Optimize Peak Detection & Integration with ApexTrack/Processing Theory

Noise and Drift Calculations
Managing Manually Integrated results
System Suitability

Rune Buhl Frederiksen,
Nordic Customer Education Manager
Presentation overview of

- Integration Theory
- Noise & Drift Calculations
- Managing Manually Integrated Results in a Result Set
- System Suitability Calculations and Limits

Exercise
Integration

Integration requires three operations:
1. Find the peak (peak detection)
2. Find the baseline of the peak
3. Compute the peak’s area and height

The first two are the challenge

Empower has two different algorithms to perform integration
- Traditional
- Apex Track
Traditional Integration

- Peak Width and Threshold work together to detect the peaks from the baseline.

4 Global Parameters

- Peak Width
- Threshold
- Minimum Area
- Minimum Height
**Peak Width**

- Peak width is measured at the baseline of the narrowest peak of interest and is used to determine a bunching factor.

- Bunching Factor = \( \frac{\text{Peak Width} \times \text{Sampling Rate}}{15} \) = 4

---

Traditional Integration  
Peak Width Determination
Traditional Integration
Determining peak start

**Threshold**

- Specifies the liftoff and touchdown values (minimum rate of change of the detector signal) for peak detection.
- Empower averages the signal slope across 3 data bunch intervals and compares to the liftoff threshold.
- When the average slope of the signal between the 3 bunches is ≥ the liftoff threshold value, point B1 is flagged as possible peak start.
- Individual points in bunch B1 is then examined to determine peak start = data point with lowest Y-value.

\[
\text{slope 1} = \frac{B2 - B1}{t_2 - t_1} \quad \text{slope 2} = \frac{B3 - B2}{t_3 - t_2} \quad \text{average slope} = \frac{\text{slope 1} + \text{slope 2}}{2}
\]
Traditional Integration
Determining peak apex

- Signal is monitored until slope sign changes from positive to negative.
- Bunch where the slope change occurs (B12 in the figure) is analyzed.
- Data point which is farthest away from the baseline is tentatively assigned as peak apex.
- Final apex is determined after integration and baseline assignment.

©2011 Waters Corporation
Traditional Integration
Determining peak end

- Slope of the signal is compared to the touchdown threshold
- When 2 consecutive slopes are < threshold, last point in the last bunch is flagged as possible peak end
- Individual points in this bunch and the next bunch to determine actual peak end = data point with lowest Y-value
Minimum Height or Minimum Area

- Defines minimum peak area (mV*sec) or minimum peak height (µV) that Empower will report
- Used to reject unwanted peaks once integration has been optimized
- Empower use 95% of the peak’s area/ height so that it can report peaks that approach the selected peak’s size
Timed Events

- a **time-based action** to adjust peak detection and/or integration in specified sections of a chromatogram

- There are 20 integration events that can be used to fine-tune integration across selected regions of a chromatogram

- You might need to apply one or more timed events when the default peak detection and integration parameters do not adequately detect and integrate all peaks in the chromatogram.
Traditional Integration
Timed Events

- II – Inhibit Integration
- SPW – Set Peak Width
- SLO – Set Liftoff
- STD – Set Touchdown
- SMA – Set Minimum Area
- SMH – Set Minimum Height
- SMxA – Set Maximum Area
- SMxH – Set Maximum Height
- VV – Valley to Valley
- ES – Exponential Skim

- TS – Tangential Skim
- ANP – Allow Negative Peaks
- FDL – Force Drop Line
- FBT – Force Baseline by Time
- FBP – Force Baseline by Peak
- FHP – Forward Horizontal by Peak
- FHT – Forward Horizontal by Time
- RHP – Reverse Horizontal by Peak
- RHT – Reverse Horizontal by Time
- FP – Force Peak
ApexTrack Integration

A New Approach to the Integration of Chromatographic Peaks

- Easier than traditional integration
- Better than traditional integration
- Based on measuring the curvature (the rate of change of slope) of the chromatogram (2\textsuperscript{nd} derivative)
- Traditional integration detects peaks by initially looking for a peak start
- ApexTrack integration detects peaks by initially looking for the peak apex
ApexTrack Integration

Easier:
- Automatically determines appropriate integration parameter settings
  - Auto Peak Width
  - Auto Threshold
- Usually integrates well at first pass using default and automatic parameters

Better:
- Integrates negative peaks effectively
- Integrates small peaks in noisy or drifting baseline effectively
- Peak shoulders are easily detected
- Gaussian skimming available
System Policies

Data Processing Technique

- Allow the use of ApexTrack Integration
- Default Settings Used When Creating New Projects
  - Enable ApexTrack Integration

Default Integration Algorithm:
- Traditional
- ApexTrack

If you are implementing FDA Electronic Records, Waters recommends setting the options marked with (ER). If you are implementing FDA Electronic Signatures, Waters recommends setting the options marked (ES).
New Project Wizard

Data Processing Techniques

- Enable ApexTrack Integration
- Default Algorithm: ApexTrack

- Enable ApexTrack Integration
- Default Algorithm: ApexTrack
Basis of ApexTrack: Curvature Threshold

- Detects the peak apex when the *curvature* is above the threshold
- **Effective:**
  - Detects shoulders
  - Baseline slope does not affect detection of peaks
  - Peak detection and baseline determination are decoupled
    - Baseline placement can be modified without affecting the number of peaks detected and vice versa
Second Derivative Measures Curvature

- Apex High (-)
- Inflection points Zero
- Liftoff/TD High (+)
- Baseline Zero

### Gaussian peak

1. Apex
2. Inflection points
3. Liftoff/TD
4. Baseline

### Curvature

- High (-)
- Zero
- High (+)
- Zero
Apex Track Integration

**Apex detection parameters**
- Start (min) (Start Detection/Integration Time)
- End (min) (End Detection/Integration Time)
- Peak Width (sec) (Peak Width @ 5% Height)—**AUTO**
  - Recommended range = 0.5 to 2 times Auto PW value
- Detection Threshold (Peak Detection Threshold)—**AUTO**

**Baseline determination parameters**
- Liftoff %
  - Baseline start threshold %. Default: 0
- Touchdown %
  - Baseline end threshold %. Default: 0.5

**Peak acceptance criteria**
- Minimum Area (works in the same way as in traditional int.)
- Minimum Height (works in the same way as in traditional int.)
ApexTrack Peak Detection

- Peak detection is controlled by the Peak Width and Threshold parameters
- Peak Width: measured in seconds, Auto Peak width sets it to 5% height of the largest peak in the second derivative (determined by using the inflection point width and calculating the gaussian peak width); used as a filter similar to traditional integration.
- Threshold: measured in units of height, Auto Threshold sets it to the peak to peak noise; used as a threshold for peak detection in the 2nd derivative
2nd derivative plot
Threshold

2nd derivative plot

AutoWidth
AutoThreshold
Peak to peak noise
Apex detection

2nd derivative plot

Apex Detection

Apex Detection

AutoWidth

AutoThreshold

Considered as noise
Apex Track Integration

**What happens?**

1. Acquire the data
2. Obtain chromatogram’s second derivative
3. Determine peak width (AutoPeakWidth)
4. Determine threshold (AutoThreshold)
5. Detect peaks - Second Derivative
6. Identify inflection points

![2nd derivative plot]

- Apex Detection
- AutoWidth
- AutoThreshold
- Considered as noise
What about Baseline determination?

- ApexTrack uses *percentage* slope threshold.
  - The slope threshold depends on peak height
  - The baseline is the same for all peaks

Why?

- Baselines change when concentration changes and the location of touchdown is most sensitive.

What happens?

- User specifies baseline threshold as a *percentage* of peak height.
- Algorithm computes a separate slope threshold for each peak
- Slope threshold is then *proportional* to peak height
  - Big peaks have big threshold
  - Small peaks have small threshold
Baseline Determination

1. Initially draws baseline between the inflection points
2. Determines slope differences ($\Delta m$) using tangents to the inflection points

3. Determines slope thresholds using Baseline % Thresholds from processing method and slope differences.
Baseline % Thresholds scale inflection point slope differences to determine liftoff and touchdown points.

\[
\text{Peak}_{\text{start}} = \Delta m_1 \times \text{Liftoff\%}/100 \\
\text{Peak}_{\text{stop}} = \Delta m_2 \times \text{Tuchdown\%}/100
\]
Baseline Determination

4. Baselines start at the "inflection point" baseline
5. Baselines are expanded until the slope threshold criteria are met

6. A Baseline % Threshold of 100 % yields baseline at inflection points
7. A Baseline % Threshold of 0 % yields baseline that is tangent to baseline noise
Concentration Change: Traditional Approach

- Height ratios of 1: 1/10 : 1/100
- Times of liftoff and touchdown change
- Biggest peak: Touchdown far down in tail
Focus on 1/10 peak
Middle peak: Touchdown is well positioned
Concentration Change: Zoom In Again

- Focus on 1/100 peak
- Smallest peak: Touchdown is high up the tail
- Relative area of smallest peak is reduced!
Concentration Change: ApexTrack

- Height ratios of 1: 1/10 : 1/100
- Liftoff is the same for each peak.
- Touchdown is the same for each peak
- Biggest peak: Touchdown is well positioned
Concentration Change: Zoom In

- Focus on 1/10 peak
- Middle peak: Touchdown is well positioned
Concentration Change: Zoom In Again

- Focus on 1/100 peak
- Smallest peak: Touchdown is well positioned
- Note different slope thresholds
Changing %Touchdown

- Focus on Big peak
- A small change in the %Touchdown will have a big impact on the slope for the big peak because it is a percentage of the peak height
- This will have very little effect on the middle peak and NO effect on the small peak
Apex Track Integration
Timed Events

- ANP - Allow Negative Peaks
- DS - Detect Shoulders
- GS - Gaussian Skim
- TS - Tangential Skim
- II - Inhibit Integration
- MP - Merge Peaks (for GPC only)
- SL% - Set Liftoff %
- ST% - Set Touchdown %

- SMA - Set Minimum Area
- SMH - Set Minimum Height
- SMxH - Set Maximum Height
- SMxA - Set Maximum Area
- VV - Valley-to-Valley
- SPW - Set Peak Width
- SDT - Set Detection Threshold
## Integration events

**Comparison: Traditional – Apex Track**

<table>
<thead>
<tr>
<th>Traditional Integration Event</th>
<th>ApexTrack Integration Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhibit Integration</td>
<td>Inhibit Integration</td>
</tr>
<tr>
<td>Allow Negative Peaks</td>
<td>Allow Negative Peaks</td>
</tr>
<tr>
<td>Set Liftoff</td>
<td>Set Liftoff %</td>
</tr>
<tr>
<td>Set Touchdown</td>
<td>Set Touchdown %</td>
</tr>
<tr>
<td>Set Peak Width (sec)</td>
<td>Set Peak Width (sec)</td>
</tr>
<tr>
<td>Set Threshold</td>
<td>Set Detection Threshold</td>
</tr>
<tr>
<td>Set Minimum Area</td>
<td>Set Minimum Area</td>
</tr>
<tr>
<td>Set Minimum Height</td>
<td>Set Minimum Height</td>
</tr>
<tr>
<td>Set Maximum Width (sec)</td>
<td>Set Maximum Width (sec)</td>
</tr>
<tr>
<td>Set Maximum Height</td>
<td>Set Maximum Height</td>
</tr>
<tr>
<td>Valley to Valley</td>
<td>Valley to Valley</td>
</tr>
<tr>
<td>Exponential Skim</td>
<td></td>
</tr>
<tr>
<td>Force Drop Line</td>
<td></td>
</tr>
<tr>
<td>Force Baseline by Peak</td>
<td></td>
</tr>
<tr>
<td>Force Baseline by Time</td>
<td></td>
</tr>
<tr>
<td>Force Peak</td>
<td></td>
</tr>
<tr>
<td>Forward Horizontal by Peak</td>
<td></td>
</tr>
<tr>
<td>Forward Horizontal by Time</td>
<td></td>
</tr>
<tr>
<td>Reverse Horizontal by Peak</td>
<td></td>
</tr>
<tr>
<td>Reverse Horizontal by Time</td>
<td></td>
</tr>
<tr>
<td>Tangential Skim</td>
<td>Tangential Skim</td>
</tr>
<tr>
<td></td>
<td>Merge Peaks (GPC option only)</td>
</tr>
<tr>
<td></td>
<td>Detect Shoulders</td>
</tr>
<tr>
<td></td>
<td>Gaussian Skim</td>
</tr>
</tbody>
</table>
Conclusions

Advantages over other Integration Packages

1. Automatic parameter determination, for rapid method development
2. Default parameters superior to those of Traditional
3. Curvature detection, for reproducible detection of difficult peaks and shoulders
4. Internally adjusted slope threshold, for accurate baseline determination, does not affect peak detection
5. Gaussian Skimming
System Suitability Calculations
System Suitability Tab

- New in Empower3

- Calculate Suitability Results
- Calculate Suitability Results for Unknown Peaks

- System and Separation Efficiency:
  - Void Volume Time (min): 0.230
  - US Pharmacopoeia
  - European Pharmacopoeia
  - Japanese Pharmacopoeia
  - All

- Tangent Percent for USP Resolution: 50
- Tangent Percent for USP Plate Count: 61

- Calculate USP, EP, and JP s/n
  - Use noise centered on peak region in blank injection
  - Half Height Multiplier for USP s/n Noise Region: 5
  - Half Height Multiplier for EP s/n Noise Region: 20
  - Half Height Multiplier for JP s/n Noise Region: 20

- Noise Value for s/n: Peak to Peak Noise

- Baseline Noise and Drift Measurements:
  - Maximum Allowable Noise
  - Maximum Allowable Drift
  - % Run Time Over Which to Average: 5.0
  - Baseline Noise Minimum: 30 Seconds
  - Baseline Start Time (min)
  - Baseline End Time (min)
<table>
<thead>
<tr>
<th>Name</th>
<th>Calculate Suit Results</th>
<th>Flag Outside Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaminophen</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Impurity 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impurity 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caffeine</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Acelaminide</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Acetylmalic Acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impurity 3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Phenacetin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>API 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impurities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown Impurities</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Suitability Limits**

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Target</th>
<th>Error %</th>
<th>Lower Error Limit (LCL)</th>
<th>Upper Error Limit (UCL)</th>
<th>Warning %</th>
<th>Lower Warning Limit</th>
<th>Upper Warning Limit</th>
<th>Ignore Blank Values</th>
<th>Check Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Setting System Suitability Limits

- List of compounds:
  1. Acetaminophen
  2. Impurity 1
  3. Impurity 2
  4. Caffeine
  5. Acetanilide
  6. Acetylsalicylic Acid
  7. Impurity 3
  8. Phenacetin
  9. API’s
  10. Impurities
  11. Unknown Impurities

- Table headings:
  - Field Name
  - Target
  - Error %
  - Lower Error Limit (LCL)

- Fields:
  - Retention Time
  - Area
  - % Area
  - Height
  - Amount
  - Relative RT
  - RT Ratio
  - Start Time
  - End Time
  - Baseline Start
<table>
<thead>
<tr>
<th>Name</th>
<th>Retention Time</th>
<th>USP Tailing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>0.900</td>
<td>1.050</td>
</tr>
<tr>
<td>Error %</td>
<td>5.00</td>
<td></td>
</tr>
<tr>
<td>Lower Error Limit (LCL)</td>
<td>0.855</td>
<td></td>
</tr>
<tr>
<td>Upper Error Limit (UCL)</td>
<td>0.945</td>
<td></td>
</tr>
</tbody>
</table>
Noise & Drift Calculations
Empower Noise and Drift Calculations

- There are 8 different calculations that can be performed:
  - Detector Noise
  - Peak-to-Peak Noise
  - Detector Drift
  - Average Detector Noise
  - Average Peak-to-Peak Noise
  - Average Drift
  - Baseline Noise
  - Baseline Drift
Enabling

System Suitability test in Empower3 Course Data as System/Administrator - Review - [LC Processing Method]

Noise and Drift Parameters

- Calculate Detector Noise and Drift

- Start Time: (min) 3.200
- Stop Time: (min) 3.900
- Segment Width: (sec) 10
Visual Representation of the Least-squares line

Value of data point predicted by the line

Value of data point
Detector drift is the slope of the least-squares line. Drift is expressed in detector units per hour.

- For example, the drift calculation for a UV detector would be expressed in absorbance units (AU) per hour. Average Drift is calculated by dividing the data into segments (specified in the processing method) and averaging the values for each segments.
The root mean square (RMS) noise of the data is calculated using the least-squares line. The formula for Detector Noise is:

\[
\sqrt{\frac{\sum (y_i - y_{pi})^2}{n - 2}}
\]

Where

- \(y_i\) = the y value of the data point
- \(y_{pi}\) = the y value of the data point predicted by the line
- \(n\) = the number of datapoints

Residual = \(y_i - y_{pi}\)

3.49X10^{-6} AU
Average Detector Noise

2.43x10^-6 AU

1.7x10^-6

2.7x10^-6

3.4x10^-6

1.9x10^-6
Peak to Peak Noise is defined to be the algebraic difference of the maximum and minimum residuals between each data point and the least-square line. The “residual” is determined by subtracting the $y$ value of the data point predicted by the line from the $y$ value of the data point.

The formula for Peak to Peak Noise is:

$$\text{Peak to Peak Noise} = \text{Max residual} - \text{Min. residual}$$

Where Residual $= y_i - y_{pi}$
Average Peak to Peak Noise?

0.7\times 10^{-5}

1.5\times 10^{-5}

1.4\times 10^{-5}

1.0\times 10^{-5}

1.15\times 10^{-5} AU
Baseline Noise and Drift Calculations

- Baseline Noise setup in System Suitability
  - Peak to Peak Calculation
  - 30 second segments (not adjustable)
  - Set time range and percent of run time.
Baseline Noise and Drift Measurements

Maximum Allowable Noise

Maximum Allowable Drift

% Run Time Over Which to Average 5.0

Baseline Noise Minimum 30 Seconds

Baseline Start Time (min) 1.0  Baseline End Time (min) 9.0

Averaging Baseline Noise Example

Averaged region 1 to 1.5 min. Averaged region 0.5 to 9 min.

Baseline start (BL start + avg. time)  (BL end - avg. time)

1 min. 1.5 min. 0.5 min. 9 min.

Baseline end
Viewing the Calculated Results

<table>
<thead>
<tr>
<th>Detector Noise (Plot Units)</th>
<th>Detector Drift (Plot Units/hour)</th>
<th>Peak to Peak Noise (Plot Units)</th>
<th>Average Detector Noise (Plot Units)</th>
<th>Average Detector Drift (Plot Units/hour)</th>
<th>Average Peak to Peak Noise (Plot Units)</th>
<th>Baseline Drift (mV)</th>
<th>Baseline Noise (mV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.007074</td>
<td>-0.887558</td>
<td>0.048889</td>
<td>0.002777</td>
<td>-2.922607</td>
<td>0.012021</td>
<td>-0.011</td>
<td>33.406</td>
</tr>
</tbody>
</table>
Managing Manually Integrated Results in a Result Set
Managing Manually Integrated Results in a Result Set, Samples and Controls

Process Sample Set

• Uses Sample Set Information (bracketing)
• Use Acquisition Method Set (recommended)

Review Result Set

• Manual changes to integration
• Click **Quantitate** only
• DO NOT change standards or click Calibrate
• Save Result

Print Result Set

• Will ensure that only latest results will be printed
Managing Manually Integrated Results in a Result Set, Samples and Controls

Sample Set
- Channel 1046
- Channel 1054
- Channel 1073

Result Set 1
- Result 2033
- Result 2039
- Result 2063

Review & Manual integrate Result Set
- Result 2033
- Result 2039
- Manual Result 3741
- Result 2063

New Version of Result Set
Managing Manually Integrated Results in a Result Set, Standards (& Samples/Controls)

1. **Process Sample Set**
   - Uses Sample Set Information (bracketing)
   - Use Acquisition Method Set (recommended)

2. **Review Result Set**
   - Manual changes to integration only
   - DO NOT **Calibrate** or change proc. method
   - Save Result

3. **Reprocess Result Set**
   - Select **Use Existing Integration**
   - **Calibrate** and **Quantitate**

4. **Print Result Set**
   - Will ensure that only latest results will be printed

©2011 Waters Corporation