thermoscientific

TECHNICAL NOTE 70698

Intelligent integration using Cobra and SmartPeaks

Authors
Peter Zipfell,
Product Marketing Specialist
Darren Barrington-Light,
Sr Manager, Product Marketing

Keywords
Chromeleon software,
Chromatography Data System,
CDS, Cobra, SmartPeaks

1. Introduction

Peak detection and integration are fundamental tasks in chromatography, most often done using a chromatography data system (CDS). Enabling software to detect and integrate the peaks as desired (or as required by laboratory rules) is challenging and time-consuming. Common challenges in peak detection include:

- Distinguishing peaks from noise
- Correctly identifying the underlying baseline
- Maintaining correct peak and baseline detection throughout a sequence of chromatograms
- Correctly handling rider peaks and other unresolved peaks

Ideally, these challenges should be addressed using detection parameters in the processing method so that the same treatment can be applied across multiple chromatograms automatically, thereby minimizing variations introduced by different operators. However, finding parameter combinations that produce the desired results has often been a tedious process, causing many chromatographers to give up and resort to manual integration, which is subjective and labor-intensive.

ThermoFisher SCIENTIFIC

Table of contents

2. Peak integration

Reliable chromatographic quantification depends upon accurate and reproducible peak integration. Integration of chromatographic peaks determines the area under the peak, the height of the peak and the peak's retention time. This information is used for all subsequent calculations, such as calibration or analysis of unknowns.

In simplistic terms, integration involves summing the detector output from peak start to peak end. To achieve this, integration is composed of two distinct events: it's start and it's termination. Peak-start and -end points define where integration commences and terminates and are identified by significant changes in detector output or by the rate of change of the detector output. Defining peak boundaries however is extremely challenging.

In an ideal world chromatographic peaks (Figure 1) would be perfectly symmetrical with pronounced detector signal changes that clearly identify peak start and end.

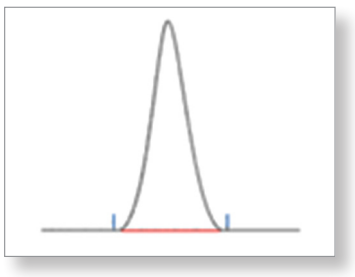

Figure 1. Ideal chromatographic peak.

In reality, integration is extremely complex and diverse. Chromatographic data systems have to decipher many different variants and effects such as peaks of varying symmetries, overlapping peaks, valleys, varying sizes of peaks and size ratios, shallow peak rises and declines, extreme fronting or tailing, shifting apexes and valleys of unresolved peaks, baselines with large sloping background absorption and background noise, etc. (Figure 2).

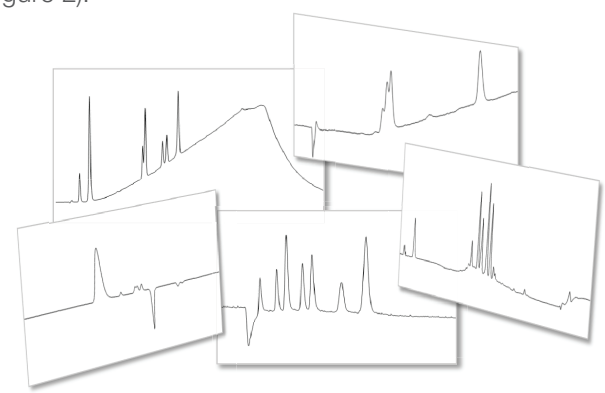

Figure 2. Typical chromatography.

3. Cobra peak detection algorithm

The Cobra peak detection algorithm in the Thermo Scientific™ Chromeleon™ Chromatography Data System (CDS) simplifies integration for all chromatographers and enables the processing of chromatograms with minimal effort providing consistent and reliable peak detection across multiple chromatograms.

The Cobra algorithm uses advanced signal processing to distinguish true peaks from noise, and sophisticated curve-fitting techniques to accurately locate peak maxima and inflection points. It can integrate all types of chromatograms using just a few simple parameters. Values for these parameters are set quickly and easily using an intelligent wizard and the Cobra algorithm then instantly integrates every chromatogram in the sequence.

A fundamental requirement of Cobra is having an optimal number of data points of the order of 10 to 40 per peak to sufficiently characterize the chromatography detection signal of a component peak. For this reason, it is not advisable to use Cobra to re-integrate data imported from Chromeleon 6 that was recorded using a data collection rate set to 'Auto'.

Cobra uses a series of algorithms to analyze the chromatogram and determine the true underlying baseline and correctly identify peak starts and ends. These algorithms are only used in the integration of the chromatogram and do not affect the underlying data, they are described in Section 7.

The Cobra wizard performs integration using an adaptive Savitsky-Golay smoothing function followed by transformation into a second derivative, assessing this second derivative against automatic thresholds based on signal characteristics.

Since the noise is amplified when derivatives are calculated, it is important that sufficient smoothing is used to optimize the signal-to-noise ratio and ensure correct identification of peak starts and ends. Cobra adopts an adaptive Savitsky-Golay smoothing function to remove the noise without losing valuable information.

Cobra uses the second derivative of a chromatographic signal as this can recognize compound peaks. Also, assessing the second derivative of a peak reduces the effects of background absorption and ensures that points of inflection and peak maxima are accurately and consistently identified wherever the background has lower curvature than the analyte peak.

For example, the following two chromatograms show a peak of the same analyte, but one (in blue) has a large sloping background absorption (Figure 3). The first derivative of these two signals is shown in the center (Figure 4). You can see that the difference between the pure analyte chromatogram (grey) and the chromatogram with interference (blue) is reduced. This effect is considerably enhanced in the second derivative (Figure 5). In this case, the chromatogram of the pure analyte and the chromatogram with interference are almost identical.

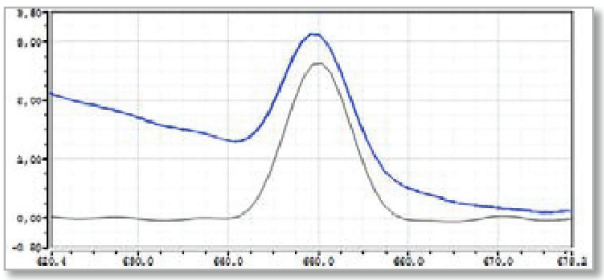

Figure 3. Overlaid chromatograms.

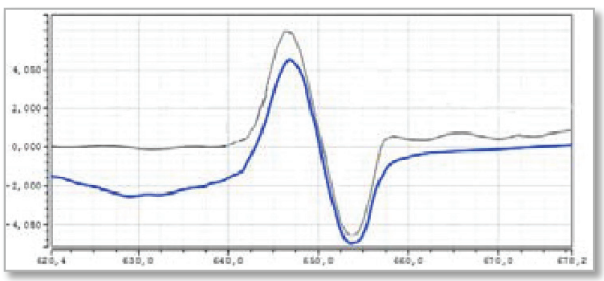

Figure 4. First derivative.

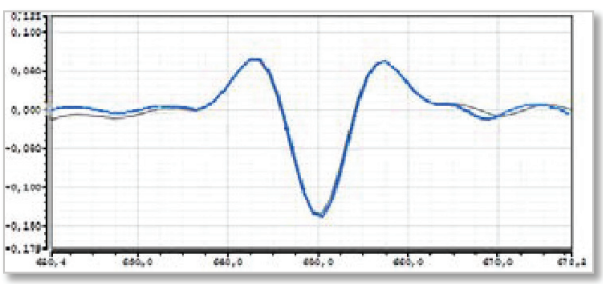

Figure 5. Second derivative.

In the second derivative of the chromatogram, noise thresholds are automatically determined (as represented by the red dashed lines in Figures 6 and 7 below). The local minimum below the threshold's lower limit is the peak apex. The points of inflection or local maxima above the threshold are the peak start and end (Figure 6). The baseline is interpolated between the points where the curvature crosses the upper noise threshold limit (Figure 7).

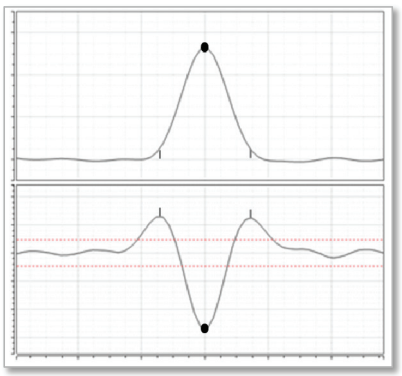

Figure 6. Local minimum and maximum.

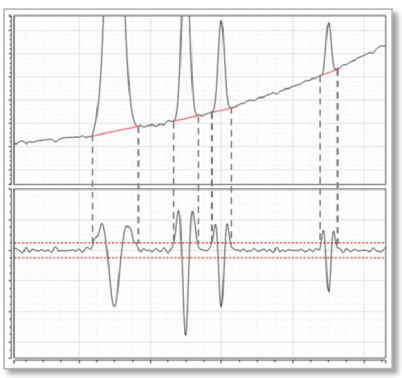

Figure 7. Baseline interpolation.

4. Cobra Wizard

The Cobra Wizard intelligently improves the analyst's experience with minimal user input. The wizard is provided to guide the operator in correctly assigning initial peak detection parameters to help refine the integration if required. Integration is more constructive and intuitive, and can be performed with confidence, thus eliminating trial-and-error methods and manual integration. Details on how Chromeleon CDS determines the values used for the settings of these parameters can be found in section 7.

The Cobra Wizard assists the analyst in defining the four key parameters required by Cobra:

1) Integration range

By default the entire chromatogram will be integrated unless the analyst specifies otherwise. Simply clicking and dragging the mouse across the chromatogram will highlight the range that should be integrated (as shown in Figure 8). Three detection setting parameters are created: Inhibit Integration On, Inhibit Integration Off and a second Inhibit Integration On.

For ion chromatography an additional chromatographic problem exists in the so-called ‘water dip’ or ‘void peak’. This negative peak (related to water) may immediately precede peaks of interest and special handling may be required to avoid interference with these other peaks. The user can define whether to consider a void peak or not via the checkbox.

2) Baseline noise range

Cobra will automatically try to select a quiet, ‘peak-free’ area of the chromatogram to measure the baseline noise (see Section 7 for more information). However the analyst may specify a different range by simply clicking and dragging to highlight the area that should be used. In most cases the default selection is acceptable (as shown in Figure 9). The detection parameter, Baseline Noise Range is created.

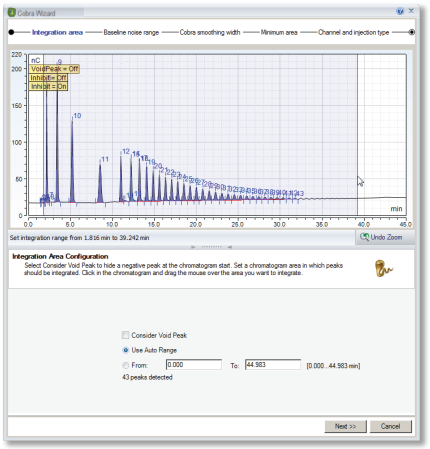

Figure 8. Defining Integration Range.

3) Cobra smoothing width

Cobra will automatically determine the width of the narrowest peak in the chromatogram and use this value to set the Cobra smoothing width (see section 7 for more information). However the user may specify a different value by simply selecting another peak in the chromatogram (as shown in Figure 10) or by clicking and dragging to highlight the width that should be used. In most cases the automatic selection is acceptable. One detection parameter setting is created: Smoothing Width.

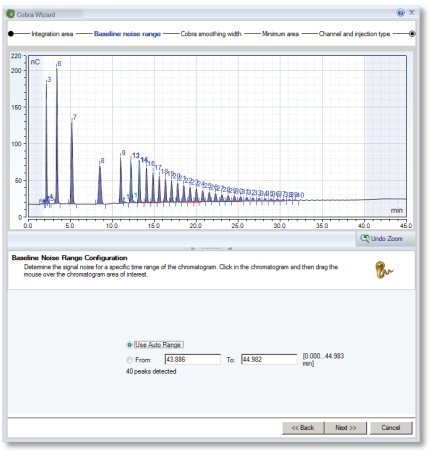

Figure 9. Defining Baseline Noise Range.

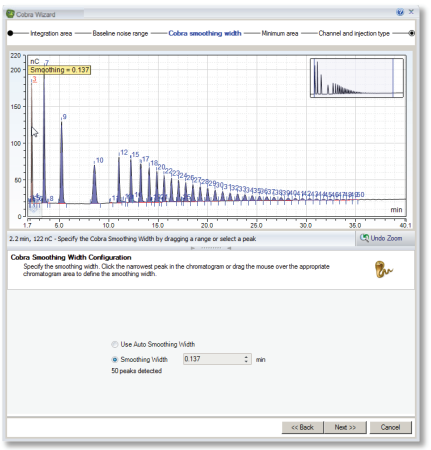

Figure 10. Defining Cobra Smoothing Width.

4) Minimum area

Cobra will automatically determine the area of the smallest integrated peak and set the minimum area accordingly (see section 7 for more information). However the operator can easily specify a different value by simply selecting the smallest peak in the chromatogram that should be integrated (as shown in Figure 11). In most cases the automatically determined value is acceptable. One detection parameter is added to the detection table: Minimum Area. The detection settings support more than one minimum area setting at different retention times however, the wizard sets and supports the first one only.

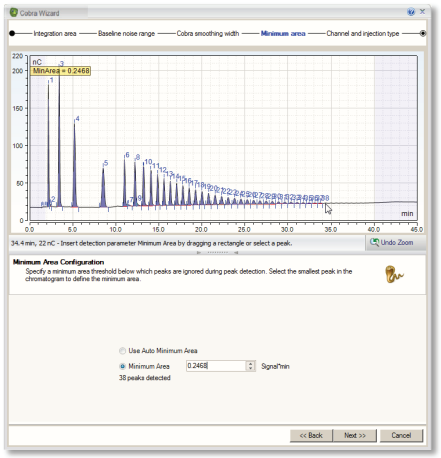

Figure 11. Defining Minimum Area.

Finally the user can decide to apply these Cobra detection parameters to all or only specified channels or injection types (Figure 12). This gives the ability to easily and quickly apply different detection parameters to different detectors or injection types.

Using these simple steps any analyst can quickly, easily, and confidently assign initial peak detection parameters to ensure consistent and accurate integration without going through laborious trial-and-error processes, thus eliminating the need for manual integration.

In addition to adding detection parameters using the Cobra Wizard, it's also possible to insert them into the Chromeleon Processing Method manually, set at specific retention times. These entries can be singular or multiple events, providing greater control for chromatographic integrations.

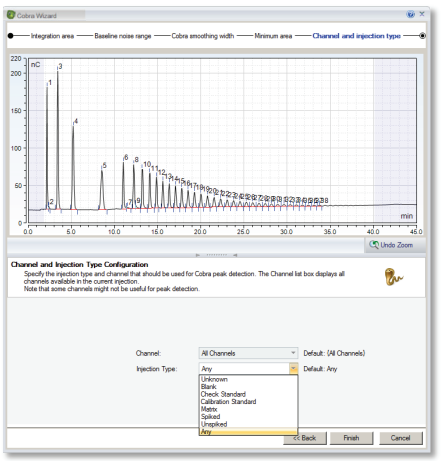

Figure 12. Defining Channel and Injection Type.

5. Common integration issues

There are some common integration issues that occur where baselines are unstable or peaks are poorly separated. For example, in Figure 13A, the data system chose to draw the baseline from the bottom of the dip before the first peak to the baseline after the peak. The problem, of course, is that the negative peak just before the peak of interest is falsely identified as true baseline. This is easily corrected by simply turning on the 'Consider Void Peak' parameter option in the Cobra algorithm (Figure 13B).

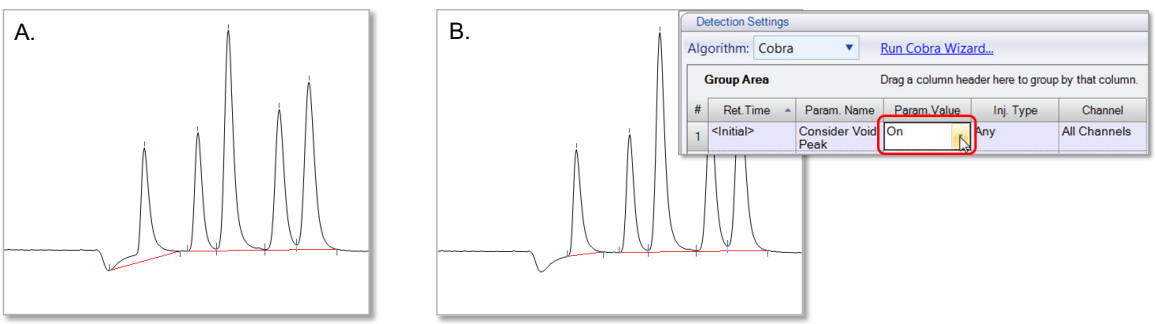

Figure 13. A: Incorrect baseline due to negative dip, B: Correct integration with Cobra.

Another common integration error is shown in Figure 14A. Here the peak endpoint is mis-identified causing the baseline of the peak to be drawn incorrectly. This often occurs when the peak tails strongly, the peak is small and the baseline drifts. When this situation exists, not only can the peak be integrated improperly, but also the peak endpoint may vary significantly from one injection to the next. The Cobra algorithm achieves consistent proper integration by automatically calculating a threshold value based on the slope of the peak baseline, as can be seen in Figure 14B.

A further source of integration issues is unresolved peak groups where there are several choices for determining the location of the chromatogram's baseline profile. The most common options for drawing the baseline between two peaks are drop perpendicular, valley-to-valley, tangential skim, exponential skim and Gaussian skim (Figure 15). Each option has benefits and limitations.

Drop perpendicular

The drop perpendicular method (Figure 15A) involves the addition of a vertical line from the valley between the peaks to the horizontal baseline. If the peaks are approximately the same size and there is no tailing or fronting, the calculated peak areas should be reasonably accurate.

However, if there is significant tailing of the first peak or fronting of the second peak, the areas will be distorted, with corresponding errors in the integrated areas. Also, if the peak ratio is large (e.g. 20:1) the smaller peak will have significant excess area contributed by the major peak. In this case, the accuracy for the larger peak would be much better than for the minor peak. Finally, when the resolution between the peaks is so small that a clear valley is not present, the perpendicular drop method will grossly over-integrate the smaller 'shoulder' peak. In these situations, one of the other options may be more appropriate.

Valley-to-valley

The valley-to-valley method (Figure 15B) sets start and stop points at the valley between the peaks, thus integrating each peak separately. This method is not normally the best approach as it often underestimates the true peak area; however it can be used if there is a known baseline disturbance present under a set of eluted peaks or, in the case of some gradient runs, there is a small, consistent, broad rise in the baseline.

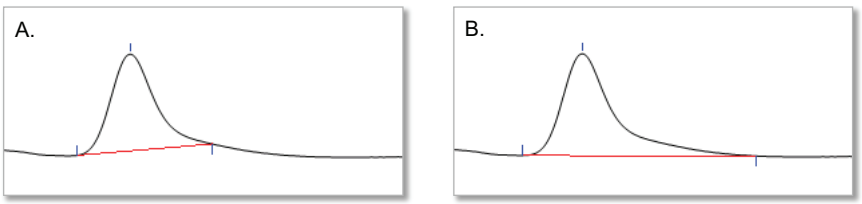

Figure 14. A: Incorrect baseline on tailing peak, B: Correct integration with Cobra

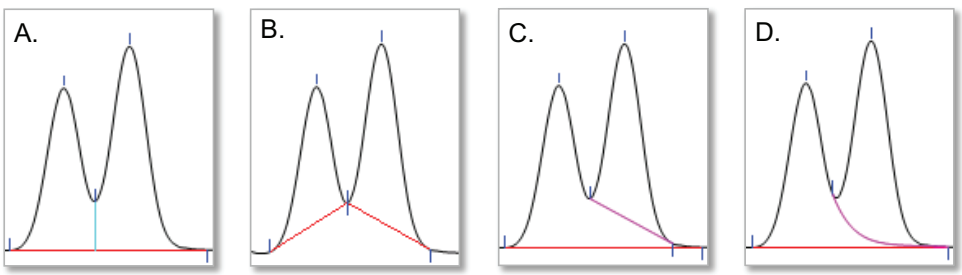

Figure 15. Baseline Profiles. A: drop perpendicular, B: valley-to-valley, C: tangential skim, D: exponential skim

Tangential or exponential skim

The two skim procedures in Figure 15 (C and D) separate the second peak from the first with individual baselines. The first peak is integrated from its starting point to the apparent end of the peak group. The second peak's baseline starts at the valley between the peaks and ends when the signal nears the baseline. The area "under" the skimmed peak is added to the first peak, not the skimmed peak. This approach has been described also as a tangent integration method and the small peak variously labeled a skim, shoulder or rider peak.

Several variations of the skim procedure are possible:

- A tangential skim: a straight line is drawn from the valley to the end (or start) of the peak (Figure 15C). There are commonly two options: Tangential at the lower peak end or tangential at both peak ends.
  - Tangential at the lower peak end means that the line is drawn tangentially to the chromatogram at the peak start for ascending rider peaks, or at the peak end for descending rider peaks (i.e. the part of the peak with the lowest signal).
  - Tangential at both peak ends means that both the peak start and peak end are determined such that the line is tangential to the chromatogram.
- An exponential skim: an exponential function is used to create curvature in the skim line in an attempt to approximate the underlying baseline of the first peak (Figure 15D).

Whichever skim is used, it is always an estimate of the peak area and in general, consistency of integration methodology is more important than whether the skim is tangential or curved. As a rule of thumb, if the minor peak is <10% of the height of the major one, skimming the peak is an appropriate integration technique. If the minor peak is >10% of the height of the major one, a perpendicular drop is usually best.

Understanding all these different methods and techniques and being able to assign the correct detection parameters to achieve the desired integration are extremely difficult tasks, even for the most experienced chromatographer. Chromeleon CDS addresses these issues by providing chromatographers with the Thermo Scientific™ SmartPeaks™ Integration Assistant.

6. SmartPeaks Integration Assistant

The SmartPeaks Integration Assistant provides analysts with the simplest of interfaces to integrate unresolved peak groups quickly and intuitively. The user simply activates the SmartPeaks tool and selects a region of the chromatogram. SmartPeaks then graphically displays the available integration options, such as valley-to-valley baselines or exponential skims (Figure 16).

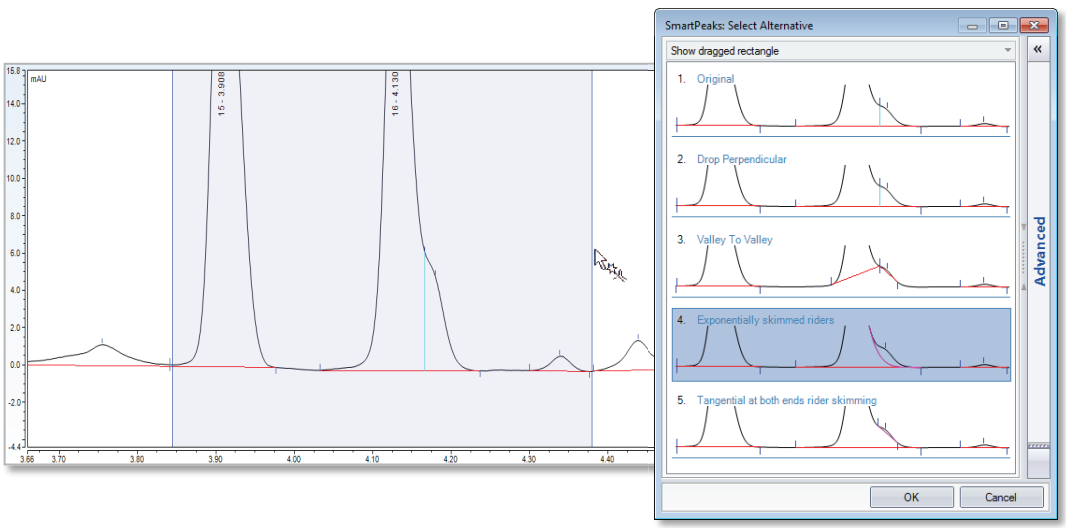

Figure 16. SmartPeaks Integration Assistant – displays integration options.

The user then selects the desired option and SmartPeaks adds the corresponding parameters and values to the processing method so that the required integration is automatically applied to all chromatograms in the sequence (Figure 17). Further options (via the Advanced side bar, Figure 16) allow the user to define which parts of the chromatogram, detection channels and/or injection types these parameters will be applied to.

This process takes just a few seconds, eliminating the trial-and-error process of adjusting integration and enabling any user to set appropriate parameters to ensure correct integration of unresolved peaks.

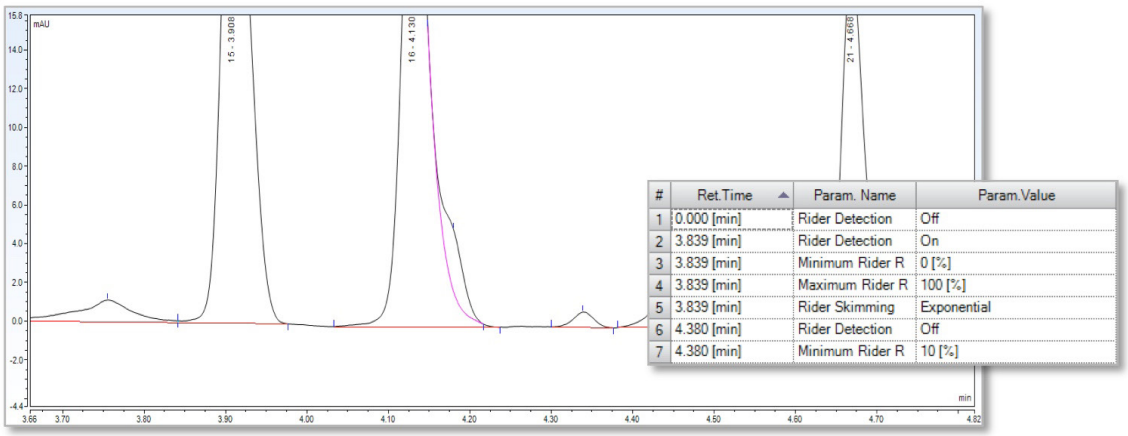

Figure 17. SmartPeaks Integration Assistant – parameters added to processing method.

7. Cobra detection parameters

Integration is extremely complex and challenging but critical to accurate quantitative analysis. Chromeleon software's Cobra detection algorithm, Cobra Wizard, and the SmartPeaks Integration Assistant simplify the peak detection and ntegration process. However, with some complex chromatograms in may still be necessary to provide even more integration parameters to achieve correct integration. Chromeleon CDS provides a comprehensive set of adaptable and flexible parameters to assist chromatographers in achieving their integration goals.

These parameters can be classed into four categories:

- Those that inhibit peak integration
- Those that influence the baseline
- Those that help define the peak start and peak end
- Those that influence the detection of rider peaks and shoulders

The detection parameters influence the integration of all chromatograms to which they apply. Each detection parameter has a default value assigned and even those that are not explicitly defined in the Processing Method are always applied to the chromatogram using this default value. Each parameter can assume new values at any retention time, which will be effective for the remainder of the chromatogram. Unlimited changes can be applied to the parameter at later retention times to allow fine tuning of the integration. These changes will not affect the default value for the parameter.

Understanding the purpose and application of these parameters is essential for any chromatographer to be able to successfully apply these to a processing method to achieve the best results.

Using the ‘Auto’ setting of Cobra parameters

Cobra smoothing width

When a user enables the “Auto” option, the smoothing width is dynamically set over the whole chromatogram, so that for each data point a separate width is used. A report variable in the Chromatogram report category called ‘Effective Smoothing Width (Cobra)’ can be used to report these values for a specific definable retention time. The following algorithm is executed to automatically calculate an initial filter width which is then used in further calculations:

- First, Cobra estimates the slope for each data point by using the current signal value and subtracting the signal value of a previous data point (e.g. signal value of data point 4 – signal value of data point 1, etc.).
- Next, Cobra analyzes this set of slopes and looks for maxima that are closely followed by minima. If there is a zero-crossing between the two (i.e. the maxima value is above 0 and the minima value is below 0), a point is computed that corresponds to a peak apex. This point gives an estimate of the width of the peak between its two inflection points (*w*) and also the height of the peak (*h*).
- These points are then sorted according to their width (*w*) and a line connecting the point with the smallest width to the one with largest width is constructed as shown in Figure 18. The point with largest deviation below this line is identified (*j*) and all smaller points are ignored as these points are probably related to signal noise. The mean width weighted by height (*h*) is computed for the remaining points (shaded area Figure 18).
- Finally, the filter width is computed by multiplying this mean width by a factor (the default is 1.5).

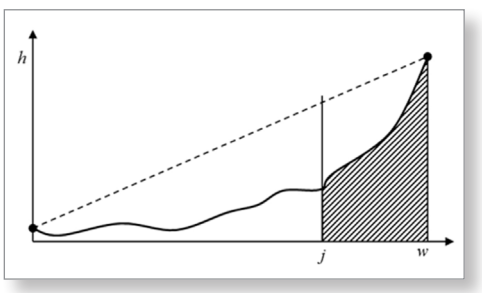

Figure 18. Selection of filter width from peak estimates.

This filter width is then used in two further steps: Resampling and smoothing.

Resampling

The resampling has several purposes:

- To remove and suppress noise
- To reduce the amount of data, speeding up further processing
- To ensure equidistant data points, which is a requirement of subsequent algorithms

Cobra takes the minimum time gap (step) between the existing data points and calculates a new step. A resampling factor is used to limit the number of data points considered (the default is 50):

New step = original step ÷ resampling factor × *maximum of either filter width/original step* or 3

If the original data is not equidistant or the calculated new step is larger than the original step, the chromatogram is then resampled using a cubic interpolation to compute the new signal values at the new step distance.

Smoothing

The detection of the baseline and the peaks is based on the curvature of the (resampled) chromatogram, but the chromatogram may contain a significant amount of noise so an additional processing step to smooth the data is required to differentiate the signal from the noise.

Cobra uses a Savitzky-Golay algorithm to smooth the data. This algorithm applies a second degree polynomial to a set of data points (defined by the previously calculated filter width and [possibly new] step size) neighboring the current data point. However, using a fixed filter width has the effect that some peaks can be smoothed too much or the smoothed signal can overshoot, as shown in Figure 19. This is especially problematic for the skimming of peaks because these processes use slopes and the smoothed signal intensively. As can be seen, the skimmed baseline (red) stops prematurely because the tangent becomes parallel to the baseline at the minimum of the smoothed signal.

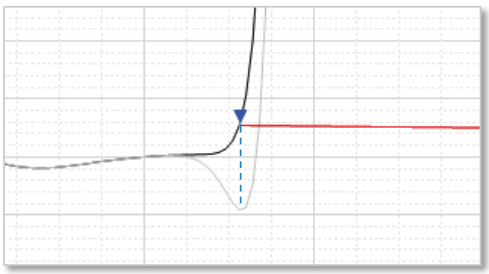

Figure 19. The original signal (black) is not reproduced by the smoothed one (gray).

Reducing the filter width to avoid this problem may lead to small peaks being missed during peak detection because smaller filter widths lead to much higher noise which can hide the curvature of real peaks. The solution is therefore to detect overshooting and adapt the filter width accordingly for each data point in the chromatogram. The filter width is therefore dynamically set over the whole chromatogram, so that for the evaluation (slope and curve) of each data point a separate width is used.

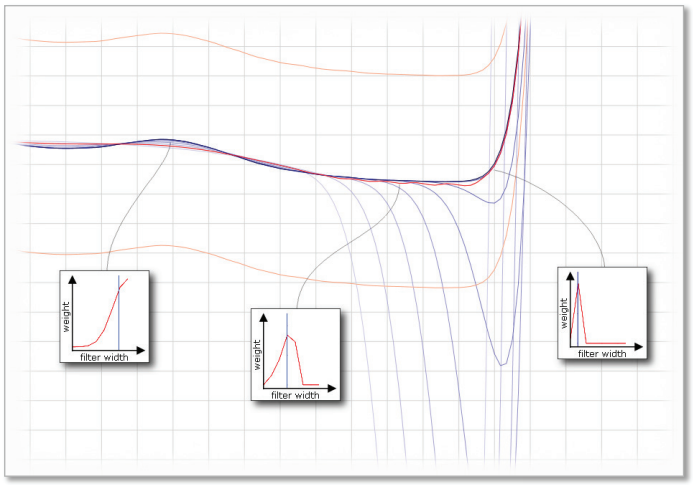

Figure 20. Overview of the adaptive smoothing process.

In Figure 20, the blue curves show several smoothed signals with fixed filter widths that are blended together for the adaptive smoothed signal (red) of the chromatogram (black). The weighting for blending is shown in the small boxes for three different points in time of the chromatogram, the weighted-mean filter width is marked as vertical blue line.

Figure 21 shows how the filter (or smoothing) width is automatically adapted to smooth the chromatogram. This is also applied to the first and second derivatives.

In conclusion, if the user has enabled the “Auto” option, the smoothing width is dynamically set over the whole chromatogram. It is important to note that the smoothing of the chromatogram and the derivatives is only used internally during peak detection and integration. All other integration results (e.g. height or area) are always computed against the original signal.

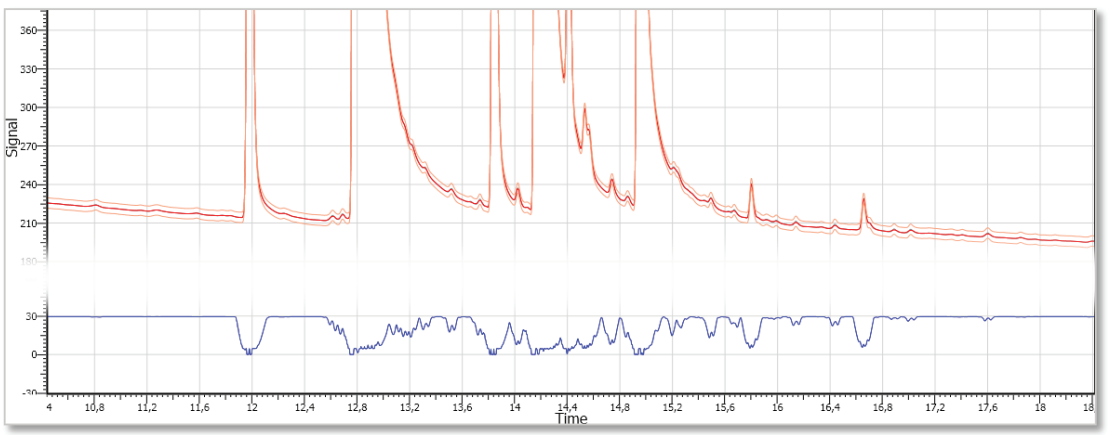

Figure 21. Variation of the mean smoothing width (blue) across a chromatogram (black, almost hidden by the red curve of the smoothed signal). The bounds of the noise level are shown in orange.

Baseline noise range

If the user selects the “Auto” option, an algorithm is used to automatically compute a suitable time range.

Using the American Society for Testing and Materials (ASTM) method E 685-93, the signal noise of the chromatogram is computed multiple times in windows with a width that is a multiple of the calculated Cobra smoothing width (dt in Figure 22), giving a series of estimates for the level of signal noise that are contaminated with outliers for cases where a window overlaps with an actual peak.

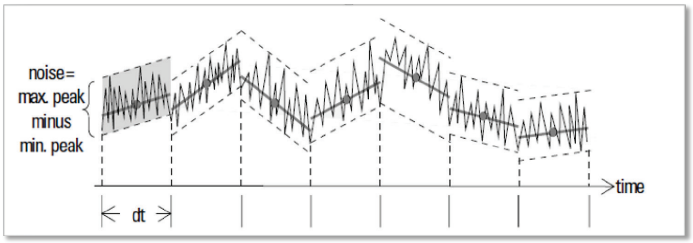

Figure 22. ASTM E 685-93 for signal to noise calculation.

Statistically a significant proportion of these estimates belong to the baseline, so a value with a configurable percentage or ‘rank’ (e.g. 50% for the median) can be used as a robust estimate of the level of the signal noise. This rank requires that the given percentage of the chromatogram is not covered by peaks. However, this percentage varies with the length of the chromatogram (e.g. short chromatograms have a lower percentage of peak-free time).

To compensate for this, the following formula is used to compute the rank:

$$rank = 30 \times chromatogram\ length(min)/10.0$$
$$(min.\ 30,\ max.\ 100)$$

By applying this rank to the set of signal noise estimates, Cobra can calculate a robust number for the signal noise.

Cobra next uses a Ramer-Douglas-Peucker algorithm to approximate the chromatogram (Figure 23). This algorithm recursively divides the chromatogram ‘line’. Initially it includes all the data points between the first and last points (Figure 23A). It automatically marks the first and last point as ones to be kept. It then finds the point that is furthest from the line segment with the first and last points as end points (marked in green). If this point is further from the line segment than the previously calculated signal noise (blue area) then that point must be kept.

The algorithm then restarts with the first point and the worst point (Figure 23B). If the most deviant point is now closer than the previously calculated signal noise to the line segment, any points not currently marked as ‘kept’ between these points can be discarded (marked in red) without the approximated line being worse than the signal noise.

The algorithm continues with the worst point and the last point (Figure 23C) and acts recursively until the entire chromatogram is processed (Figure 23D to G). When the recursion is complete a new chromatogram can be generated consisting of only those points that have been marked as kept (blue line Figure 23G).

Finally the time range (start and finish times) of the longest segment computed by this algorithm is taken as the range for the Cobra Baseline Noise Range.

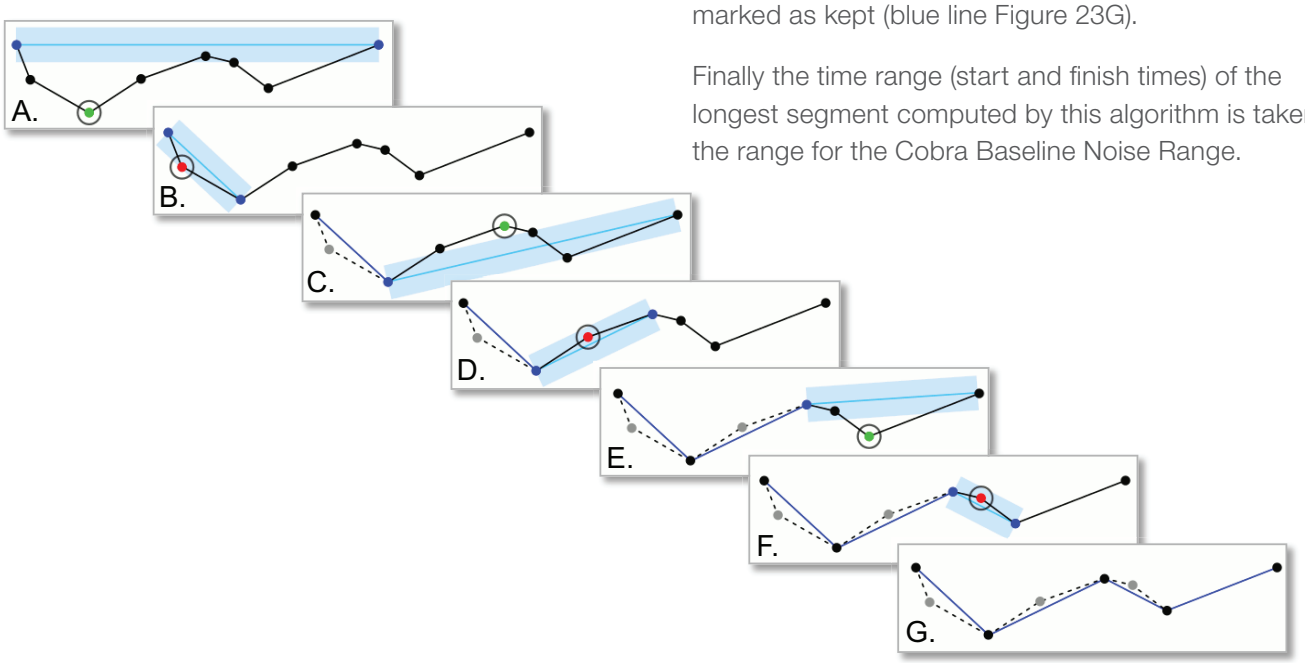

Figure 23. Ramer-Douglas-Peucker algorithm.

Minimum Area

If the 'Use Auto Minimum Area' option is selected in the Cobra Wizard, the 'Minimum Signal-to-Noise Ratio' parameter is used to determine the limit of detection. In this case, Chromeleon CDS calculates the threshold value for the minimum area by using the following formula:

Minimum Signal-to-Noise Ratio × baseline signal noise × Minimum Width

The baseline signal noise value is derived from the portion of the chromatogram specified in the Cobra Baseline Noise Range. If no Minimum Width is specified, the Cobra Smoothing Width parameter is used instead.

Minimum Height

If this parameter is set to 'Auto', the 'Minimum Signal-to-Noise Ratio' parameter is used to determine the limit of detection. In this case, Chromeleon CDS calculates the threshold value for the Minimum Height by using the following formula:

Minimum Signal-to-Noise Ratio × baseline signal noise

The baseline signal noise value is derived from the portion of the chromatogram specified in the Cobra Baseline Noise Range.

7.1. Parameters for inhibiting peak integration

Detect Negative Peaks

The Detect Negative Peaks parameter enables and disables detection of negative peaks. When the parameter is enabled, negative as well as positive peaks are detected (Figure 24). There are three options for this parameter:

- **Off:** Negative peaks are not detected (default).
- **On:** Negative peaks are detected and the Lock Baseline parameter is automatically enabled. In the results table, the area of negative peaks is indicated as a positive value.
- **On Don't Label:** Negative peaks are detected but the peaks are not labeled or included in the results table.

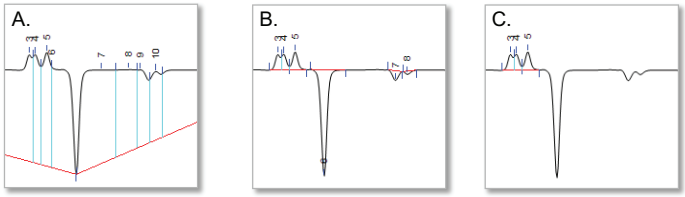

Figure 24. Detect negative peaks. A: Off, B: On, C: On Don't Label.

Detect Shoulder Peaks

This parameter can be used to enable/disable shoulder peak detection. While there is no strict numerical way to describe a shoulder peak, most chromatographers would recognize that a shoulder peak is a type of rider peak where there is no valley between the major and minor peaks (i.e. the minor peak apex is below its peak start [Figure 25]). There are only two options available:

- **On:** Shoulder peaks are treated like normal peaks (default).
- **Off:** No identification of shoulder peaks occurs. Switching shoulder detection off does not disable all rider peak detection.

If a shoulder peak is detected, how the integration is (or is not) performed is determined by the settings of the various rider and minimum parameters. For more information on rider peak parameters see Section 7.4.

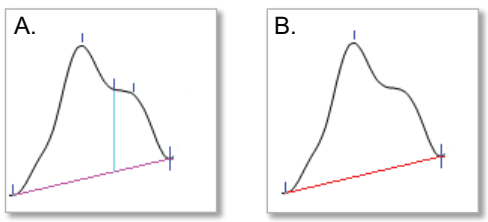

Figure 25. Detect Shoulder Peaks. A: On, B: Off

Inhibit Integration

The Inhibit Integration parameter is used to exclude certain chromatogram areas from integration. This can be useful to ignore areas of the chromatogram that are not of interest or contain peaks that do not require analysis, such as a large solvent peak at the start or a gradient wash at the end of an injection. There are two options available:

- **On:** Peak detection is disabled.
- **Off:** Peak detection is enabled (default).

When set to On, peak detection will only be re-activated when the parameter is reset to Off. In the area between Inhibit Integration = On and Inhibit Integration = Off, peak detection is disabled, that is, no peaks are recognized in this area. The chromatogram is displayed on screen, but it is not integrated (Figure 26).

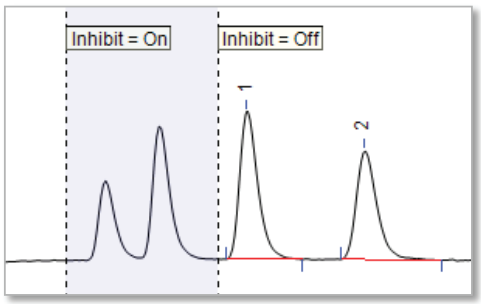

Figure 26. Inhibit Integration.

Minimum Area

This detection parameter is used to determine the minimum area threshold below which peaks are ignored and is the most commonly used parameter for eliminating small peaks from the integrated chromatogram. It is one of the parameters used by the Cobra algorithm and the settings available are:

- **Auto:** If auto is selected (default), Chromeleon CDS automatically calculates a detection limit as described in section 7.
- **0 – 1E+15:** Sets an absolute threshold in <signal unit> x minutes (e.g. mAU*min)

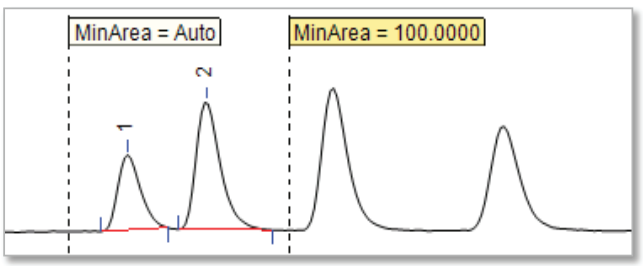

Figure 27. Minimum Area.

Minimum Height

The Minimum Height detection parameter is used to determine the minimum height threshold below which peaks are ignored. The peak height of each peak is measured relative to its own individual baseline. For rider peaks, this means the height measurement is relative to the skim baseline and therefore the baseline type of unresolved peaks can strongly influence the height. The settings available are:

- **Auto:** If auto is selected (default), the detection limit is calculated as described in section 7.
- **0 – 1E+12:** Sets an absolute threshold in <signal unit> (e.g. mAU)

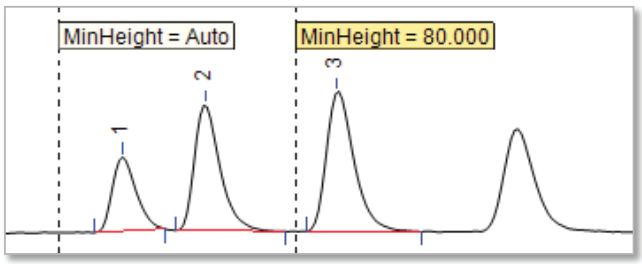

Figure 28. Minimum Height.

Minimum Relative Area

Peaks are rejected if their area is below a particular threshold. The peak area is measured, on a first pass basis, relative to the sum of all peaks (total area) in the chromatogram. This parameter can be useful within a sequence where the analyte concentration can vary and the main analyte peaks area could drop below an absolute Minimum Area threshold (used to reject smaller, unwanted peaks). Using the Minimum Relative Area parameter, peaks of interest that drop below the prescribed minimum area value will still be detected and integrated because they are greater than the minimum relative area threshold of the current chromatograms total area. The settings available are:

- **Off:** All peaks are detected (default).
- **0 – 100** in percent (%).

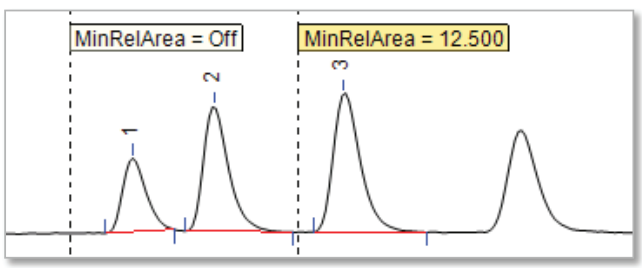

Figure 29. Minimum Relative Area.

Minimum Relative Height

Peaks are rejected if their minimum relative height is below a particular threshold. The minimum height is measured relative to the highest peak in the chromatogram and the height is measured as per the minimum height. This parameter can be used in a similar way to Minimum Relative Area. The settings available are:

- **Off:** All peaks are detected (default).
- **0 – 100** percent (%).

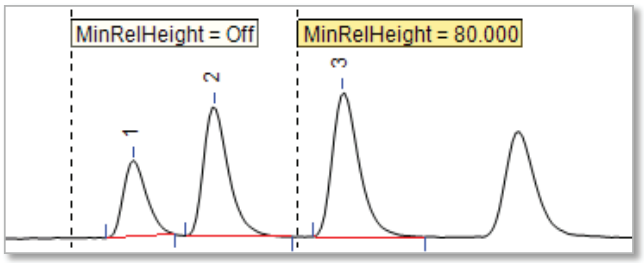

Figure 30. Minimum Relative Height.

Minimum Signal-to-Noise Ratio

This parameter sets the signal-to-noise ratio required for calculating the limit of detection and is used in the Minimum Area and Minimum Height parameters. The limit of detection is the lowest signal height that is just distinguishable from zero. The lower the Minimum Signal-to-Noise Ratio, the lower are the detection limits, and thus, the more peaks are identified. This parameter is automatically effective if **Minimum Area** or **Minimum Height = Auto**. The settings available are:

- **1 – 100:** (Default = 2).

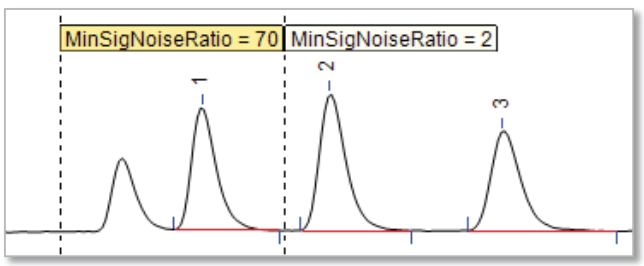

Figure 31. Minimum Signal-to-Noise Ratio.

Maximum Width

The Maximum Width detection parameter defines the maximum peak width in minutes and any peaks wider than this will be ignored during peak detection. The peak width is measured on the baseline. For peaks that do not reach the baseline, the width is extrapolated. The settings available are:

- **Off:** All peaks are considered (default).
- **0 – 1000:** Value in minutes for maximum width.

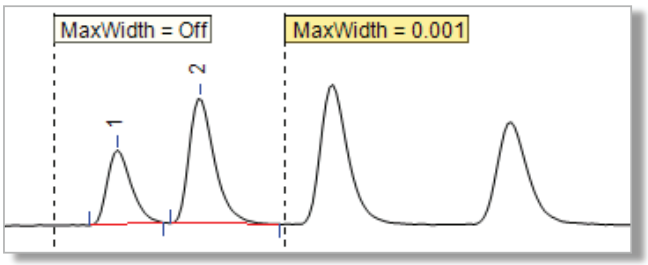

Figure 32. Maximum Width.

Minimum Width

The Minimum Width detection parameter defines the minimum width below which peaks are ignored during peak detection. The peak width is measured on the baseline. In the case of peaks that do not reach the baseline, the width is extrapolated. The settings available are:

- **Off:** All peaks with width > 0 minutes are detected (default).
- **0 – 1000** minutes.

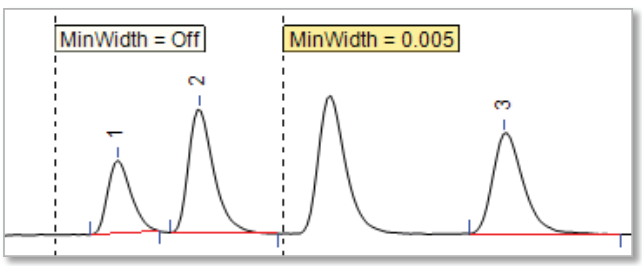

Figure 33. Minimum Width.

7.2. Parameters that influence the baseline

Baseline Point

The Baseline Point Detection parameter is used to set a baseline point at a specific time. Modification of the baseline may be required if integration starts too early or if the peak end is delayed. In these cases, you can force the peak to start later or to end earlier by inserting an appropriate baseline point. The settings available are:

- **Fixed:** A fixed baseline point is added at the selected retention time (default).

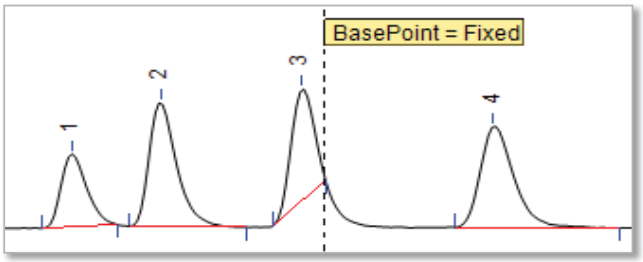

Figure 34. Baseline Point.

Baseline Type

The Baseline Type parameter combines the functions of the Valley-to-Valley and Lock Baseline parameters. The settings available are:

- **Valley-to-Valley:** Enables baseline treatment from valley-to-valley, that is, from peak minimum to peak minimum, in a series of unresolved peaks.
- **Drop Perpendicular:** (Default) For unresolved peaks, the baseline is drawn from the start to the end of the peak group and the not pulled up to the relative minima (valleys). The baseline connects the start of the first with the end of the last unresolved peak. If one of the valleys in between is located below the baseline, the baseline is connected with this minimum to avoid cutting off a peak foot.
- **Lock at current level:** The baseline is fixed at the current signal level and is extrapolated horizontally to the right (Figure 35A).
- **Lock at global minimum:** Chromeleon CDS searches to the right for the minimum absolute signal value in the chromatogram. The search is performed either to the end of the chromatogram or to the next Baseline Point (see previous parameter) and the baseline is locked horizontally at the found minimum level (Figure 35B).

Minimum Baseline Box Width

A crucial part of the whole peak detection process is the determination of the parts of the chromatogram that are baseline (i.e. outside of the peaks). In order to do this an additional algorithm, that is independent of the main peak detection algorithm, is used to identify which parts of the chromatogram are baseline. This additional algorithm calculates a threshold on the second derivative of the chromatogram to select regions with low absolute curvature. To avoid spurious detections, only regions with a width above the threshold that is set by the "Minimum

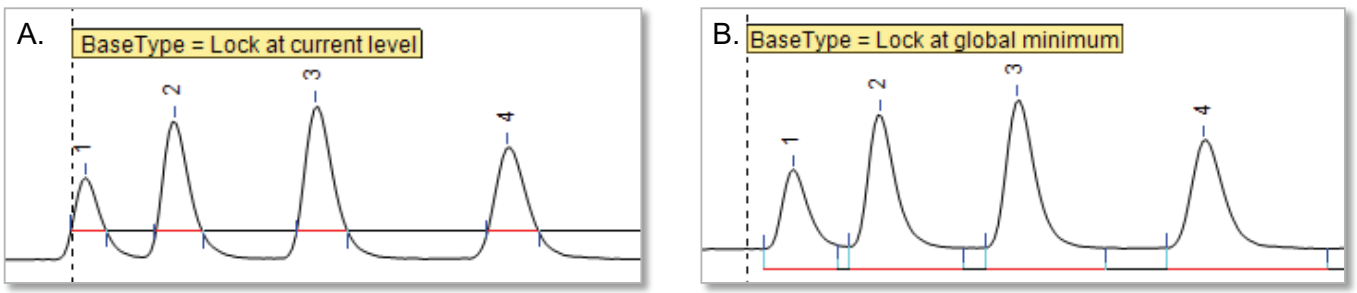

Figure 35. Baseline Type. A: Lock at current level, B: Lock at global minimum.

Baseline Box Width" parameter are used. Setting the parameter to Off turns off the peak-independent detection of baseline segments. As shown in Figure 36, the start and end points of the baseline detected in this way are used to construct the line segments of the baseline under each peak.

The settings available are:

- **Off** (default).
- **0 – 1000** minutes.

Snap Baseline

When set to On, the baseline is automatically 'pulled up' to connect the start and end points of the selected peak when a neighboring peak is no longer integrated. As can be seen in Figure 37A, the first peak is not integrated and the baseline of second peak is connected (or snapped) to the peak start. With the parameter set to off (Figure 37B), the original baseline is maintained. The settings available are:

- **On:** Baseline is snapped to peak start (default).
- **Off:** Baseline is maintained at original level.

Fixed Baseline

In complex chromatograms there can often be a need to force a straight line between two specific data points. The task is made even more onerous and time consuming when inconsistent or sloping baselines are encountered. Often a manual integration is seen as the only solution. A pair of two detection parameters 'Fixed Baseline = Start' and 'Fixed Baseline = End' defines such a straight baseline based on their corresponding retention time values. Figure 38A shows the integration without using the Fixed Baseline parameter. Note the difficulty of finding an optimum baseline. Figure 38B illustrates how the Fixed Baseline Parameter can simplify the process without the need for manual integration.

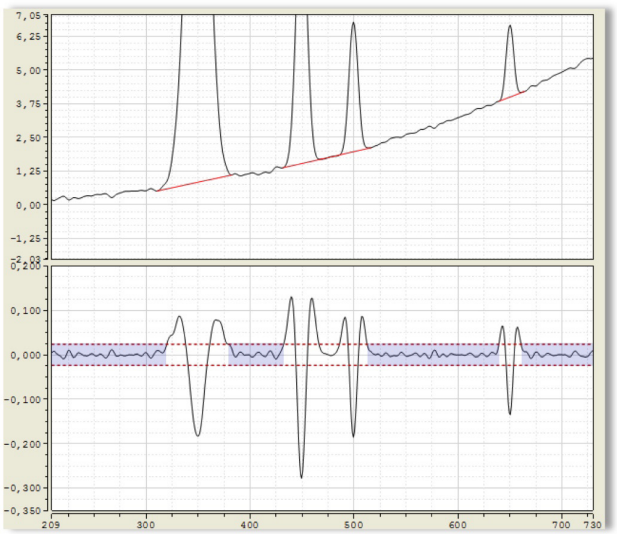

Figure 36. Detection of baseline segments (blue regions – 2nd derivative) and filling the gaps between by linear segments (red, solid lines – chromatogram).

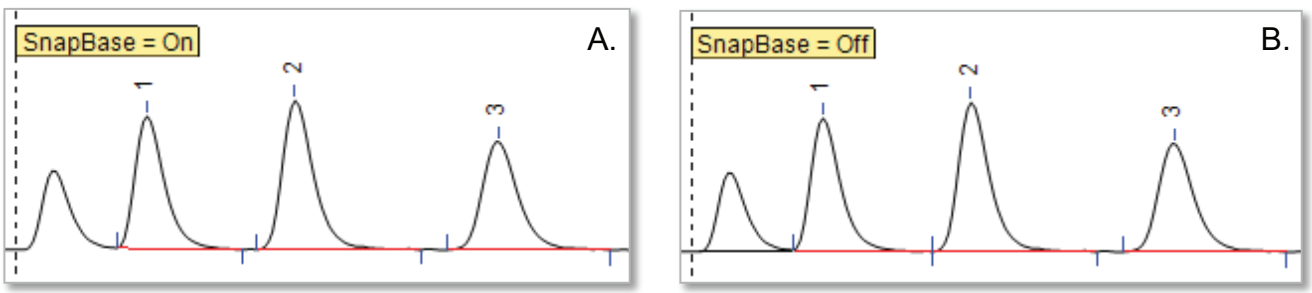

Figure 37. Snap baseline. A: On. B: Off.

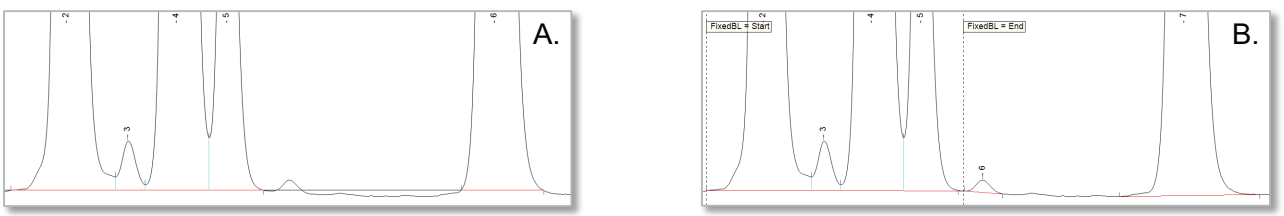

Figure 38. A: Lock Current Baseline. B: Fixed Baseline Start/End defined.

7.3. Parameters for defining the Peak Start and Peak End

Peak Group Start/Peak Group End

The Peak Group Start and Peak Group End detection parameters are used to identify several successive peaks as one peak group. The peak group is then treated as one single peak. Within the group, any other detection parameters (e.g. detecting rider peaks), do no apply. The peak maximum of the largest peak becomes the peak maximum of the entire group. Chromeleon CDS displays only one name and number in the chromatogram and component table. Results (e.g. area value and peak height) are calculated for the entire group.

The Peak Group Start parameter marks the beginning of a peak group, while the Peak Group End parameter marks the end of the group. The settings available are:

- **Auto:** The baseline of the peak group is drawn from the peak start of the first peak to the peak end of the last peak (Figure 39A). The peak delimiters and the level of the baseline are determined by the peak detection algorithm.
- **Fixed:** The baseline is drawn from the signal value at the retention time of the Peak Group Start to the signal value at the retention time of the Peak Group End (Figure 39B).

Fronting/Tailing Sensitivity Factor

The Fronting Sensitivity Factor and Tailing Sensitivity Factor are implicit thresholds for setting the peak start and end, respectively. If the peak start is set too early or the peak end is too late (usually seen in the case of peaks with significant fronting or tailing), these parameters can be set to modify the sensitivity for detecting peak start and/or end.

The larger the value of the Fronting/Tailing Sensitivity Factor, the closer the peak start/end will be to the peak apex. The settings available are:

- **Off:** The peak start and end are detected automatically (default).
- **1 – 100:** A value of 1 is often an appropriate starting point for finding the best Fronting/Tailing Sensitivity Factor.

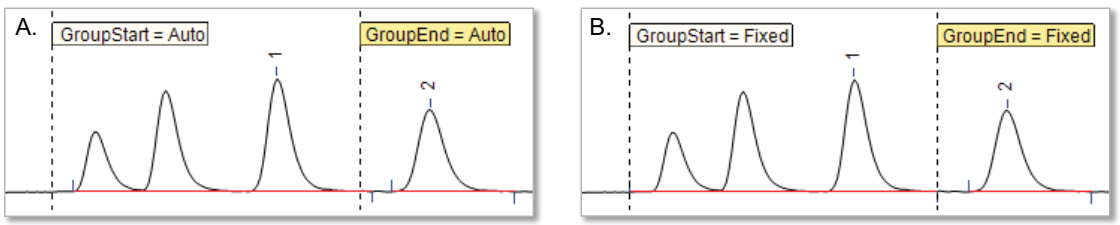

Figure 39. Peak Group Start/End. A: Auto. B: Fixed.

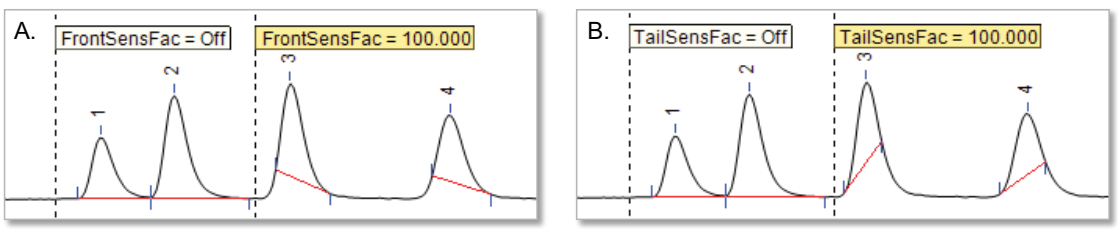

Figure 40. A: Fronting Sensitivity Factor. B: Tailing Sensitivity Factor.

7.4. Parameters that influence detection of rider peaks and shoulders

Maximum Rider Ratio

If one or more peaks (h_1 to h_4) are above the Minimum Rider Ratio (Figure 41) in a series of unresolved peaks, the Maximum Rider Ratio detection parameter is used to determine whether a peak is to be classified as main peak or rider peak.

The height of the peak to classify (e.g. h_2) is compared to the height of the tallest peak in the series (i.e. b in Figure 41). If this ratio produces a value larger than the maximum rider ratio, the peak is a main peak. If the ratio produces a value smaller, the peak is a rider peak. Starting with the largest peak in the series, all adjacent peaks are then classified.

The settings available are:

- **0 - 100** percent (%). (Default = 20).

Minimum Rider Ratio

The ratio between the heights of the single peaks (h_1 to h_5 in Figure 42) and the height of the largest peak (b in Figure 42) determines whether a peak is classified as rider peak or main peak. If the result of the height ratio is below the defined Minimum Rider Ratio, the peak is defined as a main peak. In Figure 42, this is the case for only the h_5 peak. If the result of the height ratio is above the defined minimum ratio, the Maximum Rider Ratio parameter is calculated and the resulting value is used to classify the remaining peaks as rider peaks or main peaks

The settings available are:

- **0 - 100** percent (%). (Default = 10).

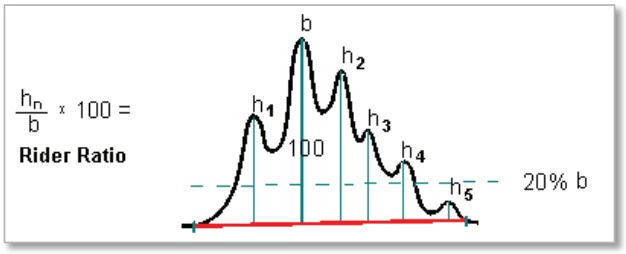

Figure 41. Maximum Rider Ratio.

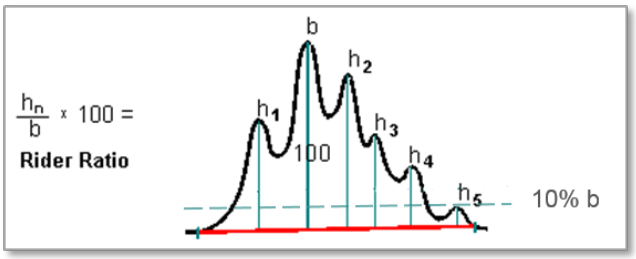

Figure 42. Minimum Rider Ratio.

Rider Detection

This parameter simply enables or disables rider detection. The settings available are:

- **On:** All riders are detected that match the Minimum/ Maximum Rider Ratio parameters (default).
- **Off:** Disables rider detection and identifies all peaks as main peaks.
- **Front Only:** Only riders at the leading edge of a peak are identified.
- **Tail Only:** Only riders at the tailing edge of a peak are identified.

Rider Skimming

The Rider Skimming parameter determines how rider peaks are skimmed. The settings available are:

- **Tangential at lower peak end:** Depending on whether the peak is ascending or descending, either the peak start or end is defined such that rider skimming is tangential to the chromatogram.
- **Tangential at both peak ends:** Peak start and peak end are determined such that rider skimming is tangential at both chromatogram ends.
- **Exponential:** The chromatogram is approximated by an exponential function. In most cases, Exponential maps the actual course of the curve very accurately and the rider peak will usually receive a more realistic area.

8. Adding and editing Cobra detection parameters

Cobra detection parameters are used to determine how chromatograms are integrated, for example, in which areas peaks are suppressed or how peak starts and peak ends are detected. Thus, the integration can be adapted as required by the analyst to virtually eliminate the need for manual re-integration effort for individual chromatograms. It also feasible to place parameters at specific/multiple time point(s) and to assign them to channels/injections.

Chromeleon CDS automatically determines and integrates peak areas, based on the Cobra detection parameters defined in the processing method. However, adding or editing parameters in the processing method can present chromatographers with some challenges:

- How to ensure parameters are added at the correct retention time (e.g. between two peaks).
- How to see the effects on the chromatogram of adding or changing a parameter.
- How to determine a sensible value for the parameter rapidly and effortlessly.

Chromeleon CDS provides a simple and intuitive interface to aid the analyst in adding and editing detection parameters directly in the chromatogram. This ensures the adding, positioning and editing of parameters can be achieved with minimal effort and delay.

Adding detection parameters

Detection parameters can be added through direct graphical interaction with the chromatogram. This significantly facilitates the procedure for defining detection parameters.

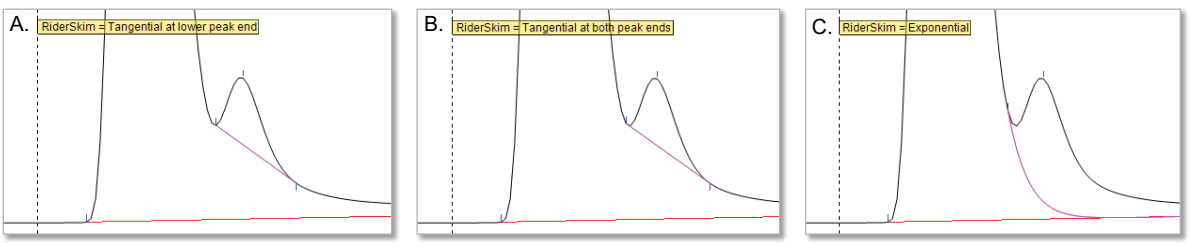

Figure 43. A: Tangential at lower peak end, B: Tangential at both peak ends, C: Exponential.

In order to add detection parameters directly in the chromatogram, the analyst simply selects the 'Insert' option in the Detection Parameters group of the Processing ribbon (Figure 44A). The most commonly used parameters are then directly available in a drop down menu (Figure 44B). The analyst selects the parameter to be inserted and the mouse pointer changes its appearance to look as follows:

The analyst then moves the pointer exactly to the location where the parameter is required and clicks in the chromatogram (Figure 44C). The defined parameter is directly added to the processing method and is also displayed as a flag in the chromatogram at the selected retention time. Dynamic data linking in the Chromleon software ensures that this new parameter is instantly applied to all chromatograms in the sequence.

An alternative method to add detection parameters directly in the chromatogram would be for the analyst to simply move the mouse pointer to the location where the parameter is to be inserted and click the right mouse button. A popup menu then offers the option to insert a detection parameter (Figure 45A). On selection of this option, the analyst has immediate access to all available detection parameters and can specify the required settings (Figure 45B). Again, the defined parameter is directly added to the processing method and is also displayed as a flag in the chromatogram at the selected retention time (Figure 45C).

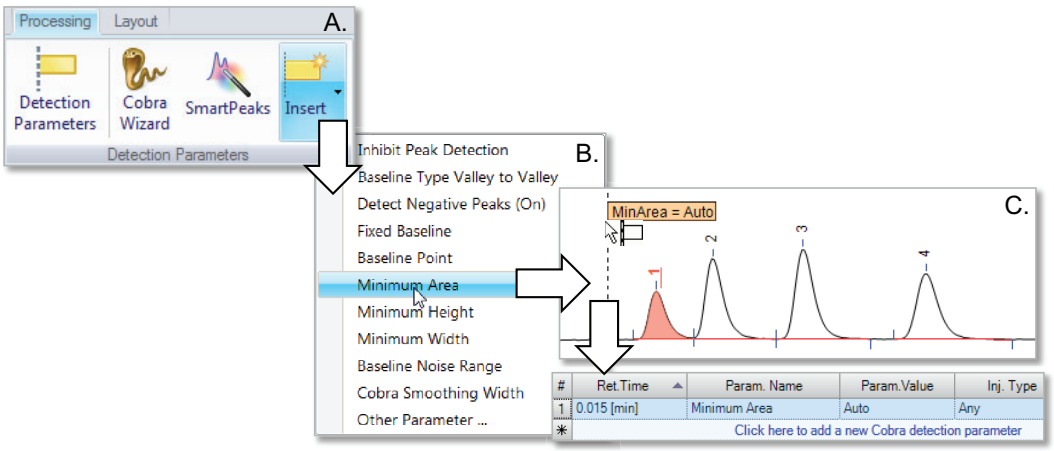

Figure 44. A: Insert option on the Processing ribbon, B: List of parameters, C: Click to insert.

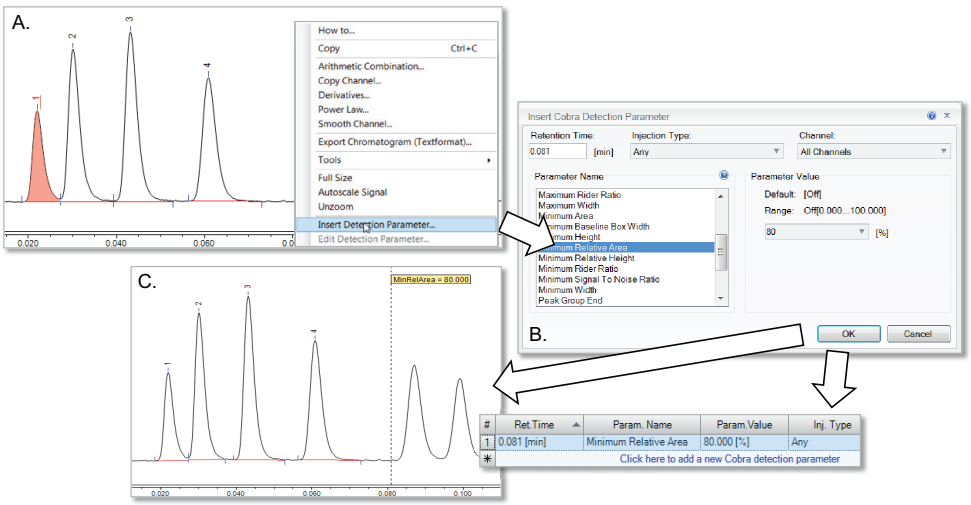

Figure 45. A: Right click to insert a detection parameter, B: Define parameter, C: Parameter applied.

By using the right mouse button and dragging a box over an area of the chromatogram it is possible to add other detection parameters. For example, setting the baseline noise range is as simple as dragging a box over an area of baseline (Figure 46A) and selecting the relevant parameter (Figure 46B) in order to add the parameters (Figure 46C).

If after adding parameters the result is not as expected, the settings can simply be removed using the undo button on the quick access toolbar:

This allows the operator to sequentially step back through all changes made.

Editing detection parameters

Editing of existing detection parameters can be performed directly in the chromatogram, that is, modification of parameter values or moving a parameter to a different location in the chromatogram are possible.

Editing parameter values

To edit an existing parameter, the analyst simply double-clicks the flag of the required detection parameter (Figure 47A and Figure 47B). As the parameter value in the flag is changed, the Chromeleon software instantly updates the chromatogram view to visualize the effect that the changing value would have (Figure 47C) without the user needing to commit this change, allowing the analyst to quickly find the correct value to achieve the desired integration and simultaneously avoiding unnecessary versions of the processing method. Pressing the Escape key will discard any changes while, to commit the new value, the analyst simply hits Enter or clicks away from the flag (Figure 47D).

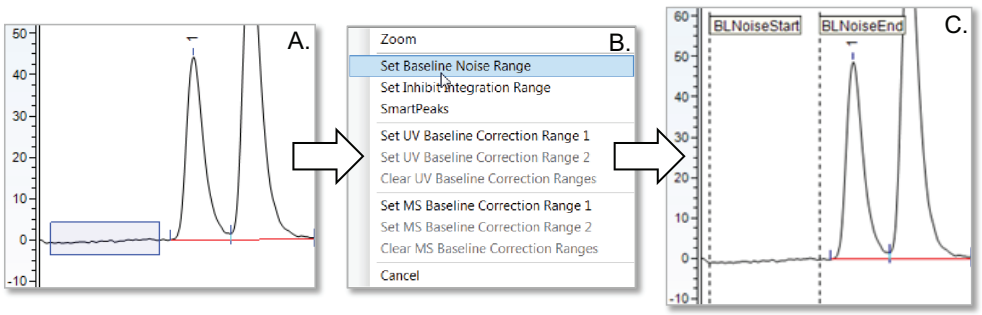

Figure 46. A: Select region, B: Select parameter, C: Parameters are inserted.

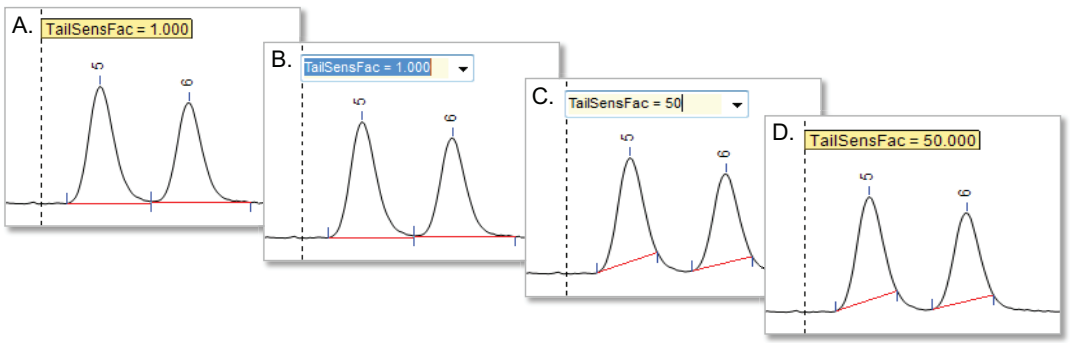

Figure 47. A: Detection parameter flag, B: Double-click to edit, C: Instant update before commit, D: New value committed.

For some parameters there are fixed options for the value that can be accessed by clicking the arrow of the flag (which can be seen in Figure 47B) and selecting a parameter value from the dropdown list (Figure 48).

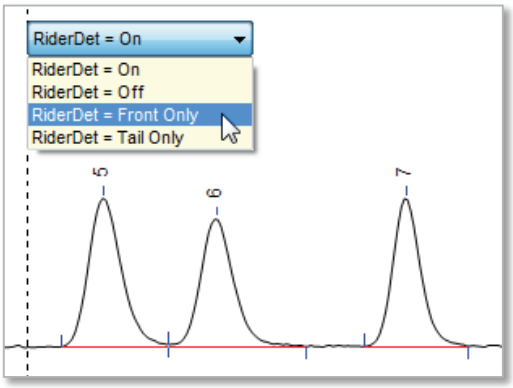

Figure 48. Dropdown list in flag.

Editing parameter locations

In order to change the retention time of a particular parameter, it is simply a case of moving the mouse over the dotted line of the flag, and the mouse pointer changes its appearance to look as follows: ⇿

Then just click and drag the parameter to a new retention time (Figure 49). Chromeleon CDS updates the retention time in the processing method and instantly shows the revised chromatogram.

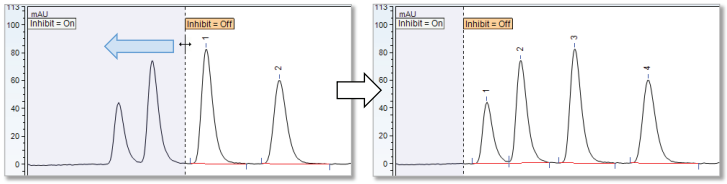

Figure 49. Moving a single parameter.

If there are several parameters at the same retention time, it is possible to move either all or individual parameters. To move all the parameters together use the procedure described above to select the dotted line and drag to the required retention time (Figure 50).

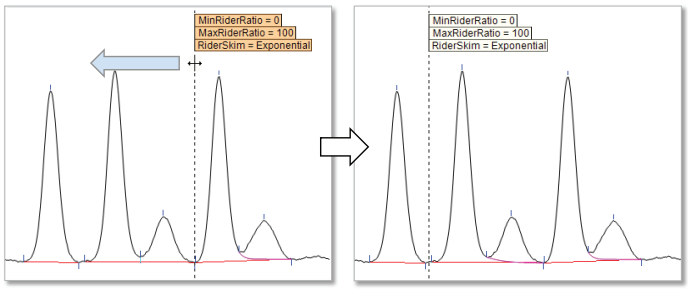

Figure 50. Moving a group of parameters.

To move only one flag from the group, move the mouse over the required flag until it changes color, then simply click and drag to reposition the parameter (Figure 51).

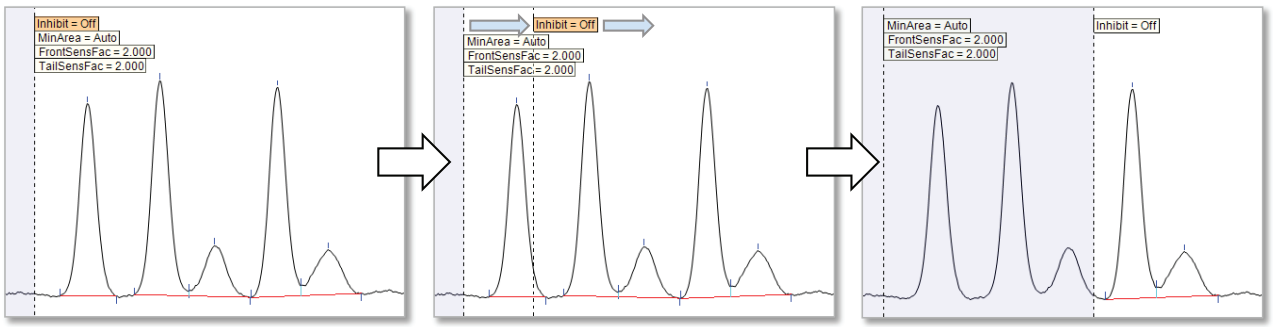

Figure 51. Moving one parameter from a group.

9. Summary

The powerful and intelligent peak detection and integration features of Chromeleon CDS address the common challenges in peak detection:

- The Cobra algorithm successfully distinguishes peaks from noise and correctly identifies the underlying baseline.
- The Cobra Wizard guides the operator through correctly assigning initial peak detection parameters.
- The SmartPeaks Integration Assistant aids and guides the chromatographer through correctly handling rider peaks and other unresolved peaks.
- Using simple in-chromatogram editing tools, any chromatographer can easily, intuitively, and quickly insert new and modify existing parameters.

All of these integration tasks can be performed without the chromatographer needing to open the processing method - although the parameters are added to the processing method, thereby maintaining correct peak and baseline detection throughout a sequence of chromatograms and minimizing variations introduced by different operators and manual integration manipulations.

Using Chromeleon CDS, laboratories gain a significant boost in overall efficiency and productivity.

Find out more at **thermofisher.com/chromeleon**

ThermoFisher SCIENTIFIC