Supporting Information

Precision of Taylor Dispersion

Patricia Taladriz-Blanco[a], Barbara Rothen-Rutishauser[a], Alke Petri-Fink,[a,b] and Sandor Balog[a]*

[a]Adolphe Merkle Institute, University of Fribourg, Chemin des Verdiers 4, 1700 Fribourg, Switzerland

[b]Chemistry Department, University of Fribourg, Chemin du Musée 9, 1700 Fribourg, Switzerland

ABSTRACT: Taylor dispersion is capable of measuring accurately the hydrodynamic radius over several orders of magnitude. Accordingly, it is now a highly competitive technique dedicated to characterizing small molecules, proteins, macromolecules, nanoparticles and their self-assembly. Regardless, an in-depth analysis addressing the precision of the technique, being a key indicator of reproducibility, is not available. Benefiting from analytical modeling and statistical analysis, we address error propagation and present a comprehensive theoretical study of the precision of Taylor dispersion. Theory is then compared against experiment, and we find full consistency. Our results are most helpful when either the design, objectives, or control of analytical quality is in focus.

*Corresponding author: Sandor Balog (sandor.balog@unifr.ch)

Contents

The signal-to-noise ratio of taylorgrams

It was shown before that the signal-to-noise ratio (SN) of A, defined as the ratio of the mean and standard deviation of $P(A)$, is $SN(A) \cong \sqrt{N}(1-T)$ when $T > 0.5$. $N = \alpha\, I\, \tau$ is the number of the photons illuminating the flow during τ integration time (temporal resolution), α a detector-and wavelength-specific constant, and I the intensity of the illuminating light.[1]

This result is obtained via basic principles considering the quantized nature of light and the fundamentals of photon detectors and UV-Vis spectroscopy. To measure the transmission of an analyte and to determine its absorbance, the intensity of the transmitted light is measured. Measuring the intensity of light is never instantaneous, but involves a detection time interval $\tau > 0$. Detecting photons is intrinsically random, and the consequence is that the number of photons detected during τ is a random variable. In other words, if one measures the absorbance of a sample k times under the very same conditions, one tends to obtain k different results. This is because even if the intensity of the illumination is completely stable, the probability density of the photon counts follows a Poisson distribution. This randomness is an inherent property of classical linear spectroscopy and referred to as shot noise. Accordingly, measuring transmission (T) and absorbance (A) is also probabilistic, and thus noisy. Due to the relatively fine temporal resolution, this stochastic character becomes relevant to Taylor dispersion where a long continuous sequence of short and single measurements is required to resolve the dynamics of the band dispersion, while measuring the transmission of the solvent background may be a single and considerably longer measurement. By starting from the Poisson distribution of the photon counts, and by applying the rule of transforming random variables, we calculated the probability density of the absorbance when the intensity of illumination is precisely known and the influence of particle number density fluctuations[2] is negligible: $P(A) = Ln10 \cdot \mathrm{e}^{-N \cdot T} \cdot N^{u} \cdot T^{u} / \Gamma[u]$ where $u = N \cdot 10^{-A}$, $N = \alpha \cdot I \cdot \tau$ is the number of the photons illuminating the flow during τ integration time, α a detector-and wavelength-specific constant, I the intensity of the illuminating light reaching the flow, T the value of the transmission, and Γ the gamma function. It can be shown that when $N > 100$, the signal-to-noise ratio (SN) of A—defined as the ratio of the mean and standard deviation of $P(A)$—is

$$SN(A) = -Log_{10}T \cdot (Log_{10}2 + 2) \cdot \sqrt{T \cdot N}\,. \qquad \text{(S1)}$$

The mean value of the absorbance is equal to $-Log_{10}T$, which is in fact the 'true' value one would always measure in the absence of shot noise. When $T > 0.5$, eq 15 simplifies:

$$SN(A) \cong \sqrt{N} \cdot (1-T). \qquad \text{(S2)}$$

Therefore, the expected value of SN is not a constant over the course of a measurement, and it is the highest at the peak of the Taylorgram. The signal-to-noise ratio can be improved by increasing either a) the initial concentration of the sample, b) the capillary diameter, c) the detection area, d) the intensity of the illuminating lamp or by decreasing the temporal resolution. The concentration can be increased as long as a) the Lambert-Beer law remains applicable and b) inter-particle interactions remain negligible. The maximum SN values of the taylorgrams of Figure 1 are shown below.

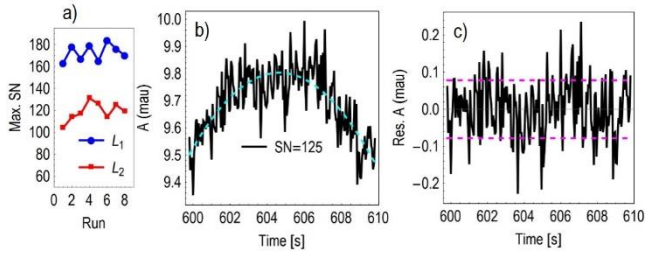


Figure S1. a) The maximum signal-to-noise ratio of the taylorgrams shown in Figure 1. b) Each SN value was estimated by the best fit of eq 1 (dashed line) using a ±5s interval around the center of each taylogram. c) The residuals represent noise, and the dashed lines indicate the ±1 standard deviation. The maximum value of the signal-to-noise ratio varies from run to run, and as expected, the absorbance is decreasing upon dispersion, and thus SN higher for the shorter residence time ($L_1 < L_2$).

The uncertainty of determining the width and center of a taylorgram

To determine the parameters in eq 3, we simulated and analyzed 100000 fully realistic taylorgrams. To construct each taylorgram, each parameter was drawn randomly and independently from uniform distributions corresponding to realistic scenarios: $0.1\ s \le \kappa \le 10\ s$, $300\ s \le\ t_0 \le 1000\ s$, $0.05\ s \le \tau \le 1\ s$, $10 \le SN \le 200$, and $0.75 \le\ T \le 0.95$. Each noiseless taylorgram was created with eq 1 and a temporal resolution of τ considering that the measurement of absorbance is not instantaneous but involves values averaged over τ. Noise was included for each data point of the taylorgram independently by replacing the 'clean' absorbance value with a single variable drawn randomly from the probability density distribution of the absorbance $P(A)$ corresponding to the minimum transmission value and signal-to-noise ratio. Representative examples are shown in (Figure S2). Eq 1 was fitted against the simulated noisy taylorgrams, via minimizing residual least squares without any constraint. The best-fitting parameters and their standard errors were recorded. When the two five-dimensional $(4+1)$ data sets were completed, eq 3 was fit without any constraints again via minimizing residual least squares. In both cases, the fit converged quickly and the quality of the fit was high ($R^2 > 0.998$). These indicate that the two expressions describe very well the results obtained by the simulation experiments. Next, we blind-tested the validity of the parameters and eq 3 by simulating and analyzing 'random' taylorgrams as described above, and by comparing the uncertainties predicted by eq 3 with the uncertainties obtained after fitting eq 1 to the taylorgrams. As shown Figure S3, the agreement between the two sets of data is very good, and therefore, this approach is a reliable source of estimating uncertainty of κ and t_0.

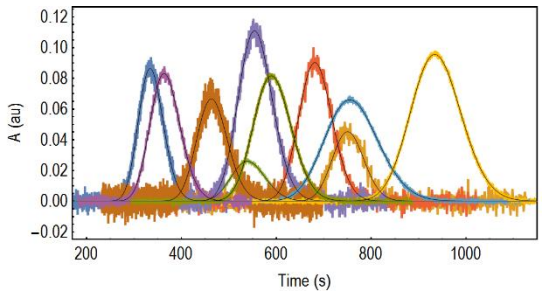


Figure S2. Ten simulated taylorgrams of the 100000 with varying width, center, temporal resolution, noise level, and amplitude. We used such taylorgrams to determine the expectable precision of measuring κ and t_0 via fitting eq 1 (black lines).

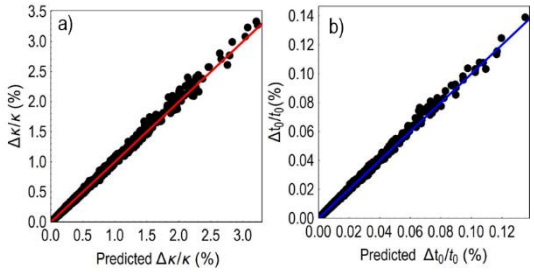
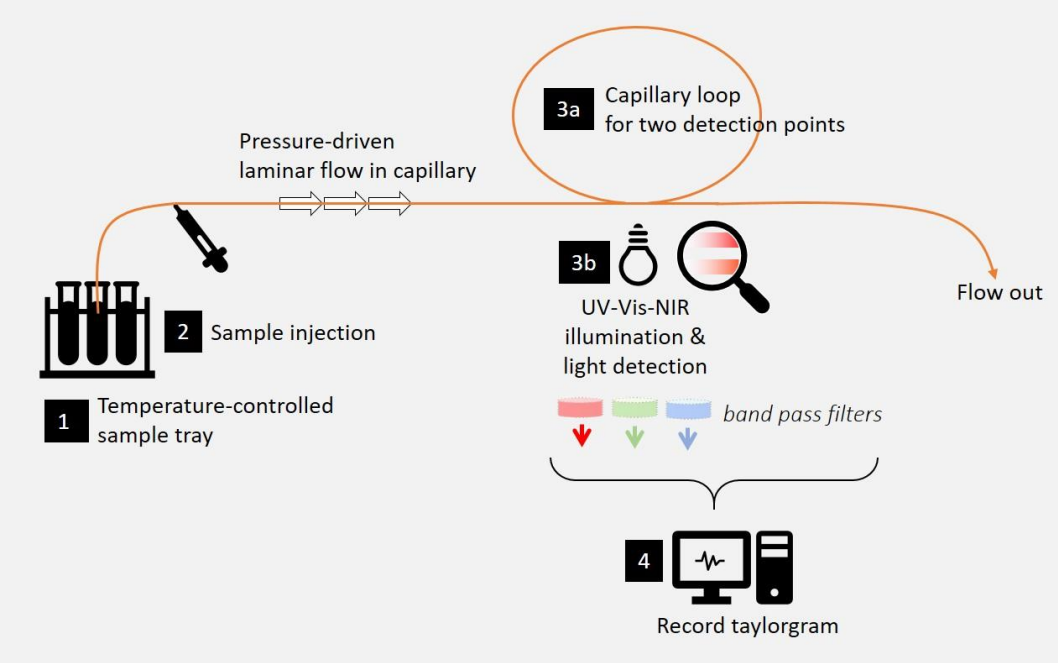
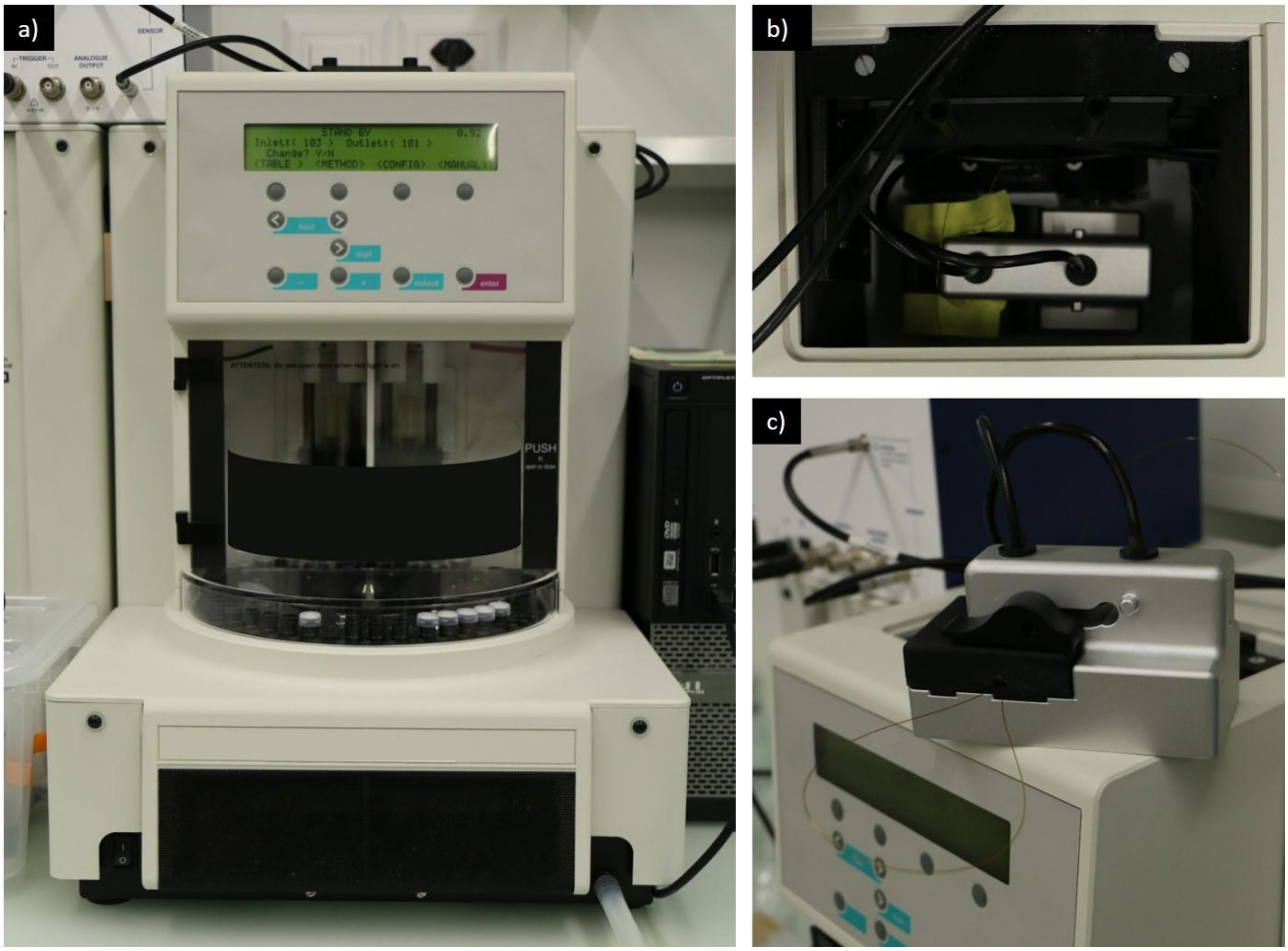


Figure S3. Predicted and actual precision obtained via simulation experiments. The precision of determining κ and t_0 agree with a Pearson correlation coefficient of 0.998 and 0.997, respectively.

Taylor dispersion experiments of BSA

For the experiments, taylorgrams of a dilute aqueous dispersion (0.5g/L) of bovine serum albumin (BSA, ≥ 98%, lyophilized powder, product number: Sigma-Aldrich-A7906) were collected at a sampling rate of 20 Hz (τ=0.05s), using an ActiPix D100 UV-Vis area imaging detector (Paraytec, York, UK). The detector automatically corrected the intensity values for dark current, controlled the intensity of illumination (by a pulsed xenon lamp) and performed the background measurement on the respective running buffer before each run. A band-pass filter, 214 nm center wavelength with a 22 nm FWHM, was used (Edmund Optics, York, UK). The absorbance of the dispersion was therefore measured in the linear range of the detector throughout the measurements, by considering variations in illumination, background, and dark current. Aliquots with a volume of nearly 32 nL and at a nominally constant temperature of 25 °C were injected into a fused silica capillary ($2Y$=74.5±3 μm inner diameter), Polymicro Technologies, Phoenix, USA) using a capillary electrophoresis injection system (Prince 560 CE Autosampler, Prince Technologies B.V., Netherlands). The running buffer was MilliQ water. After sample injection a pressure of P=90±0.9 mbar drove the samples through the capillary. The total capillary length was L=145±0.05 cm, and the distances between injection and the two detection points were L_1=37±0.05 and L_2=72±0.05 cm. Accordingly, in our experiments, the duration of the pressure ramp (1.8s) was negligible compared to the residence time reached at the second detection window, and we did not apply any correction.[3] Yet, the hydrodynamic radius determined at this window via eq 2 is expected to be highly accurate. The layout of our instrument is shown below.





Scheme S1. Top: The layout of our TDA instrument, where the basic operation is alike to that of liquid chromatography. A pressure-driven laminar flow carries the injected sample taken from a temperature-controlled sample tray, and the optical extinction of the dispersed band is observed at fixed position along the capillary, using a selected wavelength by means of a set of exchangeable filters. In many cases, one obtains two detection points by looping the capillary. Bottom: a) The main housing and front panel of the instrument. b) The illumination and detection unit inserted from top. c) The illumination and detection unit securing the capillary loop.

Model parameters

To fit eq 1 against the experimental data, we used unconstrained nonlinear model fit with three parameters: amplitude, residence time, and width parameter (κ). The taylograms and the corresponding best fits are shown in Figure 1. The tables below also list the coefficient of determination (R^2) and the maximum signal-to-noise ratio for each taylorgram.

Meas./Window: 1/1
Coefficient of determination: 0.999432
Amplitude: 0.509726
Residence time: 343.999 s
Δ(Residence time): 0.00924004 s
κ: 1.52675 s
Δκ: 0.0019916 s
Max. signal-to-noise ratio: 163

Meas./Window: 1/2
Coefficient of determination: 0.999281
Amplitude: 0.500775
Residence time: 629.498 s
Δ(Residence time): 0.0127473 s
κ: 1.55351 s
Δκ: 0.00210706 s
Max. signal-to-noise ratio: 104

Meas./Window: 2/1
Coefficient of determination: 0.99992
Amplitude: 0.514149
Residence time: 333.144 s
Δ(Residence time): 0.00869571 s
κ: 1.46516 s
Δκ: 0.00185109 s
Max. signal-to-noise ratio: 178

Meas./Window: 2/2
Coefficient of determination: 0.999934
Amplitude: 0.518693
Residence time: 613.346 s
Δ(Residence time): 0.0123471 s
κ: 1.56265 s
Δκ: 0.00206759 s
Max. signal-to-noise ratio: 114

Meas./Window: 3/1
Coefficient of determination: 0.999973
Amplitude: 0.511152
Residence time: 339.243 s
Δ(Residence time): 0.00892978 s
κ: 1.49932 s
Δκ: 0.0019139 s
Max. signal-to-noise ratio: 167

Meas./Window: 3/2
Coefficient of determination: 0.999979
Amplitude: 0.503094
Residence time: 619.746 s
Δ(Residence time): 0.0123322 s
κ: 1.5242 s
Δκ: 0.00202475 s
Max. signal-to-noise ratio: 117

Meas./Window: 4/1
Coefficient of determination: 0.999986
Amplitude: 0.509312
Residence time: 330.199 s
Δ(Residence time): 0.0087343 s
κ: 1.45483 s
Δκ: 0.00185794 s

Max. signal-to-noise ratio: 179

Meas./Window: 4/2
Coefficient of determination: 0.99999
Amplitude: 0.512447
Residence time: 608.022 s
Δ(Residence time): 0.0121762 s
κ: 1.53822 s
Δκ: 0.00202479 s
Max. signal-to-noise ratio: 131

Meas./Window: 5/1
Coefficient of determination: 0.999992
Amplitude: 0.504491
Residence time: 329.805 s
Δ(Residence time): 0.00860659 s
κ: 1.44819 s
Δκ: 0.00182662 s
Max. signal-to-noise ratio: 165

Meas./Window: 5/2
Coefficient of determination: 0.999995
Amplitude: 0.50832
Residence time: 606.942 s
Δ(Residence time): 0.0115633 s
κ: 1.54251 s
Δκ: 0.00192755 s
Max. signal-to-noise ratio: 126

Meas./Window: 6/1
Coefficient of determination: 0.999995
Amplitude: 0.507374
Residence time: 328.783 s
Δ(Residence time): 0.00865999 s
κ: 1.44868 s
Δκ: 0.00184061 s
Max. signal-to-noise ratio: 184

Meas./Window: 6/2
Coefficient of determination: 0.999996
Amplitude: 0.511283
Residence time: 605.543 s
Δ(Residence time): 0.0120396 s
κ: 1.53114 s
Δκ: 0.00199913 s
Max. signal-to-noise ratio: 114

Meas./Window: 7/1
Coefficient of determination: 0.999996
Amplitude: 0.514707
Residence time: 328.137 s
Δ(Residence time): 0.00853297 s
κ: 1.44155 s
Δκ: 0.00180967 s
Max. signal-to-noise ratio: 176

Meas./Window: 7/2
Coefficient of determination: 0.999997
Amplitude: 0.518736
Residence time: 604.791 s
Δ(Residence time): 0.0115645 s
κ: 1.53007 s
Δκ: 0.00192025 s
Max. signal-to-noise ratio: 125

Meas./Window: 8/1
Coefficient of determination: 0.999997
Amplitude: 0.502297
Residence time: 328.039 s
Δ(Residence time): 0.00857453 s
κ: 1.45404 s
Δκ: 0.00182816 s
Max. signal-to-noise ratio: 170

Meas./Window: 8/2
Coefficient of determination: 0.999998
Amplitude: 0.506398
Residence time: 604.613 s
Δ(Residence time): 0.0116744 s
κ: 1.53219 s
Δκ: 0.00194041 s
Max. signal-to-noise ratio: 119

The viscosity of water as a function of temperature

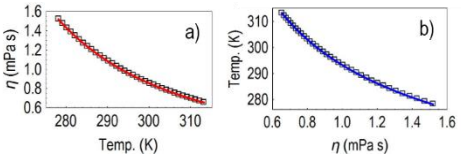


Figure S4. The viscosity of water on the temperature range of 5-40 °C. a) The viscosity as a function of temperature. The solid line is an unconstrained nonlinear least squares fit of eq 6 ($R^2 = 1$) with a $= 38.41 \pm 0.44\ 10^{-6}$ Pa s, $b = 432.91 \pm 3.02\ K^{-1}$, $c = 160.4 \pm 0.46\ K$ (mean ± standard error). b) Temperature against viscosity. The solid line is the inverted phenomenological relationship: $f^{-1}(\eta) = c + b/Ln(\eta/a)$.

Capillary flow temperature by light absorption and heat dissipation

We construct a simple a model where the sample temperature is influenced by two phenomena: 1) light absorption generating heat and 2) loss of heat owing to an imperfect insulation of the system from its environment. Given a moderate temperature difference between the capillary flow and its environment (room, housing), the loss of heat can be described by the Fourier law. Furthermore, it is a fair assumption that the temperature of the environment was constant during the measurements, and so were and the parameters describing heat absorption (A) and heat loss (k). According to these assumptions, we construct the following differential equation

(S3) $$\frac{dT(t)}{dt} = A - k(T(t) - T_0)$$

with the initial condition of $T(0) = T_0$.

The solution describes the temperature as a function of time:

(S4) $$T(t) = \frac{A}{k}(1 - e^{-kt}) + T_0.$$

The total length of a single measurement was nearly 15 minutes, and thus the measurement numbers (1-8) were converted into time accordingly. The two parameters $A[°Ch^{-1}]$ and $k[h^{-1}]$ describe the rate of temperature increase and heat loss. The maximum temperature this system can reach is $A/k + T_0$.

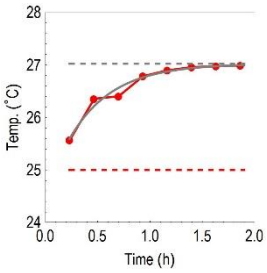


Figure S5. The temperature as a function of time, and the best unconstrained nonlinear fit of eq S4 (solid gray line). The parameters are $A = 6.07\ °C\ h^{-1}$, $k = 2.43\ h^{-1}$, and $T_0 = 24.52\ °C$. The *nominal* temperature is indicated by the dashed red line. The dashed gray line indicates the maximum temperature this system could reach in this series of experiments.

Sampling distributions of normally distributed random variables

Another important concept relevant to the precision of measurements relates to inferential statistics (Figure S6). From this point of view, each set of individual experiments performed with nominally identical parameters, usually referred to as repeats, represents a random sample drawn from the population. Therefore, the sample statistics, such as the mean and standard deviation vary from sample to sample, and the magnitudes of these variations are functions of the sample size, *i.e.*, the number of repeats. These variations can be described accurately by the sampling distributions of population mean (μ) and population standard deviation (σ). (Here $\mu = r$ and $\sigma = \Delta r$.) When μ and σ describe a normal distribution, the sampling distribution of the measured mean is Gaussian with a mean equal to the population mean, $\mu_s = \mu$, and with a standard deviation equal to $\sigma_s = \sigma/\sqrt{n}$. The latter is called as the standard error of the mean, where n is the size of the sample. Given that the Gaussian function is symmetric, and the most probable value equals exactly μ, the sample mean is an accurate estimate of the population mean, even if the sample size is small. However, this is not the

case for precision. Regarding the sampling distribution of the standard deviation $p(\sigma_s)$, it is shown by transforming the χ^2-distribution with $n-1$ degrees of freedom that $p(\sigma_s)$ is generally asymmetric:[4]

$$p(\sigma_s) = \frac{2^{\frac{3-n}{2}} \cdot e^{-\frac{n\sigma_s^2}{2\sigma^2}} \cdot n^{\frac{1}{2}(n-1)} \cdot \sigma^{1-n} \cdot \sigma_s^{n-2}}{\Gamma\left(\frac{n-1}{2}\right)} \tag{S5}$$

where Γ is the Euler gamma function. The expected mean and standard deviation (STD) of $p(\sigma_s)$ are

$$\langle \sigma_s \rangle = \int_0^\infty \sigma_s \, p(\sigma_s) \, d\sigma_s = \sigma \sqrt{\frac{2}{n}} \frac{\Gamma(n/2)}{\Gamma\left(\frac{n-1}{2}\right)} \tag{S6}$$

and

$$\sqrt{\langle \sigma_s^2 \rangle - \langle \sigma_s \rangle^2} = \sigma \sqrt{\left(\frac{s-1}{s} - \frac{2\,\Gamma(n/2)^2}{s\,\Gamma\left(\frac{n-1}{2}\right)^2} \right)} \tag{S7}$$

As shown in Figure S7, sample-to-sample variations have important consequences. First, the expectable measure of precision $\langle \sigma_s \rangle$ is a function of the sample size. For example, in the case of triplicates one measures three times, and the standard deviation is used as a metric of precision. By calculating the cumulative probability of $p(\sigma_s)$, it is little effort to show that at this sample size the measure of precision is not reliable at all. When $n = 3$ the probability that $\sigma_s < \sigma$ is nearly 78%. In other words, the chance of underestimating the standard deviation and overestimating the precision is considerable. The chance is 65% when $n = 10$, 57% when $n = 50$, and 55% when $n = 100$. Second, both repeatability and reproducibility are poor when the sample size is not particularly large, because the width of the distribution of σ_s is generally large. For example, when $n = 3$ the width is more than twice as large as the mean value (Figure S7b).

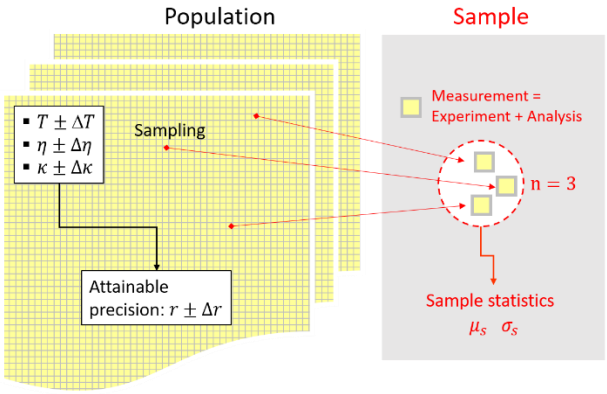


Figure S6. The view of inferential statistics on measurement analysis. A single measurement consists of an experiment and its analysis. The population is the ensemble of all the realizable measurements described by the given uncertainties. These parameters ultimately define the attainable precision of measuring the hydrodynamic radius. In contrast to the population, a sample is usually a small set of individual measurements (replicates), whose statistics display sample-to-sample variations following the corresponding sampling distributions. These sample-to-sample variations become negligible only when the sample size (n) is large.

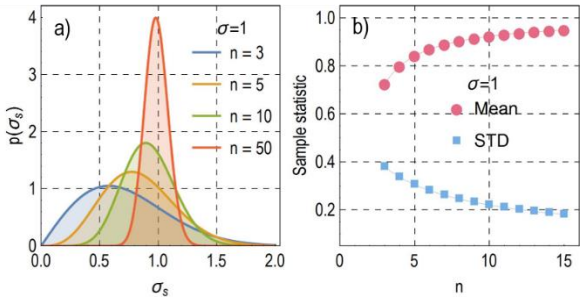


Figure S7. Sampling distribution of the standard deviation. a) The probability density of the sample standard deviation at different sample sizes when the population standard deviation is one. b) The expected sample mean value and sample standard deviation (STD) of $p(\sigma_s)$ as a function of the sample size.

The sampling distribution of the relative standard deviation as a function of sample size ($\delta_s = \sigma_s/\mu_s$) was estimated via McKay's approximation.[5-7] The approximation addresses finite-size samples drawn randomly from a normally distributed population, and defines a variable K_s being the following function of the sample statistic:

$$K_s(\delta, n;\ \delta_s) = \left(1 + \frac{1}{\delta^2}\right) \frac{(n-1)\delta_s^2}{1 + \frac{n-1}{n}\delta_s^2}. \tag{S8}$$

δ and δ_s are the population and sample coefficient of variation, respectively, and n is the sample size. McKay showed that when $\delta < 1/3$, the PDF of K_s can be described by a central χ^2-distribution with $s-1$ degrees of freedom:

(S9) $$f_{n-1}(K_s) = \begin{cases} \frac{2^{\frac{1-n}{2}} \cdot e^{-\frac{K_s}{2}} \cdot K_s^{\frac{n-3}{2}}}{\int_0^\infty K_s^{\frac{n-3}{2}} e^{-K_s}\, dK_s} & K_s > 0 \\ 0 \end{cases}$$

To calculate the probability density of the sample coefficient of variation $p(\delta_s)$, we transform McKay's approximation, and we obtain the sampling distribution of the relative standard deviation of normally distributed random variables where δ is the population parameter and n is the sample size.

(S10) $$p(\delta_s) = f_{s-1}\left(K_s(\delta_s)\right) \times \left|\frac{\partial K_s}{\partial \delta_s}\right|$$

(S11) $$p(\delta, n; \delta_s) = \frac{2^{\frac{3}{2}-\frac{n}{2}} \cdot \sqrt{\frac{\left(\delta^2+1\right)^2 (n-1)^2 n^4 \delta_s^2}{\delta^4 \left((n-1)\delta_s^2+n\right)^4}} \cdot \left(\frac{\left(\frac{1}{\delta^2}+1\right)(n-1)n\delta_s^2}{(n-1)\delta_s^2+n}\right)^{\frac{n-3}{2}} \cdot e^{-\frac{\left(\frac{1}{\delta^2}+1\right)(n-1)n\delta_s^2}{2\left((n-1)\delta_s^2+n\right)}}}{\int_0^\infty x^{\frac{n-3}{2}} e^{-x} dx}$$

Two-window combination of determining the hydrodynamic radius

When the hydrodynamic radius is measured via the combination of two windows

(S12) $$r = \frac{4\, k_B T}{\pi\, \eta\, Y^2} \frac{\sigma_2^2 - \sigma_1^2}{t_2 - t_1}$$

where $\sigma_i^2 = \kappa_i t_i / 2$, the corresponding relative error is

(S13) $$\frac{\Delta r}{r} = \sqrt{\left(\frac{\Delta T}{T}\right)^2 + \left(\frac{\Delta \eta}{\eta}\right)^2 + h^2}$$

where

(S14) $$h^2 = \frac{t_1^2}{(t_1\kappa_1 - t_2\kappa_2)^2} \Delta\kappa_1^2 + \frac{t_2^2}{(t_1\kappa_1 - t_2\kappa_2)^2} \Delta\kappa_2^2 + \frac{t_2^2(\kappa_1-\kappa_2)^2}{(t_1-t_2)^2(t_1\kappa_1 - t_2\kappa_2)^2} \Delta t_1^2 + \frac{t_1^2(\kappa_1-\kappa_2)^2}{(t_1-t_2)^2(t_1\kappa_1 - t_2\kappa_2)^2} \Delta t_2^2.$$

Impact of noise on numerical integration and temporal moments

By adapting the concept of statistical moments, the temporal moments (mean and variance) are frequently used when characterizing multimodal and polydisperse samples and their optical extinction-weighted average radius.[8-10] In this case, noise affects the attainable precision differently than it affects model-based nonlinear regression. To describe this difference, here we outline a straightforward theoretical approach.

The n^{th} temporal moment of a taylorgram calculated as a normalized temporal average on the closed interval capped by t_a and t_b is defined as

(S15) $$\langle t^n A(t) \rangle = \frac{\int_{t_a}^{t_b} t^n A(t) dt}{\int_{t_a}^{t_b} A(t) dt}$$

where t_a and t_b are chosen that way that they contain the peak of the taylorgram (Figure S8a). Furthermore, t_a and t_b are chosen symmetrically around the center of the peak, *i.e.* they are at an equal distance from t_0. In experimental practice the taylorgram is not continuous in time for $A(t)$ is recorded with discrete timepoints, and thus, the integration becomes a summation, *e.g.*

(S16) $$\int_{t_a}^{t_b} A(t)\, dt \cong \tau \sum_{i=1}^{n} A(t_i)$$

where τ is the temporal resolution, $t_1 = t_a$ and $t_1 + n\,\tau = t_b$. The temporal mean and temporal variance are

(S17) $$M = \langle t\, A \rangle = t_0 + \kappa/2$$

and

(S18) $$V = \langle (t - M)^2 A \rangle = \langle t^2 A \rangle - \langle t\, A \rangle^2 = (t_0 + \kappa)\kappa/2.$$

where from $\kappa \cong 2\, V/M$ when $t_0 \gg \kappa$.

In the absence of noise, the precision and accuracy are perfect, and there is neither uncertainty nor bias in determining the value of κ. Accordingly, $\Delta\kappa = 0$. However, $\Delta\kappa$ does not vanish in the presence of noise. To show this, we consider that the experimentally recorded taylorgram may be decomposed into two terms:

(S19) $$A_\epsilon = A + \epsilon$$

where ϵ represents additive noise. The origin of ϵ is the shot noise and the related Poisson distribution,[1] and its probability density function $p(\epsilon)$ is practically a gaussian. It is easy to show that $p(\epsilon)$ is practically stationary in time, *i.e.* the variance of $p(\epsilon)$ is basically constant along the taylorgram when the peak absorbance is not too high (*i.e.* $A < 0.1$). The value of ϵ varies randomly along the taylorgram, with a zero mean $\bar{\epsilon} = 0$ and variance $\overline{\epsilon^2} - \bar{\epsilon}^2 = \overline{\epsilon^2}$. The noise in TDA is also uncorrelated, that is $\overline{\epsilon_i \epsilon_j} = \overline{\epsilon^2}\delta_{ij}$, where δ_{ij} is the Kronecker delta. The overline denotes ensemble average, which is the average over the distribution of the values ϵ can take:

(S20) $$\bar{\epsilon} = \int_{-\infty}^{\infty} \epsilon\, p(\epsilon)\, d\epsilon$$

and

(S21) $\overline{\epsilon^2} = \int_{-\infty}^{\infty} \epsilon^2 p(\epsilon)\, d\epsilon\,.$

Following the definition of the signal-to noise ratio (eq S1), it is easy to show that $\sqrt{\overline{\epsilon^2}} = A_{Max}\, SN^{-1} \cong A(t_0)\, SN^{-1}$. The temporal mean and temporal variance of a noisy taylorgram are calculated the same way as above

(S22) $M_\epsilon = \langle t\, A_\epsilon \rangle = \langle t(A+\epsilon) \rangle = \underbrace{\langle t\, A \rangle}_{M} + \langle t\, \epsilon \rangle = M + \langle t\, \epsilon \rangle$

(S23) $V_\epsilon = \langle t^2 A_\epsilon \rangle - M_\epsilon^2$

and the expected values are calculated via ensemble averring

(S24) $\overline{M_\epsilon} = \int_{-\infty}^{\infty} M_\epsilon\, p(\epsilon)\, d\epsilon$

(S25) $\overline{V_\epsilon} = \int_{-\infty}^{\infty} V_\epsilon\, p(\epsilon)\, d\epsilon.$

To calculate the expected precision attainable via numerical integration, we need to calculate the ensemble variance of κ_ϵ via the ensemble variance of M_ϵ and V_ϵ. These two are practically uncorrelated variables, and thus, we can write

(S26) $\Delta\kappa^2 = \overline{\kappa_\epsilon^2} - \overline{\kappa_\epsilon}^2 \cong \overline{(2V_\epsilon)^2}\ \overline{M_\epsilon^{-2}} - (\overline{2V_\epsilon})^2 \left(\overline{M_\epsilon^{-1}}\right)^2.$

We take advantage of the fact that the ensemble variance of the temporal mean is negligible compared to the ensemble variance of the temporal variance, that is, $\overline{V_\epsilon^2} - \overline{V_\epsilon}^2 \gg \overline{M_\epsilon^2} - \overline{M_\epsilon}^2$, and thus, we can express the ensemble average as

(S27) $\Delta\kappa^2 \cong \frac{\overline{(2V_\epsilon)^2} - (2\overline{V_\epsilon})^2}{\overline{M_\epsilon}^2}.$

Accordingly, the precision of determining the width parameter is

(S28) $\frac{\Delta\kappa}{\kappa} \cong \sqrt{\frac{\overline{V_\epsilon^2}}{\overline{V_\epsilon}^2} - 1}.$

To evaluate further eq S28, we perform the following steps:

(S29) $M_\epsilon = M + \langle t\, \epsilon \rangle$

(S30) $V_\epsilon = \langle t^2 A_\epsilon \rangle - M_\epsilon^2 = V + \underbrace{\langle t^2 \epsilon \rangle - \langle t\, \epsilon \rangle^2}_{V_N} = V + V_N$

where the term

(S31) $V_N = \langle t^2 \epsilon \rangle - \langle t\, \epsilon \rangle^2$

may be considered as an analogue of the temporal variance of the noise. It is important to point out, that the temporal noise cannot be treated as a proper probability density function, because it can take negative values as well. Thus, the term is merely an analogy. The ensemble averages are

(S32) $\overline{V_\epsilon} = V + \overline{V_N}$

(S33) $\overline{V_\epsilon}^2 = V^2 + 2V\overline{V_N} + \overline{V_N}^2$

and

(S34) $\overline{V_\epsilon^2} = V^2 + 2V\overline{V_N} + \overline{V_N^2}$

thus

(S35) $\frac{\Delta\kappa}{\kappa} \cong \frac{\overline{V_\epsilon^2} - \overline{V_\epsilon}^2}{\overline{V_\epsilon}^2} = \frac{\overline{V_N^2} - \overline{V_N}^2}{(V + \overline{V_N})^2}.$

Finally, we can write the expression of uncertainty into three distinct groups

(S36) $\overline{V_N^2} - \overline{V_N}^2 = \underbrace{\overline{\langle t^2\epsilon \rangle^2} - \left(\overline{\langle t^2 \epsilon \rangle}\right)^2}_{1.} + \underbrace{\overline{\langle t\, \epsilon \rangle^4} - \left(\overline{\langle t\, \epsilon \rangle^2}\right)^2}_{2.} - \underbrace{\left(2\overline{\langle t^2\epsilon \rangle \cdot \langle t\, \epsilon \rangle^2} - 2\overline{\langle t^2 \epsilon \rangle} \cdot \overline{\langle t\, \epsilon \rangle^2}\right)}_{3.}$

where via eq S15

(S37) $\overline{\langle t^m \epsilon \rangle^n} = \int_{-\infty}^{\infty} \left(\frac{\int_{t_a}^{t_b} t^m \epsilon\, dt}{\int_{t_a}^{t_b} (A+\epsilon)\, dt} \right)^n p(\epsilon)\, d\epsilon.$

As it can be seen, the precision of determining κ is a nontrivial function of the noise level and the width of the integration interval. We could not find a straightforward closed-form expression to evaluate eq S37, and therefore, used a numerical approach. Our results indicate that the precision scales as

(S38) $\frac{\Delta\kappa}{\kappa} \propto \sqrt{\overline{\epsilon^2} \cdot (t_b - t_a)^5}.$

This scaling can be understood by considering that $\overline{\langle t^m \epsilon \rangle^n} = \overline{\langle t^{m \cdot n} \epsilon^n \rangle} = \langle t^{m \cdot n} \overline{\epsilon^n} \rangle = \overline{\epsilon^n} \langle t^{m \cdot n} \rangle$, and $\langle t^{m \cdot n} \rangle \propto t^{m \cdot n + 1}$.

Eq S38 means that the precision is improving when the interval of integration becomes shorter, but the precision of numerical integration is inferior to model-based parameter fitting. This is because the uncertainty of model-based parameter fitting is weakly dependent on the interval chosen—provided that the peak is within the interval capped by t_a and t_b—while numerical integration is not. Indeed, when comparing the model-based fit with the temporal moments obtained via numerical integration, the scaling is evident (Figure S8). Finally, it is important to point out that this result is of general validity, and does not depend on the nature of the particle system under study, such as multimodality, polydispersity, and nature of optical extinction.

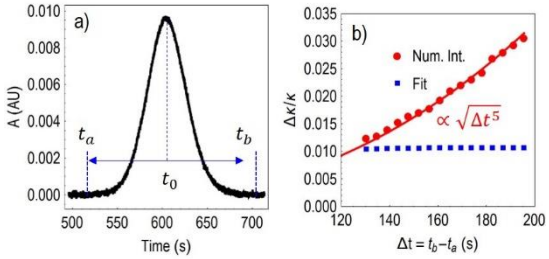


Figure S8. a) A BSA taylorgram and the symmetric interval used for analysis. Being symmetric means that $|t_0 - t_a| = |t_0 - t_b|$. b) The precision of determining κ via the temporal moments (Num. Int.) and model-based fitting (Fit). While the latter is practically independent of the interval used, numerical integration exhibits a strong increase with the width of the integral. Accordingly, the integration interval should be kept as narrow as possible.

References

(1) Balog, S. Taylor Dispersion of Polydisperse Nanoclusters and Nanoparticles: Modeling, Simulation, and Analysis*, Analytical Chemistry,* **2018**, *90*, 4258-4262.
(2) Spears, K. G.; Robinson, T. J. Particle Number Densities by Light-Scattering Fluctuation Analysis*, The Journal of Physical Chemistry,* **1988**, *92*, 5302-5305.
(3) Sharma, U.; Gleason, N. J.; Carbeck, J. D. Diffusivity of Solutes Measured in Glass Capillaries Using Taylor's Analysis of Dispersion and a Commercial Ce Instrument*, Analytical Chemistry,* **2005**, *77*, 806-813.
(4) Stockburger, D. W. In *International Encyclopedia of Statistical Science*, Lovric, M., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011, pp 1274-1277.
(5) Bossert, D.; Natterodt, J.; Urban, D. A.; Weder, C.; Petri-Fink, A.; Balog, S. Speckle-Visibility Spectroscopy of Depolarized Dynamic Light Scattering*, The Journal of Physical Chemistry B,* **2017**, *121*, 7999-8007.
(6) McKay, A. T. Distribution of the Coefficient of Variation and the Extended "T" Distribution*, Journal of the Royal Statistical Society,* **1932**, *95*, 695-698.
(7) Forkman, J.; Verrill, S. The Distribution of Mckay's Approximation for the Coefficient of Variation*, Statistics & Probability Letters,* **2008**, *78*, 10-14.
(8) Cottet, H.; Martin, M.; Papillaud, A.; Souaïd, E.; Collet, H.; Commeyras, A. Determination of Dendrigraft Poly-L-Lysine Diffusion Coefficients by Taylor Dispersion Analysis*, Biomacromolecules,* **2007**, *8*, 3235-3243.
(9) Chamieh, J.; Oukacine, F.; Cottet, H. Taylor Dispersion Analysis with Two Detection Points on a Commercial Capillary Electrophoresis Apparatus*, J Chromatogr A,* **2012**, *1235*, 174-177.
(10) Balog, S.; Urban, D. A.; Milosevic, A. M.; Crippa, F.; Rothen-Rutishauser, B.; Petri-Fink, A. Taylor Dispersion of Nanoparticles*, Journal of Nanoparticle Research,* **2017**, *19*, 287.