



OpenLab ChemStation

## Data Analysis Reference

# Notices

## Document Information

Document No: D0013749 Rev. A  
Edition: 07/2022

## Copyright

© Agilent Technologies, Inc. 2010-2022

No part of this manual may be reproduced in any form or by any means (including electronic storage and retrieval or translation into a foreign language) without prior agreement and written consent from Agilent Technologies, Inc. as governed by United States and international copyright laws.

Agilent Technologies  
Hewlett-Packard-Strasse 8  
76337 Waldbronn, Germany

## Software Revision

This guide is valid for revision LTS 01.11 of Agilent OpenLab ChemStation.

## Warranty

The material contained in this document is provided "as is," and is subject to being changed, without notice, in future editions. Further, to the maximum extent permitted by applicable law, Agilent disclaims all warranties, either express or implied, with regard to this manual and any information contained herein, including but not limited to the implied warranties of merchantability and fitness for a particular purpose. Agilent shall not be liable for errors or for incidental or consequential damages in connection with the furnishing, use, or performance of this document or of any information contained herein. Should Agilent and the user have a separate written agreement with warranty terms covering the material in this document that conflict with these terms, the warranty terms in the separate agreement shall control.

## Technology Licenses

The hardware and/or software described in this document are furnished under a license and may be used or copied only in accordance with the terms of such license.

## Restricted Rights Legend

U.S. Government Restricted Rights. Software and technical data rights granted to the federal government include only those rights customarily provided to end user customers. Agilent provides this customary commercial license in Software and technical data pursuant to FAR 12.211 (Technical Data) and 12.212 (Computer Software) and, for the Department of Defense, DFARS 252.227-7015 (Technical Data - Commercial Items) and DFARS 227.7202-3 (Rights in Commercial Computer Software or Computer Software Documentation).

## Safety Notices

### CAUTION

A **CAUTION** notice denotes a hazard. It calls attention to an operating procedure, practice, or the like that, if not correctly performed or adhered to, could result in damage to the product or loss of important data. Do not proceed beyond a **CAUTION** notice until the indicated conditions are fully understood and met.

### WARNING

A **WARNING** notice denotes a hazard. It calls attention to an operating procedure, practice, or the like that, if not correctly performed or adhered to, could result in personal injury or death. Do not proceed beyond a **WARNING** notice until the indicated conditions are fully understood and met.

## In This Guide...

This guide addresses the advanced users, system administrators and persons responsible for validating Agilent OpenLab ChemStation. It contains reference information on the principles of operation, calculations and data analysis algorithms used in revision LTS 01.11 of OpenLab ChemStation.

Use this guide to verify system functionality against your user requirements specifications and to define and execute the system validation tasks defined in your validation plan. The following resources contain additional information.

- For context-specific task ("How To") information, a tutorial, reference to the User Interface, and troubleshooting help: The ChemStation online help system.
- For concepts and workflows of OpenLab ChemStation: The manual OpenLab ChemStation Basic Concepts and Workflows.
- For details on system installation and site preparation: The OpenLab ChemStation Workstation Installation Guide, or the OpenLab ChemStation Networked Workstations and Distributed System Installation Guide respectively
- For details on system administration principles and tasks: the OpenLab ChemStation Workstation Administrator's Guide.

In the following the term ChemStation refers to Agilent OpenLab ChemStation.

### 1 Integration

This chapter describes the concepts and integrator algorithms of the ChemStation integrator.

### 2 Peak Identification

This chapter describes the concepts of peak identification.

### 3 Calibration

This chapter contains details of the calculations used in the calibration process.

### 4 Quantitation

This chapter describes how compounds are quantified, and explains the calculations used in quantitation.

## **5 System Suitability**

This chapter describes what OpenLab CDS can do to evaluate the performance of both the analytical instrument and the analytical method.

## **6 CE specific Calculations**

This chapter is relevant only if you use ChemStation to control CE instruments.

## **7 System Verification**

This chapter describes the verification function and the GLP verification features of the ChemStation.

# Contents

## 1 Integration 7

What is Integration?	8
The Integrator Algorithms	10
Principle of operation	15
Peak Recognition	16
Peak Area Measurement	26
Baseline Allocation	29
Integration Events	41
Manual Integration	62

## 2 Peak Identification 65

What is Peak Identification?	66
Peak Matching Rules	67
Types of Peak Identification	68
Absolute Retention/Migration Time	69
Corrected Retention/Migration Times	71
Peak Qualifiers	73
The Identification Process	75

## 3 Calibration 77

What is Calibration?	78
Calibration Curve	79
Calibration Curve Calculation	80
Group Calibration	84
Recalibration Options	85

## 4 Quantitation 86

What is Quantitation?	87
Correction Factors	88
Area% and Height%	90

Quantitation of Calibrated Compounds 91

Quantitation of Uncalibrated Peaks 96

Norm% Calculation 98

## **5 System Suitability 99**

Peak Performance Evaluation 100

Evaluating System Suitability 100

Noise Determination 103

Calculation of Peak Asymmetry and Symmetry 114

System Suitability Formulas and Calculations 116

General Definitions 117

Performance Test Definitions 118

Definitions for Reproducibility 130

Internally Stored Double Precision Number Access 135

## **6 CE specific Calculations 138**

Calibration Tables 139

Calibration using Mobility Correction 142

Special Report Styles for Capillary Electrophoresis 148

Corrected Peak Areas 149

System Suitability for Capillary Electrophoresis 150

CE-MSD 151

## **7 System Verification 152**

Verification and Diagnosis Views 153

The GLPsave Register 156

DAD Test Function 157

# 1

## Integration

What is Integration?	8
What does integration do?	8
Integrator Capabilities	9
The Integrator Algorithms	10
Overview	10
Defining the Initial Baseline	11
Tracking the Baseline	12
Allocating the Baseline	13
Definition of Terms	14
Principle of operation	15
Peak Recognition	16
Peak Width	16
Peak Recognition Filters	17
Bunching	18
The Peak Recognition Algorithm	19
Merged Peaks	21
Shoulders	22
Default Baseline Construction	23
Peak Separation Codes	24
Peak Area Measurement	26
Determination of the area	26
Units and Conversion Factors	28
Baseline Allocation	29
The Start of the Baseline	29
The End of the Baseline	29
Baseline Correction Modes	30
Peak-to-Valley Ratio	32
Tangent Skimming	33
Integration Events	41
Integration Events for all Signals	41
Initial Events	44
Timed Events	49
Autointegrate	60
Manual Integration	62
Saving Manual Integration Events	63

This chapter describes the concepts and integrator algorithms of the ChemStation integrator.

## What is Integration?

Integration locates the peaks in a signal and calculates their size.

Integration is a necessary step for:

- Identification
- qualification
- calibration
- quantification,
- peak purity calculations, and
- spectral library search.

## What does integration do?

When a signal is integrated, the software:

- identifies a start and an end time for each peak
- finds the apex of each peak; that is, the retention/migration time,
- constructs a baseline, and
- calculates the area, height, peak width, and symmetry for each peak.

This process is controlled by parameters called integration events.



## Integrator Capabilities

The integrator algorithms include the following key capabilities:

- an autointegrate capability used to set up initial integrator parameters,
- the ability to define individual integration event tables for each chromatographic/electropherographic signal if multiple signals or more than one detector is used,
- interactive definition of integration events that allows users to graphically select event times,
- graphical manual integration of chromatograms or electropherograms requiring human interpretation (these events may also be recorded in the method and used as part of the automated operation),
- annotation of integration results,
- integrator parameter definitions to set or modify the basic integrator settings for area rejection, height rejection, peak width, slope sensitivity, shoulder detection, baseline correction and front /tail tangent skim detection,
- baseline control parameters, such as force baseline, hold baseline, baseline at all valleys, baseline at the next valley, fit baseline backwards from the end of the current peak,
- area summation control,
- negative peak recognition,
- solvent peak definition detection
- integrator control commands defining retention/migration time ranges for the integrator operation.
- peak shoulder allocation through the use of second derivative calculations,
- improved sampling of non-equidistant data points for better performance with DAD LC data files that are reconstructed from DAD spectra.

## The Integrator Algorithms

### Overview

To integrate a chromatogram/electropherogram the integrator:

- 1 defines the initial baseline,
- 2 continuously tracks and updates the baseline,
- 3 identifies the start time for a peak,
- 4 finds the apex of each peak,
- 5 identifies the end time for the peak,
- 6 constructs a baseline, and
- 7 calculates the area, height, and peak width for each peak.

This process is controlled by **integration events**. The most important events are initial slope sensitivity, peak width, baseline correction, area reject, and height reject. The software allows you to set initial values for these and other events. The initial values take effect at the beginning of the chromatogram. In addition, the auto integration function provides a set of initial events that you can optimize further.

In most cases, the initial events will give good integration results for the entire chromatogram, but there may be times when you want more control over the progress of an integration.

The software allows you to control how an integration is performed by enabling you to program new integration events at appropriate times in the chromatogram.

For more information, see ["Initial Events"](#) on page 44.

## Defining the Initial Baseline

### Cardinal points

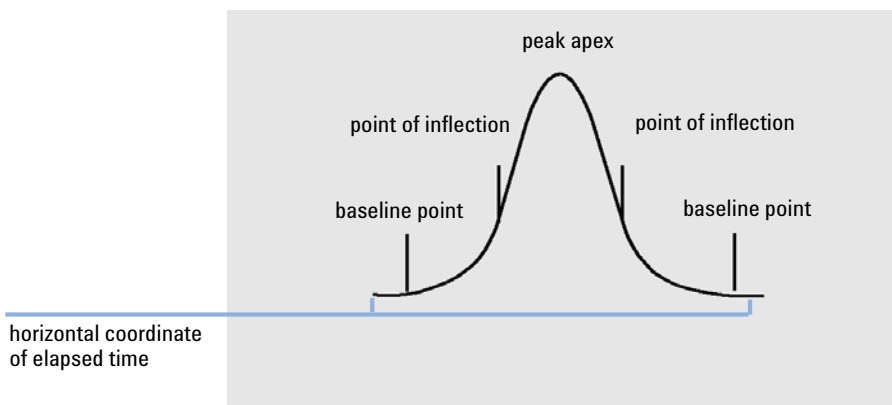


Figure 1 Cardinal points

### Defining the initial baseline

Because baseline conditions vary according to the application and detector hardware, the integrator uses parameters from both the integration events and the data file to optimize the baseline.

Before the integrator can integrate peaks, it must establish a **baseline point**. At the beginning of the analysis, the integrator establishes an initial baseline level by taking the first data point as a tentative baseline point. It then attempts to redefine this initial baseline point based on the average of the input signal. If the integrator does not obtain a redefined initial baseline point, it retains the first data point as a potential initial baseline point.

### Identifying the cardinal points of a peak

The integrator determines that a peak may be starting when potential baseline points lie outside the baseline envelope, and the baseline curvature exceeds a certain value, as determined by the integrator's slope sensitivity parameter. If this condition continues, the integrator recognizes that it is on the up-slope of a peak, and the peak is processed.

## Integration

### The Integrator Algorithms

Start

- 1 Slope and curvature within limit: continue tracking the baseline.
- 2 Slope and curvature above limit: possibility of a peak.
- 3 Slope remains above limit: peak recognized, peak start point defined.
- 4 Curvature becomes negative: front inflection point defined.

Apex

- 1 Slope passes through zero and becomes negative: apex of peak, apex point defined.
- 2 Curvature becomes positive: rear inflection point defined.

End

- 1 Slope and curvature within limit: approaching end of the peak.
- 2 Slope and curvature remain within limit: end of peak defined.
- 3 The integrator returns to the baseline tracking mode.

## Tracking the Baseline

The integrator samples the digital data at a rate determined by the initial peak width or by the calculated peak width, as the run progresses. It considers each data point as a potential baseline point.

The integrator determines a *baseline envelope* from the slope of the baseline, using a baseline-tracking algorithm in which the slope is determined by the first derivative and the curvature by the second derivative. The baseline envelope can be visualized as a cone, with its tip at the current data point. The upper and lower acceptance levels of the cone are:

- + upslope + curvature + baseline bias must be lower than the threshold level,
- - upslope - curvature + baseline bias must be more positive (i.e. less negative) than the threshold level.

As new data points are accepted, the cone moves forward until a break-out occurs.

## Integration

### The Integrator Algorithms

To be accepted as a baseline point, a data point must satisfy the following conditions:

- it must lie within the defined baseline envelope,
- the curvature of the baseline at the data point (determined by the derivative filters), must be below a critical value, as determined by the current slope sensitivity setting.

The initial baseline point, established at the start of the analysis is then continuously reset, at a rate determined by the peak width, to the moving average of the data points that lie within the baseline envelope over a period determined by the peak width. The integrator tracks and periodically resets the baseline to compensate for drift, until a peak up-slope is detected.

## Allocating the Baseline

The integrator allocates the chromatographic/electropherographic baseline during the analysis at a frequency determined by the peak width value. When the integrator has sampled a certain number of data points, it resets the baseline from the initial baseline point to the current baseline point. The integrator resumes tracking the baseline over the next set of data points and resets the baseline again. This process continues until the integrator identifies the start of a peak

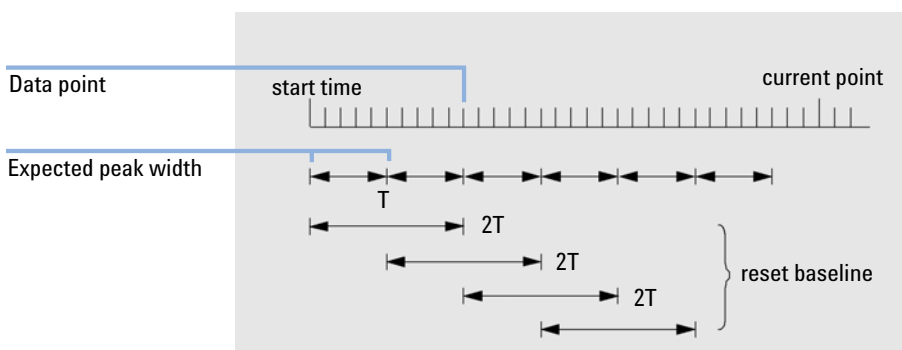


Figure 2 Baseline

At the start of the run the first data point is used. This baseline point is periodically reset according to the following formula:

Areas are summed over a time  $T$  (expected peak width). This time can never be shorter than one data point. This continues as long as baseline condition exists. Slope and curvature are also taken. If both slope and curvature are less than the

threshold, two summed areas are added together, and compared with the previous baseline. If the new value is less than the previous baseline, the new value immediately replaces the old one. If the new value is greater than the previous value, it is stored as a tentative new baseline value and is confirmed if one more value satisfies slope and curvature flatness criteria. This latter limitation is not in effect if negative peaks are allowed. During baseline, a check must also be made to examine fast rising solvents. They may be too fast for upslope detection. (By the time upslope is confirmed, solvent criterion may no longer be valid.) At first time through the first data point is baseline. It is replaced by the 2 T average if signal is on base. Baseline is then reset every T (see [Figure 2](#) on page 13).

## Definition of Terms

### Solvent peak

The solvent peak, which is generally a very large peak of no analytical importance, is not normally integrated. However, when small peaks of analytical interest elute close to the solvent peak, for example, on the tail of the solvent peak, special integration conditions can be set up to calculate their areas corrected for the contribution of the solvent peak tail.

### Shoulder (front, rear)

Shoulders occur when two peaks elute so close together that no valley exists between them, and they are unresolved. Shoulders may occur on the leading edge (front) of the peak, or on the trailing edge (rear) of the peak. When shoulders are detected, they may be integrated either by tangent skim or by drop-lines.

### Slope

The slope of a peak, which denotes the change of concentration of the component against time, is used to determine the onset of a peak, the peak apex, and the end of the peak.

## Principle of operation

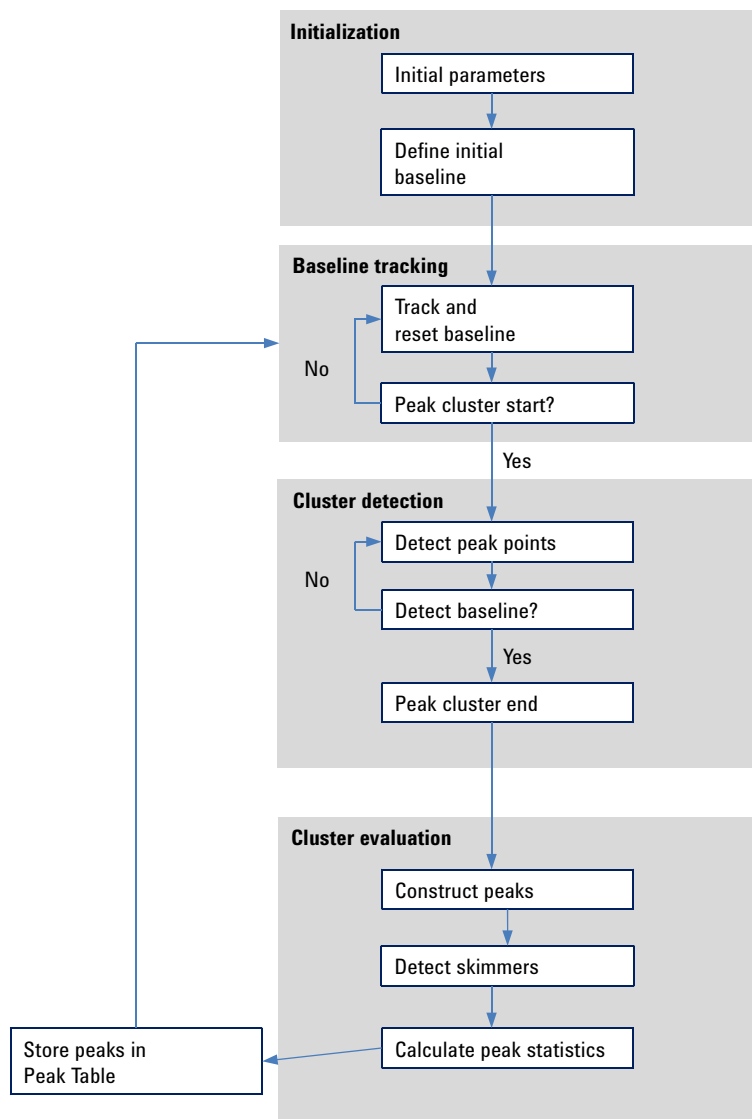


Figure 3 Integrator Flow Diagram

## Peak Recognition

The integrator uses several tools to recognize and characterize a peak:

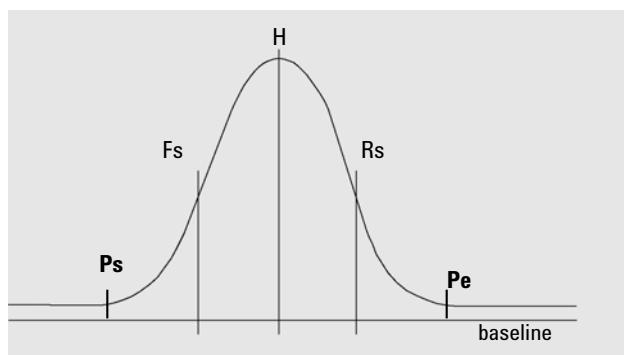
- peak width,
- peak recognition filters,
- bunching,
- peak recognition algorithm,
- peak apex algorithm, and
- non-Gaussian calculations (for example tailing, merged peaks).

### Peak Width

During integration, the peak width is calculated from the adjusted peak area and height:

$$\text{Width} = \text{adjusted area} / \text{adjusted height}$$

or, if the inflection points are available, from the width between the inflection points.



**Figure 4** Peak width calculation

In the figure above, the total area,  $A$ , is the sum of the areas from peak start ( $P_s$ ) to peak end ( $P_e$ ), adjusted for the baseline.  $F_s$  is the front slope at the inflection point,  $R_s$  is the rear slope at the inflection point.



## Integration

### Peak Recognition

The peak width setting controls the ability of the integrator to distinguish peaks from baseline noise. To obtain good performance, the peak width must be set close to the width of the actual chromatographic peaks.

There are three ways the peak width is changed:

- before the integration process, you can specify the initial peak width,
- during the integration process, the integrator automatically updates the peak width as necessary to maintain a good match with the peak recognition filters,
- during the integration process, you can reset or modify the peak width using a time-programmed event.

## Peak Recognition Filters

The integrator has three peak recognition filters that it can use to recognize peaks by detecting changes in the slope and curvature within a set of contiguous data points. These filters contain the first derivative (to measure slope) and the second derivative (to measure curvature) of the data points being examined by the integrator. The recognition filters are:

- Filter 1** Slope (curvature) of two (three) contiguous data points
- Filter 2** Slope of four contiguous data points and curvature of three non-contiguous data points
- Filter 3** Slope of eight contiguous data points and curvature of three non-contiguous data points

The actual filter used is determined by the peak width setting. For example, at the start of an analysis, Filter 1 may be used. If the peak width increases during the analysis, the filter is changed first to Filter 2 and then to Filter 3. To obtain good performance from the recognition filters, the peak width must be set close to the width of the actual chromatographic/electropherographic peaks. During the run, the integrator updates the peak width as necessary to optimize the integration.

The integrator calculates the updated peak width in different ways, depending on the instrument configuration:

For LC/CE configurations, the default peak width calculation uses a composite calculation:

$$0.3 \times (\text{Right Inflection Point} - \text{Left Inflection point}) + 0.7 \times \text{Area/Height}$$

## Integration

### Peak Recognition

For GC configurations, the default peak width calculation uses area/height. This calculation does not overestimate the width when peaks are merged above the half-height point.

In certain types of analysis, for example isothermal GC and isocratic LC analyses, peaks become significantly broader as the analysis progresses. To compensate for this, the integrator automatically updates the peak width as the peaks broaden during the analysis. It does this automatically unless the updating has been disabled with the fixed peak width timed event.

The peak width update is weighted in the following way:

$$0.75 \times (\text{existing peak width}) + 0.25 \times (\text{width of current peak})$$

## Bunching

Bunching is the means by which the integrator keeps broadening peaks within the effective range of the peak recognition filters to maintain good selectivity.

The integrator cannot continue indefinitely to increase the peak width for broadening peaks. Eventually, the peaks would become so broad that they could not be seen by the peak recognition filters. To overcome this limitation, the integrator bunches the data points together, effectively narrowing the peak while maintaining the same area.

When data is bunched, the data points are bunched as two raised to the bunching power, i.e. unbunched = 1x, bunched once = 2x, bunched twice = 4x etc.

Bunching is based on the data rate and the peak width. The integrator uses these parameters to set the bunching factor to give the appropriate number of data points (see [Table 1](#) on page 19).

Bunching is performed in the powers of two based on the expected or experienced peak width. The bunching algorithm is summarized in [Table 1](#) on page 19.

**Table 1** Bunching criteria

Expected Peak Width	Filter(s) Used	Bunching Done
0 - 10 data points	First	None
8 - 16 data points	Second	None
12 - 24 data points	Third	None
16 - 32 data points	Second	Once
24 - 48 data points	Third	Once
32 - 96 data points	Third, second	Twice
64 - 192 data points	Third, second	Three times

## The Peak Recognition Algorithm

The integrator identifies the start of the peak with a baseline point determined by the peak recognition algorithm. The peak recognition algorithm first compares the outputs of the peak recognition filters with the value of the initial slope sensitivity, to increase or decrease the up-slope accumulator. The integrator declares the point at which the value of the up-slope accumulator is  $\geq 15$  the point that indicates that a peak has begun.

### Peak Start

In [Table 2](#) on page 20 the expected peak width determines which filter's slope and curvature values are compared with the Slope Sensitivity. For example, when the expected peak width is small, Filter 1 numbers are added to the up-slope accumulator. If the expected peak width increases, then the numbers for Filter 2 and, eventually, Filter 3 are used.

When the value of the up-slope accumulator is  $\geq 15$ , the algorithm recognizes that a peak may be starting.

**Table 2 Incremental Values to Upslope Accumulator**

Derivative Filter 1 - 3 Outputs against Slope Sensitivity	Filter 1	Filter 2	Filter 3
Slope > Slope Sensitivity	+8	+5	+3
Curvature > Slope Sensitivity	+0	+2	+1
Slope < (-) Slope Sensitivity	-8	-5	-3
Slope <  Slope Sensitivity	-4	-2	-1
Curvature < (-) Slope Sensitivity	-0	-2	-1

#### Peak End

In [Table 3](#) on page 20 the expected peak width determines which filter's slope and curvature values are compared with the Slope Sensitivity. For example, when the expected peak width is small, Filter 1 numbers are added to the down-slope accumulator. If the expected peak width increases, then the numbers for Filter 2 and, eventually, Filter 3 are used.

When the value of the down-slope accumulator is  $\geq 15$ , the algorithm recognizes that a peak may be ending.

**Table 3 Incremental Values for Downslope Accumulator**

Derivative Filter 1 - 3 Outputs against Slope Sensitivity	Filter 1	Filter 2	Filter 3
Slope < (-) Slope Sensitivity	+8	+5	+3
Curvature < (-) Slope Sensitivity	+0	+2	+1
Slope > Slope Sensitivity	-11	-7	-4
Slope >  Slope Sensitivity	-28	-18	-11
Curvature > Slope Sensitivity	-0	-2	-1

#### The Peak Apex Algorithm

The peak apex is recognized as the highest point in the chromatogram by constructing a parabolic fit that passes through the highest data points.

## Merged Peaks

Merged peaks occur when a new peak begins before the end of peak is found. The figure illustrates how the integrator deals with merged peaks.

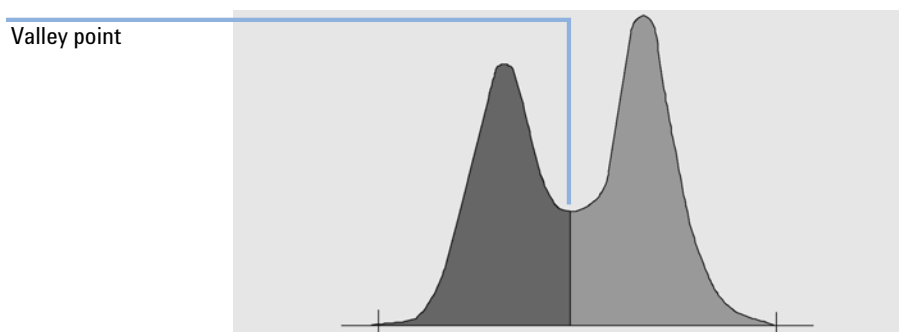


Figure 5 Merged Peaks

The integrator processes merged peaks in the following way:

- 1 it sums the area of the first peak until the valley point.
- 2 at the valley point, area summation for the first peak ends and summation for the second peak begins.
- 3 when the integrator locates the end of the second peak, the area summation stops. This process can be visualized as separating the merged peaks by dropping a perpendicular from the valley point between the two peaks.

## Shoulders

Shoulders are unresolved peaks on the leading or trailing edge of a larger peak. When a shoulder is present, there is no true valley in the sense of negative slope followed by positive slope. A peak can have any number of front and/or rear shoulders.

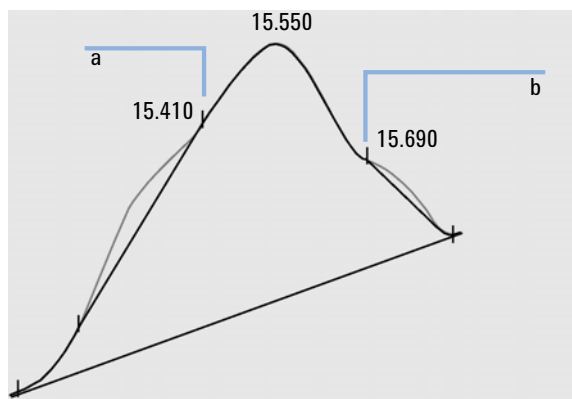


Figure 6 Peak Shoulders

Shoulders are detected from the curvature of the peak as given by the second derivative. When the curvature tends to zero, the integrator identifies a point of inflection, such as points a and b in Figure 6 on page 22.

- A potential front shoulder exists when a second inflection point is detected before the peak apex. If a shoulder is confirmed, the start of the shoulder point is set at the maximum positive curvature point before the point of inflection.
- A potential rear shoulder exists when a second inflection point is detected before the peak end or valley. If a shoulder is confirmed, the start of the shoulder point is set at the point of the first minimum of the slope after the peak apex.

Retention/migration time is determined from the shoulder's point of maximum negative curvature. With a programmed integration event, the integrator can also calculate shoulder areas as normal peaks with drop-lines at the shoulder peak points of inflection.

The area of the shoulder is subtracted from the main peak.

Peak shoulders can be treated as normal peaks by use of an integrator timed event.

## Default Baseline Construction

After any peak cluster is complete, and the baseline is found, the integrator requests the baseline allocation algorithm to allocate the baseline using a pegs-and-thread technique. It uses trapezoidal area and proportional height corrections to normalize and maintain the lowest possible baseline. Inputs to the baseline allocation algorithm also include parameters from the method and data files that identify the detector and the application, which the integrator uses to optimize its calculations.

In the simplest case, the integrator constructs the baseline as a series of straight line segments between:

- the start of baseline,
- peakstart, valley, end points,
- the peak baseline

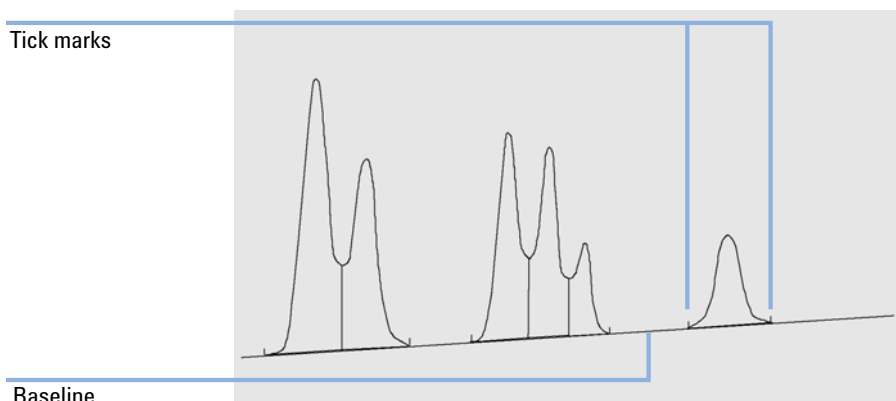


Figure 7 Default Baseline Construction

## Peak Separation Codes

In the integration results of a report, each peak is assigned a two-, three- or four-character code that describes how the signal baseline was drawn.

**Table 4** Four character code

First character	Second character	Third character	Fourth character
Baseline at start	Baseline at end	Error/peak flag	Peak type

### Characters 1 and 2

The first character describes the baseline at the start of the peak and the second character describes the baseline at the end of the peak.

- B** The peak started or stopped on the baseline.
- P** The peak started or stopped while the baseline was penetrated.
- V** The peak started or stopped with a valley drop-line.
- H** The peak started or stopped on a forced horizontal baseline.
- F** The peak started or stopped on a forced point.
- M** The peak was manually integrated.
- U** The peak was unassigned.

Additional flags may also be appended (in order of precedence).

### Character 3

The third character describes an error or peak flag:

- A** The integration was aborted. For example due to the integration events ON/OFF, or due to the end of signal run time.
- D** The peak was distorted (bad peak shape).

**Blank space** The peak is a normal peak.



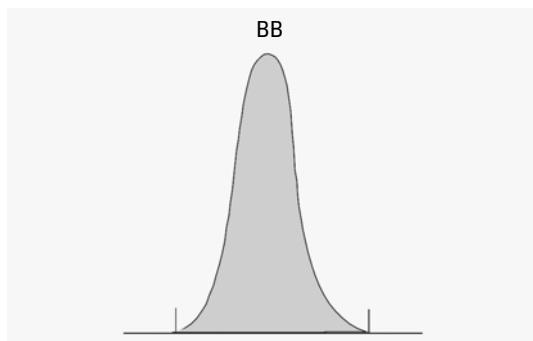
**Character 4**

The fourth character describes the peak type.

- S** The peak is a solvent peak.
- N** The peak is a negative peak.
- +** The peak is an area summed peak.
- T** Tangent-skimmed peak (standard skim).
- X** Tangent-skimmed peak (old mode exponential skim).
- E** Tangent-skimmed peak (new mode exponential skim).
- m** Peak defined by manual baseline.
- n** Negative peak defined by manual baseline.
- t** Tangent-skimmed peak defined by manual baseline.
- x** Tangent-skimmed peak (exponential skim) defined by manual baseline.
- R** The peak is a recalculated peak (for example a parent peak, see “Tangent Skim Modes” on page 36).
- f** Peak defined by a front shoulder tangent.
- b** Peak defined by a rear shoulder tangent.
- F** Peak defined by a front shoulder drop-line.
- B** Peak defined by a rear shoulder drop-line.
- U** The peak is unassigned.

## Peak Area Measurement

The final step in peak integration is determining the final area of the peak.



**Figure 8** Area measurement for Baseline-to-Baseline Peaks

In the case of a simple, isolated peak, the peak area is determined by the accumulated area above the baseline between peak start and stop.

### Determination of the area

The area that the integrator calculates during integration is determined as follows:

- for baseline-to-baseline (BB) peaks, the area above the baseline between the peak start and peak end, as in [Figure 8](#) on page 26,
- for valley-to-valley (VV) peaks, the area above the baseline, segmented with vertical dropped lines from the valley points, as in [Figure 9](#) on page 27,

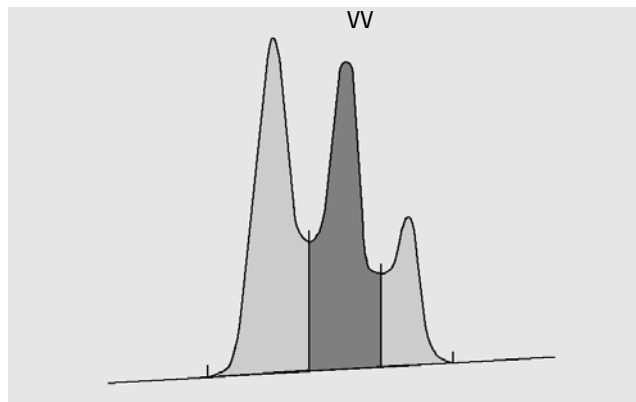


Figure 9 Area Measurement for Valley-to-Valley Peaks

- for tangent (T) peaks, the area above the reset baseline,
- for solvent (S) peaks, the area above the horizontal extension from the last-found baseline point and below the reset baseline given to tangent (T) peaks. A solvent peak may rise too slowly to be recognized, or there may be a group of peaks well into the run which you feel should be treated as a solvent with a set of riders. This usually involves a merged group of peaks where the first one is far larger than the rest. The simple drop-line treatment would exaggerate the later peaks because they are actually sitting on the tail of the first one. By forcing the first peak to be recognized as a solvent, the rest of the group is skimmed off the tail,
- negative peaks that occur below the baseline have a positive area, as shown in Figure 10 on page 27.

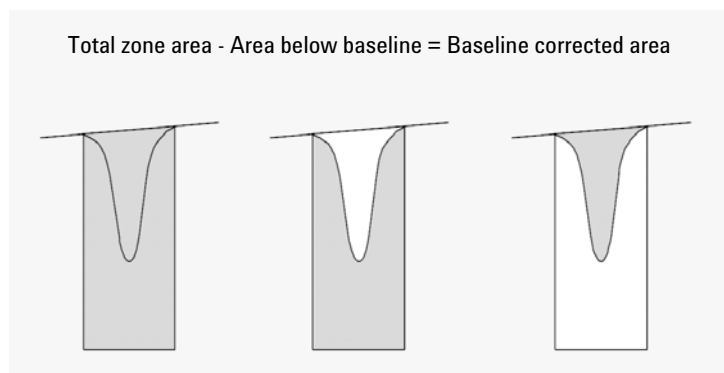


Figure 10 Area Measurement for Negative Peaks

## Units and Conversion Factors

Externally, the data contains a set of data points; they can be either sampled data or integrated data. In the case of integrated data, each data point corresponds to an area, which is expressed as *Height × Time*. In the case of sampled data, each data point corresponds to a height.

Therefore, in the case of integrated data, height is a calculated entity, obtained by dividing area by the time elapsed since the preceding data point. In the case of sampled data, area is calculated by multiplying the data by the time elapsed since the preceding data point.

The integration calculation makes use of both entities. The units carried internally inside the integrator are: *detector response × seconds* for area, and **detector response** as height. This is done to provide a common base for integer truncations when needed. The measurements of time, area and height are reported in real physical units, irrespective of how they are measured, calculated and stored in the software.

## Baseline Allocation

### The Start of the Baseline

If no baseline is found at the start of the run, the start of the baseline is established in one of the following ways:

- from the start of the run to the first baseline point, if the start of run point is lower than the first baseline point,
- from the start of the run to the first valley point, if the start of run point is lower than the first valley,
- from the start of the run to the first valley point, if the first valley penetrates an imaginary line drawn from the start of run to the first baseline,
- from the start of the run to a horizontal baseline extended to the first baseline point.

### The End of the Baseline

The last valid baseline point is used to designate the end of the baseline. In cases where the run does not end on the baseline, the end of the baseline is calculated from the last valid baseline point to the established baseline drift.

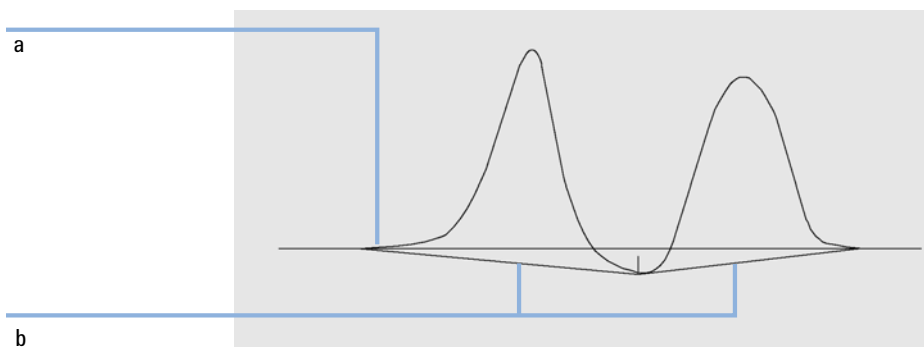
If a peak ends in an apparent valley but the following peak is below the area reject value as you have set it, the baseline is projected from the beginning of the peak to the next true baseline point. If a peak starts in a similar way, the same rule applies.

## Baseline Correction Modes

In OpenLab ChemStation, several baseline correction modes are available. They are described in the following sections.

### Baseline Correction Mode: Classical

A penetration occurs when the signal drops below the constructed baseline (point a in [Figure 11](#) on page 30).



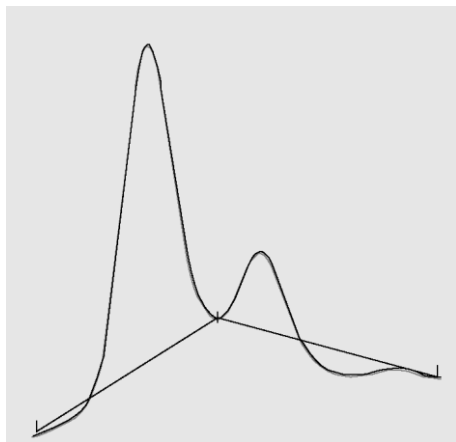
**Figure 11** Baseline Penetration

If a baseline penetration occurs, that part of the baseline may be reconstructed, as shown by points b in [Figure 11](#) on page 30. You can use the following correction modes to remove all baseline penetrations:

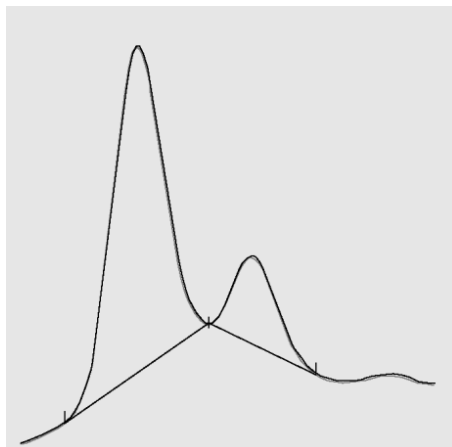
- **No penetration**
- **Advanced**

#### Baseline Correction Mode: No Penetration

When this option is selected, each peak cluster is searched for baseline penetrations. If penetrations are found, the start and/or end points of the peak are shifted until there are no penetrations left.



Baseline correction mode **Classical**



Baseline correction mode **No penetration**

#### NOTE

The baseline correction mode **No penetration** is not available for solvent peaks, with their child peaks and shoulders.

#### Baseline Correction Mode: Advanced

In the advanced baseline correction mode, the integrator tries to optimize the start and end locations of the peaks, re-establishes the baseline for a cluster of peaks, and removes baseline penetrations (see "[Baseline Correction Mode: No Penetration](#)" on page 31). In many cases, advanced baseline correction gives a more stable baseline, which is less dependent on slope sensitivity.

## Peak-to-Valley Ratio

The Peak to valley ratio is a measure of quality, indicating how well the peak is separated from other substance peaks. This user-specified parameter is a constituent of advanced baseline tracking mode. It is used to decide whether two peaks that do not show baseline separation are separated using a drop line or a valley baseline. The integrator calculates the ratio between the baseline-corrected height of the smaller peak and the baseline-corrected height of the valley. When the peak valley ratio is lower than the user-specified value, a drop-line is used; otherwise, a baseline is drawn from the baseline at the start of the first peak to the valley, and from the valley to the baseline at the end of the second peak (compare “Baseline Correction Mode: No Penetration” on page 31 with Figure 12 on page 32).

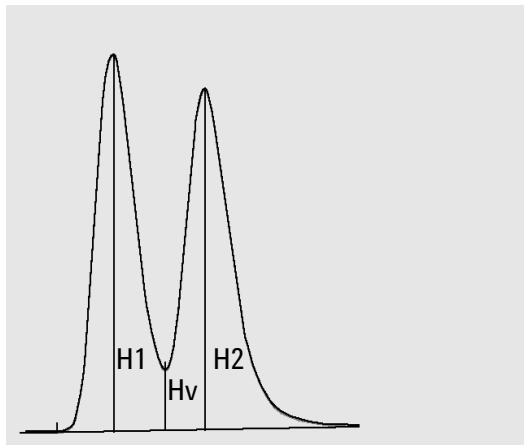


Figure 12 Peak Valley Ratio

The peak-to-valley ratio is calculated using the following equations:

$$H1 \geq H2, \text{ Peak valley ratio} = H2/Hv$$

and

$$H1 < H2, \text{ Peak valley ratio} = H1/Hv$$

Figure 13 on page 33 shows how the user-specified value of the peak valley ratio affects the baselines.



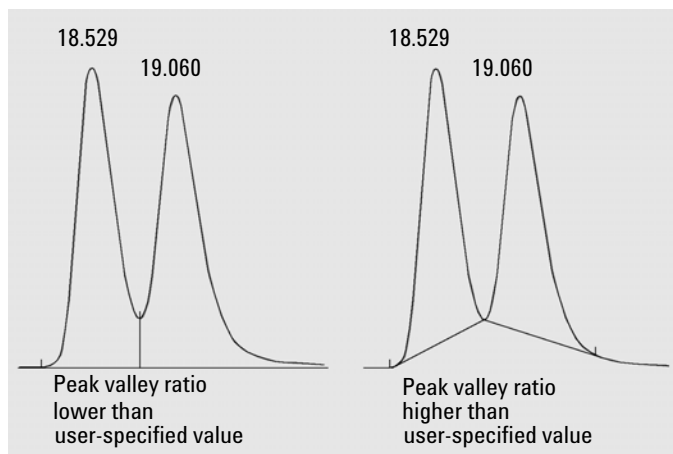


Figure 13 Effect of peak valley ratio on the baselines

## Tangent Skimming

Tangent skimming is a form of baseline constructed for peaks found on the upslope or downslope of a peak. The prerequisite is that the two peaks are not baseline-separated.

The following figures illustrate the principle of tangent skimming:

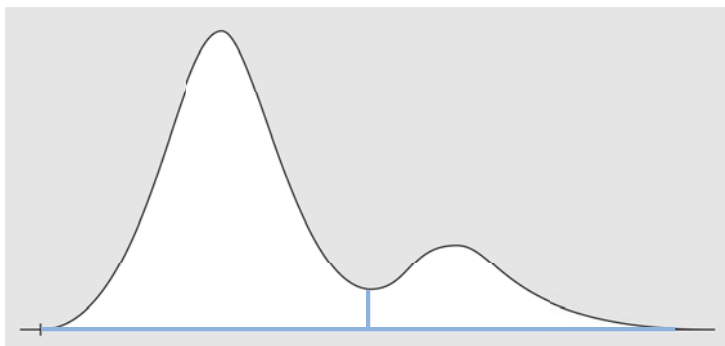


Figure 14 Peaks without skimming, separated by a drop line

## Integration

### Baseline Allocation

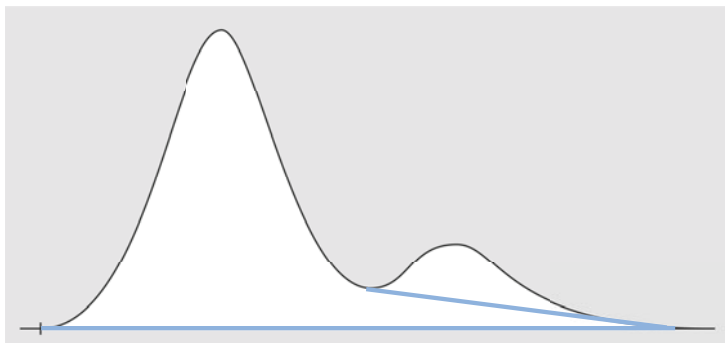


Figure 15 Tail skimming

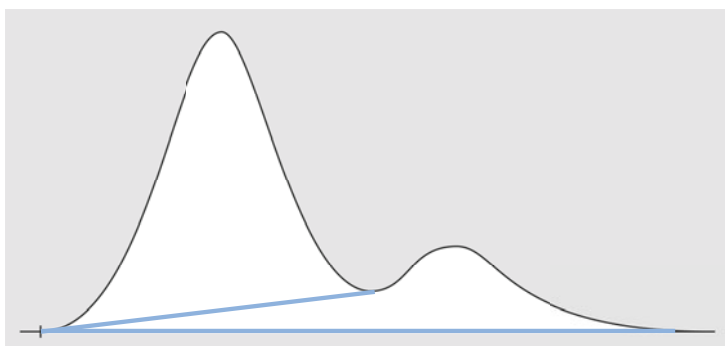


Figure 16 Front skimming

## Skim Criteria

The following criteria determine whether a skim line is used to calculate the area of a child peak eluting on the leading or trailing edge of a parent peak:

- Skim height ratio (**Front skim height ratio** or **Tail skim height ratio**)
- **Skim valley ratio**

The *skim height ratio* is the ratio of the baseline-corrected height of the parent peak ( $H_p$  in the figure below) to the baseline-corrected height of the child peak ( $H_c$ ). To have the child peak skimmed, use a value lower than this ratio. To disable exponential skimming throughout a run, you can set this parameter to a high value or to zero.

The *skim valley ratio* is the ratio of the height of the child peak above the baseline ( $H_c$  in the figure below) to the height of the valley above the baseline ( $H_v$ ). To have the child peak skimmed, use a value greater than this ratio.

**NOTE**

If one of these criteria is not met for a set of child peaks at the tail of the parent peak, all child peaks after the last child peak that met both criteria are not skimmed anymore but use a drop line

**NOTE**

These criteria are not used if a timed event for an exponential is in effect, or if the parent peak is itself a child peak. The baseline code between parent peak and child peak must be of type **Valley** (see “Peak Separation Codes” on page 24).

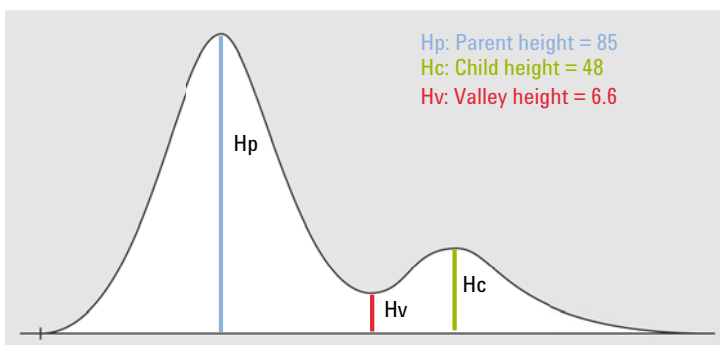


Figure 17 Example for calculating the skim criteria values

$$\text{Skim height ratio} = H_p / H_c$$

$$\text{Skim valley ratio} = H_c / H_v$$

where

## Integration

### Baseline Allocation

Hp	Baseline-corrected height of parent peak
Hv	Height of valley above the baseline
Hc	Baseline-corrected height of child peak

**Tail Skimming** To use tail skimming, you would set the parameters as follows:

- Tail skim height ratio =  $85 / 48 = 1.77$   
In the integration events, use a value  $< 1.77$ .
- Skim valley ratio =  $48 / 6.6 = 7.3$   
In the integration events, use a value  $> 7.3$ .

**Front Skimming** With front skimming, the first peak is the child peak, and the second peak is the parent peak. Thus, to use front skimming, you would set the parameters as follows:

- Front skim height ratio =  $48 / 85 = 0.56$   
In the integration events, use a value  $< 0.56$ .
- Skim valley ratio =  $85 / 6.6 = 12.9$   
In the integration events, use a value  $> 12.9$ .

### Tangent Skim Modes

When tangent skimming is enabled, four models are available to calculate suitable peak areas:

- Exponential curve
- New exponential skim
- Straight line skim
- Combined exponential and straight line calculations for the best fit (standard skims)

#### Exponential Curve

This skim model draws a curve using an exponential equation through the start and end of the child peak. The curve passes under each child peak that follows the parent peak; the area under the skim curve is subtracted from the child peaks and added to the parent peak.

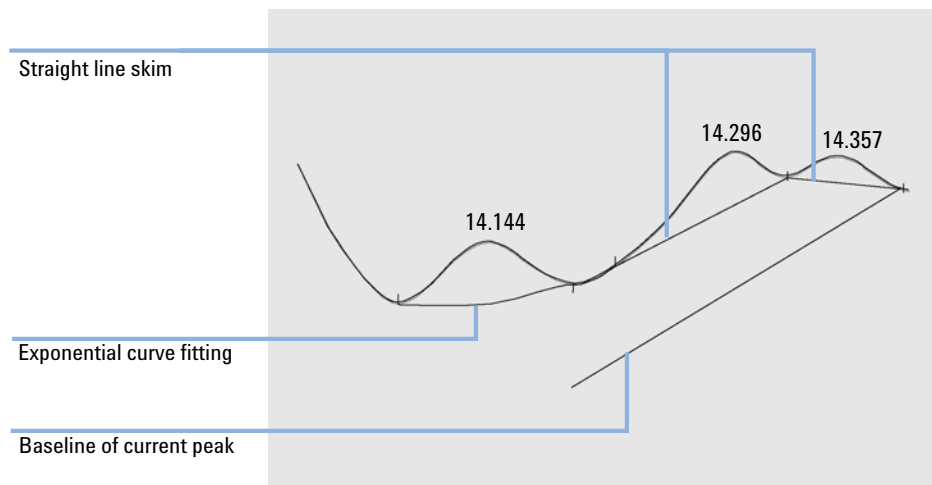


Figure 18 Exponential skim

#### New exponential curve

This skim model draws a curve using an exponential equation to approximate the leading or trailing edge of the parent peak. The curve passes under one or more peaks that follow the parent peak (child peaks). The area under the skim curve is subtracted from the child peaks and added to the main peak. More than one child peak can be skimmed using the same exponential model; all peaks after the first child peak are separated by drop lines, beginning at the end of the first child peak, and are dropped only to the skim curve.

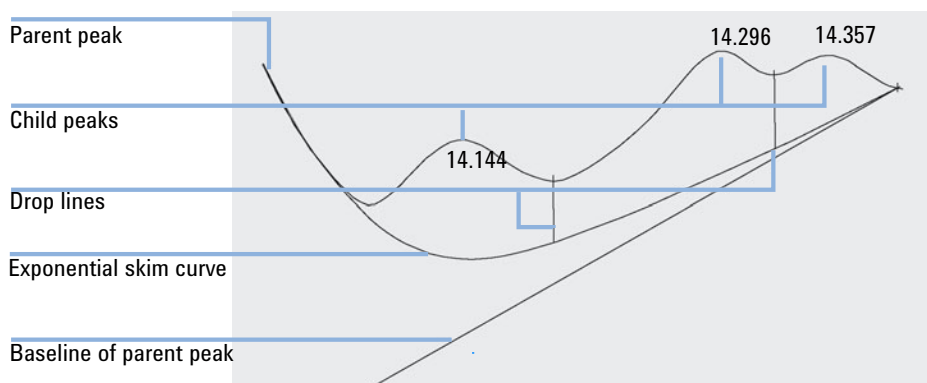
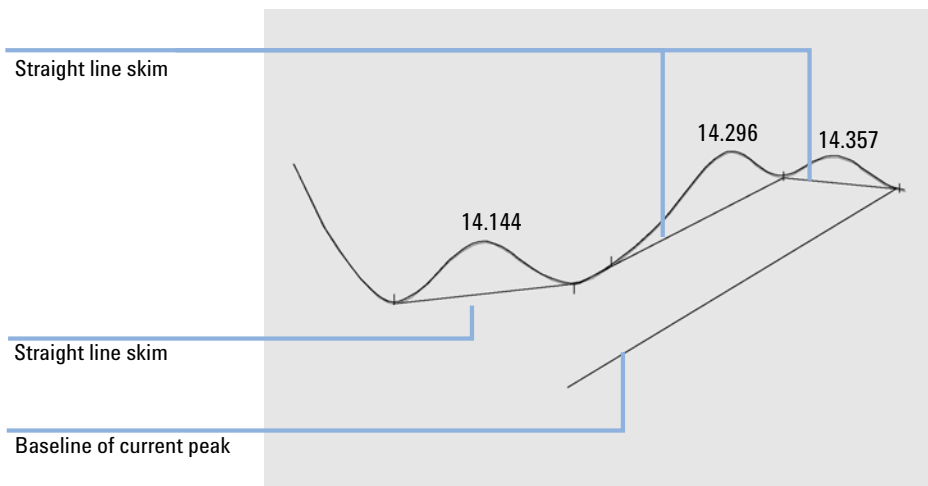


Figure 19 New exponential skim

#### Straight Line Skim

This skim model draws a straight line through the start and end of a child peak. The height of the start of the child peak is corrected for the parent peak slope. The area under the straight line is subtracted from the child peak and added to the parent peak.



**Figure 20** Straight line skim

#### Standard Skims

This default method is a combination of exponential and straight line calculations for the best fit.

The switch from an exponential to a linear calculation is performed in a way that eliminates abrupt discontinuities of heights or areas.

- When the signal is well above the baseline, the tail-fitting calculation is exponential.
- When the signal is within the baseline envelope, the tail fitting calculation is a straight line.

The combination calculations are reported as exponential or straight tangent skim.

## Integration

### Baseline Allocation

Calculation of Exponential Curve for Skims

The following equation is used to calculate an exponential skim:

$$H_b(t_R) = H_0 * \exp(-B * (t_R - t_0)) + A * t_R + C$$

where

$H_b$	Height of the exponential skim at time $t_R$
$H_0$	Height (above baseline) of the start of the exponential skim
$B$	Decay factor of the exponential function
$t_0$	Time corresponding to the start of the exponential skim
$t_R$	Retention time
$A$	Slope of the baseline of the parent peak
$C$	Offset of the baseline of the parent peak

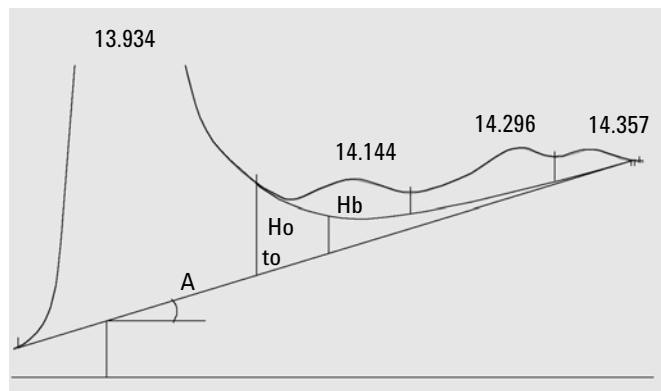


Figure 21 Values used to calculate an exponential skim



## Integration Events

The integrator provides you with a number of initial and timed integrator events. Many events are on/off or start/stop pairs.

### Integration Events for all Signals

#### Tangent Skim Mode

Define the type of baseline construction for peaks found on the upslope or downslope of a peak. See “[Tangent Skim Modes](#)” on page 36.

<b>Exponential</b>	Draws an exponential curve through the height-corrected start and end of each child peak.
<b>New Exponential</b>	Draws an exponential curve to approximate the trailing edge of the parent peak.
<b>Standard</b>	Combines exponential and straight line calculations for best fit.
<b>Straight</b>	Draws a straight line through the height-corrected start and end of each child peak.

#### Tail Peak Skim Height Ratio

Together with the **Skim valley ratio**, sets the conditions for tangent skimming a small peak on the tail of a solvent or other large peak. See “[Skim Criteria](#)” on page 35.

It is the ratio of the height of the baseline-corrected parent peak ( $H_p$ ) to the height of the baseline-corrected child peak ( $H_c$ ). Ratios higher than the specified value will enable skimming.

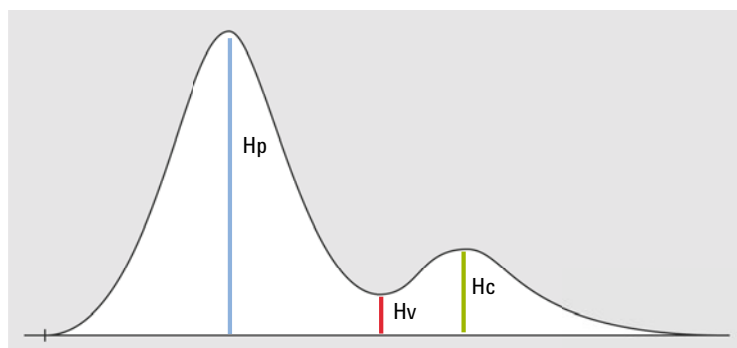


Figure 22 Example: Peak with tail skimming

**Front Peak  
Skim Height  
Ratio**

Together with the **Skim Valley Ratio**, sets the conditions for tangent skimming a small peak on the front of a solvent or other large peak. See “[Skim Criteria](#)” on page 35.

It is the ratio of the height of the baseline-corrected parent peak ( $H_p$ ) to the height of the baseline-corrected child peak ( $H_c$ ). Ratios higher than the specified value will enable skimming.

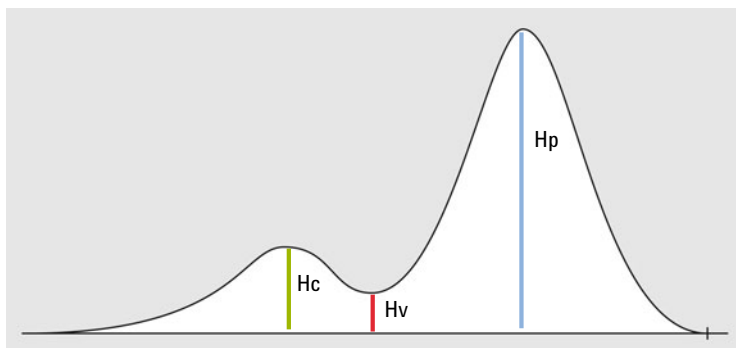


Figure 23 Example: Peak with front skimming

**Skim Valley  
Ratio**

Together with the **Tail Skim Height Ratio** or **Front Skim Height Ratio**, sets the conditions for tangent skimming a small peak on the tail or front of a solvent or other large peak. See “[Skim Criteria](#)” on page 35.

It is the ratio of the height of the baseline-corrected child peak ( $H_c$ ) to the height of the baseline-corrected valley ( $H_v$ ). Ratios lower than the specified value will enable skimming.

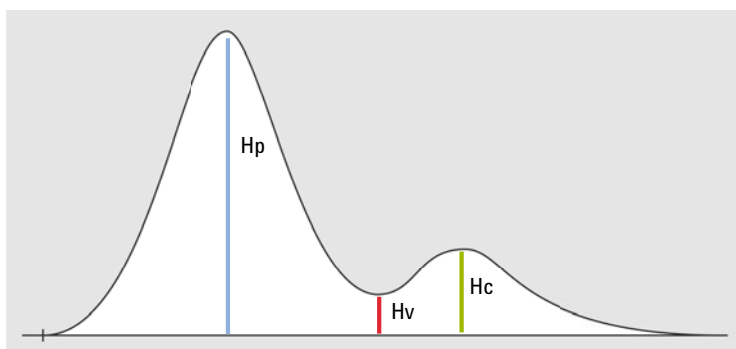


Figure 24 Example: Peak with tail skimming

#### Baseline Correction

Sets the type of baseline correction. See “Baseline Correction Modes” on page 30.

You can choose between the following parameters:

<b>Classical</b>	Accepts baseline penetrations.
<b>No penetrations</b>	Removes baseline penetrations by reconstructing the baseline.
<b>Advanced</b>	The integrator tries to optimize the start and end locations of the peaks, re-establishes the baseline for a cluster of peaks and removes baseline penetrations.

#### Peak to Valley Ratio

Used to decide whether two peaks that do not show baseline separation are separated using a drop line or a valley baseline, it is the ratio of the baseline-corrected height of the smaller peak to the baseline-corrected height of the valley. See “Peak-to-Valley Ratio” on page 32.

When the peak to valley ratio is lower than the specified value, a drop line is used (A); otherwise, a baseline is drawn from the baseline at the start of the first peak to the valley, and from the valley to the baseline at the end of the second peak (B).

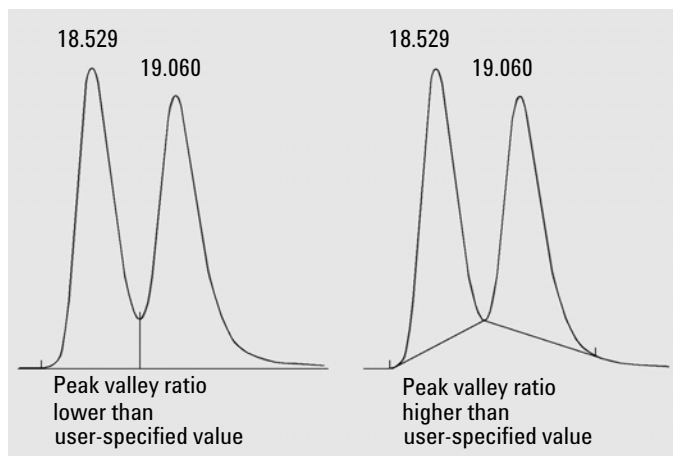


Figure 25 Effect of peak valley ratio on the baselines

## Initial Events

**Slope Sensitivity**

Slope sensitivity is the setting for peak sensitivity. When the signal slope exceeds the **Slope Sensitivity** value, a peak start point is established; when the signal slope decreases below the **Slope Sensitivity** value, a peak end point is established.

**Peak Width**

Controls the selectivity of the integrator to distinguish peaks from baseline noise. You specify the peak width in units of time that correspond to the peak width at half-height of the first expected peak (excluding the solvent peak).

The integrator updates the peak width when necessary during the run to optimize the integration:

If the selected initial peak width is too low, noise may be interpreted as peaks. If broad and narrow peaks are mixed, you may decide to use runtime programmed events to adjust the peak width for certain peaks. Sometimes, peaks become significantly broader as the analysis progresses, for example in isothermal GC and isocratic LC analyses. To compensate for this, the integrator automatically updates the peak width as peaks broaden during an analysis unless disabled with a timed event.

The Peak Width update is weighted in the following way:

$$0,75 \times (\text{existing peak width}) + 0,25 \times (\text{width of current peak})$$

**Area reject**

Sets the area of the smallest peak of interest.

Any peaks that have areas less than the minimum area are not reported: The integrator rejects any peaks that are smaller than the **Area Reject** value after baseline correction. The **Area Reject** value must be greater than or equal to zero.

## Integration

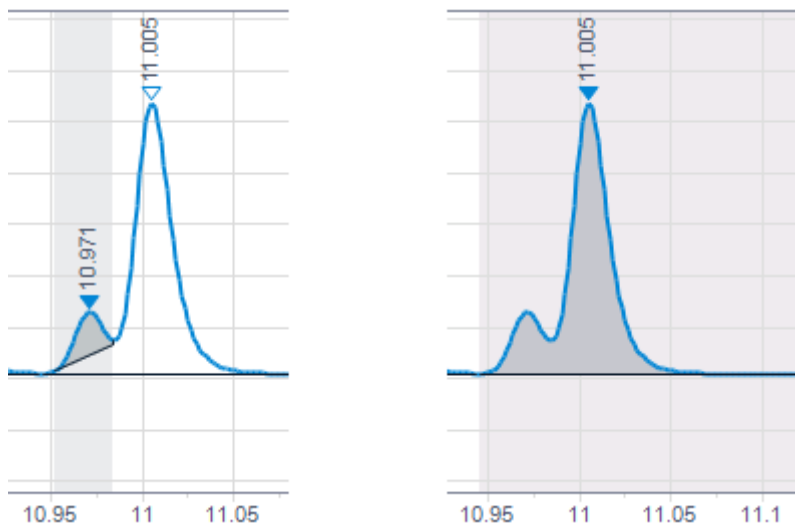
### Integration Events

#### Area Percent reject

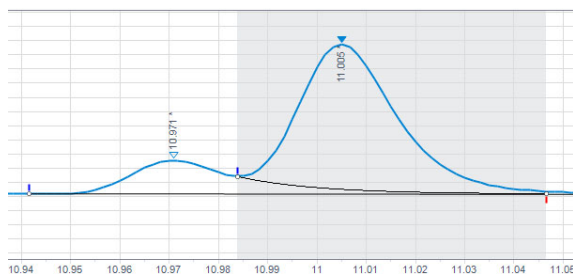
Sets the area% of the smallest peak of interest.

Any peaks with an area% less than the minimum area% are not reported. The integrator rejects any peaks with an area% smaller than the given value after baseline correction.

If a peak that is not integrated due to low area% is a rider peak, it will be merged with the parent peak.



If the parent peak is below the area% threshold, but the rider peak is above the threshold, the parent peak is kept, as the rider peak's calculation and baseline construction would otherwise be based on an excluded peak.



#### Height reject

Sets the height of the smallest peak of interest.

Any peaks that have heights less than this minimum height are not reported: The integrator rejects any peaks that are smaller than the **Height Reject** value after baseline correction.

**Shoulders** Sets the initial method of detecting shoulders on peaks.

When shoulder detection is on, the integrator detects shoulders using the curvature of the peak as given by the second derivative. When the curvature goes to zero, the integrator identifies this point of inflection as a possible shoulder. If the integrator identifies another point of inflection before the apex of the peak, a shoulder has been detected.

You can choose from:

- Off  
Shoulders are not detected.
- Drop Baseline  
Shoulders are integrated with a drop line.
- Tangent Baseline  
Shoulders are integrated with a tangent baseline. For more information on tangent skimming, see ["Tangent Skimming"](#) on page 33. If you use a tangent baseline, you can choose between different modes (see ["Tangent Skim Modes"](#) on page 36).

## Choosing Peak Width

Choose the setting that provides just enough filtering to prevent noise being interpreted as peaks without distorting the information in the signal.

- To choose a suitable initial peak width for a single peak of interest, use the peak's time width at the base as a reference.
- To choose a suitable initial peak width when there are multiple peaks of interest, set the initial peak width to a value equal to or less than the narrowest peak width to obtain optimal peak selectivity.

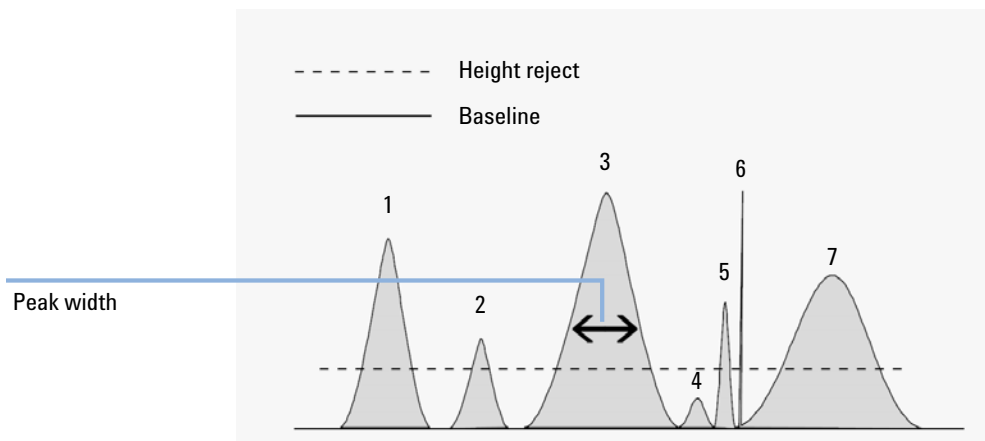
## Height Reject and Peak Width

Both **peak width** and **height reject** are very important in the integration process. You can achieve different results by changing these values.

- Increase both the height reject and peak width where relatively dominant components must be detected and quantified in a high-noise environment. An increased peak width improves the filtering of noise and an increased height reject ensures that random noise is ignored.
- Decrease height reject and peak width to detect and quantify trace components, those whose heights approach that of the noise itself. Decreasing peak width decreases signal filtering, while decreasing height reject ensures that small peaks are not rejected because they have insufficient height.
- When an analysis contains peaks with varying peak widths, set peak width for the narrower peaks and reduce height reject to ensure that the broad peaks are not ignored because of their reduced height.

## Tuning Integration

It is often useful to change the values for the slope sensitivity, peak width, height reject, and area reject to customize integration. The figure below shows how these parameters affect the integration of five peaks in a signal.



**Figure 26** Using Initial Events

A peak is integrated only when all of the four integration parameters are satisfied. Using the peak width for peak 3, the area reject and slope sensitivity shown, only peaks 1, 3, and 7 are integrated.

- Peak 1** is integrated as all four integration parameters are satisfied.
- Peak 2** is rejected because the area is below the set area reject value.
- Peak 3** is integrated as all four integration parameters are satisfied.
- Peak 4** is not integrated because the peak height is below the Height Reject.
- Peak 5** is rejected because the area is below the set area reject value.
- Peak 6** is not integrated; filtering and bunching make the peak invisible.
- Peak 7** is integrated.

**Table 5** Height and Area Reject Values

Integration Parameter	Peak 1	Peak 2	Peak 3	Peak 4	Peak 5	Peak 7
Height reject	Above	Above	Above	Below	Above	Above
Area reject	Above	Below	Above	Below	Below	Above
Peak integrated	Yes	No	Yes	No	No	Yes



## Timed Events

ChemStation offers a set of timed events, that allow a choice between the integrator modes of internal algorithm baseline definition and the user's definition. These timed events can be used to customize signal baseline construction when default construction is not appropriate. E.g. the user can create a new area sum event type, which does not alter the results of the default AreaSum. These events can be useful for summing final peak areas and for correcting short- and long-term baseline aberrations.

**Area reject** See "Initial Events" on page 44.

**Area Sum** Sets points (On/Off) between which the integrator sums the areas.

The retention/migration time of a peak created with area summing is the average of the start and end times. If an **Area sum on** event occurs after the beginning of a peak but before the apex, the entire peak is included in the sum. If it occurs after the peak apex, but before the end of the peak, the peak is truncated and the area sum begins immediately.

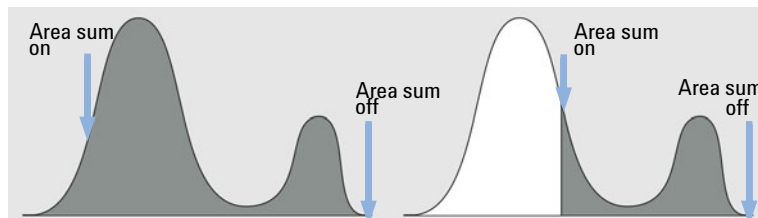


Figure 27 Area sum on event after peak apex, but before end of peak

If an **Area sum off** event occurs after the beginning of a peak but before the apex, the area sum ends immediately. The point on the signal where this occurs becomes a Valley Point. If the **Area sum off** event occurs after the apex, the event is postponed until the end of the peak.

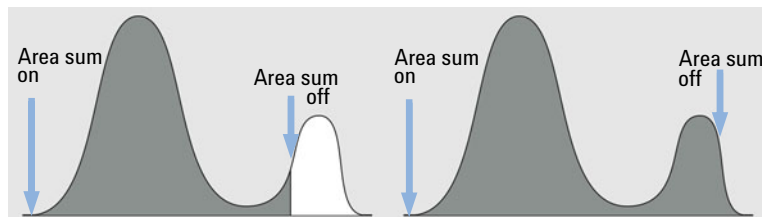


Figure 28 Area sum off event after beginning of peak, but before apex

**Area Sum Slice** This event allows you to define consecutive area sum intervals without any loss in area or time intervals.

This event is similar to the **Area Sum** event. However, with this event you can define contiguous area sum intervals without any loss in time intervals and integrated peak areas. A peak is split at the point where you set this event; area summing starts and ends exactly where the **Area Sum Slice** intervals are specified.

The retention time of the area sum slice peak is the middle of the slice time interval. The retention time does not change with identification or recalibration. It may only be shifted slightly, as the integrator only starts taking data points with the area sum slice start event, and ends with the area sum slice end event. Thus, the retention time may at most vary by the time between two data points.

Use the **Start** parameter to define the starting times for each area sum slice. The next start time is used as the end time for the preceding time-slice, so you can use several start events after each other.

The **Start-negA** parameter defines the start of integration of a time-slice where any negative area (below the set baseline) is subtracted from the area of the time-slice.

The **End** parameter defines the end of the last time-slice. The area of the time-slice is calculated ignoring any area below the set baseline. If no other area sum slice events follow, the integrator resumes its regular peak detection again.

Within the range from a **Start** event to the next **End** event, the baseline is always one straight line with no changes in direction in between. Only after the end point (at least 0.001 min later) long term baseline changes can be applied again by using the events **Set Baseline from Range**, **Set Low Baseline from Range** or **Use Baseline from Range**.

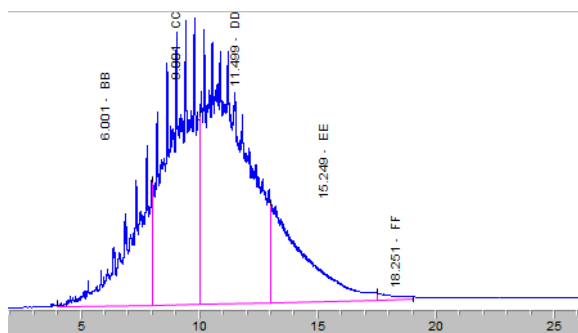


Figure 29 Example: Area Sum Slice

## Integration

### Integration Events

The figure above shows an example with the following timed events:

**Table 6** Baseline construction

Time	Event	Parameter
4 min	Set BL from Range	+2 min
22 min	Set BL from Range	+4 min

**Table 7** Area sum slices

Time	Event	Parameter
4 min	Area Sum Slice	Start
8 min	Area Sum Slice	Start
10 min	Area Sum Slice	Start
13 min	Area Sum Slice	Start
17.5 min	Area Sum Slice	Start
19 min	Area Sum Slice	End

#### **Auto Peak Width**

Turns on the automatic update of the peak width for the next peaks. It will resume with whatever the peak width is at that time and resume peak width tracking based on the previous found peak widths.

#### Baseline at Valleys

Sets points (On/Off) between which the integrator resets the baseline at every valley between peaks.

The repeated resetting of the baseline can cut off corners of peaks. Such corners become negative area, they reduce the total measured area of the peaks.

This function is useful when peaks are riding on the back of a broad, low peak and you want the baseline to be reset to all the valley points.

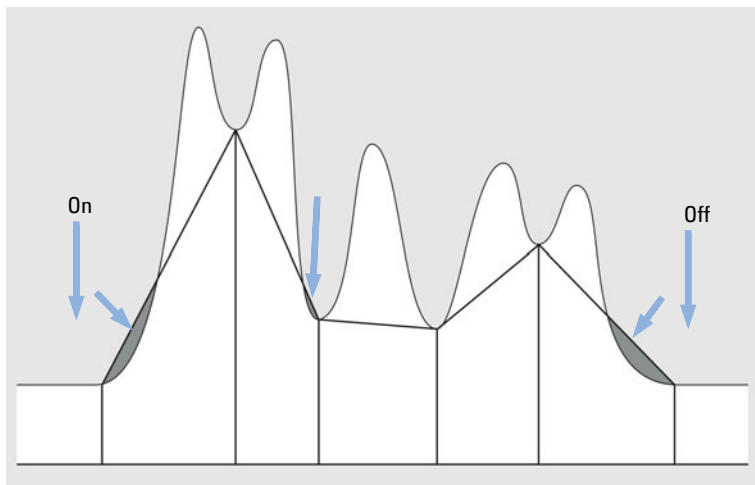


Figure 30 Baseline at valleys event

**Baseline Backwards** Sets a point at which the standard integrator extends the baseline, horizontally backward from the declared baseline point to this point.

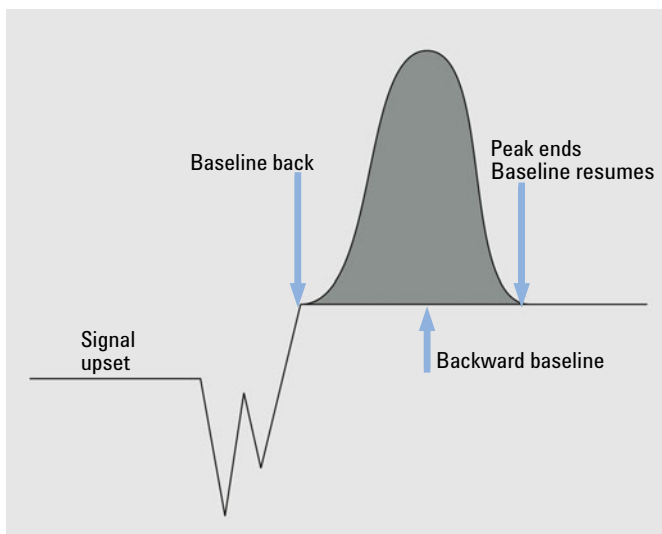


Figure 31 Baseline backwards event

**Baseline Hold** A horizontal baseline is drawn at the height of the established baseline from where the baseline hold event is switched on until where the baseline hold event is switched off.

**Baseline Next Valley** Sets a point at which the integrator resets the baseline at the next valley between peaks, and then cancels this function automatically.

This function is useful in groups of merged peaks, which you assume are riding on the back or are in separate clusters close together. The function is ignored during area summing.

**Baseline Now** Sets a point (time) at which the integrator resets the baseline to the current height of the data point, if the signal is on a peak.

**Detect Shoulders** Sets points (On/Off) between which the integrator starts and stops detecting shoulders.

If the signal is on the baseline, the function is ignored and the detected baseline is used.

Shoulders are detected according to the specified Shoulders Mode. See "Initial Events" on page 44.

## Integration

### Integration Events

- Fixed Peak Width** Sets the peak width and disables the automatic update of the peak width for the next peaks. To obtain good performance, set the peak width close to the width at half-height of the actual peaks.
- Height Reject** See “Initial Events” on page 44.
- Integration** Sets points (On/Off) between which the integrator stops and starts integrating. Peaks between the times where the integrator is turned off and on are ignored.

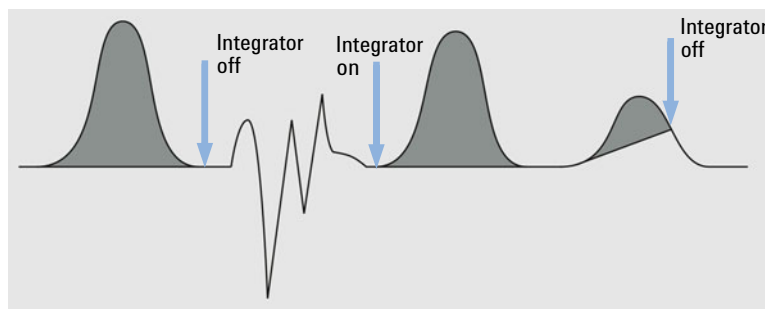


Figure 32 Integration event

The baseline is drawn from the last declared point including any resets for penetration. All other integrator functions together with set changes of peak width, threshold and area reject are ignored when the integrator is turned off. At the **On** and **Off** points, the baseline point is re-established.

When the integrator is set to restart, a new baseline point is reset at the current signal level.

This function is useful for ignoring parts of the chromatogram/electropherogram or to eliminate baseline disturbances.

- Maximum Area** Sets the area of the largest peak of interest. Any peaks that have areas greater than the maximum area are not reported: The integrator rejects any peaks that are greater than the maximum area value after baseline correction.
- You can use this event, for example, to exclude the solvent peak of a GC chromatogram from the integration results, but include its rider peaks.

#### Maximum Height

Sets the height of the largest peak of interest.

Any peaks that have heights greater than the maximum height are not reported: The integrator rejects any peaks that are higher than the maximum height value after baseline correction.

You can use this event, for example, to exclude the solvent peak of a GC chromatogram from the integration results, but include its rider peaks.

#### Negative Peak

Sets points (On/Off) between which the integrator recognizes negative peaks.

When negative peaks are recognized, the integrator no longer automatically resets the baseline after penetration. From now on any penetration of the baseline will be integrated using the established baseline as zero. Areas are constructed relative to this baseline and are given an absolute value.

The negative peak function can only be used with confidence when the baseline drift is small compared with peak size, since the baseline is constructed from the declared baseline point at the start of the peak cluster up to the established baseline at the end of the peak.

#### NOTE

Area Summation is automatically deactivated if the **Negative Peaks On** event is activated.

Tangent skimming is also deactivated during negative peak detection; such peaks are separated by dropline.

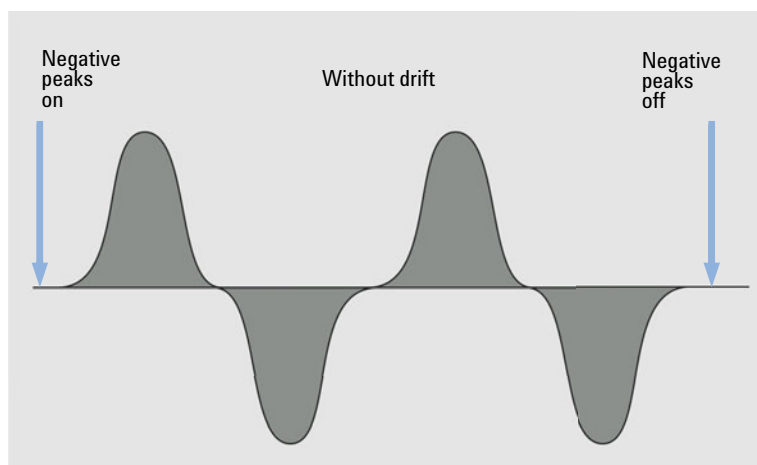


Figure 33 Negative peak event

#### Set Baseline from Range

Uses a range of data points to calculate a statistically meaningful baseline point at the midpoint of a time-range.

The value that you provide is the time interval around a specified point in time. It defines the range to be used to determine the baseline point. See “[Baseline Correction Modes](#)” on page 30 for details of the statistical calculations of the baseline.

If you set the value =0, the nearest chromatogram data point is used as a baseline point; no statistics is done at all. If you set a negative value, the setting does the same as **Use baseline from range=Clear**. It stops the usage of the statistical baseline algorithm.

You can specify any area in the chromatogram for the baseline calculation. Ideally, it should be an area that is free from chemical background and contains only noise. [Figure 34](#) on page 56 illustrates the setting of the baseline range interval which is shown as a shaded gray area.

If you specify two **Set Baseline from Range** points (for example at the beginning and end of a chromatogram), the baseline between them is connected with a straight line.

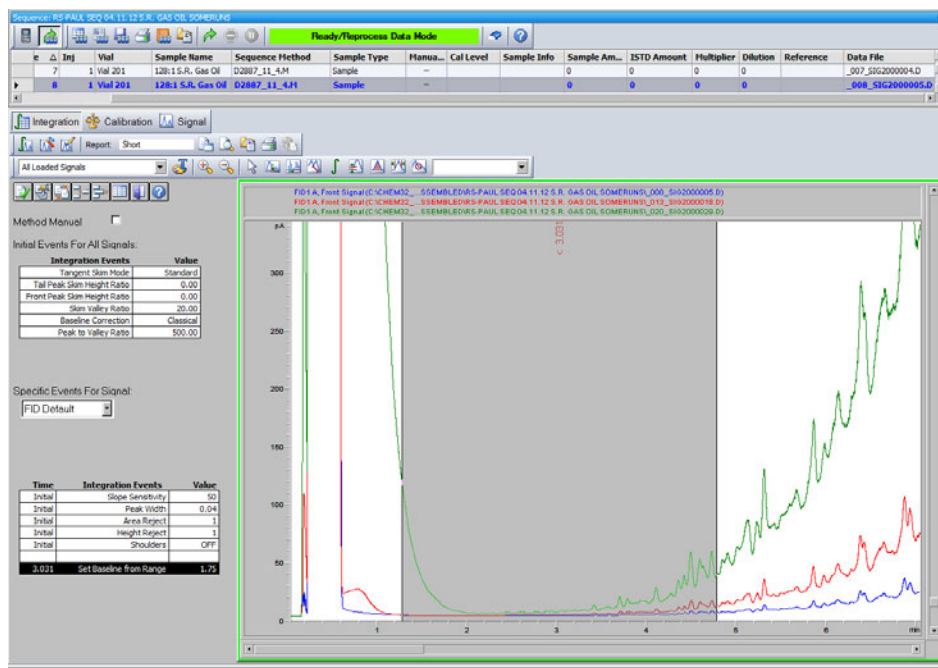


Figure 34 "Set Baseline from Range" : The baseline range interval is indicated by a gray shade



#### **Set Low Baseline from Range**

Similar to **Set Baseline from Range**, but uses the lowest likely baseline point which allows 30 % more noise data points to be above it. Thus, baseline penetration is minimized.

The value that you provide is the time interval around the specified time of the event. It defines the range to be used to determine the baseline point. See “[Baseline Correction Modes](#)” on page 30 for details of the statistical calculations of the baseline.

If you set the value =0, the nearest chromatogram data point is used as a baseline point; no statistics is done at all. If you set a negative value, the setting does the same as **Use baseline from range=Clear**. It stops the usage of the statistical baseline algorithm.

You can specify any time and interval in the chromatogram for the baseline calculation. Ideally, it should be free from chemical background and contain only noise.

If you specify two **Set Baseline from Range** points (for example at the beginning and end of a chromatogram), the baseline between them is connected with a straight line.

Use **Set Low Baseline from Range** instead of **Set Baseline from Range** when the area of the chromatogram used for the calculation contains excessive chemical noise or electronic noise spikes.

#### **Shoulders**

See “[Initial Events](#)” on page 44.

#### **Slope Sensitivity**

See “[Initial Events](#)” on page 44.

#### **Solvent Peak**

Peaks above a specific slope in units of mV/s are detected as solvent peaks that lie outside the range of the analog-to-digital conversion.

The trailing peaks are automatically tangent-skimmed; you do not need to switch on the tangent skim event.

If solvent peak detection is off, droplines are drawn from the trailing peak instead of tangents.

#### **Split Peak**

Specifies a point at which to split a peak with a dropline.

#### **NOTE**

You cannot use **Split Peak** while **Area Sum** is switched on. To split a peak while **Area Sum** is switched on, use the corresponding manual integration event.

You cannot split skimmed peaks using the **Split Peak** event.

**Tail Tangent Skim** Specifies where to start or end tangent skimming.

**On**

Sets a point at which the integrator sets a tangent skim on the trailing edge of the next peak. All peaks above the tangent are integrated to the reset baseline. The tangent is drawn from the valley before the small peak to the point after it where the detector signal gradient is equal to the tangent gradient. The tangent skim event time can be entered any time during the peak. Designates peak also as a solvent peak.

**Off**

Ends tangent skimming after current peak is completed or if in the designated interval no peaks are found (and a solvent will not inadvertently be designated in the next cluster).

**Tangent Skim Mode** The following models are available to calculate suitable peak areas:

- Exponential (see [Figure 18](#) on page 37)
- New Exponential (see [Figure 19](#) on page 38)
- Straight (see [Figure 20](#) on page 39)
- Standard (see “[Standard Skims](#)” on page 39)

**Unassigned Peaks** With some baseline constructions, there are small areas that are above the baseline and below the signal, but are not part of any recognized peaks. Normally, such areas are neither measured nor reported. If unassigned peaks is turned on, these areas are measured and reported as unassigned peaks. The retention/migration time for such an area is the midpoint between the start and end of the area.

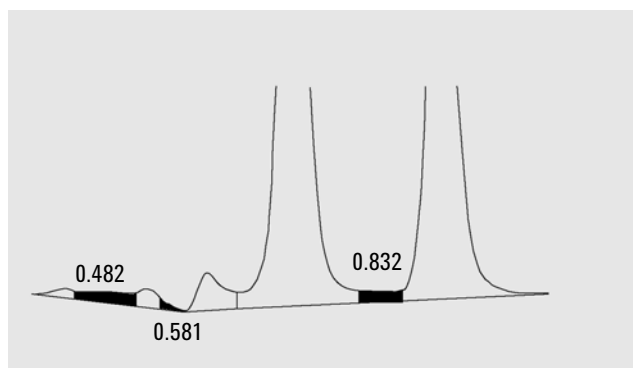


Figure 35 Unassigned Peaks

#### Update Peak Height

This event forces the integrator to use the absolute height of the highest data point as the peak height. Without this event, the maximum of an interpolated curve is used. The Update peak height event is useful especially in signals with extremely steep and angular peaks, or peaks preceded with a downward dip. Peaks like this are typical for MSD signals.

The start time of the Update peak height is not evaluated. The event always affects the entire chromatogram.

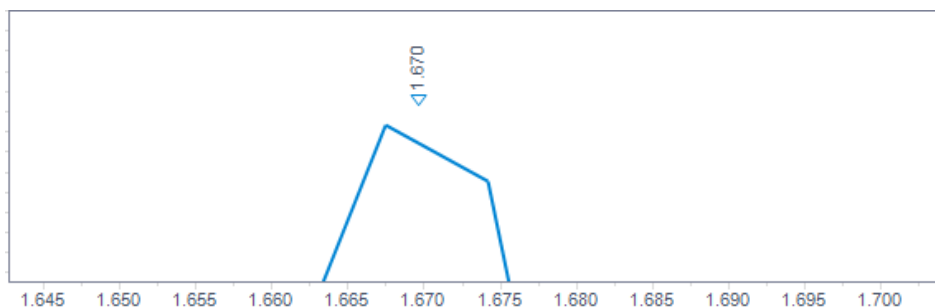


Figure 36 Default integration

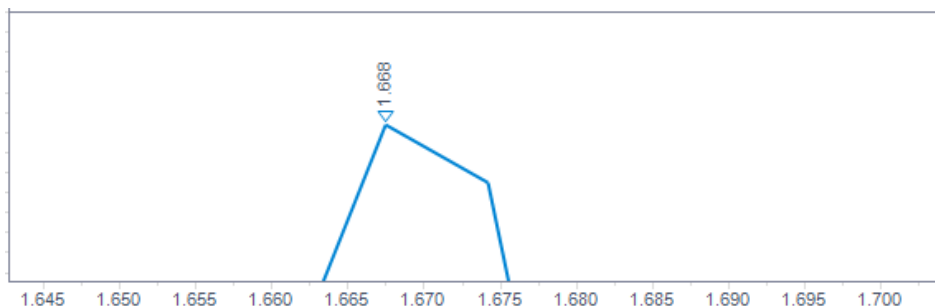


Figure 37 Integration with Update peak height event

#### Use Baseline from Range

Allows to project a baseline value to a later or earlier time to minimize baseline penetrations.

If the **Set Baseline from Range** or **Set Low Baseline from Range** value is calculated in an area with no chromatographic peaks, it can be advantageous to project the calculated baseline to the time immediately before the first peak of interest elutes (or to the time immediately after the last peak of interest has eluted). **Use Baseline from Range** allows you to make up to three such projections in either direction.

This event can be advantageous to use when you have constructed an upslope or downslope baseline, since otherwise the straight baseline might cut through

the chromatogram curve unintentionally. The parameter tells the integrator from which of the baseline ranges to pick the baseline point and project the baseline to the baseline point at the given time interval.

You can use the following parameters:

- **Clear:** Clear the new baseline behavior and return to the traditional algorithm from this point.
- **Left:** Use the baseline value from the baseline range nearest to the left of this point in time.
- **Right:** Use the baseline value from the baseline range nearest to the right of this point in time.
- **Range 1—Range 9:** Use the baseline value from the given baseline range. Baseline ranges are counted from the beginning of the chromatogram.

See also the example under **Area Sum Slice** (Figure 29 on page 50).

## Autointegrate

The **Autointegrate** function provides a starting point for setting initial events. This is particularly useful when you are implementing a new method. You start with a default integration events table that contains no timed events; you can then optimize the parameters proposed by the Autointegrate function for general use.

### Principles of Operation

The **Autointegrate** function reads the chromatogram data and calculates the optimal values for the initial integration parameters for each signal in the chromatogram object.

The algorithm examines 1% at the start and end of the chromatogram and determines the noise and slope for this part. Noise is determined as 3 times the standard deviation of the linear regression divided by the square root of the percent number of points used in the regression. These values are used to assign appropriate values to the height reject & slope sensitivity for the integration. The algorithm then assigns a temporary value for the peak width, depending on the length of the chromatogram, using 0.5% for LC and 0.3% to 0.2% for GC. The initial area reject is set to zero and a trial integration is performed. The trial is repeated several times if necessary, adjusting the parameters each time until at least 5 peaks are detected or integration is performed with an initial height reject of 0. The trial integration is terminated if the above conditions are not met after 10 trials.

## Integration

### Integration Events

The results of the integration are examined and the peak width is adjusted based on the peak widths of the detected peaks, biasing the calculation towards the initial peaks. The peak symmetry of the detected peaks is used to include only those peaks with symmetry between 0.8 and 1.3 for the peak width calculation. If not enough symmetric peaks are found, this limit is relaxed to  $minSymmetry/1.5$  and  $maxSymmetry \times 1.5$ . The baseline between the peaks is then examined to refine the earlier values of height reject & slope sensitivity. The area reject is set to 90% of the minimum area of the most symmetric peak detected during the trial integration.

The chromatogram is re-integrated using these final values for the integration parameters, and the results of the integration are stored.

#### Autointegrate Parameters

The following parameters are set by the autointegrate function:

- Initial slope sensitivity
- Initial height
- Initial peak width
- Initial area reject

## Manual Integration

This type of integration allows you to integrate selected peaks or groups of peaks. Except for the initial area reject value, the software's event integration is ignored within the specified range of manual integration. If one or more of the peaks resulting from manual integration is below the area reject threshold, it is discarded. The manual integration events use absolute time values. They do not adjust for signal drift.

**Manual Integration** enables you to define the peak start and stop points, and then include the recalculated areas in quantification and reporting. Each of these points is labeled in reports with the peak separation code M.

Manual Integration offers the following features:

- Draw Baseline** specifies where the baselines are to be drawn for a peak or set of peaks. With menuitem **Integration >all valleys** you can also specify whether peaks in the range given should be automatically separated at all valley points.
- Negative Peaks** specifies when to treat any areas below the baseline as negative peaks. You can also specify whether peaks in the range given should be automatically separated at all valley points.
- Tangent Skim** calculates the areas of peaks tangentially skimmed off a main peak. The area of the tangent skimmed peak is subtracted from the area of the main peak.
- Split Peak** specifies a point where to split a peak with a drop-line.
- Delete Peak(s)** deletes one or more peaks from the integration results.

### Peak Separation Codes for Manually-Integrated Peaks

Manually-integrated peaks are labeled in the integration reports by the peak code *MM*.

If there is a peak before the manually-integrated peak, and the end of this peak changes because of the manual integration, it is given the code *F* (forced). When valley points are detected they are set to code *V*.

A solvent on main peak which has been affected by manual integration, such as tangent skim, is labeled *R* (re-calculated solvent).

## Saving Manual Integration Events

Manual integration events, e.g. a manually drawn baseline, are even more data file and signal specific than timed integration events. In case of complicated chromatograms, it is highly desirable to be able to use these events for reprocessing. Therefore manual integration events can be stored directly in the data file per signal rather than with the method.

Each time the data file is reviewed or reprocessed, the manual events in the data file are automatically applied. A run containing manual integration events is marked in the **Navigation Table** in the corresponding column.

In addition to the tools for drawing a baseline and deleting a peak manually, three additional tools are available in the user interface to

- Save manual events of the currently shown chromatograms into the data file,
- Remove all events from the currently shown chromatograms,
- Undo the last manual integration events (available until the event is saved).

When continuing to the next data file during review in the **Navigation Table**, ChemStation will check for unsaved manual integration events and ask the user whether he wants to save the events.

Manual events stored in the data file during review in the **Navigation Table** do not interfere with manual integration events stored during review in the **Batch** mode. These two ways of review are completely separated with regard to the manual events of a data file.

In ChemStation revisions prior to B.04.01, manual integration events were stored in the method instead of the individual data file. This workflow can still be used. The **Integration** menu in **Data Analysis** view provides the following items in order to handle manual integration events with the method:

- **Update Manual Events of Method:** Save newly drawn manual events to the method.
- **Apply Manual Events from Method:** Apply the manual events currently saved in the method to the currently loaded data file.
- **Remove Manual Events from Method:** Delete the manual events from the method.

## Integration

### Manual Integration

In order to convert manual events stored in a method to storage in the data file, apply the events from the method and store the results in the data file. If wanted, remove the events from the method.

In case the **Manual Events** checkbox of the **Integration Events Table** of a method is enabled, the manual events of the method are always applied when loading a data file using this method. If the data file contains additional manual events, they are applied after the events of the method. When the **Manual Events** checkbox is enabled, the user is never asked to save the events to the data file.

In order to convert manual events stored in a method to storage in the data file, apply the events from the method and store the results in the data file. You may remove the events from the method now.

In case the **Manual Events** checkbox of the **Integration Events Table** of a method is enabled, the manual events of the method are always applied when loading a data file using this method. If the data file contains additional manual events, they are applied after the events of the method. When the **Manual Events** checkbox is enabled, the user is never asked to save the events to the data file.



## 2 Peak Identification

What is Peak Identification?	66
Peak Matching Rules	67
Types of Peak Identification	68
Absolute Retention/Migration Time	68
Relative Retention Time	68
Corrected Retention/Migration Time	68
Peak Qualifiers	68
Amount Limits	68
Absolute Retention/Migration Time	69
Corrected Retention/Migration Times	71
Single Reference Peaks	71
Multiple Reference Peaks	72
Peak Qualifiers	73
Signal Correlation	74
Qualifier Verification	74
Qualifier Ratio Calculation	74
The Identification Process	75
Finding the Reference Peaks	75
Finding the ISTD Peaks	76
Finding the Remaining Calibrated Peaks	76
Classification of Unidentified Peaks	76

This chapter describes the concepts of peak identification.

## What is Peak Identification?

Peak identification identifies the components in an unknown sample based on their chromatographic/electropherographic characteristics determined by the analysis of a well-defined calibration sample.

The identification of these components is a necessary step in quantification if the analytical method requires quantification. The signal characteristics of each component of interest are stored in the calibration table of the method.

The function of the peak identification process is to compare each peak in the signal with the peaks stored in the calibration table.

The calibration table contains the expected retention/migration times of components of interest. A peak that matches the retention/migration time of a peak in the calibration table is given the attributes of that component, for example, the name and response factor. Peaks that do not match any of the peaks in the calibration table are classified as unknown. The process is controlled by:

- the retention/migration time in the calibration table for peaks designated as time reference peaks,
- the retention/migration time windows specified for reference peaks,
- the retention/migration times in the calibration table for the calibrated peaks that are not time reference peaks,
- the retention/migration time window specified for these non-reference peaks, and
- the presence of any additional qualifying peaks in the correct ratios.

## Peak Matching Rules

The following rules apply to the peak matching process:

- if a sample peak falls within the peak matching window of a component peak from the calibration table, the peak is given the attributes of that component,
- if more than one sample peak falls within the peak matching window, then, the peak closest to the expected retention/migration time is identified as that component,
- if a peak is a time reference or internal standard, then the largest peak in the window is identified as that component,
- if peak qualifiers are also used then the peak ratio is used in combination with the peak matching window to identify the component peak,
- if the peak is a qualifier peak, the measured peak closest to the main peak of the compound is identified, and
- if a sample peak does not fall in any peak matching window, it is listed as an unknown component.

## Types of Peak Identification

There are different techniques that can be used to match sample peaks with those in the calibration table of the ChemStation software.

### **Absolute Retention/Migration Time**

The retention/migration time of the sample peak is compared with the expected retention/migration time specified for each component in the calibration table.

### **Relative Retention Time**

The system calculates Relative retention time (EP) and Relative retention time (USP) as  $(R_r = t_2/t_1)$  both for calibrated peaks and for uncalibrated peaks.

### **Corrected Retention/Migration Time**

The expected retention/migration times of component peaks are corrected using the actual retention/migration times of one or more reference peaks, and the matching process is done using these corrected (relative) retention/migration times. The reference peak or peaks must be specified in the calibration table.

### **Peak Qualifiers**

In addition to identifying peaks by retention/migration time, you can use peak qualifiers to allow a more precise result. If more than one peak occurs in a retention/migration time window then qualifiers should be used to identify the correct compound.

### **Amount Limits**

The amount limits defined in the Compound Details dialog box are used to qualify the peak identification. If the amount of the identified compound is inside the amount limits the peak identification is indicated in the report (classic reporting only, not in intelligent reporting).

## Absolute Retention/Migration Time

A retention/migration time window is used in the peak matching process. The retention/migration time window is a window which is centered on the retention/migration time for an expected peak. Any sample peak that falls within this window may be considered as a candidate for component identification.

Figure 38 on page 69 shows a retention/migration time window for peak 2 which is between 1.809 and 2.631 minutes where the expected retention/migration time is 2.22 minutes. There are two possibilities for peak 2. One is at 1.85 minutes and the other at 2.33 minutes. If the expected peak is a non-reference peak, the peak closest to the expected retention/migration time of 2.22 minutes is selected.

If the expected peak is a time reference or internal standard, the largest peak in the window is selected.

In both cases the ChemStation selects the peak at 2.33 minutes. If the two peaks were the same size then the peak closest to the center of the window is chosen.

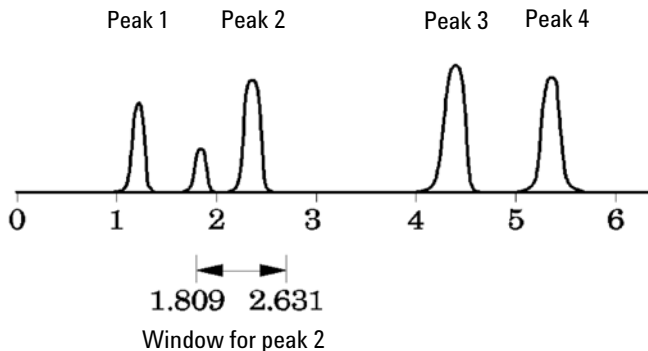


Figure 38 Retention/Migration Time Windows

## Peak Identification

### Absolute Retention/Migration Time

Three types of windows are used when trying to locate peaks.

- reference peak windows which apply to reference peaks only,
- non-reference peak windows which apply to all other calibrated peaks, and
- specific window values for individual components which are set in the **Compound Details** dialog box.

The default values for these windows are entered in the Calibration Settings dialog box. The width on either side of the retention/migration time that defines the peak matching window is the sum of the absolute and percentage windows.

A window of 5 % means the peak must have a retention/migration time between less than 2.5 % and more than 2.5 % of the calibrated retention/migration time for that peak. For example, a peak with a retention/migration time of 2.00 in the calibration run must appear between 1.95 and 2.05 minutes in subsequent runs.

For example, an absolute window of 0.20 minutes and a relative window of 10 % gives a retention/migration time window of between 1.80 and 2.20 minutes.

$1.80 \text{ min} = 2.00 \text{ min} - 0.10 \text{ min} (0.20 \text{ min} / 2) - 0.10 \text{ min} (5\% \text{ of } 2.00 \text{ min}).$

$2.20 \text{ min} = 2.00 \text{ min} + 0.10 \text{ min} (0.20 \text{ min} / 2) + 0.10 \text{ min} (5\% \text{ of } 2.00 \text{ min}).$

## Corrected Retention/Migration Times

To match peaks by absolute retention/migration times may be simple but not always reliable. Individual retention/migration times may vary slightly due to a small change in conditions or technique. As a result peaks may occur outside the peak matching windows and therefore are not identified.

A technique to deal with the inevitable fluctuations that occur in absolute retention/migration times is to express component retention/migration times relative to one or more reference peaks.

Reference peaks are identified in the calibration table with an entry in the reference column for that peak. The relative peak matching technique uses the reference peak or peaks to modify the location of the peak matching windows in order to compensate for shifts in the retention/migration times of sample peaks.

If no reference peak is defined in the method or the ChemStation cannot identify at least one reference peak during the run, the software will use absolute retention/migration times for identification.

### Single Reference Peaks

A retention/migration time window for the reference peak is created around its retention/migration time. The largest peak falling within this window is identified as the reference peak. The expected retention/migration times of all other peaks in the calibration table are corrected, in proportion to the ratio of the expected retention/migration time to the actual retention/migration time of the reference peak.

## Multiple Reference Peaks

Correcting retention/migration times with a single reference peak is based on the assumption that the deviation of actual retention/migration time from the expected retention/migration times changes uniformly and linearly as the run progresses. Often during a long run the retention/migration time changes non-uniformly. In such cases better results are obtained using multiple reference peaks spaced at intervals across the run. This splits the signal into separate zones. Within each zone the deviation between retention/migration times is assumed to change linearly, but the rate of change is determined separately for each zone.

**NOTE**

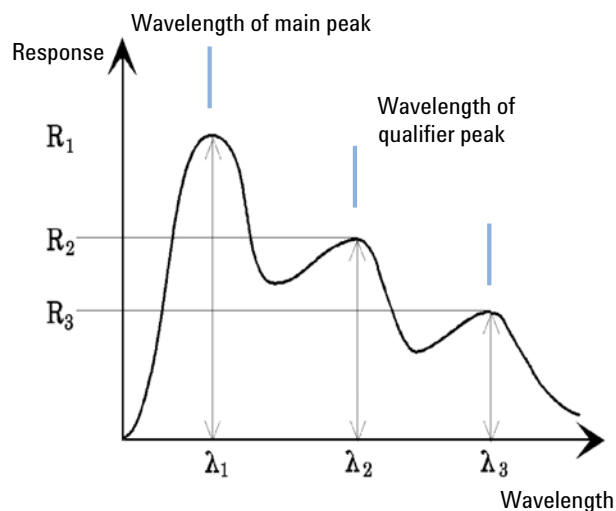
The time correction algorithm may fail if the retention times of multiple reference peaks are too close to each other and are not distributed across the total run time.



## Peak Qualifiers

A component can be detected with more than one signal. Although applicable to all forms of chromatography using multiple detectors or detectors capable of producing multiple signals, multisignal detection is most commonly used in liquid chromatography with multiple wavelength or diode array detectors. Such detectors are normally set up so that the wavelength closest to the greatest absorbance (area) is used to define the main peak in the calibration table. In Figure 39 on page 73 this is  $\lambda_1$ .

The two other wavelengths that were acquired as signals can be used as peak qualifiers. In the figure these are  $\lambda_2$  and  $\lambda_3$ .



**Figure 39** Peak Qualifiers

Peaks of a compound have a constant response ratio over different wavelengths.

The qualifier peak response is a certain percentage of the main peak response. Limits which determine the acceptable range for the expected response can be set in the calibration table when the Identification Details option is selected. If the ratio between the main peak qualifier  $\lambda_1$  and the qualifier peak, for example,  $\lambda_3$  is within the allowed limits then the compound identity can be confirmed.

## Signal Correlation

Signal correlation means that two peaks measured in different detector signals within a defined time window are assigned to the same compound. The signal correlation window can be controlled by the **SignalCorrWin** parameter in the **QuantParm** table of the **\_DaMethod** register. Signal correlation is disabled when setting the signal correlation window to 0.0 minutes (see the Online Help for more information). When signal correlation is off, peaks eluting at the same retention/migration time in different detector signals are treated as different compounds.

The default signal correlation window for LC, CE, CE/MS and LC/MS data is 0.03 minutes and 0.0 minutes for GC data.

## Qualifier Verification

If signal correlation is enabled, qualifier verification is active for all data file types by default. It can be disabled by setting the **UseQualifiers** flag in the **Quantification Parameters** table of the method. Qualifier verification is also disabled when signal correlation is switched off.

## Qualifier Ratio Calculation

When qualifiers verification is enabled for a compound, the ratio of the qualifier size and the main peak size is verified against the calibrated limits. The size may be height or area according to the calculation base setting in Specify Report.

The qualifier peaks can be calibrated in the same way as the target compounds. The user does not need to specify the expected qualifier ratio. The expected qualifier ratio is calculated automatically:

both measured at the retention time of the compound.

The QualTolerance parameter defines the acceptable range of the qualifier ratio, for example,  $\pm 20\%$ .

The tolerance can be set in the calibration table user interface (Identification Details) and is an absolute percentage.

## Peak Identification

### The Identification Process

For multilevel calibrations, the ChemStation calculates a minimum qualifier tolerance based on the measured qualifier ratios at each calibration level. The minimum qualifier tolerance is calculated using the following equation:

$$\text{minimum qualifier tolerance} = \frac{\sum_{i=1}^n (q_i - \bar{q})}{\bar{q} \times i} \times 100$$

where  $q_i$  is the measured qualifier ratio at level  $i$ .

## The Identification Process

When attempting to identify peaks, the software makes three passes through the integration data.

### Finding the Reference Peaks

The first pass identifies the time reference peaks. The software searches peak retention/migration times from a run for matches within the retention/migration windows of the reference peaks in the calibration table. A peak from the run is identified as a reference peak in the calibration table if the run peak's retention/migration time is within the window constructed for the calibration table peak.

If more than one peak is found within a window, the peak with the largest area or height followed by a positive signal qualifier match, if set up, is chosen as the reference peak.

After each time reference peak is found, the difference between its retention/migration time and that given in the calibration table is used to adjust the expected retention/migration times of all other peaks in the Calibration table.

## Finding the ISTD Peaks

The second pass identifies any defined internal standard peaks. If they have not already been identified as ISTD, peaks may be identified as time reference peaks. ISTD peaks are identified by peak retention/migration time windows and peak qualifiers. If more than one peak is found in the same ISTD window, the largest peak is chosen.

## Finding the Remaining Calibrated Peaks

The third pass identifies all remaining peaks listed in the calibration table. The non-reference peaks in the calibration table are matched to the remaining run peaks by using their RT window.

Each non-reference calibrated peak has its own retention/migration time in the calibration table. This is adjusted for the particular run based on the pre-identification of the time reference peaks. The retention/migration time window of the calibrated peak is adjusted based on the corrected retention/migration time of the calibrated peak.

If more than one peak is found in the same window, the peak with a retention/migration time which is closest to the expected retention/migration time and also meets the optional qualifier specifications is chosen.

## Classification of Unidentified Peaks

If there are remaining peaks, which are still not identified, they are classified as unknown. The ChemStation attempts to group the unknown peaks that belong to the same compound. If a peak has been detected in more than one signal, the peaks with the same retention/migration time in each signal are grouped to one compound.

Unknown peaks are reported in classic reporting if the corresponding selection has been made in the **Specify Report** dialog box.



## 3

# Calibration

What is Calibration?	78
Calibration Curve	79
Calibration Curve Calculation	80
Linear Fit	80
Quadratic Fit	80
Relative Residuals	83
Group Calibration	84
Recalibration Options	85

This chapter contains details of the calculations used in the calibration process.

## What is Calibration?

After the peaks have been integrated and identified, the next step in the quantitative analysis is the calibration. The amount and response is rarely in direct proportion to the actual mass of the sample to be analyzed. This makes the calibration with reference materials necessary. Quantitation uses peak area or height to determine the amount of a compound in a sample.

A quantitative analysis involves many steps which are briefly summarized as follows:

- Know the compound you are analyzing.
- Establish a method for analyzing samples containing a known amount of this compound, which is called the calibration sample or standard.
- Analyze the calibration sample to obtain the response due to that amount. You may alternatively analyze a number of these standards with different amounts of the compounds of interest if your detector has a non-linear response. This process is referred to as *multi-level calibration*.

With the following calibration methods you can perform quantitation:

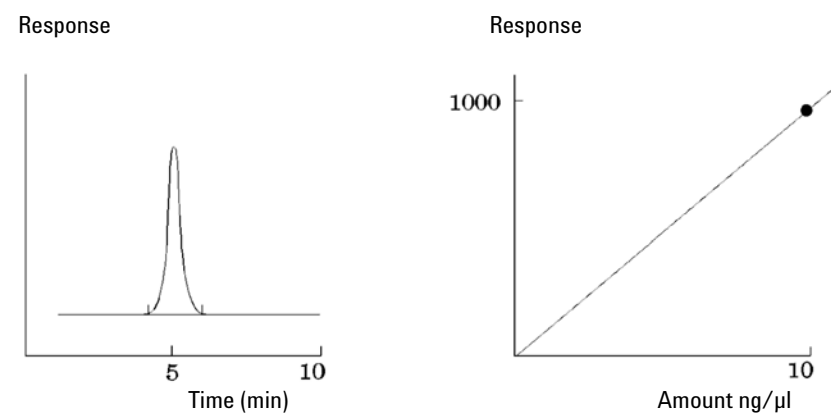
- Compound specific calibration (ESTD, ISTD)
- Indirect quantitation using calibration or response factor from another compound or group
- Fixed response factor (**Manual Factor**)

The ESTD calibration curves and calculations are based on measured responses (area or height) of given amounts. The ISTD calibration curves and calculations are based on relative responses and relative amounts.

## Calibration Curve

A calibration curve is a graphical presentation of the amount and response data for one compound obtained from one or more calibration samples.

Normally an aliquot of the calibration sample is injected, a signal is obtained, and the response is determined by calculating the area or height of the peak, similar to [Figure 40](#) on page 79.



**Figure 40** Calibration Sample (10 ng/μl) Signal and Calibration Curve

## Calibration Curve Calculation

### Linear Fit

N = number of discrete observations

$X_i$  = independent variable,  $i^{\text{th}}$  observation

$Y_i$  = dependent variable,  $i^{\text{th}}$  observation

Curve formula:

$$y_{(X)} = a + bX$$

Coefficients:

$$a = \frac{1}{\Delta x} \left( \sum_{i=1}^N X_i^2 * \sum_{i=1}^N Y_i - \left( \sum_{i=1}^N X_i * \sum_{i=1}^N X_i Y_i \right) \right)$$

$$b = \frac{1}{\Delta x} \left( N * \sum_{i=1}^N X_i Y_i - \left( \sum_{i=1}^N X_i * \sum_{i=1}^N Y_i \right) \right)$$

where:

$$\Delta x = N * \sum_{i=1}^N X_i^2 - \left( \sum_{i=1}^N X_i \right)^2$$

### Quadratic Fit

Quadratic curve formula:

$$y = a + (b * x) + (c * x^2)$$

At least three calibration points are required for the quadratic fit. Two points are required if the origin is included or forced.



## Calibration

### Calibration Curve Calculation

Calculation of coefficients for quadratic fit

The coefficients result from the below simultaneous linear equations. Crout's algorithm is used to solve the corresponding normal matrix equation ( $A^T A x = A^T y$ ). In the given formula, sums are abbreviated as:

$$\begin{aligned} W &= \sum(wt) \\ XW &= \sum(x * wt) \\ X2W &= \sum(x^2 * wt) \\ X3W &= \sum(x^3 * wt) \\ X4W &= \sum(x^4 * wt) \\ YW &= \sum(y * wt) \\ XYW &= \sum(x * y * wt) \\ X2YW &= \sum(x^2 * y * wt) \end{aligned}$$

In order to avoid overflow, the x-values are normalized before entering calculation:

$$\text{Norm} = \sum(x)$$

$$x = x / \text{Norm}$$

Normal equations for quadratic curve:

$$\begin{aligned} \sum(wt) * a + \sum(x * wt) * b + \sum(x^2 * wt) * c &= \sum(y * wt) \\ \sum(x * wt) * a + \sum(x^2 * wt) * b + \sum(x^3 * wt) * c &= \sum(x * y * wt) \\ \sum(x^2 * wt) * a + \sum(x^3 * wt) * b + \sum(x^4 * wt) * c &= \sum(x^2 * y * wt) \end{aligned}$$

Or written as matrix equation:

$$\begin{bmatrix} W & XW & X2W \\ XW & X2W & X3W \\ X2W & X3W & X4W \end{bmatrix} * \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} YW \\ XYW \\ X2YW \end{bmatrix}$$

Crout's decomposition:

$$\begin{bmatrix} W & XW & X2W \\ XW & X2W & X3W \\ X2W & X3W & X4W \end{bmatrix} = \begin{bmatrix} L11 & & \\ L21 & L22 & \\ L31 & L32 & L33 \end{bmatrix} * \begin{bmatrix} 1 & U12 & U13 \\ & 1 & U23 \\ & & 1 \end{bmatrix}$$

With value abbreviations:

$$\begin{aligned}
 L_{11} &= W \\
 U_{12} &= \frac{XW}{L_{11}} \\
 L_{21} &= XW \\
 U_{13} &= \frac{X^2W}{L_{11}} \\
 L_{31} &= X^2W \\
 L_{22} &= X^2W - L_{21} * U_{12} \\
 U_{23} &= \frac{X^3W - L_{21} * U_{13}}{L_{22}} \\
 L_{32} &= X^3W - L_{31} * U_{12} \\
 L_{33} &= X^4W - (L_{31} * U_{13}) - (L_{32} * U_{23})
 \end{aligned}$$

$$\begin{aligned}
 z_0 &= \frac{YW}{L_{11}} \\
 z_1 &= \frac{XYW - (L_{21} * z_0)}{L_{22}} \\
 z_2 &= \frac{X^2YW - (L_{31} * z_0) - (L_{32} * z_1)}{L_{33}} \\
 c' &= z_2 \\
 b' &= z_1 - (U_{23} * c') \\
 a' &= z_0 - (U_{12} * b') - (U_{13} * c')
 \end{aligned}$$

Finally, the normalization must be reversed:

$$\begin{aligned}
 a &= a' \\
 b &= \frac{b'}{\text{Norm}} \\
 c &= \frac{c'}{\text{Norm}^2}
 \end{aligned}$$

## Relative Residuals

For each calibration level the *relative residual* is displayed. It is calculated using the following formula:

$$relRES = \frac{Response_{calibrated} - Response_{calculated}}{Response_{calculated}} \cdot 100$$

where:

relRES = relative residual in percent

The calculated response represents the point on the calibration curve.

The *residual standard deviation*, which is printed on some reports and when selecting Print calibration table and curves is calculated using the following formula:

$$ResSTD = \sqrt{\frac{\sum_{i=1}^n (Resp_{calibratedi} - Resp_{calculatedi})^2}{n - 2}}$$

where:

ResSTD = residual standard deviation

Resp<sub>calibratedi</sub> = calibrated response for point i

Resp<sub>calculatedi</sub> = calculated response for point i

n = number of calibration points

## Group Calibration

Group calibration can be applied for compounds where the individual concentrations are not known but the sum of concentrations for a group of compounds is known. An example are isomers. Complete compound groups are calibrated. The following formulae are used:

Calibration

$$Conc_{AB} = RF_A \cdot Response_A + RF_B \cdot Response_B$$

where:

$Conc_{AB}$  is the concentration of the compound group consisting of compound A and B

$Response_A$  is the area (or height) of compound A

$RF_A$  is the response factor

For compounds within a compound group we assume equal response factors:

$$RF_A = RF_B$$

Therefore the concentration of a compound within a compound group is calculated as follows:

$$Conc_A = \frac{Conc_{AB} \cdot Resp_A}{Resp_A + Resp_B}$$

## Recalibration Options

You have several ways to update the responses in the calibration table with the new calibration data.

### Average

The average from all calibration runs are calculated using the following formula

$$Response = \frac{n \cdot Response + MeasResponse}{n + 1}$$

### Floating Average

A weighted average for all calibration runs is calculated. The updated weight is set in the **Recalibration Settings** dialog box.

$$Response = \left(1 - \frac{Weight}{100}\right) \cdot Response + \left(\frac{Weight}{100}\right) \cdot MeasResponse$$

### Replace

The new response values replace the old values.

# 4

## Quantitation

What is Quantitation?	87
Quantitation Calculations	88
Correction Factors	88
Absolute Response Factor	88
Multiplier	89
Dilution Factor	89
Sample Amount	89
Area% and Height%	90
Quantitation of Calibrated Compounds	91
ESTD Calculation	91
ISTD Calculation	93
Quantitation of Uncalibrated Peaks	96
Indirect quantitation using a calibrated compound	96
Quantitation using a manual factor	97
Norm% Calculation	98

This chapter describes how compounds are quantified, and explains the calculations used in quantitation.

## What is Quantitation?

After the peaks have been integrated and identified, the next step in the analysis is quantification. Quantification uses peak area or height to determine the concentration of a compound in a sample.

A quantitative analysis involves many steps which are briefly summarized as follows:

- Know the compound you are analyzing.
- Establish a method for analyzing samples containing this compound.
- Analyze a sample or samples containing a known concentration or concentrations of the compound to obtain the response due to that concentration.

You may alternatively analyze a number of these samples with different concentrations of the compounds of interest if your detector has a non-linear response. This process is referred to as *multi-level calibration*.

- Analyze the sample containing an unknown concentration of the compound to obtain the response due to the unknown concentration.
- Compare the response of the unknown concentration to the response of the known concentration to determine how much of the compound is present.

To obtain a valid comparison for the unknown sample response to that of the known sample, the data must be acquired and processed under identical conditions.

## Quantitation Calculations

The ChemStation offers the following calculation procedures for determining the concentration of each component present in a mixture:

- Percent
- Normalization
- External standard (ESTD)
- ESTD%
- Internal standard (ISTD)
- ISTD%

The calculations used to determine the concentration of a compound in an unknown sample depend on the type of quantification. Each calculation procedure uses the peak area or height for the calculation and produces a different type of report.

## Correction Factors

The quantification calculations use four correction factors, the *absolute response factor*, the *multiplier*, the *dilution factor*, and the *sample amount*. These factors are used in the calibration procedures to compensate for variations in detector response to different sample components, concentrations, sample dilutions, sample amounts, and for converting units.

## Absolute Response Factor

The absolute response factor for a sample component represents the amount of the component divided by the measured area or height of the component's peak in the analysis of a calibration mixture. The absolute response factor, which is used by each calibrated calculation procedure, corrects for detector response to individual sample components.



## Multiplier

The multiplier is used in each calculation formula to multiply the result for each component. The multiplier may be used to convert units to express amounts.

## Dilution Factor

The dilution factor is a number by which all calculated results are multiplied before the report is printed. You can use the dilution factor to change the scale of the results or correct for changes in sample composition during pre-analysis work. You can also use the dilution factor for any other purposes that require the use of a constant factor.

## Sample Amount

If the ESTD% or ISTD% calculations are selected, the ESTD and ISTD reports give relative values rather than absolute values, that is, the amount of each component is expressed as a percentage of the sample amount. The sample amount is used in ESTD% and ISTD% reports to convert the absolute amount of the components analyzed to relative values by dividing by the value specified.

## Area% and Height%

The **Area%** calculation procedure reports the area of each peak in the run as a percentage of the total area of all peaks in the run. **Area%** does not require prior calibration and does not depend upon the amount of sample injected within the limits of the detector. No response factors are used. If all components respond equally in the detector, then **Area%** provides a suitable approximation of the relative amounts of components.

**Area%** is used routinely where qualitative results are of interest and to produce information to create the calibration table required for other calibration procedures.

The **Height%** calculation procedure reports the height of each peak in the run as a percentage of the total height of all peaks in the run.

The multiplier and dilution factor from the **Calibration Settings**, from the **Sample Information** dialog box, or from the **Sequence Table** are not applied in Area% or Height% calculation.

## Quantitation of Calibrated Compounds

The external standard (ESTD), normalization, and internal standard (ISTD) calculation procedures require response factors and therefore use a calibration table. The calibration table specifies conversion of responses into the units you choose by the procedure you select.

### ESTD Calculation

The ESTD procedure is the basic quantification procedure in which both calibration and unknown samples are analyzed under the same conditions. The results from the unknown sample are then compared with those of the calibration sample to calculate the amount in the unknown.

The ESTD procedure uses absolute response factors unlike the ISTD procedure. The response factors are obtained from a calibration and then stored. In following sample runs, component amounts are calculated by applying these response factors to the measured sample amounts. Make sure that the sample injection size is reproducible from run to run, since there is no standard in the sample to correct for variations in injection size or sample preparation.

When preparing an ESTD report, the calculation of the amount of a particular compound in an unknown sample occurs in two steps:

- 1 An equation for the curve through the calibration points for this compound is calculated using the type of fit specified in the Calibration Settings or Calibration Curve dialog box.
- 2 The amount of the compound in the unknown is calculated using the equation described below. This amount may appear in the report or it may be used in additional calculations called for by Multiplier, Dilution Factor, or Sample Amount values before being reported.

## Quantitation

### Quantitation of Calibrated Compounds

If the ESTD report is selected, the equation used to compute the absolute amount of component x is:

$$\text{Absolute Amt of } x = \text{Response}_x \cdot RF_x \cdot M \cdot D$$

where:

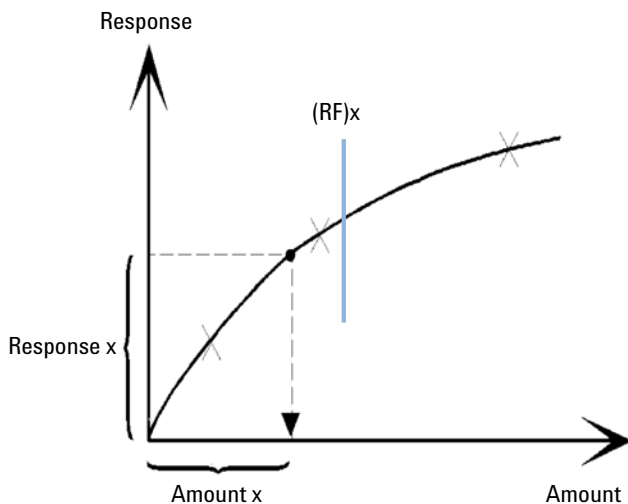
$\text{Response}_x$  is the response of peak x;

$RF_x$  is the response factor for component x, calculated as:

$$RF_x = \frac{\text{Amount}_x}{\text{Response}_x}$$

$M$  is the multiplier.

$D$  is the dilution factor.



**Figure 41** Response Factor

The multiplier and dilution factor are read either from the **Calibration Settings** or from the **Sample Information** dialog box.

## Quantitation

### Quantitation of Calibrated Compounds

If the ESTD% report is selected and sample amount is not zero, the relative amount (%) of a component x is calculated as shown below:

$$\text{Relative Amt of } x = \frac{(\text{Absolute Amt of } x) \cdot 100}{\text{Sample Amount}}$$

where:

*Absolute amount of x* is calculated as shown above in the ESTD calculation;

*Sample amount* is obtained from the Sample Information box, or from the Quantitation Settings dialog box for single runs. If sample amount is zero, the ESTD is calculated.

## ISTD Calculation

The ISTD procedure eliminates the disadvantages of the ESTD method by adding a known amount of a component which serves as a normalizing factor. This component, the *internal standard*, is added to both calibration and unknown samples.

The software takes the appropriate response factors obtained from a previous calibration stored in the method. Using the internal standard concentration and peak areas or heights from the run, the software calculates component concentrations.

The compound used as an internal standard should be similar to the calibrated compound, both chemically and in retention/migration time, but it must be chromatographically distinguishable.

**Table 8** ISTD procedure

Advantages	Disadvantages
Sample-size variation is not critical.	The internal standard must be added to every sample.
Instrument drift is compensated by the internal standard.	
The effects of sample preparations are minimized if the chemical behavior of the ISTD and unknown are similar.	

If the ISTD procedure is used for calibrations with a non-linear characteristic, care must be taken that errors which result from the calculation principle do not cause

## Quantitation

### Quantitation of Calibrated Compounds

systematic errors. In multi-level calibrations, the amount of the ISTD compound should be kept constant, i.e. the same for all levels if the calibration curve of the compound is non-linear.

In the internal standard analysis, the amount of the component of interest is related to the amount of the internal standard component by the ratio of the responses of the two peaks.

In a two-run ISTD calibration, the calculation of the corrected amount ratio of a particular compound in an unknown sample occurs in the following stages:

#### Run 1: Calibration

- 1 The calibration points are constructed by calculating an amount ratio and a response ratio for each level of a particular peak in the calibration table.

The amount ratio is the amount of the compound divided by the amount of the internal standard at this level.

The response ratio is the area of the compound divided by the area or height of the internal standard at this level.

- 2 An equation for the curve through the calibration points is calculated using the type of curve fit specified in the Calibration Settings dialog box or Calibration Curve dialog box.

$$RF_x = \frac{\text{Amount Ratio}}{\text{Response Ratio}}$$

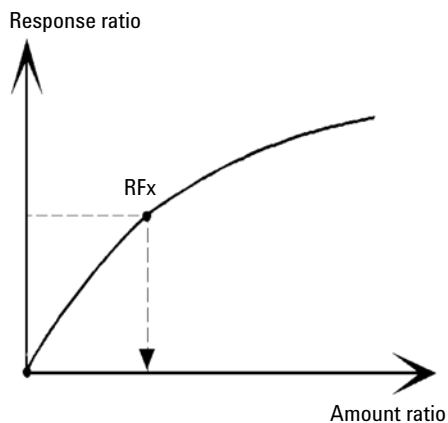


Figure 42 Amount Ratio

### Run 2: Unknown Sample

- 1 The response of the compound in the unknown sample is divided by the response of the internal standard in the unknown sample to give a response ratio for the unknown.
- 2 An amount ratio for the unknown is calculated using the curve fit equation determined in step 2 above, and the actual amount of ISTD in the sample.

### ISTD Calculation of Calibrated Compounds

The equations used to calculate the actual amount of a calibrated compound x for a single-level calibration are:

$$\text{Response Ratio} = \frac{\text{Response}_x}{\text{Response}_{\text{ISTD}}}$$

$$\text{Actual Amount of } x = RF_x \cdot (\text{Response Ratio})_x \cdot \text{Actual Amount of ISTD} \cdot M \cdot D$$

where:

$RF_x$  is the response factor for compound x;

The actual amount (*Actual Amt*) of ISTD is the value that was entered in the Calibration Settings dialog box or the Sample Info dialog box for the internal standard added to the unknown sample;

M is the multiplier.

D is the dilution factor.

If the ISTD% report type is selected, the following equation is used to calculate the relative (%) amount of component x:

$$\text{Relative Amt of } x = \frac{[\text{Absolute Amt of } x] \cdot 100}{\text{Sample Amount}}$$

## Quantitation of Uncalibrated Peaks

Uncalibrated peaks can be quantified either with a fixed response factor or using the calibration data of one of the calibrated compounds. Quantitation using a fixed response factor or calibrated compound data is signal-specific. In the latter case, if the calibrated compound is quantified by an ISTD method, the ISTD is used for the not identified peaks in the same way as for the calibrated compound.

### Indirect quantitation using a calibrated compound

If the calibration data of a calibrated compound is to be used to quantify uncalibrated peaks, the calibrated compound is selected from the drop-down list **Using Compound** in the **Calibration Settings** dialog box. Calculations are the same as for calibrated compounds. If the reference compound is quantified by an ISTD method, the ISTD is used for the uncalibrated compound in the same way as for the reference compound.

A missing reference peak results in a zero amount of the uncalibrated peak.



## Quantitation using a manual factor

The software allows you to quantify an identified compound that is based on a fixed response factor (**With Rsp Factor** in the **Calibration Settings** dialog box). In this case, the compound amount is calculated using the fixed response factor:

$$\text{Amount} = \text{Response} * \text{RspFactor} * M * D$$

where

Manual Factor

Fixed response factor

Response

Response can be Area or Height

Using a manual factor with an ISTD method

If the compound amount is quantified using the fixed response factor and ISTD, the formula is read as follows:

$$\text{Area ratio} = \text{Area} / \text{Area}_{\text{ISTD}}$$

or:

$$\text{Height ratio} = \text{Height} / \text{Height}_{\text{ISTD}}$$

The amount is then calculated as follows:

$$\text{Amount} = \text{Area ratio} * \text{Manual Factor} * \text{Amount}_{\text{ISTD}}$$

or:

$$\text{Amount} = \text{Height ratio} * \text{Manual Factor} * \text{Amount}_{\text{ISTD}}$$

Dependency of manual factor and response factor (RF)

With RF defined as **Response per amount** (default setting):

$$\text{RF} = 1 / \text{Manual Factor}$$

With RF defined as **Amount per response**:

$$\text{RF} = \text{Manual Factor}$$

## Norm% Calculation

In the normalization method, response factors are applied to the peak areas (or heights) to compensate for changes that occur in detector sensitivity for the different sample components.

The Norm% report is calculated in the same way as an ESTD report except that there is an additional step to calculate the relative rather than absolute amounts of compounds.

The Norm% report has the same disadvantage as the Area% and Height% reports. Any changes that affect the total peak area will affect the concentration calculation of each individual peak. The normalization report should only be used if all components of interest are eluted and integrated. Excluding selected peaks from a normalization report will change the reported results in the sample.

The equation used to calculate the **Norm%** of a component x is:

$$\text{Norm\% of } x = \frac{\text{Response}_x \cdot RF_x \cdot 100 \cdot M \cdot D}{\sum(\text{Response} \cdot RF)}$$

where:

$\text{Response}_x$	is the area (or height) of peak x,
$RF_x$	is the response factor,
$\sum(\text{Response} \cdot RF)$	is the total of all the $(\text{Response} \cdot RF)$ products for all peaks including peak x,
$M$	is the multiplier,
$D$	is the dilution factor.

The multiplier and dilution factor are read either from the **Quantitation Settings** available in the **Specify Report** dialog box, or from the Sequence Table.

## 5 System Suitability

Peak Performance Evaluation	100
Evaluating System Suitability	100
Noise Determination	103
Noise Calculation Using Six Times the Standard Deviation	104
Noise Calculation Using the Peak-to-Peak Formula	105
Noise Calculation by the ASTM Method	106
Signal-to-noise calculation	109
Drift and Wander	112
Calculation of Peak Asymmetry and Symmetry	114
System Suitability Formulas and Calculations	116
General Definitions	117
Void Volume void volume	117
Retention Time of Unretained Compound $t(m)$ [min]	117
Performance Test Definitions	118
Overview Performance Tests	118
Statistical Moments, Skew and Excess	120
True Peak Width $W_x$ [min]	122
Capacity Factor (USP)	122
Tailing Factor (USP) $t$	123
Number of Theoretical Plates per Column $n$	124
Number of Theoretical Plates per Meter $N$ [1-m]	125
Relative Retention, Selectivity	126
Resolution (USP, ASTM) $R$	127
Resolution (EP/JP) $R_s$	127
Resolution (ChemStation classic definitions)	127
Peak to Valley Ratio (EP/JP)	128
Definitions for Reproducibility	130
Sample Mean $M$	130
Sample Standard Deviation $S$	131
Relative Standard Deviation RSD[%] (USP)	131
Standard Deviation of the Mean $S$	131
Standard Deviation (S)	132
Confidence Interval CI	132
Correlation Coefficient	133
Internally Stored Double Precision Number Access	135

This chapter describes what OpenLab CDS can do to evaluate the performance of both the analytical instrument and the analytical method.

## Peak Performance Evaluation

Peak Performance can be calculated for any integrated peak of the data loaded, and also for new manually integrated peaks. The Interactive *Peak Performance tool* calculates peak characteristics and displays them on UI. The Peak Performance tool uses the true peak width at various heights, and a retention time evaluated by a peak modelling algorithm. These values are displayed exclusively in the Peak Performance tool UI. They may vary slightly to the integrator's values which are displayed in the reports.

## Evaluating System Suitability

Evaluating the performance of both the analytical instrument before it is used for sample analysis and the analytical method before it is used routinely is good analytical practice. It is also a good idea to check the performance of analysis systems before, and during, routine analysis. The ChemStation software provides the tools to do these three types of tests automatically. An instrument test can include the detector sensitivity, the precision of peak retention/migration times and the precision of peak areas. A method test can include precision of retention/migration times and amounts, the selectivity, and the robustness of the method to day-to-day variance in operation. A system test can include precision of amounts, resolution between two specific peaks and peak tailing.

Laboratories which have to comply with:

- Good Laboratory Practice regulations (GLP),
- Good Manufacturing Practice regulations (GMP) and Current Good Manufacturing Practice regulations (cGMP), and
- Good Automated Laboratory Practice (GALP).

Laboratories are advised to perform these tests and to document the results thoroughly. Laboratories which are part of a quality control system, for example, to comply with ISO9000 certification, will have to demonstrate the proper performance of their instruments.

The ChemStation collates results from several runs and evaluates them statistically in the sequence summary report.

The tests are documented in a format which is generally accepted by regulatory authorities and independent auditors. Statistics include:

- peak retention/migration time,
- peak area,
- amount,
- peak height,
- peak width at half height,
- peak symmetry,
- peak tailing,
- capacity factor ( $k'$ ),
- plate numbers,
- resolution between peaks,
- selectivity relative to preceding peak,
- skew, and
- excess.

The mean value, the standard deviation, the relative standard deviation and the confidence interval are calculated. You can set limits for either standard deviation, the relative standard deviation or the confidence interval for each of these parameters. Should the values exceed your limits, the report is flagged to draw your attention to them.

The quality of the analytical data can be supported by keeping records of the actual conditions at the time the measurements were made. This information is stored with the data and reported with sample data. Instrument performance curves are recorded during the entire analysis as signals, and stored in the data file. If supported by the instrument these records, overlaid on the chromatogram, can be recalled on demand, for example, during an audit.

Baseline noise and drift can be measured automatically. A minimum detectable level can be calculated from peak height data for each calibrated compound in the method.

Finally, instrument configuration, instrument serial numbers, column/capillary identification, and your own comments can be included in each report printed.

Extended performance results are calculated only for compounds calibrated for in the method, ensuring characterization by retention/migration times and compound names.

A typical system performance test report contains the following performance results:

- instrument details,
- column/capillary details,
- analytical method,
- sample information,
- acquisition information,
- signal description and baseline noise determination, and
- signal labeled with either retention/migration times, or compound names.

In addition, the following information is generated for each calibrated compound in the chromatogram:

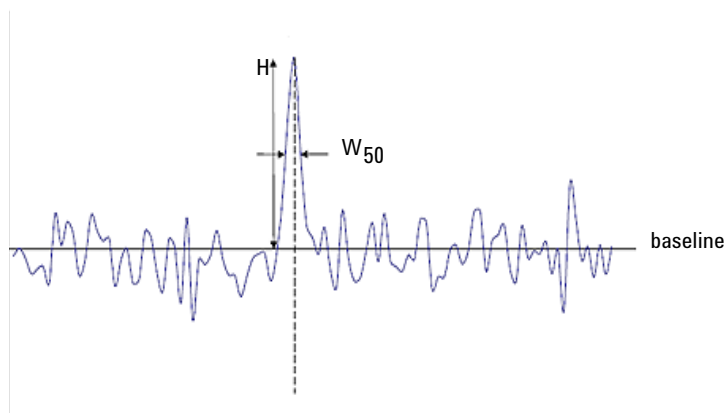
- retention/migration time,
- $k'$ ,
- symmetry,
- peak width,
- plate number,
- resolution,
- signal-to-noise ratio, and
- compound name.

## Noise Determination

Noise can be determined from the data point values from a selected time range of a signal. Noise is treated in three different ways:

- as six times the standard deviation (sd) of the linear regression of the drift,
- as peak-to-peak (drift corrected), and
- as determined by the ASTM method (ASTM E 685-93).

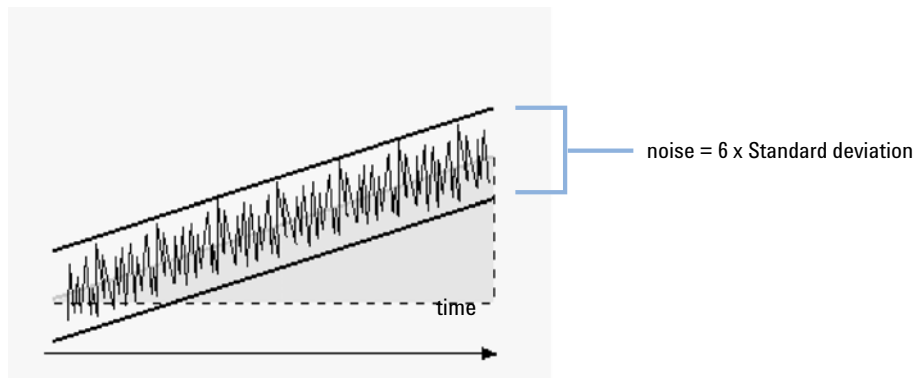
The noise can be calculated for up to seven ranges of the signal; the ranges are specified as part of the system suitability settings in the reporting parameters.



**Figure 43** Chromatogram with peak signal and noise

H	Peak height from top to baseline (best straight line through noise)
$W_{50}$	Peak width at half height

## Noise Calculation Using Six Times the Standard Deviation



The linear regression is calculated using all the data points within the time range of the current signal. The noise is given by the formula:

$$N = 6 \times Std$$

where

N	Noise based on the six time standard deviation method
Std	Standard deviation of the linear regression of all data points in the selected time range



## Noise Calculation Using the Peak-to-Peak Formula

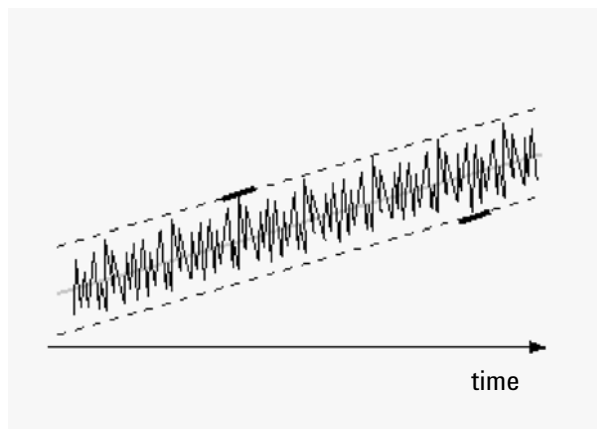


Figure 44 Illustration of peak-to-peak noise with drift

The drift is first calculated by determining the linear regression using all the data points in the time range of a peak. The linear regression line is subtracted from all data points within the time range to give the drift-corrected signal.

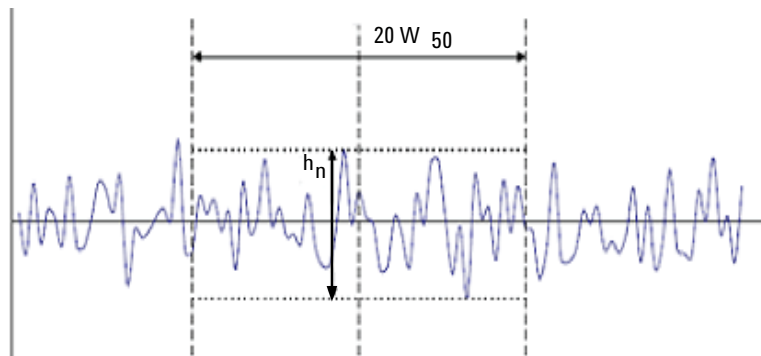
The peak-to-peak noise is then calculated using the formula:

$$N = I_{\max} - I_{\min}$$

where

N	Peak-to-peak noise
$I_{\max}$	Highest (maximum) $I_x$ value in the time range
$I_{\min}$	Lowest (minimum) $I_x$ value in the time range
$I_x$	Intensity of the signal, corrected by the drift (drift is calculated using the LSQ formula)

For European Pharmacopoeia calculations the Peak-to-Peak noise is calculated using a blank reference signal over a range of  $-10$  and  $+10$  times  $W_{50}$  flanking each peak. This region can be symmetrical to the signal of interest, or asymmetrical if required due to matrix signals.



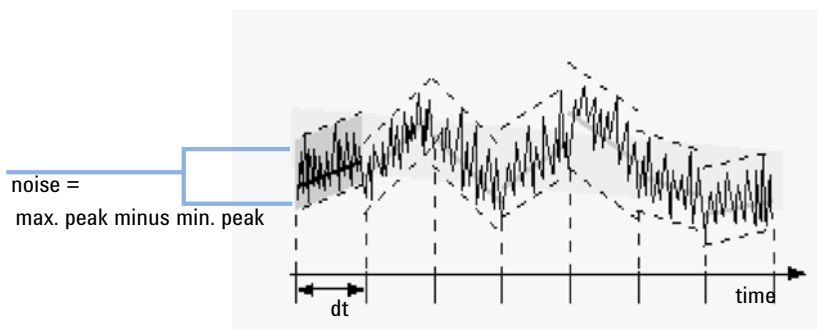
**Figure 45** Determination of noise from the chromatogram of a blank sample

Where

$20 W_{50}$  is the region corresponding to the 20 fold of  $W_{50}$ .

$h_n$  is the maximum amplitude of the baseline noise in the 20-fold  $W_{50}$  region.

## Noise Calculation by the ASTM Method



**Figure 46** Noise determination by the ASTM method

ASTM noise determination (ASTM E 685-93) is based on the standard practice for testing variable-wavelength photometric detectors used in liquid chromatography, as defined by the American Society for Testing and Materials.

Based on the size of the time range, three different types of noise can be distinguished. Noise determination is based on peak-to-peak measurement within defined time ranges.

- *Cycle Time, t*

*Long-term noise*, the maximum amplitude for all random variations of the detector signal of frequencies between 6 and 60 cycles per hour. Long-term noise is determined when the selected time range exceeds one hour. The time range for each cycle (dt) is set to 10 minutes which will give at least six cycles within the selected time range.

*Short-term noise*, the maximum amplitude for all random variations of the detector signal of a frequency greater than one cycle per minute. Short-term noise is determined for a selected time range between 10 and 60 minutes. The time range for each cycle (dt) is set to one minute which will give at least 10 cycles within the selected time range.

*Very-short-term noise (not part of ASTM E 685-93)*, this term is introduced to describe the maximum amplitude for all random variations of the detector signal of a frequency greater than one cycle per 0.1 minute.

Very-short-term noise is determined for a selected time range between 1 and 10 minutes. The time range for each cycle (dt) is set to 0.1 minute which will give at least 10 cycles within the selected time range.

- *Number of Cycles, n*

The number of cycles is calculated as:

$$n = \frac{t_{\text{tot}}}{t}$$

where t is the cycle time and  $t_{\text{tot}}$  is the total time over which the noise is calculated.

- *Peak-to-Peak Noise in Each Cycle*

The drift is first calculated by determining the linear regression using all the data points in the time range. The linear regression line is subtracted from all data points within the time range to give the drift-corrected signal. The peak-to-peak noise is then calculated using the formula:

$$N = I_{\max} - I_{\min}$$

where N is the peak-to-peak noise,  $I_{\max}$  is the highest (maximum) intensity peak and  $I_{\min}$  is the lowest (minimum) intensity peak in the time range.

- *ASTM Noise*

The ASTM noise is calculated as:

$$N_{\text{ASTM}} = \frac{\sum_{i=1}^n N}{n}$$

where  $N_{\text{ASTM}}$  is the noise based on the ASTM method.

An ASTM noise determination is not done if the selected time range is below one minute. Depending on the range, if the selected time range is greater than, or equal to one minute, noise is determined using one of the ASTM methods previously described. At least seven data points per cycle are used in the calculation. The cycles in the automated noise determination are overlapped by 10 %.

## Signal-to-noise calculation

ChemStation has the following options to calculate the noise for the signal-to-noise ratio:

- *6 Sigma*: Noise is calculated using six times the standard deviation of the linear regression (6 sigma). The data for the noise calculation is taken from a specific time interval in the current signal. If you defined multiple time intervals, the interval closest to the peak will be used.
- *USP* (according to the definition of the United States Pharmacopoeia): Noise is calculated using the Peak-to-Peak formula. The data for the noise calculation is taken from a specific time interval in the current signal. If you defined multiple time intervals, the interval closest to the peak will be used.
- *EP* (according to the definition of the European Pharmacopoeia): Noise is calculated using the Peak-to-Peak formula. The data for the noise calculation is taken from a blank signal. The time range for the noise calculation is a time range interval of 20 times the peak width, centered around the retention time of the peak.

Signal-to-Noise Calculation without reference signal (6 Sigma, USP)

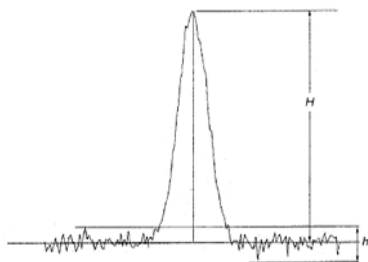
The range closest to the peak is selected from the ranges as specified in the system suitability settings.

The noise is calculated either using six times the standard deviation of the linear regression or using the Peak-to-Peak formula (USP).

The signal-to-noise is calculated for each peak in the signal. If no noise value can be found, the signal-to-noise is reported as "-".

The signal-to-noise is calculated using the formula:

$$\text{Signal-to-Noise} = \frac{\text{Height of the peak}}{\text{Noise of closest range}}$$



**Figure 47** Signal-to-noise ratio

Signal-to-Noise Ratio Calculation according to the EP Definition

The signal-to-noise ratio (S/N) can be calculated per the European Pharmacopoeia definition. S/N is calculated using the equation:

$$S/N = 2H/h$$

Where:

*H* is the height of the peak corresponding to the component concerned in the chromatogram obtained with the prescribed reference solution,

*h* is the absolute value of the largest noise fluctuation from the baseline in a chromatogram obtained after injection of a blank and observed over a distance equal to twenty times the width at half-height of the peak in the chromatogram obtained with the prescribed reference solution, and situated equally around the place where this peak would be found.

The noise value used is calculated using the "Peak To Peak" method (see ["Noise Calculation Using the Peak-to-Peak Formula"](#) on page 105).

S/N is reported in for all peaks present in the chromatogram signal, provided there exists a corresponding reference signal. For a particular chromatogram signal the reference signal is assigned automatically if you specify the reference datafile. If no reference signal can be assigned to a chromatogram signal, signal-to-noise ratio will not be calculated for the peaks in that particular signal.

## Determination of Noise Range

The noise range in the reference signal is determined according to one of the following algorithms

- If the reference signal is not long enough:  $StartTime - EndTime < 20 * W_{50}$ 
  - $StartTime = starttime$  (of reference signal), and
  - $EndTime = endtime$  (of the reference signal)
- If the reference signal is long enough, but the peak is situated such, that  $(RT - 10 * W_{50})$  is less than the start point of reference signal
  - $StartTime = starttime$  (of reference signal), and
  - $EndTime = StartTime + 20 * W_{50}$
- If the reference signal is long enough, but the peak is situated such, that  $RT$  or  $RT + 10 * W_{50}$  is greater than the end point of the reference signal
  - $EndTime = endtime$  (of the reference signal), and
  - $StartTime = EndTime - 20 * W_{50}$
- If the peak is situated such, that  $RT$  or  $RT + 10 * W_{50}$  is greater than the end point of the reference signal
  - $StartTime = RT - 10 * W_{50}$ , and
  - $EndTime = RT + 10 * W_{50}$

Where:

$RT$  is the Retention Time, and

$W_{50}$  is the peak width at half height.

## Drift and Wander

Drift and wander are calculated if **Signal to noise** is selected in the processing method. They are calculated regardless of which noise calculation type you select.

**Drift** Drift is given as the slope of the linear regression. The drift is first calculated by determining the linear regression using all the data points in the time range. The linear regression line is subtracted from all data points within the time range to give the drift-corrected signal.

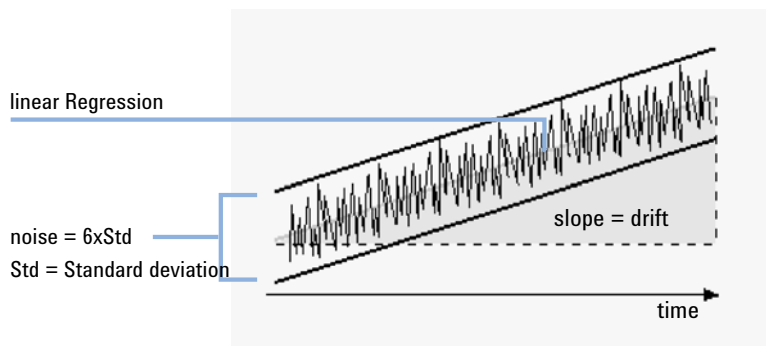


Figure 48 Drift for noise as Six Times the Standard Deviation

Curve Formula:

$$y(x) = a + bX$$

where

N	Number of discrete observations
$X_i$	Independent variable, $i^{\text{th}}$ observation
$Y_i$	Dependent variable, $i^{\text{th}}$ observation

Coefficients:

$$a = \frac{1}{\Delta X} \left( \sum_{i=1}^N X_i^2 * \sum_{i=1}^N Y_i - \left( \sum_{i=1}^N X_i * \sum_{i=1}^N X_i Y_i \right) \right)$$

$$b = \frac{1}{\Delta X} \left( N * \sum_{i=1}^N X_i Y_i - \left( \sum_{i=1}^N X_i * \sum_{i=1}^N Y_i \right) \right)$$

$$\Delta X = N * \sum_{i=1}^N X_i^2 - \left( \sum_{i=1}^N X_i \right)^2$$



**Wander** Wander is determined as the peak-to-peak noise of the mid-data values in the ASTM noise cycles, see “Noise Calculation Using Six Times the Standard Deviation” on page 104.

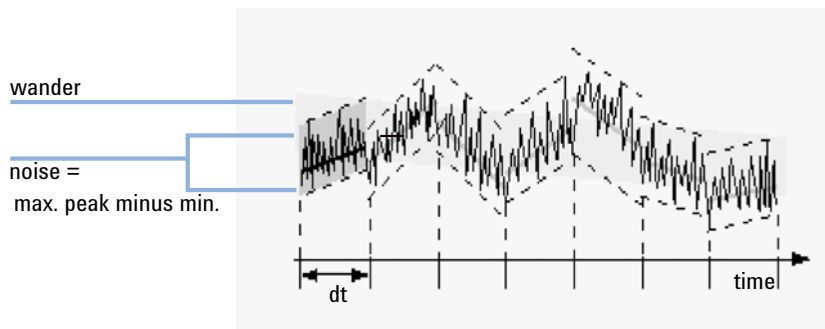


Figure 49 Wander of noise as determined by the ASTM Method

## Calculation of Peak Asymmetry and Symmetry

**Asymmetry** The ChemStation determines the asymmetry ratio of a peak by comparing the peak half-widths at 5% (see "Tailing Factor (USP)  $t$ " on page 123) or 10% (see equation below) of the peak height.

$$A_s = \frac{W_{10}}{2 W_{f, 10}}$$

where

$A_s$  Asymmetry 10%

$W_{10}$  Peak width at 10% of the peak height

$W_{f, 10}$  Front half of the peak width at 10% of the peak height.

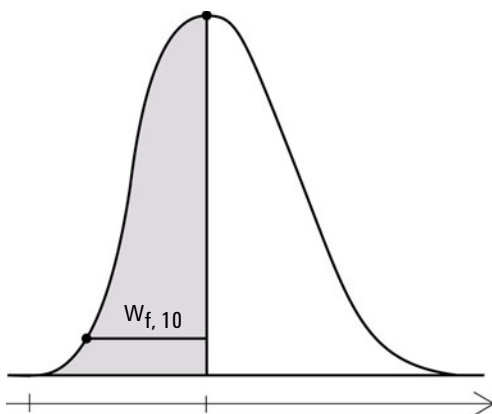


Figure 50 Calculation of peak asymmetry

**Symmetry** Peak symmetry is calculated as a pseudomoment by the integrator using the following moment equations:

$$m_1 = a_1 \left( t_2 + \frac{a_1}{1.5 H_f} \right)$$

$$m_2 = \frac{a_2^2}{0.5 H_f + 1.5 H}$$

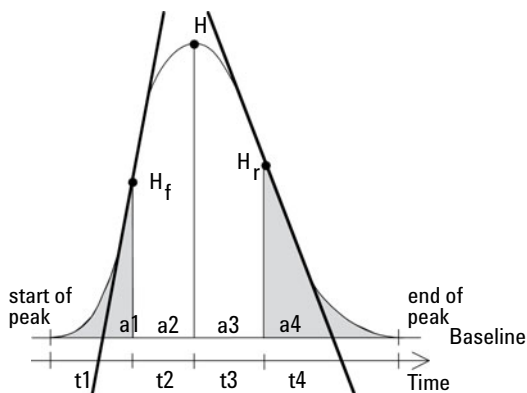
$$m_3 = \frac{a_3^2}{0.5 H_r + 1.5 H}$$

$$m_4 = a_4 \left( t_3 + \frac{a_4}{1.5 H_r} \right)$$

$$\text{Peak symmetry} = \sqrt{\frac{m_1 + m_2}{m_3 + m_4}}$$

If no inflection points are found, or only one inflection point is reported, then the peak symmetry is calculated as follows:

$$\text{Peak symmetry} = \frac{a_1 + a_2}{a_3 + a_4}$$



**Figure 51** Calculation of the Peak Symmetry Factor

where:

$a_i$  = area of slice

$t_i$  = time of slice

$H_f$  = height of front inflection point

$H_r$  = height of rear inflection point

$H$  = height at apex

## System Suitability Formulas and Calculations

The ChemStation uses the following formulae to obtain the results for the various System Suitability tests. The results are reported using the **Performance**, **Performance+Noise**, **Performance+LibSearch**, and **Extended Performance** report styles.

When ASTM or USP is specified for a given definition, then the definition conforms to those given in the corresponding reference. However, the symbols used here may not be the same as those used in the reference.

The references used in this context are:

- *ASTM: Section E 682 – 93, Annual Book of ASTM Standards, Vol. 14.01*
- *USP: The United States Pharmacopeia, First Supplement to USP 37-NF32*
- *EP: European Pharmacopoeia, 7th Edition*
- *JP: Japanese Pharmacopoeia, 15th Edition*

## General Definitions

**Void Volume void volume**

$$V = d^2 \pi l \left( \frac{f}{4} \right)$$

where

d	diameter of column [cm]
$\pi$	constant, ratio of circumference to diameter of a circle
l	length of column [cm]
f	fraction of column volume that is not taken up by stationary phase but available for mobile phase; default value for f = 0.68 (for Hypersil)

**Retention Time of Unretained Compound t (m) [min]**

(Also referred to as dead time or void time)

$$T_m = \frac{V}{F}$$

where

F	Flow rate of LC [ml/min]
---	--------------------------

## Performance Test Definitions

### Overview Performance Tests

Factors from the peak table as defined in the USP, the EP, and JP are available for use in ChemStation reports. The table below provides an overview of the available factors, their definitions and the value names. For more details on the calculations please refer to the respective sections in this guide.

**Table 9** Pharmacopoeia values in ChemStation reporting

USP	EP	JP	Definition	Classic reporting (RLE)	Intelligent reporting (RTE)
Symmetry factor or tailing Factor	Symmetry factor	Symmetry factor	$S = W_5/2f$	USP Tailing	Peak_TailFactor
-	-	-	$S=W_{10}/2f$	USP asymmetry at 10% height	Peak_Asymmetry_10Perc
Separation factor	-	Separation factor	$\alpha = k'_{(a)} / k'_{(b)}$ $T_R$ of peak a < $T_R$ of peak b	Selectivity	Peak_Selectivity
Relative retention	Relative retention	-	With RRT compounds: $r = (t_{R2}-t_0)/(t_{R1}-t_0)$	-	Peak_RelativeRetTime_EP
Relative retention time (RRT)	Unadjusted relative retention	Unadjusted relative retention	$Rr = t_2/t_1$	-	Peak_RelativeRetTime
-	Resolution	Resolution	$R_s = 1.18 \cdot \frac{t_{R2} - t_{R1}}{W_{50(1)} + W_{50(2)}}$	Resolution (EP) Resolution (JP)	Peak_Resolution_EP Peak_Resolution_JP
-	-	-	$R = \frac{\left(\frac{2.35}{2}\right) (T_{R(b)} - T_{R(a)})}{W_{50(b)} + W_{50(a)}}$	Resolution	Peak_Resolution_Classic
Resolution	-	-	$R = 2 \cdot \frac{t_{R2} - t_{R1}}{W_{t(2)} + W_{t(1)}}$	-	Peak_Resolution_USP
Number of theoretical plates (Efficiency)	-	-	$n = 16 \left(\frac{t_R}{W_t}\right)^2$	Plates Tangent method	Peak_TheoreticalPlates_USP

**Table 9** Pharmacopoeia values in ChemStation reporting

USP	EP	JP	Definition	Classic reporting (RLE)	Intelligent reporting (RTE)
-	Plate number (Efficiency)	Plate number (Efficiency)	$N = 5.54 \times t_R^2 / W_{50}^2$	Plates halfheight method	Peak_TheoreticalPlates_EP Peak_TheoreticalPlates_JP
	S/N ratio	S/N ratio	S/N = 2H/h Noise: Peak-to-Peak calculation; blank reference signal; time interval 20x peak width.	-	Peak_SignalToNoise_EP
S/N ratio			S/N = 2H/h Noise: Peak-to-Peak calculation; time interval in current signal.		Peak_SignalToNoise_USP
			S/N = 2H/h Noise: 6 Sigma calculation; time interval in current signal.		Peak_SignalToNoise_6Sigma a
Peak-to-valley ratio	Peak-to-valley ratio	Peak-to-valley ratio	$p/v = H_p/H_v$	Peak to Valley Ratio (Front and Tail)	Peak_PeakValleyRatio
-	-	-	$S=B/A$	Foley-Dorsey Asymmetry at 10% height	-
-	-	-	$N_{sys} = \frac{41.7 (T_R/W_{10})^2}{1.25 + (\max(A,B) / \min(A,B))}$	Foley-Dorsey plates	-

## Statistical Moments, Skew and Excess

Statistical moments are calculated as an alternative to describe asymmetric peak shapes. There is an infinite number of peak moments, but only the first five are used in connection with chromatographic peaks. These are called 0<sup>th</sup> Moment, 1<sup>st</sup> Moment, ... 4<sup>th</sup> Moment.

The 0<sup>th</sup> Moment represents the peak area.

The 1<sup>st</sup> Moment is the mean retention time, or retention time measured at the center of gravity of the peak. It is different from the chromatographic retention time measured at peak maximum unless the peak is symmetrical.

The 2<sup>nd</sup> Moment is the peak variance which is a measure of lateral spreading. It is the sum of the variance contributed by different parts of the instrument system.

The 3<sup>rd</sup> Moment describes the vertical symmetry or skew. It is a measure of the departure of the peak shape from the Gaussian standard. The skew given additionally in the Performance & Extended report is its dimensionless form. A symmetrically peak has a skew of zero. Tailing peaks have positive skew and their 1. Moment is greater than the retention time. Fronting peaks have negative skew and their 1. Moment is less than the retention time.

The 4<sup>th</sup> Moment or excess is a measure of the compression or stretching of the peak along a vertical axis, and how this compares to a Gaussian standard for which the 4. Moment is zero. It can be visualized by moving in or pulling apart the sides of a Gaussian peak while maintaining constant area. If the peak is compressed or squashed down in comparison, its excess is negative. If it is taller, its excess is positive. Also the excess is given in the Performance & Extended report in its dimensionless form.



### Calculation of Statistical Moments

$$M0 = d_t \cdot X$$

$$M1 = t_0 + d_t \cdot \frac{X}{Y}$$

$$M2 = \frac{d_t^2}{X} \cdot \sum_{i=1}^N \left( \left( i - 1 - \frac{Y}{X} \right)^2 \cdot A_i \right)$$

$$M3 = \frac{d_t^3}{X} \cdot \sum_{i=1}^N \left( \left( i - 1 - \frac{Y}{X} \right)^3 \cdot A_i \right)$$

$$M4 = \frac{d_t^4}{X} \cdot \sum_{i=1}^N \left( \left( i - 1 - \frac{Y}{X} \right)^4 \cdot A_i \right)$$

where:

N = Number of area slices

A<sub>i</sub> = Value (Response) of area slice indexed by i

d<sub>t</sub> = Time interval between adjacent area slices

t<sub>0</sub> = Time of first area slice

$\sum_{i=1}^N$  = Sum of starting index 1 to final index N for discrete observations

$$X = \sum_{i=1}^N (A_i)$$

$$Y = \sum_{i=1}^N ((i-1)A_i)$$

## True Peak Width $W_x$ [min]

$W_x$  = width of peak at height  $x$  % of total

where

$W_t$	Tangent peak width, 4 sigma, obtained by intersecting tangents through the inflection points with the baseline
$W_{4.4}$	Width at 4.4% of height (5 sigma width)
$W_5$	Width at 5% of height (tailing peak width), used for USP tailing factor
$W_{10}$	Width at 10% of height
$W_{50}$	Width at 50% of height (true half-height peak width or 2.35 sigma).

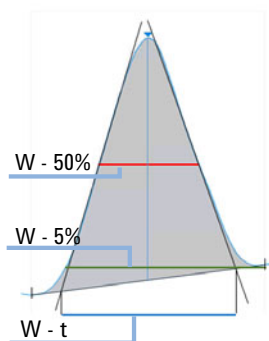


Figure 52 Peak width at  $x$  % of the height

## Capacity Factor (USP)

$$k' = \frac{t_R - t_0}{t_0}$$

where:

$t_R$  = retention time of peak [min]

$t_0$  = void time [min]

## Tailing Factor (USP) t

### NOTE

The Symmetry Factor (USP, EP, JP) is identical with the Tailing Factor (USP). Both are available as "Peak\_TailFactor" in Intelligent Reporting. See also "Overview Performance Tests" on page 118.

$$S = \frac{W_5}{2f}$$

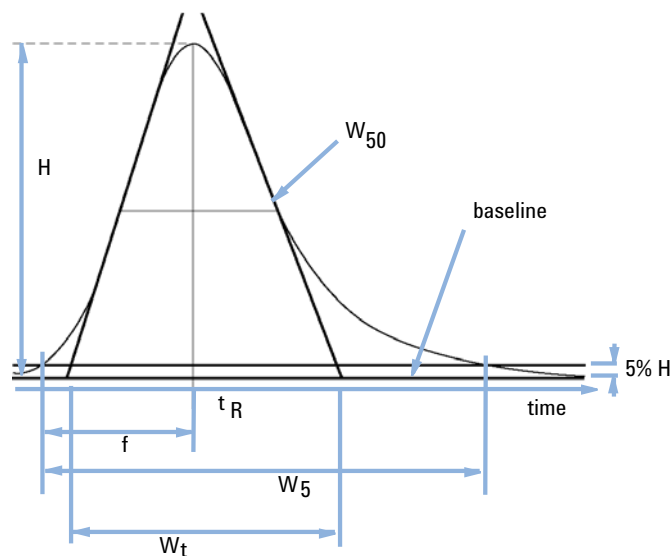


Figure 53 Performance Parameters

S	Symmetry factor, tailing factor (USP)
H	Peak height
t <sub>R</sub>	Retention time
f	Distance in min between peak front and t <sub>R</sub> , measured at 5% of the peak height
W <sub>50</sub>	Peak width at 50% of height [min]
W <sub>5</sub>	Peak width at 5% of peak height [min]
W <sub>t</sub>	Tangent peak width

## Number of Theoretical Plates per Column n

Tangent method (USP, ASTM):

$$n = 16 \left( \frac{t_R}{W_t} \right)^2$$

where

$T_R$	Retention time
$W_t$	Tangent peak width [min]

Half-width method (ASTM, EP, JP):

$$n = 5.54 \left( \frac{t_R}{W_{50}} \right)^2$$

where

$T_R$	Retention time
$W_{50}$	Peak width at half-height [min]

5 Sigma method:

$$n = 25 \left( \frac{T_R}{W_{4.4}} \right)^2$$

where

$T_R$	Retention time
$W_{4.4}$	Peak width at 4.4% of peak height [min]

Statistical method:

$$n = \frac{M1^2}{M2}$$

where:

$M_x = x^{\text{th}}$  statistical moment

Foley Dorsey method

The Foley-Dorsey equation is used for asymmetrical peaks. It corrects plate count for peak tailing and broadening.

$$N_{\text{sys}} = \frac{41.7(T_R/W_{10})^2}{1.25 + (\max(A,B) / \min(A,B))}$$

Where

- $W_{10}$  = peak width at 10% peak height
- A: fronting and B: tailing, with  $A+B = W_{10}$

## Number of Theoretical Plates per Meter N [1-m]

$$N = 100 \cdot \frac{n}{l}$$

where

n

Number of theoretical plates

l

Length of column [cm] (as provided in the processing method)

## Relative Retention, Selectivity

**Selectivity** Selectivity calculates the alpha value for all signal peaks except the first one. For every pair of adjacent peaks (peaks 1 and 2,  $t_R$  of peak 1 <  $t_R$  of peak 2), the selectivity is calculated as follows:

$$\alpha = \frac{k'_{2}}{k'_{1}} = \frac{t_{R2} - t_0}{t_{R1} - t_0}, \alpha > 1$$

where

$k'_{(x)}$  capacity factor for peak x:  $(t_{Rx} - t_0)/t_0$

**Relative Retention (EP)** Relative Retention (adjusted) according to EP can be calculated only if an RT Ratio Reference has been defined and identified. *Alpha* values are < 1 if the peak is to the left of the reference, and > 1 if the peak is to the right of the reference.

$$r = \frac{t_{Ri} - t_M}{t_{Rst} - t_M}$$

Where

$t_{Ri}$  = retention time of the peak of interest

$t_{Rst}$  = retention time of the reference peak

$t_M$  = hold-up time

Relative Retention (unadjusted) according to EP is calculated as

$$r_G = t_{Ri} / t_{Rst}$$

## Resolution (USP, ASTM) R

Tangent method (pertaining to peaks 1 and 2,  $t_{R}$  of peak 1 <  $t_{R}$  of peak 2;  $t_{R}$  in min)

$$R = 2 \cdot \frac{t_{R2} - t_{R1}}{W_{t(2)} + W_{t(1)}}$$

where

$t_{R}$  Retention time

$W_t$  Tangent width [min]

## Resolution (EP/JP) Rs

Resolution (JP) and Resolution (EP) are calculated with the following definition:

$$R_s = 1.18 \cdot \frac{t_{R2} - t_{R1}}{W_{50(1)} + W_{50(2)}}$$

### NOTE

In addition, Classic Resolution  $(2.35/2)^*$ ... is available for in Intelligent Reporting as `Peak_Resolution_Classic`. For a complete list of the values see [“Overview Performance Tests”](#) on page 118

## Resolution (ChemStation classic definitions)

Half-width method:

$$R = \frac{\left(\frac{2.35}{2}\right)(T_{R(b)} - T_{R(a)})}{W_{50(b)} + W_{50(a)}}$$

5 Sigma method:

$$R = \frac{2.5(T_{R(b)} - T_{R(a)})}{W_{4.4(b)} + W_{4.4(a)}}$$

Statistical method:

$$R = \frac{M1_{(b)} - M1_{(a)}}{W_{S(b)} + W_{S(a)}}$$

where:

$M1_{(x)}$  = mean retention time for peak x (1st Statistical Moment) [min]

$W_{B(x)}$  = base width for peak x [min]

$W_{4.4(x)}$  = width at 4.4% height for peak x [min]

$W_{50(x)}$  = width at 50% height for peak x [min]

$W_S(x)$  = width derived from statistical moments =  $\sqrt{(M2)}$  for peak x [min] (see also "Calculation of Statistical Moments" on page 121)

## Peak to Valley Ratio (EP/JP)

The peak to valley ratio (**p/v ratio** in the injection results) is calculated to indicate the quality of peak separation. It is calculated with the European and Japanese Pharmacopeia (EP, JP).

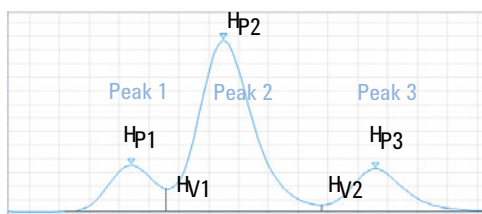
### NOTE

This value is calculated differently compared to the threshold used for the integrator's advanced baseline peak split operation.

The peak to valley ratio is computed for peaks separated by a valley:

$$PV = \text{peak height} / \text{valley height}$$

If there are valleys to both left and right of a peak, the peak to valley ratio is computed for front and tail. With Intelligent Reporting, the minimum p/v will be displayed. With Classic Reporting, the *Classic extended performance* report will display both values.





For peak 1:

$$PV = \frac{H_{P1}}{H_{V1}}$$

For peak 2:

$$PV_F = \frac{H_{P2}}{H_{V1}}$$

$$PV_T = \frac{H_{P2}}{H_{V2}}$$

For peak 3:

$$PV = \frac{H_{P3}}{H_{V2}}$$

where

PV	Peak to valley ratio
PV <sub>F</sub>	Peak to valley ratio, front
PV <sub>T</sub>	Peak to valley ratio, tail
H <sub>P<sub>x</sub></sub>	Height of peak x
H <sub>V<sub>x</sub></sub>	Height of valley x

If a peak has multiple shoulders that are separated by valley, the peak to valley ratio is calculated for each shoulder.

Definition of a valley:

- Its height and time are shared between two consecutive peaks
- Its baseline is shared between two consecutive peaks
- The absolute baseline height is greater than 10<sup>-5</sup>.

## Definitions for Reproducibility

For the statistical review of analytical data in terms of reproducibility the sequence is considered as a small random sample taken out of an infinite number of possible experimental results. To accomplish a complete set of results, an unlimited amount of sample material as well as time would be required. Strictly statistical data does only apply to a complete self-contained set or population of data. Therefore a prerequisite for such a treatment is that the selected sample can be assumed as representative for all data.

### Sample Mean M

The mean value M of a random sample consisting of N measurements is calculated from this limited set of N single observed values  $X_i$  indexed with a consecutive counter i according to the formula:

$$M = \frac{\sum_{i=1}^N X_i}{N}$$

where:

N = number of discrete observations

$X_i$  = value of discrete observations indexed by i

## Sample Standard Deviation S

Consider a random sample of size N. The sample standard deviation S for the selected finite sample taken out of the large population of data is determined by

$$S = \sqrt{\frac{\sum_{i=1}^N (X_i - M)^2}{N - 1}}$$

The sample standard deviation S differs in two points from the standard deviation s for the whole population:

- instead of the real mean value only the sample mean value M is used and
- division by N-1 instead of N.

## Relative Standard Deviation RSD[%] (USP)

The relative standard deviation is defined as

$$RSD = 100 \frac{S}{M}$$

## Standard Deviation of the Mean S

Let M be the sample mean and S the sample [or (N-1)] standard deviation. The standard deviation  $S_M$  of the sample mean M is determined by

$$S_M = \frac{S}{\sqrt{N}}$$

This can be further illustrated by an example:

While the retention time of a certain compound may deviate slightly from the calculated mean value during one sequence, the data from another sequence may differ much more due to e.g. ambient temperature changes, degradation of the column material over time etc. To determine this deviation the standard deviation of the sample mean  $S_M$  can be calculated according to the above formula.

## Standard Deviation (S)

$$S = \sqrt{\frac{\sum_{i=1}^N (Y_i - a - bX_i)^2}{N - 2}}$$

## Confidence Interval CI

The confidence interval is calculated to give information on how good the estimation of a mean value is, when applying it to the whole population and not only to a sample.

The  $100 \times (1 - \alpha)$  % confidence interval for the overall mean is given by

$$CI = t_{(\alpha/2); N-1} \cdot S_M$$

where:

$$t_{(\alpha/2); N-1}$$

percentage point of the t distribution table at a risk probability of  $\alpha$

For the extended statistics in the sequence summary report the 95% confidence interval may be used ( $\alpha = 0.05$ ).

The t distribution (or 'student distribution') must be used for small sample volumes. In case of large sample volumes the results for the t distribution and the normal (gaussian) distribution do not differ any more. Therefore in case of 30 or more samples the normal distribution can be used instead (it would be very difficult to calculate the t-distribution for large numbers, the normal distribution is the best approximation of it).

95% Confidence Interval for 6 samples:

$$1 - \alpha = 0.95$$

$$N = 6$$

The correct value for t has to be taken from the t distribution table for 5 (N-1) degrees of freedom and for the value  $\alpha/2$ , being 0.025. This would give the following calculation formula for CI:

$$CI = 2.571 \cdot \frac{1}{\sqrt{6}} \cdot S_M$$

## Correlation Coefficient

A *correlation coefficient* is displayed with the graphic of the calibration curve. The correlation coefficient ( $r$ ) gives a measure of the fit of the calibration curve between the data points. For calibration curves with a positive slope the value of the coefficient is given to five decimal places, in the range:

0.00000 to 1.00000

where:

0.00000 = no fit

1.00000 = perfect fit

### NOTE

The correlation coefficient does not directly correlate to a good precision and accuracy of the analytical method. It should mainly be used to determine the best curve type.

When evaluating a curve type you need to work with sufficient calibration levels to get sufficient statistical relevance.

The correlation coefficient ( $r$ ) is calculated using the following equation:

$$r = \frac{\sum (y_i - \bar{y}) * (Y_i - \bar{Y}) * wt_i}{(\sum (y_i - \bar{y})^2 * wt_i) * \sum (Y_i - \bar{Y})^2 * wt_i)^{\frac{1}{2}}}$$

where

$r$	Correlation coefficient
$wt_i$	Weight of the data point
$\bar{y}$	Mean values of the measured responses or amounts If the calibration curve is forced through the origin ( <b>Origin=Force</b> in the processing method), OpenLab ChemStation calculates the uncentered determination coefficient. In this case, $\bar{y}$ is omitted.
$y_i$	Measured response (Area, AreaRatio (ISTD method), Height or HeightRatio (ISTD method)) or amount (Amount, AmountRatio (ISTD Method)), depending on calibration mode
$\bar{Y}$	Mean values of the predicted responses or amounts
$Y_i$	Predicted response or amount (using the calibration curve)

## System Suitability

### Definitions for Reproducibility

$\bar{y}$  and  $\bar{Y}$  are mean values of the measured and predicted responses or amounts, calculated as follows:

$$\bar{y} = \frac{\sum(y_i * wt_i)}{\sum(wt_i)}$$

where

$wt_i$	Weight of the data point
$\bar{y}$	Mean values of the measured responses or amounts
$y_i$	Measured response (Area, AreaRatio (ISTD method), Height or HeightRatio (ISTD method)) or amount (Amount, AmountRatio (ISTD Method)), depending on calibration mode

and

$$\bar{Y} = \frac{\sum(Y_i * wt_i)}{\sum(wt_i)}$$

where

$wt_i$	Weight of the data point
$\bar{Y}$	Mean values of the predicted responses or amounts
$Y_i$	Predicted response or amount (using the calibration curve)

For **Forced Origin** it is assumed that the points are centered on zero (mirrored to third quadrant) and the mean values are substituted with zero.

The correlation coefficient is 1 for a perfect fit or when points are distributed symmetrically around the curve. It reduces as the calibration points are distributed less symmetrically. Typical values are between 0.99 and 1. The correlation coefficient is not a very sensitive measure of curve quality.

## Internally Stored Double Precision Number Access

For validation purposes, it might become necessary to manually recalculate the ChemStation results such as calibration curves, correlation coefficients, theoretical plates, etc. When doing so the number format used in the ChemStation has to be taken into account.

For all numbers stored internally within the ChemStation, the "C" data type DOUBLE is used. This means that 14 significant digits are stored for each number. The implementation of this data type adheres to the Microsoft implementation of the IEEE standard for "C" data type and the associated rounding rules (see Microsoft documents Q42980, Q145889 and Q125056).

Due to the non-limited number of parameters that might be used for the calculation of the calibration table, it is not possible to calculate the exact error possibly introduced by the propagation and accumulation of rounding errors. Thorough testing with different calibration curve constructions however has shown that the accuracy of up to 10 digits can be guaranteed. Whereas the area, height and retention time repeatability of a chromatographic analysis usually has 3 significant digits, 10 significant digits within the calculations is sufficient. For this reason, the calibration, and other tables, display a maximum of 10 significant digits.

If an external (manual) calculation for validation is required, it is recommended that all digits used for the internal calculations are utilized. Using the displayed and/or rounded data for the external calculations might give results differing from the ChemStation due to rounding errors.

The following paragraphs describe how to access all internally stored digits for numbers typically required for manual calculations. In all cases, a data file must be loaded and reported with the appropriate report style prior to execution of the listed command. All commands are entered on the ChemStation command line which may be enabled from the view menu.

The following examples generate a .TXT file in the public instruments folder (for example, C:\Users\Public\Documents\ChemStation\1) To get the actual path, use the command line entry

```
print _instpath$. Use other file and folder names as appropriate. The information in this file may be viewed using NOTEPAD or a suitable TEXT editor.
```

## System Suitability

### Internally Stored Double Precision Number Access

Raw Peak Information:

- Retention Time
- Area
- Height
- Width (integrator)
- Symmetry
- Peak Start Time
- Peak End Time

Use Command Line Entry:

```
DUMPTABLE CHROMREG, INTRESULTS, _instpath$ + "INTRES.TXT"
```

Processed Peak Information:

- Measured Retention Time
- Expected Retention Time
- Area
- Height
- Width (integrator)
- Symmetry
- Half Width - Half Peak Height (Performance & Extended Performance)
- Tailing Factor (Performance & Extended Performance)
- Selectivity (Performance & Extended Performance)
- $K'$  (Extended Performance)
- Tangent Peak Width (Extended Performance)
- Skew (Extended Performance)
- Theoretical Plates - Half Width (Performance & Extended Performance)
- Theoretical Plates - Tangent (Extended Performance)
- Theoretical Plates - 5-Sigma (Extended Performance)
- Theoretical Plates - Statistical (Extended Performance)
- Resolution - Half Width (Performance & Extended Performance)
- Resolution - Tangent (Extended Performance)
- Resolution - 5-Sigma (Extended Performance)
- Resolution - Statistical (Extended Performance)



## System Suitability

### Internally Stored Double Precision Number Access

Use Command Line Entry:

```
DUMPTABLE CHROMRES, PEAK, _instpath$ + "PEAK.TXT"
```

Processed Compound Information:

- Calculated Amount

Use Command Line Entry:

```
DUMPTABLE CHROMRES, COMPOUND, _instpath$ + "COMPOUND.TXT"
```

Calibration Table Information:

- Level Number
- Amount
- Area
- Height

Use Command Line Entry:

```
DUMPTABLE _DAMETHOD, CALPOINT, _instpath$ + "CALIB.TXT"
```

Linear Regression Information:

- Y-Intercept (CurveParm1)
- Slope (CurveParm2)
- Correlation Coefficient

Use Command Line Entry:

```
DUMPTABLE _DAMETHOD, PEAK, _instpath$ + "REGRESS.TXT"
```

## 6

# CE specific Calculations

Calibration Tables	139
Standard Calibration	139
Protein Molecular Weight Calibration	140
DNA Base-Pair Calibration	140
Capillary Isoelectric Focusing	141
Calibration using Mobility Correction	142
Introduction	142
Effective Mobility Calculations	143
Relative Mobility Calculations	146
Special Report Styles for Capillary Electrophoresis	148
Corrected Peak Areas	149
System Suitability for Capillary Electrophoresis	150
Capacity Factor $k'$	150
CE-MSD	151
Background Subtraction	151

This chapter is relevant only if you use ChemStation to control CE instruments.

## Calibration Tables

Four different calibrations types are available in the drop-down list for your Calibration Table.

### Standard Calibration

Standard Calibration is based on peak area or peak height. When you select **Standard Calibration** you have the option to **Calculate Signals Separately** or **Calculate with Corrected Areas**.

Calculate Signals Separately is selected when you want to ensure that, in the calculation of Norm% reports, the amount percent of separately reported signals add up to 100% for each signal. When **Calculate signals separately** is deselected, the amount percent of all signals add up to 100%. Selecting **Calculate signals separately** is a prerequisite for sorting by signal in the calibration table.

Select **Calculate with Corrected Areas** to make a correction to the peak area based on the migration time. In this mode, the area is divided by the migration time which can improve reproducibility in quantitative analysis when migration times are unstable.

In addition to the Standard Calibration, there are 3 capillary electrophoresis specific calibrations that are migration time based on signal. The signal is defined by the signal description in the calibration method. If the data file contains multiple signals, then only one signal can be selected and is extracted from the data file. The format of the calibration table is dependent on the calibration type selected.

Quantitation tasks may be performed based on biopolymer size calibration (Ferguson Plot) for SDS-Protein.

## Protein Molecular Weight Calibration

The **Protein molecular weight calibration** requires a calibration standard with components of known molecular weights and a reference peak. The calibration equation is:

$$\log(MW) = k_1 \cdot (t_{ref}/t) + k_0$$

where:

$MW$  is the molecular weight

$t_{ref}$  is the migration time of the reference peak

$t$  is the migration time

$k_0$  and  $k_1$  are the coefficients of the linear equation

The calibration table contains the Name, Migration Time,  $t_{ref}/t$  (relative migration time), Molecular Weight and  $\log(MW)$  for each component.

## DNA Base-Pair Calibration

The **DNA base-pair calibration** is similar to the **protein molecular weight calibration**, but operates without a reference peak; it requires a calibration standard with a known number of base pairs. The calibration equation is:

$$\log(\#BP) = k_1 \cdot 1/t + k_0$$

where:

$\#BP$  is the number of base pairs

$t$  is the migration time

$k_0$  and  $k_1$  are the coefficients of the linear equation

The calibration table contains the Name, Migration Time,  $1/t$ , Base Pairs and  $\log(\text{Base Pairs})$  for each component.

## Capillary Isoelectric Focusing

The **capillary isoelectric focusing calibration** (cIEF) requires a calibration standard with standard proteins of known isoelectric points ( $pI$ ). The calibration equation is:

$$pI = k_1 \cdot t + k_0$$

where:

$pI$  is the isoelectric point

$t$  is the migration time

$k_0$  and  $k_1$  are the coefficients of the linear equation

The calibration table contains the Name, Migration Time and  $pI$ (isoelectric point) for each component.

# Calibration using Mobility Correction

## Introduction

Slight changes in buffer composition, run temperature or viscosity, as well as adsorption to the capillary wall, can influence the Electro Osmotic Flow (EOF) and cause it to be unstable. The resulting change in the EOF can create a rather high standard deviation of migration times. Corrections for mobility can significantly reduce the effect of run-to-run migration time shifts by monitoring the migration time of a mobility reference peak and in turn significantly increasing the migration time reproducibility.

The mobility reference peak should be chosen with the following priorities:

- Select peak with the highest signal
- Select the most isolated peak
- The EOF marker or internal standard can also be used as the mobility reference peak
- Enlarge the search window to always find the mobility reference peak
- If several peaks fall in the search window, the peak with the highest signal is automatically chosen as the mobility reference peak.

There are two mobility correction types available:

### Effective Mobility Correction

**Effective Mobility Correction** uses the effective mobilities of all peaks and requires the availability of the voltage ramp data together with the electropherogram. In addition, working with effective mobility correction allows the true effective mobilities for all sample components to be determined.

### Relative Mobility Correction

**Relative Mobility Correction** can operate in the absence of voltage data and would then assume a constant voltage for all measurements.

## Effective Mobility Calculations

In addition to a reference peak the requirements for effective mobility correction include a neutral marker which corresponds to the velocity of the EOF. Some commonly used markers and their associated wavelengths are:

**Table 10** Commonly Used EOF Markers

Compound	Wavelength
1-Propanol	210nm
Acetone	330nm
Acetonitrile	190nm
Benzene	280nm
Guanosine	252nm
Mesityl oxide	253nm
Methanol	205nm
Phenol	218nm
Pyridine	315nm
Tetrahydrofuran	212nm
Uracil	259nm

Voltage over time data and the capillary dimensions are either saved with the data file or they can manually be entered during the calibration table setup. Storing the voltage data during the run does this most accurately. Make sure to also store the capillary dimensions with the method. To reprocess signals that have been acquired without voltage data/capillary dimensions, enter the voltage and ramp time manually in the "Voltage and Capillary Dimensions" group of the dialog box.

From the data the effective mobility for each component is determined.

General

The apparent mobility of a sample peak is defined by the equation:

$$\mu_{app} = (l \cdot L)/(t \cdot V(t))$$

where

$l$  is the effective length of the capillary (the length from the point of injection to the point of detection)

$L$  is the total capillary length

$V(t)$  is the average voltage from time 0 to the migration time  $t$  of the peak

The average voltage is calculated from either the measured voltage or from the voltage ramp specified in the method using the following equations:

If  $t < t_R$  then

$$V(t) = V/(2 \cdot t_R) \cdot t$$

If  $t > t_R$  then

$$V(t) = V \cdot (1 - t_R/(2 \cdot t))$$

where

$t$  is the migration time of the peak

$t_R$  is the ramp time

$V$  is the end voltage

The equation for mobility can be simplified by introducing a coefficient:

$$k(t) = (l \times L)/V(t)$$

The relative or apparent mobility is then

$$\mu_{app} = k(t)/t$$

Effective or real mobility is

$$\mu_{real} = \mu_{app} - \mu_{EOF}$$

where

$\mu_{app}$  is the apparent mobility of any peak

$\mu_{EOF}$  is the apparent mobility of a neutral marker

Components with lower velocity than the EOF (usually anions) will result in negative values for the effective mobility.



#### Calibration

The real mobility of a sample peak to be used as the mobility reference peak in future measurements is calculated using the migration time of a neutral marker ( $\mu_{EOF}$ ):

$$\mu_{realref} = \mu_{appref} - \mu_{EOF} = k(t_{ref})/t_{ref} - k(t_{EOF})/t_{EOF}$$

The effective mobilities of all peaks are then calculated and stored as expected mobilities:

$$\mu_{realN} = \mu_{appN} - \mu_{EOF} = k(t_N)/t_N - k(t_{EOF})/t_{EOF}$$

The calibration table then contains the measured migration time and the calculated real mobility for each compound in the columns for the expected migration time and the expected mobility.

#### Mobility Calculation

The actual value of  $\mu_{EOF}$  is calculated using the Mobility Reference Peak:

$$\mu_{EOFact} = \mu_{appref} - \mu_{realref} = k(t_{ref})/t_{ref} - \mu_{realref}$$

The expected migration time for each peak is then adjusted:

$$t_{newexpN} = k(t_{oldexpN}) / (\mu_{realN} + \mu_{EOFact})$$

The calculated values are used for peak identification and replace the values within the calibration table.

#### Recalibration

The migration time of the mobility reference peak is used to calculate the actual value of  $\mu_{EOF}$  :

$$\mu_{EOFact} = \mu_{appref} - \mu_{realref} = k(t_{ref})/t_{ref} - \mu_{realref}$$

The expected migration time for every peak is adjusted:

$$t_{newexpN} = k(t_{oldexpN}) / (\mu_{realN} + \mu_{EOFact})$$

and the mobilities are updated:

$$\mu_{realN} = \mu_{appN} - \mu_{EOFact}$$

During a calibration run the expected values for the migration time as well as the real mobility values are updated in the calibration table.

## Relative Mobility Calculations

Migration time correction based on relative mobilities can also be performed. In this case neither an EOF marker, voltage, nor capillary dimensions are needed. The software still corrects migration time shifts but does not display mobility values.

### General

Just as in the effective mobility calculations, the coefficient

$$k(t) = (l \cdot L) / V(t)$$

is used in the relative mobility calculations to describe the relationship between mobility and migration time:

$$\mu_{app} = k(t) / t$$

The difference is that in the equations for Relative Mobility,  $k$  appears in both numerator and denominator of a fraction; this means that the capillary dimension can be eliminated. The factor  $k$  is calculated as

$$k(t) = 1 / V(t)$$

where  $V(t)$  is the average voltage from time 0 to the migration time of the peak.

When the voltage parameter is set to **Ignore**,  $k$  is a constant and can be eliminated from the equations for the expected migration time (see below).

The following equations describe the general case for  $k = k(t)$ , although the software takes all cases into account when calculating  $k$ .

### Calibration

A mobility reference peak is identified and its migration time ( $t_{refcal}$ ) is stored. The expected migration times ( $t_{expcalN}$ ) of all other peaks are saved.

## CE specific Calculations

### Calibration using Mobility Correction

#### Mobility Calculation

After detection of the reference peak, the expected migration time for each peak is adjusted according to the actual migration time of the mobility reference peak:

$$t_{new\ exp\ N} = \frac{k(t_{old\ exp\ N})}{(k(t_{exp\ cal\ N})/t_{exp\ cal\ N} - k(t_{ref\ cal})/t_{ref\ cal} + k(t_{ref\ act})/t_{ref\ act})}$$

Then, the migration time of the reference peak from the last calibration run is updated:

$$t_{ref\ cal} = t_{ref\ act}$$

## Special Report Styles for Capillary Electrophoresis

**NOTE**

CE specific calculations can only be reported with *Classic ChemStation Reporting*.

The following report style is specific for ChemStation for CE systems:

**CE Mobility**

**CE Mobility** comprises quantitative text results, especially the apparent mobility. To use this report style, you need to supply the information on the capillary used before acquisition and you store the voltage signal. The apparent mobility is calculated according to the following formula.

$$\mu_{app} = \frac{l \cdot L}{t \cdot V}$$

Where

$l$  is effective capillary length (cm)

$L$  is total capillary length (cm)

$t$  is migration time (min)

$V$  is voltage (kV)

If effective mobility correction (see “[Effective Mobility Calculations](#)” on page 143) is activated, the peak type column in simple reports (external standard reports for example) is replaced by a mobility column. The CE mobility report prints effective instead of apparent mobilities.

## Corrected Peak Areas

The Agilent ChemStation for CE systems allows you to use corrected peak areas instead of the normal area calculations. These areas are used in standard calibration and reports.

To activate this feature, select **Calculate with Corrected Areas** to make a correction to the peak area based on the migration time. In this mode, the area is divided by the migration time which can improve reproducibility in quantitative analysis when migration times are unstable.

The corrected area is calculated according to the following formula:

$$A_c = \frac{A}{60 \cdot t}$$

Where

$A_c$  is corrected peak area (mAU)

$A$  is peak area (mAU·sec)

$t$  is migration time (min)

This corrected area is sometimes also referred to as normalized area.

## System Suitability for Capillary Electrophoresis

### Capacity Factor $k'$

In capillary electrophoresis the capacity factor  $k'$  value can't be calculated automatically for all operation modes. Refer to the manual *High Performance Capillary Electrophoresis: A Primer* for the formulas respectively. The values listed in the reports are only valid for the Agilent ChemStation for LC 3D systems since the Agilent ChemStation for CE systems uses the same algorithms as the Agilent ChemStation for LC 3D systems.

## CE-MSD

## Background Subtraction

When you select the **Subtract Background** (BSB) menu item, the most recently selected mass spectrum is subtracted from each point in the current electropherogram. The resulting data is saved in the same directory and with the same name as the original data file; however, the file extension is changed to .BSB.

The new data file becomes the current data file and the background subtracted electropherogram is displayed. A record of the number of background subtractions that have been performed is kept in the Operator item of the data file header.

If you view a tabular listing of BSB data, you may observe differences due to the precision of data representation.

**NOTE**

The HELP text files in the LC/MS refer only to LC parameters and not CE. Some features that are available in the LC/MS software are either not available or not applicable to CE/MS applications but are used in LC. The function **peak matching** is not applicable for CE-MS and is therefore not active. In CE-MS, UV and MS detection occurs at different effective lengths of the separation capillary. Because of the different resolution at different effective lengths, peak matching is not possible.



# 7

## System Verification

Verification and Diagnosis Views 153

System Verification 153

The GLPsave Register 156

DAD Test Function 157

Review DAD Test Function 157

This chapter describes the verification function and the GLP verification features of the ChemStation.



## Verification and Diagnosis Views

If supported by the configured instrument, the ChemStation comprises two additional views to perform instrument verification and diagnosis tasks. For more information, see the online help system.

### System Verification

System verification is a key component in the routine use of an analytical instrument in a regulated laboratory. The GLP verification features of the ChemStation are designed to help you to prove that the software, or a relevant components of the software, are performing correctly, or were performing correctly at the time of a particular analysis.

The ChemStation verification function enables you to verify the correct operation of your ChemStation software. You can do this by reprocessing data files according to specific methods, and comparing the results with a pre-defined standard. The verification function is particularly important to prove the integrity of the integration and quantification results.

You can use the standard verification test, or define your own tests using your own method and data files to check the algorithmic software combinations used by your analysis methods. The verification test is a protected file and cannot be changed or deleted.

The Verification item in the Data Analysis view allows you to choose any of the following options:

- run a verification test in the database,
- define a new verification test and add it to the database, and
- delete a verification test from the database.

The How To section of the online help system describes how to perform these tasks. When running a ChemStation verification test, you can choose whether to run the entire test, or select a combination of parts.

## System Verification

### Verification and Diagnosis Views

Verification test results are saved in binary format to the default subdirectory: C:\Users\Public\Documents\ChemStation\1\Verify, together with the method and data files. The Verify subdirectory is at the same level as the sequence, methods and data subdirectories. You can send the results to a printer or to a file. The test results, including a combined verification test result, are rated as either pass or fail.

The following verification test components are available:

#### Digital Electronics (Agilent 1100/1200 Series DAD only)

A test chromatogram is stored in the diode-array detector. This chromatogram is sent to the ChemStation after it has gone through the same preprocessing steps as normal raw data from the photodiodes. The resulting data are compared to original result data stored in the ChemStation for this test chromatogram. If there is a mismatch the test fails. This test ensures that the DAD electronics which do the data preprocessing are still functioning correctly. As a stored test chromatogram is used, the lamp or the diode array are not part of this test. They can be checked with the [“DAD Test Function”](#) on page 157.

#### Peak Integration

The data file is integrated again using the original method. The results are compared to the original integration results stored in verification register. If they do not match, the test fails.

#### Compound Quantification

The compounds in the data file are quantified again. The results are compared to the original quantification results stored in the verification register. If they do not match, the test fails.

#### Report Printing

The original report is printed again.

The following page shows an example of a successfully completed verification test.

```
=====
ChemStation Verification Test Report
=====
```

**Tested Configuration:**

Component	Revision
ChemStation for LC 3D ChemStation	B.01.01
Microsoft Windows	Microsoft Windows XP
Processor	Processor_Architecture_Intel
CoProcessor	yes

ChemStation Verification Test Details:

```
Test Name : C:\CHEM32\1\VERIFY\DEFAULT.VAL
Data File : C:\CHEM32\1\VERIFY\DEFAULT.VAL\VERIFY.D
Method    : C:\CHEM32\1\VERIFY\DEFAULT.VAL\VERIFY.M
Original Datafile      : VERIFY.D
Original Acquisition Method : VERIFY.M
Original Operator      : Hewlett-Packard
Original Injection Date : 4/16/93 11:56:07 AM
Original Sample Name   : Isocratic Std.
```

Signals Tested:

```
Signal 1: DAD1 A, Sig=254,4 Ref=450,80 of VERIFY.D
```

ChemStation Verification Test Results:

Test Module	Selected For Test	Test Result
Digital electronics test	No	N/A
Integration test	yes	Pass
Quantification test	yes	Pass
Print Analytical Report	No	N/A

ChemStation Verification Test Overall Results:      Pass

## The GLPsave Register

The GLPsave register is saved at the end of each analysis when selected in the run time checklist. It contains the following information:

- signals,
- logbook,
- integration results table,
- quantification results table,
- instrument performance data, and
- data analysis method.

This register is a complete protected record, generated at the time of analysis. You can recall it at any time in the future as proof of your analytical methods.

The GLPsave Register option in the Data Analysis view enables you to review the GLPsave register file at any time. The file is protected by a checksum and is encoded into binary to ensure it is not changed.

In the dialog box used to select the GLPsave register for review, you can choose your review options from the following:

- load original method,
- load original signals,
- load instrument performance data,
- print original method,
- print original integration results,
- print original quantification results, and
- generate original report from the original method and signals.

You can use the GLP review function to show that chromatographic data are original, prove the quality of the analysis from the instrument performance data, and demonstrate the authenticity of the data interpretation.

For example, you can:

- reload and reprint the data analysis part of the method used at the time of the sample analysis to prove that the data evaluation, presented as the results of the analysis has not been modified in any way, and
- review without recalculating, the integration and quantification results to prove the authenticity of the report.

## DAD Test Function

Detector tests can be used as a step in the routine system validation of an analytical instrument in a regulated laboratory.

The DAD test assesses the performance of your diode array detector. When you select the DAD test from the Instrument menu (for LC3D and CE only) it checks the instrument for intensity and wavelength calibration. When you press Save the test results are automatically saved to the DADTest database, a register file called DADTest.Reg located in the default instrument directory.

### Review DAD Test Function

The **Review DAD Test** function in the data analysis View menu enables you to review the DADTest.Reg file at any time. The file is protected by a checksum, and is encoded into binary to ensure that it is not changed.

You can select any of the following parts of the DAD test for review:

- Show Holmium Spectra** Plots all Holmium spectra listed in the DAD Test review table. The active spectrum is tagged.
- Show Intensity Spectra** Plots all intensity spectra listed in the DAD Test review table. The active spectrum is tagged.
- Save as New Database** If you change the lamp in your DAD you can reset the DADTest by deleting any unwanted test results from the table and then using this function to save as new database.
- Show Selected Spectra** Displays only spectra you selected in the table.
- Show Intensity Graph** You can plot an intensity graph to give an indication of the life of the lamp in your diode array detector. The graph provides a function of maximum lamp intensity against time.

## In This Book

This guide contains the reference information on the principles of operation, calculations and data analysis algorithms used in Agilent OpenLab ChemStation.

The information contained herein may be used by validation professionals for planning and execution of system validation tasks.

[www.agilent.com](http://www.agilent.com)

© Agilent Technologies Inc. 2010-2022  
Edition: 07/2022

Document No: D0013749 Rev. A

