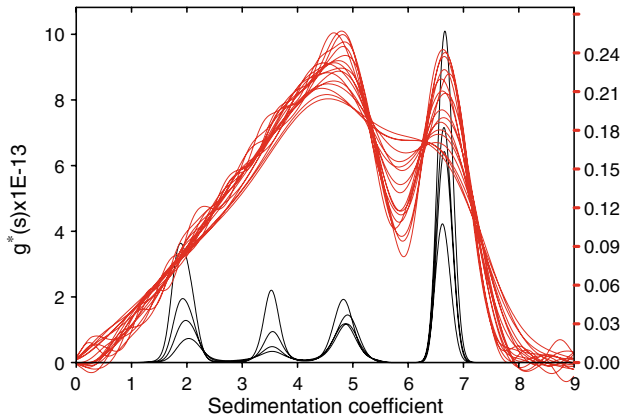

Fig. 8 Sedimentation coefficient distribution plot for four species with equal concentrations: $s_1 = 2.54$ S, $D_1 = 9.17\ 10^{-7}$ cms^{-1}; $f/f_0 = 1.2$, $s_2 = 3.96$ S, $D_2 = 4.77\ 10^{-7}$ cm^2 s^{-1}, $f/f_0 = 1.6$, $s_3 = 5.03$ S, $D_3 = 3.03\ 10^{-7}$ cm^2 s^{-1}, $f/f_0 = 2.0$, $s_4 = 6.65$ S, $D_4 = 2.00\ 10^{-7}$ cm^2 s^{-1}, $f/f_0 = 2.4$

improved successfully Stafford's (1992) time extrapolation function $g^*(s)$. In combination with the program LAMM (Behlke and Ristau 2002) that works advantageous with derivations instead the SIMPLEX METHOD, we are able to analyze two or more components with relatively narrow molecular weights *on a personal computer*. To make it more difficult with regard to the depressing diffusion effects in most cases we have considered the behavior of spherical macromolecules with a frictional ratio of 1.2. However, the empirically derived exponents for the time extrapolation in Eq. 12 are reliable for different mixtures and variable frictional ratios. In order to simulate the experimental conditions better most data sets were supplied with 1% noise.

The successful analysis of mixtures with different non-interacting components requires some important experimental conditions. The number of scans should be high enough since a small number of noisy curves did not result in reliable or expected data. Therefore, the running time interval must be small. For low-sedimentation coefficients $s < 6$ about 4 min are reasonable, while for higher s-values 3 min are suitable. Extremely important for resolution is a long path of the moving boundary and a column height of 1.2 cm might be necessary. High-rotor speed makes the bell-shaped distribution curves narrower and supports better resolution. The best resolution being obtained with profiles near the bottom were elicited by a relatively long running time. Resolution is diminished when there are many profiles near the meniscus, which do not deplete. Thus, optimal conditions can be obtained as follows: One should start with the last concentration profile near the bottom (high file number) and then with the first one near the meniscus (low file number). LAMM calculates the extrapolated distribution curve and displays the result. Next, the number of profiles included in the extrapolation can be reduced by increasing the file number for the first one until resolution appears. In our program LAMM this can be done by pressing the button "<"or vice versa ">". Important for the resolution is the choice of the correct smoothing sequence. Too intensive smoothing (to high repeat number) can prevent resolution while too low smoothing can deform the spectrum and even simulate

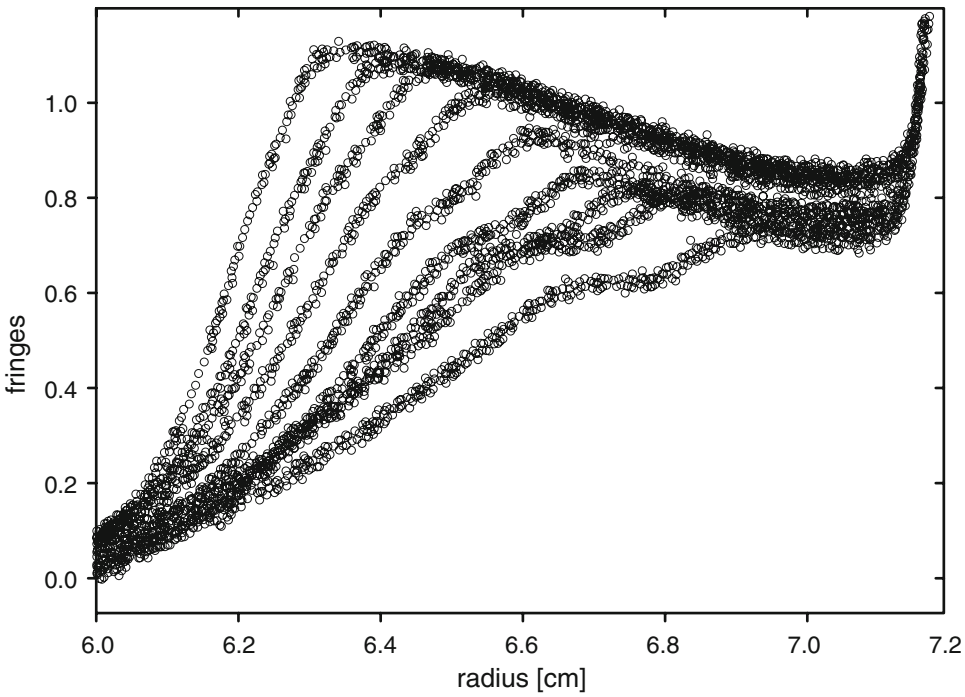

Fig. 9 The original concentration profiles used in Fig. 8. Only every second trace is shown. Random noise is 1%

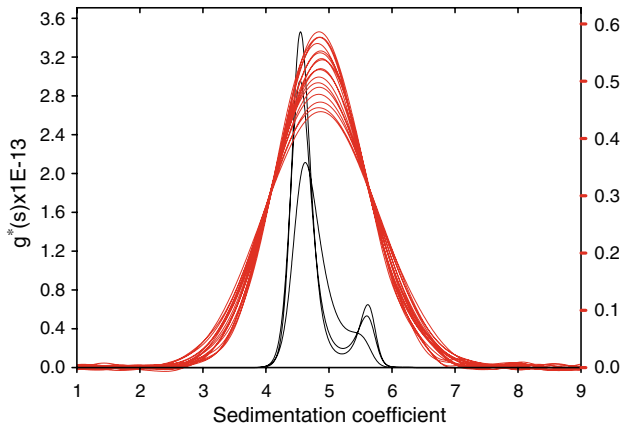

Fig. 10 Sedimentation coefficient distribution plot for two species with equal molecular weight of 75 kDa but different friction ratios. $s_1 = 4.70, D_1 = 5.65, ff_0 = 1.35, c_1/c_0 = 0.7$. $s_2 = 5.28, D_2 = 6.36, ff_0 = 1.2, c_2/c_0 = 0.3$

additional bands. Frequently the sedimentation coefficients obtained are not very accurate but sufficient as start values for our program LAMM (Behlke and Ristau 2002).

The surprising analysis of two species with equal molecular weights but different frictional ratios is shown in Fig. 10. For such mixtures we can expect only one band in experiments with SDS gel electrophoresis but clearly two substances with the correct molecular weight in sedimentation velocity experiments. The concentration ratios for this case are 0.7 and 0.3. This may be the limit concentration ratio for clear recognition.

The fit quality obtained with LAMM also helps to insure that the correct number of components will be identified. It is crucial to keep uniform deviations as small as possible. Although the time extrapolation of sedimentation distribution profiles is a useful approach to detect independent components, especially in combination with the fit to the original data with a suitable program like LAMM it does not provide absolute evidence. The resolving power is limited by the concentration ratio of the components. When the concentration ratio c_1/c_2 is lower than 0.1 resolution disappears.

More resolution as the normal differential sedimentation coefficient distribution function $g(s)$ promises the size distribution analysis $c(s)$ developed by Schuck (2000), Schuck et al. (2002) in which corrections for diffusion are made. But the $c(s)$ method cannot separate such strongly overlapping systems as presented in this article. The mixture with three different species yields only two peaks missing one component (the recommended confidence level 0.95 was used). The example borrows from Brookes and Demeler (2008) with four species (see Figs. 8, 9) shows satisfying resolution. As can be seen sedimentation coefficients are sufficiently different indicating some peaks in the normal differential sedimentation distribution plot. The $c(s)$ method of Schuck (2000) is well suitable for simultaneously fitting of many species with sufficient different sedimentation coefficients and similar friction ratios. The two-dimensional spectrum analysis performed in parallel computational techniques developed by Brookes and Demeler (2008) seems the best method to separate neighboring sedimentation profiles but this approach is not suitable for normal desk computer. The restriction to one

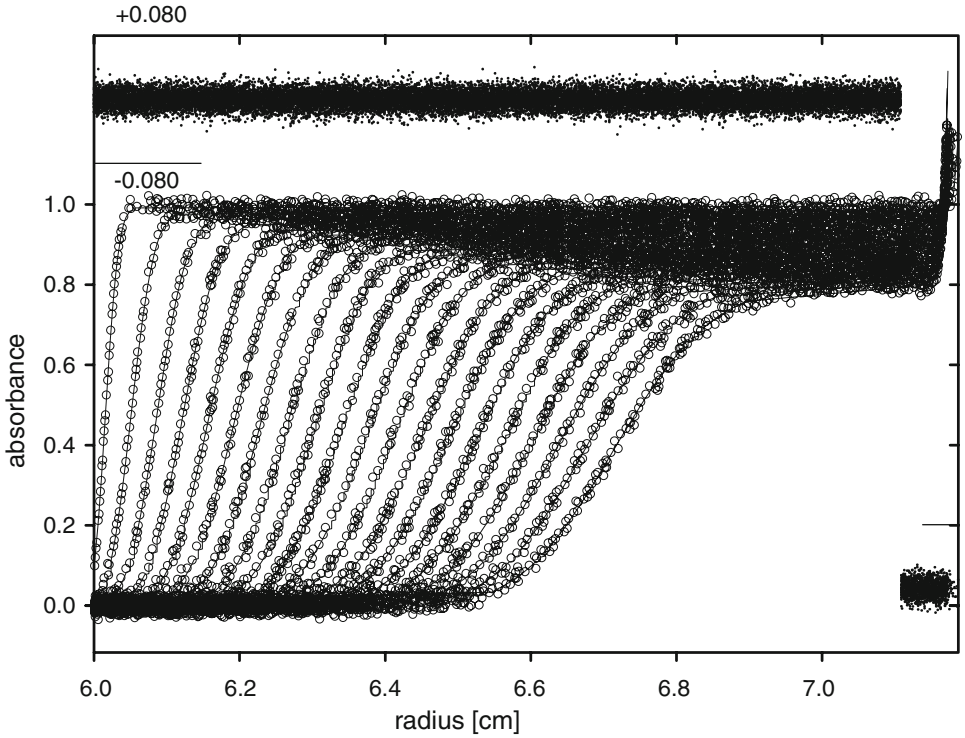

Fig. 11 The original concentration profiles used in Fig. 10. Only every second trace is shown. Random noise is 1%

common friction ratio as in $c(s)$ is abolished. The method of Van Holde and Weischet (1978) improved by Demeler and Van Holde (2004) can be used for estimating the limits of the sedimentation coefficients usefully for the two-dimensional program. In case of sufficiently different sedimentation coefficients estimation of discrete values is possible. Comparing with all the discussed programs the here presented approach is a suitable, robust, and fast working alternative variant useful for personal computer.

References

Baldwin RL, Williams JW (1950) Boundary spreading in sedimentation velocity experiments. J Am Chem Soc 72:4325. doi:10.1021/ja01165a554

Behlke J, Ristau O (2002) A new approximate whole boundary solution of the Lamm differential equation for the analysis of sedimentation velocity experiments. Biophys Chem 95:59–68. doi:10.1016/S0301-4622(01)00248-4

Bridgman WB (1942) Some physical chemical characteristics of glycogen. J Am Chem Soc 64:2349–2356. doi:10.1021/ja01262a037

Brookes E, Demeler B (2008) Parallel computational techniques for the analysis of sedimentation velocity experiments in ultrascan. Colloid Polym Sci 286:139–148. doi:10.1007/s00396-007-1714-9

Claverie J-M, Dreux H, Cohen R (1975) Sedimentation of generalized systems of interacting particles. I. Solution of systems of complete Lamm equations. Biopolymers 14:1685–1700. doi:10.1002/bip.1975.360140811

De Boor C (2001) A practical guide to splines. Springer Verlag, New York

Demeler B, Van Holde KE (2004) Sedimentation velocity analysis of highly heterogeneous systems. Anal Biochem 335:279–288. doi:10.1016/j.ab.2004.08.039

Fujita H (1962) Mathematical theory of sedimentation analysis. Academic Press, New York

Gralen N, Lagermalm G (1951) A contribution to the knowledge of some physico-chemical properties of polystyrene. J Phys Chem 56:514–523. doi:10.1021/j150496a025

Schuck P (2000) Size-distribution of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophys J 78:1606–1619

Schuck P, Rossmanith P (2000) Determination of the sedimentation coefficient distribution by least squares boundary modeling. Biopolymers 54:328–341. doi:10.1002/1097-0282(20001015)54:5<328::AID-BIP40>3.0.CO;2-P

Schuck P, Perugini MA, Gonzales NR, Howlett GJ, Schubert D (2002) Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. Biophys J 82:1096–1111

Stafford WF (1992) Methods for obtaining sedimentation coefficient distributions. In: Harding SE, Rowe AJ, Horton JC (eds) Analytical ultracentrifugation in biochemistry and polymer science. Royal Society of Chemistry, Cambridge, England, pp 359–393

Van Holde KE, Weischet WO (1978) Boundary analysis of sedimentation velocity experiments with monodisperse and paucidisperse solutes. Biopolymers 17:1387–1403