Differences Between Cumulants and the Distribution Algorithms

Dr Mike Kaszuba
Technical Support Manager
Michael.kaszuba@malvern.com

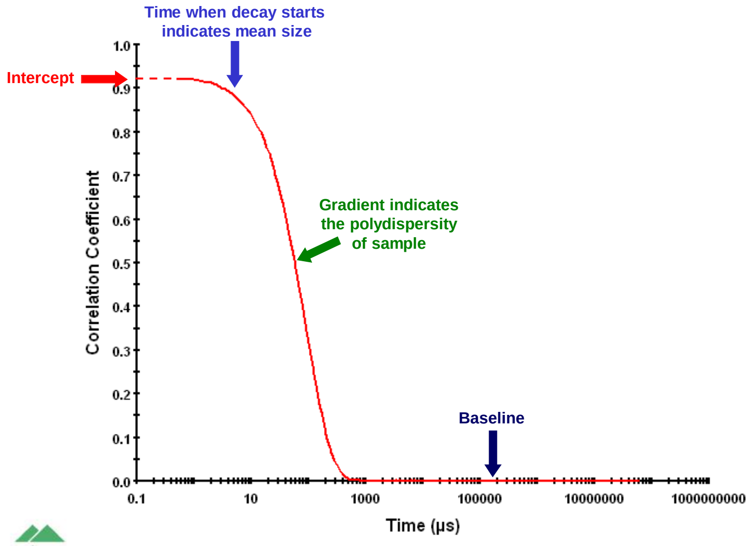
Malvern

Contents

- Correlation functions
- Analysing the correlation function
  - Cumulants
  - Distribution algorithms

Malvern

Correlation Functions



Correlation Functions

- The correlation function can be modelled with an exponential expression such as:

$$G(\tau) = B + A\sum \mathrm{e}^{-2q^2D\tau}$$

Where
B = baseline at infinite time
A = amplitude (or intercept)
q = scattering vector = (4π n/λ_o) sin(θ/2)
  n = dispersant refractive index
  λ_o = laser wavelength
  θ = detection angle
D = diffusion coefficient
τ = correlator delay time

Analysing The Correlation Function

- Correlation function contains the diffusion coefficient information required to be entered into the Stokes-Einstein equation
- The diffusion coefficients are obtained by fitting the correlation function with a suitable algorithm

Malvern www.malvern.com

Analysing The Correlation Function

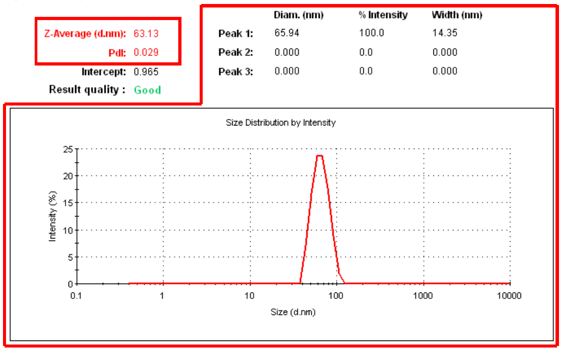
Two different analyses are performed:

Cumulants analysis

- Mean size (z-average)
- Polydispersity index

Distribution analysis

- Distribution of sizes



Z-Average (d.nm): 63.13
PdI: 0.029
Intercept: 0.965
Result quality : Good

	Diam. (nm)	% Intensity	Width (nm)
Peak 1:	65.94	100.0	14.35
Peak 2:	0.000	0.0	0.000
Peak 3:	0.000	0.0	0.000

Malvern www.malvern.com

Cumulants Analysis

- Defined in the International Standards ISO13321 (1996), ISO22412 (2008) and ISO22412 (2017)
- Only gives a mean particle size (z-average) and an estimate of the width of the distribution (polydispersity index)
- Only the dispersant refractive index and viscosity are required for this analysis

---

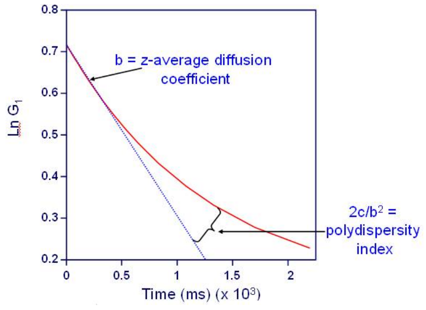
Cumulants Analysis

- 3rd order fit of a polynomial

$$\mathrm{Ln}[G_1] = a + b\tau + c\tau^2$$

where τ is the delay time of the correlator

- b = z-average diffusion coefficient
- $2c/b^2$ = polydispersity index



The z-Average Diameter

- Definition of the z-Average Diameter (Z_D):

The intensity-weighted mean diameter derived from the cumulants analysis

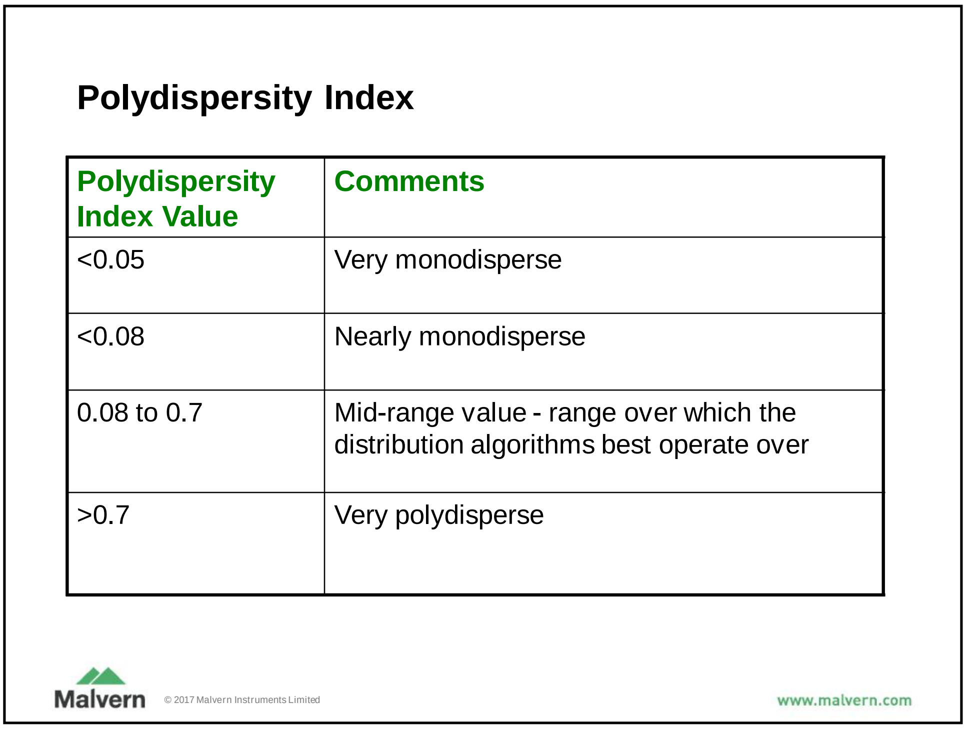
- Specific to light scattering
- Very sensitive to the presence of aggregates or large contaminants due to the inherent intensity weighting

Polydispersity Index

- Definition of the Polydispersity Index (PdI):

A dimensionless measure of the broadness of the size distribution calculated from the cumulants analysis

- Ranges from 0 to 1 in the Zetasizer software
- Values > 1 indicate that the distribution is so polydisperse, the sample may not be suitable for measurement by DLS

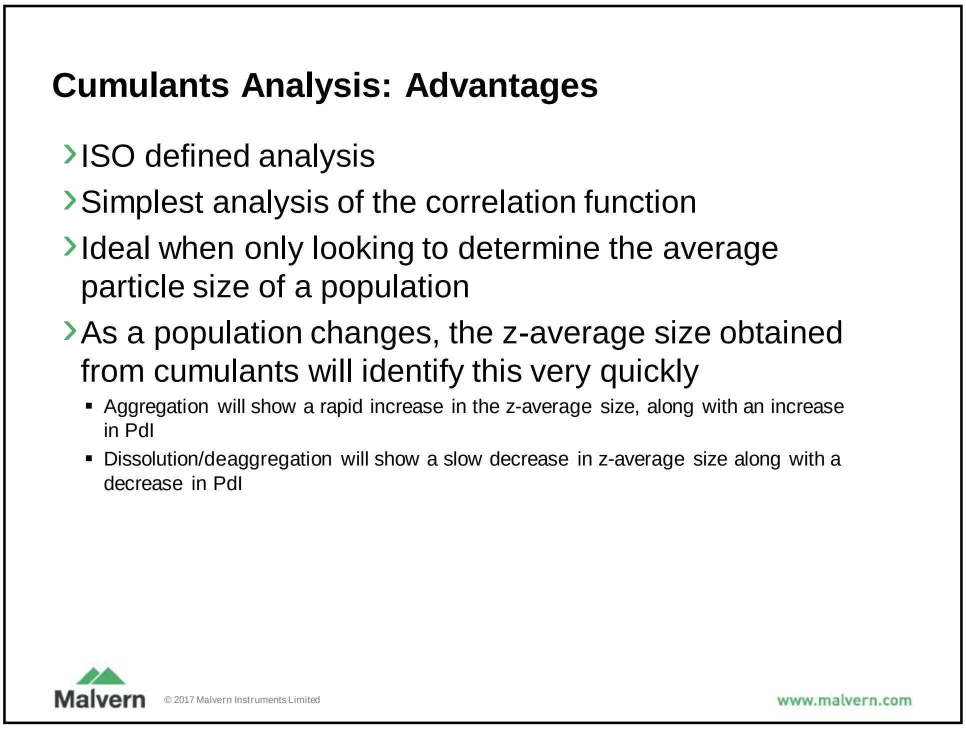


Polydispersity Index

Polydispersity Index Value	**Comments**
<0.05	Very monodisperse
<0.08	Nearly monodisperse
0.08 to 0.7	Mid-range value - range over which the distribution algorithms best operate over
>0.7	Very polydisperse

Malvern

www.malvern.com



Cumulants Analysis: Advantages

- ISO defined analysis
- Simplest analysis of the correlation function
- Ideal when only looking to determine the average particle size of a population
- As a population changes, the z-average size obtained from cumulants will identify this very quickly
  - Aggregation will show a rapid increase in the z-average size, along with an increase in PdI
  - Dissolution/deaggregation will show a slow decrease in z-average size along with a decrease in PdI

Malvern

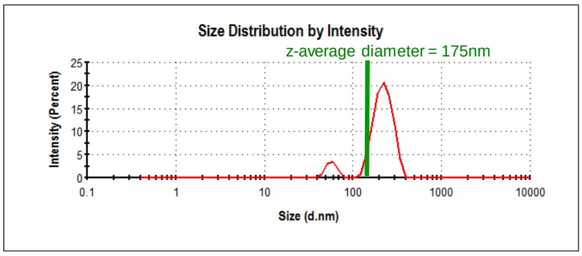
www.malvern.com

Cumulants Analysis: Disadvantages

- Cumulants only describes a single average size value, along with the PdI, it provides a good idea of changes in a population, and the broadness of that population
- Cumulants cannot identify individual modes in a population so it becomes less descriptive as the sample type moves further away from monomodal
    - i.e. as the PdI increases, the usefulness of a single size value decreases

Cumulants Analysis: Multimodal Samples

- If a sample of 60nm and 220nm latex was prepared with equal volumes of each mode present



- Cumulants cannot describe the discrete populations in a sample - it can only state that the average size is 175nm with a PdI of 0.19
- As the modality of a population increases, the ability of cumulants analysis to provide useful information about our population decreases

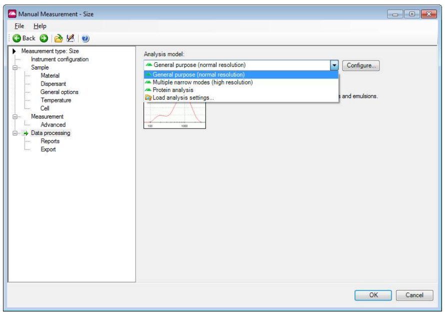
Distribution Analysis

- To investigate the modality of a sample, a mathematical analysis of the correlation function is performed which can describe the distribution
- This is know as distribution analysis
- The distribution analysis is more complex than cumulants
- The data set is fitted to a multi-exponential model to produce a distribution
- A lot more information is required to perform this analysis:
  - Size range over which the distribution should be analyzed
  - Number of size classes within this size range
  - The expected noise level within the data set (regularizer)

Distribution Analysis

- The distribution analysis becomes useful when:
  - The number of populations needs to be known
  - The relative composition of these populations needs to be known
  - The presence/absence of a large population needs to be determined (aggregate detection)
  - The presence/absence of a small population needs to be determined (fragmentation, monomer detection)
  - The abundance of different species in a formulation needs to be known

Distribution Algorithms



Analysis model: General purpose (normal resolution)

- General Purpose (non-negative least squares (NNLS))
- Multiple Narrow Modes (non-negative least squares (NNLS))
- Protein Analysis (L-curve)

- The difference between these algorithms is the regularizer used

---

Regularizer

- A small amount of noise in the correlation function can generate a large number of distributions
- Regularizer can be thought of as an estimator of the noise contained in the correlogram
- It controls the acceptable degree of “spikiness” in the size distribution obtained
  - Large regularizer values produce smooth distributions
  - Small regularizer values produce spiky distributions
- There is no ideal regularizer value; the appropriate value depends on the sample being measured

Available Algorithms and Associated Regularizer

Algorithm	Regularizer
General Purpose	0.01
Multiple Narrow Modes	0.001
Protein Analysis	Variable (appropriate value automatically determined)

Zetasizer Distribution Algorithms

- General Purpose
  - Suitable for the majority of samples where no knowledge of the distribution is available
  - Will give broad, smooth distributions
- Multiple Narrow Modes
  - Suitable for samples suspected to contain discrete populations
  - Will give narrow peaks
- Protein Analysis
  - Best suited for protein samples – will give narrow peaks
  - Automatically picks the optimal distribution

Cumulants Versus Distribution Analysis

- It is important to note that neither analysis method makes the other defunct
  - Cumulants analysis provides information on the average particle size present in a population, along with a general idea of the polydispersity of that population
  - Distribution analysis allows us to obtain more specific information relating to the total number of modes present in a population, and how they relate to one another
- Neither analysis model provides the complete solution, but together give us a greater understanding of our sample

Many thanks for your attention

Any questions?