



Neurotechnology Research and Development

# User's Guide

Version 2.7 January 2005

## Data Analysis Software **Offline Sorter**



Plexon Inc  
6500 Greenville Ave., Ste 730  
Dallas, Texas 75206  
USA

**Caution****Electrostatic Discharge**

Some devices can be damaged by improper handling. Use appropriate electrostatic discharge (ESD) procedures when handling these devices. See <http://www.esda.org/> for additional information on ESD procedures.

---

Data Analysis Software

# Offline Sorter

User's Guide

---

Document Number: 00000-C-5002-A

Document Issue: 1.6

Document Status: Standard

Software Version and Issue: 2.7

Date: January 2005

---

Copyright © 1983-2005 Plexon Inc, All Rights Reserved

Printed in the United States of America

## **Plexon Inc Proprietary**

The information contained herein is the property of Plexon Inc. and it is proprietary and restricted solely to assist Plexon Inc. customers. Neither this document nor the contents may be disclosed, copied, revealed or used in whole or in part for any other purpose without the prior written permission of Plexon Inc. This document must be returned upon request of Plexon Inc.

Information is subject to change without notice. Plexon Inc. reserves the right to make changes in equipment design or components as progress in engineering or manufacturing may warrant.

Plexon Inc. and the Plexon logo are trademarks of Plexon Inc.

---



---

# Publication History

---

## January 2005

Version 2.7 of the Offline Sorter User's Guide includes the following new items:

- Automatic artifact detection
- Automatic waveform detection
- Minor updates

## September 2004

Version 2.6 of the Offline Sorter User's Guide includes the following new items:

- Documentation of the new features described in the release notes
- Revised illustrations of interface changes
- Minor edits and updates

## March 2004

Version 2.5 of the Offline Sorter User's Guide includes the following new items:

- Revised format
- Documentation for several new features, which are described in detail in the release notes
- Additional information for the Overview, Units Display, 3D Cluster View, Spike Sorting, Unit Cleaning, Info View Statistics, and Reference sections
- Revised graphics in several areas

## September 2003

Version 2.4 of Offline Sorter includes the following new features:

- Zoom and pan capability for the 2D cluster view
- Revised zoom display and zoom toolbar buttons
- Zooming and moving along Z axis

- Zoom reset
- New sorting ability to the Feature vs Feature view, along with specialized layouts for supporting tetrode sorting.
- Support for 40-bit timestamps in plx files
- Neuralynx importer is more general
- Stereotrode/tetrode conversion support for CED Spike2 files
- Option to base an interval on start of file instead of an event
- Revised alignment for batch processing
- Option to skip reading continuous data channels when loading large PLX files
- Option to skip writing continuous data channels into NEX files
- Option to skip calculating advanced statistics (i.e. sort quality statistics)
- Support for waveform changes (e.g. do alignment) on plx0 files
- Display of slow channel frequency for PLX files in Global tab
- Faster operation of the continuous data view for large files

## April 2003

Version 2.3 of Offline Sorter includes the following new features:

- Importation of Neuralynx files
- Auto-Sort Tool—automatically sorts all channels
- Display calculates and optionally displays ellipsoids around clusters, in both 2D and 3D views.
- The Remove Outliers and Assign Unsorted Waveforms tools use Mahalanobis distance instead of simple Euclidean distance—application is optional following a waveform-limited valley-seeking sort.
- Data can be exported on a per-unit basis.
- Faster processing of large files:
  - A Defer Drawing option for the Waveforms Display plots the waveforms only when the mouse is clicked in the waveforms display.
  - By default, unit pair-wise sort quality metrics are not calculated after each change—this calculation can be re-enabled on the Stats options page.
  - Optimized statistics calculations
- An extended scan view that shows more information.
- Support for DDT Version 102 files, with per-channel gains

---

# Contents

---

---

<b>Publication History</b>	<b>iii</b>
----------------------------	------------

---

<b>About This Document</b>	<b>xi</b>
----------------------------	-----------

---

Applicability of this document	1-xi
How this document is organized	1-xi
Indication of trademarks in this document	1-xi
Related documents	1-xi
Document Conventions	1-xii

---

<b>Chapter 1</b>	
<b>Getting Started with Offline Sorter</b>	<b>1-1</b>

---

Installation	1-2
Installing Hardware Keys	1-2
Licensing	1-3
Off-line Sorter Technical Support and Updates	1-4
Offline Sorter Technical Support	1-4
Offline Sorter Updates	1-4
Web Updates	1-4
Demo Mode	1-4

---

<b>Chapter 2</b>	
<b>Overview</b>	<b>2-1</b>

---

Offline Sorter Uses	2-2
Offline Sorter Inputs	2-3
Sorting Techniques	2-4
Offline Sorter Outputs	2-7
Printing Results	2-7
Saving Results	2-7
Save As New PLX File	2-7
Save As NEX File	2-7
Exporting Per-Waveform Data	2-7
Exporting Per-Channel Data	2-7
Exporting Per-Unit Data	2-8
Reproducing Results	2-8
Export Sort Info to .TPL	2-8
Sort Using .TPL file	2-8
Sort Using .NEX file	2-8

---

Update .EXP file with Sort Info	2-8
Execute Batch Command File	2-8

<b>Chapter 3</b>	
<b>Screen Elements and Displays</b>	<b>3-1</b>
Screen Elements Overview	3-2
Toolbars	3-2
Control Table	3-2
Waveform Window	3-2
Units Display	3-3
Continuous Data Window	3-3
Multi-Display Window	3-3
Control Table	3-5
Waveform Window	3-9
Units Display	3-11
Selecting Units	3-12
Multi-Display Window	3-13
Info View	3-13
Channels Tab	3-14
Chan Params Tab	3-15
Sort Details Tab	3-16
Global Tab	3-17
Events Tab	3-19
Stats Tab	3-19
PCA Tab	3-19
Templates Tab	3-19
Scan Tab	3-19
Chan Map Tab	3-19
Cluster Views	3-19
Cluster View Common Functionality	3-21
2D Cluster View Details	3-28
3D Cluster View Details	3-31
3D Phase Space View	3-35
Waveform Inspection View	3-37
Waveform Detection View	3-38
Adjusting the Threshold for Waveform Detection	3-39
Changing the Detector	3-39
ISI Histograms View	3-41
Feature vs. Feature View	3-42
Common Functionality	3-45
Sorting Summary View	3-46
Output to PowerPoint	3-49
Outputting for All Channels	3-50
PCA Results View	3-50
Cross-Correlograms View	3-51
Rasters View	3-52
Continuous Data Window	3-54
For Continuously-Recorded Data	3-55
For Spike Waveform Data	3-56
Other Features of the Continuous Data Windows	3-57

<b>Chapter 4</b>	
<b>Spike Sorting Step-by-Step</b>	<b>4-1</b>



Procedure Summary	4-2
Waveform Acquisition From Continuously Recorded Data	4-3
Opening the Digitized Data Files	4-3
Selecting the Continuous Data Channel	4-4
Selecting the Low-Cut Digital Filter	4-6
Specifying the Waveform Detection Options	4-6
Extracting the Thresholded Waveforms	4-9
Waveform Acquisition From Spike/Waveform Data	4-10
Opening the Spike/Waveform Data File	4-10
Stereotrode and Tetrode Data Type Selection	4-10
Selecting the Data Channel	4-11
Preparation For Sorting	4-12
Data Inspection	4-12
Waveform Window	4-13
Cluster Views	4-13
Continuous Data Window	4-15
Invalidating Waveforms	4-15
Marking Artifact Waveforms as Invalid using the Waveform Window	4-16
Marking Noise Waveforms as Invalid Using the Cluster Views	4-17
Invalidating All Except	4-20
Automatically Invalidating Artifacts	4-20
Waveform Alignment	4-22
Adjusting the Sort Range	4-25
Selecting Features to Use For Sorting	4-26
Spike Sorting	4-27
Manual, Semi-Automatic, and Automatic Clustering Methods	4-27
Selecting the Sorting Method	4-28
Manual Clustering Using Contours	4-29
Adjusting the Sorting	4-31
Manual Clustering using the Waveform Crossing Method	4-36
Adjusting the Sorting	4-38
Semi-Automatic Clustering using Templates	4-41
Semi-Automatic Clustering using K-Means	4-42
Selecting Initial Cluster Centroids	4-42
Running the K-Means Clustering Algorithm	4-44
Semi-Automatic Clustering using Standard E-M	4-45
Running the Standard E-M Clustering Algorithm	4-46
Automatic Clustering using Valley-Seeking Method	4-46
Automatic Clustering using T-Distribution E-M Method	4-49
Unit Cleaning	4-50
Controlling Waveforms Displayed	4-50
Controlling waveforms viewed by Units	4-50
Controlling waveforms viewed by Time	4-52
Waveform Selection Options	4-53
Cleaning the Selection	4-54
Automatic Outlier Removal Tool	4-54
Manual Waveform Removal	4-54
Combining Units	4-56
Swapping Units	4-56
Invalidating Units	4-56
Removing Spikes with Short Inter-spike Intervals	4-56
Automatic Short ISI Spike Removal	4-58
Examining Sorting Statistics	4-58
Info View Statistics	4-59

Top Half of the Window	4-59
Bottom Half of the Window	4-60
Statistics Feedback While Sorting	4-61
Saving the Sorting Results	4-62
Saving Timestamps in a NEX File	4-62
Re-sorting from a NEX File	4-63
Saving Timestamps and Waveforms	4-63
Overwrite/Backup Control on File-Saves	4-63
Saving a New PLX File	4-64
Exporting Information	4-65
Exporting Per-Waveform Data	4-67
Exporting Per-Channel Data	4-69
Exporting Per-Unit Data	4-71

## Chapter 5

### Reference Section

5-1

Channel Types and Associated Channels	5-2
Stereotrode and Tetrode Data	5-4
Converting Data Type	5-5
Saving Converted Data	5-6
Disabling Electrodes	5-7
Automatic Waveform Detection	5-8
Features Available for Sorting	5-9
Principal Components	5-11
Voltage at Specific Slices in Time	5-12
Details of the Sorting Algorithms	5-13
Feature Space Used for Sorting	5-13
Template Algorithm	5-14
K-Means Algorithm	5-15
Standard E-M Algorithm	5-16
Valley Seeking Algorithm	5-16
T-Distribution E-M Algorithm	5-17
Running Automatic Sorting on All Channels	5-18
Special Sorting	5-18
Continue Sorts	5-19
Partial Sorts	5-19
Sorting Quality Statistics	5-19
MANOVA F and p-values	5-20
Some details on the MANOVA calculation	5-20
J3	5-21
Pseudo-F	5-21
Davies-Bouldin	5-22
Dunn	5-22
Playing a Movie of the data	5-22
Examining Events	5-23
Manual Short ISI Spike Removal	5-25
Interval Selection	5-27
Waveform Inspection	5-29
Selecting Waveforms in the Continuous Data Window	5-30
Comparing the Waveform with the Templates	5-32
Template Files	5-33
Saving Templates	5-33
Sort Using Saved Templates	5-34
Save Templates to a Sort Client EXP File	5-35

Parameter Scans	5-36
Re-thresholding Spike Data	5-38
Assigning Unsorted Points to Closest Unit	5-39
Batch File Processing	5-39
Example Batch Command File	5-42
Notes	5-43
Printing	5-44
Large File Tips	5-45
Statistics Calculations	5-45
Buffer Sizes	5-45
Ignoring Continuous Data Channels	5-46
Initial PCA Calculation	5-46
Deferring Waveform Drawing	5-46
Split PLX files (.plx0)	5-47



---

# About This Document

---

This user's guide describes Offline Sorter and its associated components.

## Applicability of this document

This document applies to Version 2.7 of Offline Sorter.

## How this document is organized

This book is organized into sections that cover the following topics:

- Getting started
- Overview
- Screen elements and displays
- Spike sorting procedures
- Reference

## Indication of trademarks in this document


The asterisk after the first use of a name indicates a trademarked item. The inside back cover acknowledges all trademarked items.







## Related documents

There are no other related documents.

## Document Conventions

The consistent use of typographic conventions in Plexon documentation helps you locate and interpret information easily. The following guidelines present some of the specific typographic conventions used by Plexon.

Item	Convention	Example
Plexon programs	Title caps	Sort Client Server Recorder
Abbreviations, acronyms, and initialisms	All uppercase	DAQ PCA
Code samples or machine output	Monospaced font	Myfile.ddt C:\Program Files\Plexon Inc\Recorder
Command-line commands and options (switches)	All lowercase, bold	<i>copy</i> command <b>/a</b> option
Commands on menus and buttons	Bold, usually with title caps	<b>Start DAQ</b> <b>Apply</b>
Data formats	All uppercase	PLX DDT
Device names	All uppercase	MAP NI DAQ
Dialog box options	Bold; capitalization follows the interface	Click the <b>Generate file name and start recording immediately</b> check box
Dialog box titles	Bold; title caps	<b>Options</b> dialog box <b>Import/Export Setup</b> dialog box
File names	Case sensitive	MyChannelOne.plx
New terms or emphasis	Italic or quotation marks	Plexon software consists of both <i>server</i> and <i>client</i> components. You “must” close the window before you exit.
Toolbar button names	Title caps (sometimes followed by button icon); bold	Click <b>Option</b> . The <b>Start DAQ</b> (  ) button ...

Item	Convention	Example
Cautions and Warnings		 <b>CAUTION</b> <b>Unpredictable Results</b> If you change card selections you must restart Recorder for the changes to take effect. Otherwise, unpredictable results can occur.
		 <b>CAUTION</b> <b>Electrostatic Discharge</b> These devices can be damaged by improper handling. Use appropriate electrostatic discharge (ESD) procedures when handling these devices.
Hints or Tips		 <b>HINT</b> <b>Too many channels?</b> All 64 channels are displayed by default. You can choose a display with fewer channels in the Options menu.





# Chapter 1

## Getting Started with Offline Sorter

---

Installation .....	1-2
Licensing.....	1-3
Off-line Sorter Technical Support and Updates .....	1-4
Demo Mode .....	1-4

---

## Installation

Before you can use Offline Sorter, you must install the Offline Sorter program files, the Sentinel System Drivers, and you must install the Sentinel parallel port or USB hardware key. The Plexon hardware key(s) are marked with the Plexon logo and they are included with your shipment of Offline Sorter.

### Running Offline Sorter Setup

**Note:** Before you begin installing Offline Sorter, exit all currently running applications.

- 1 Place the Offline Sorter setup CD in your computer's CD-ROM drive.
- 2 Navigate to the `OFSv2Setup.exe` file on the CD and double-click the file.  
*The Offline Sorter Version 2 Setup screen appears.*
- 3 Follow the prompts on the setup dialogs and complete the installation.  
*At the end of the installation process, the Sentinel System Driver Setup starts automatically.*
- 4 Follow the Sentinel Driver Setup prompts and complete the installation of the Sentinel Drivers.

By default, the Setup program places the Offline Sorter program and the help files in `C:\Program Files\Plexon Inc\OfflineSorter`.

## Installing Hardware Keys

You must install USB or parallel-port hardware keys on the computer running Offline Sorter.

### To install a USB key

- Locate an available USB port on your computer and insert the USB key in the USB port.

### To install a parallel port key

- 1 Locate an available parallel port on your computer.  
**Note:** If your computer has one parallel port, you can temporarily remove any existing parallel port devices, such as a printer or a Zip drive, to connect the key. After you have installed the key, you can reconnect these devices to the key's outside connector.
- 2 Attach the key to the parallel port connector and tighten the screws to connect the key securely to the port.
- 3 If necessary, reconnect any other parallel port devices to the outside connector of the key.

**Note:** If you are connecting a printer to your computer through a Sentinel Key, Sentinel recommends using shielded printer cables.

You can attach multiple parallel-port Sentinel keys to the same parallel port. You can cascade SentinelSuperPro keys with other Rainbow Technologies keys that

---

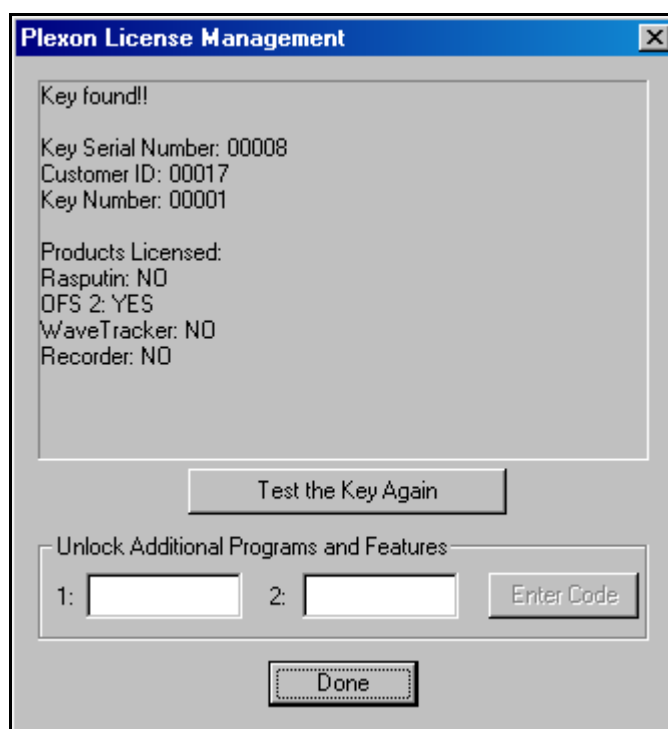
support cascading. Plexon provides SentinelSuperPro keys with your copy of Offline Sorter.

## Licensing

After you have installed the Sentinel hardware key, it requires no additional attention. However, Offline Sorter contains facilities for examining the licensing status of OfflineSorter and other Plexon products, and for performing remote upgrades to unlock additional features, or product licenses. Use the following procedure to check Plexon licenses.

Opening the License Management Window

- 1 From the main menu, select **Tools | Licensing** to bring up the **Plexon License Management** window.



**Note:** In the main text area, the window contains a summary of the licensing status of the hardware key or keys. If Offline Sorter cannot read a hardware key, an explanation appears in the window. If Offline Sorter can read a hardware key, it interrogates it and displays the information from it. The *Key Serial Number*, *Customer ID*, and *Key Number* identify a particular key. You must provide this information to Plexon to unlock additional Plexon products. The *Products Licensed* section tells you which Plexon products are currently unlocked by the hardware key. Offline Sorter is listed as **OFS 2** in this section.

- 2 Click **Test the Key Again** to re-interrogate the key and re-display the information. The **Unlock Additional Program and Features** section applies

to remote upgrades. If a remote upgrade applies, Plexon provides instructions on how to use this feature.

**Note:** Multiple Plexon hardware keys can be installed on your computer. Offline Sorter can find and use a license on any of the keys attached to your computer. If Plexon software detects multiple Plexon keys on your computer, additional **Next Key>>** and **<<Previous Key** buttons appear in the license window. You can use these buttons to step through the keys to individually examine each key.

## Off-line Sorter Technical Support and Updates

### Offline Sorter Technical Support

You can get help with Offline Sorter via e-mail at [info@plexoninc.com](mailto:info@plexoninc.com).

### Offline Sorter Updates

You can download the latest version of Offline Sorter from the Plexon website:

<http://www.plexoninc.com/Updates/OFSv2Setup.exe>

OFSv2Setup.exe is the standard Offline Sorter setup program. To install the latest version of Offline Sorter, download the program to your PC and run it from your hard drive.

### Web Updates

If you have installed a previous version of Offline Sorter, use the **Help | Web Update** menu item to automatically check for updates at the Plexon web site, and to download and install them.

## Demo Mode

If you have not connected a valid Plexon hardware key to a computer, Offline Sorter runs in **Demo Mode**. Offline Sorter opens a startup message box to inform you that it is running in Demo mode.

Plexon provides all Offline Sorter features in Demo mode, except the file saving functionality, which is limited to the first 100 waveforms on the first unit on the first channel.

---

# Chapter 2

## Overview

---

Offline Sorter Uses ..... 2-2

Offline Sorter Inputs ..... 2-3

Sorting Techniques..... 2-4

Offline Sorter Outputs..... 2-7

### Offline Sorter Uses

You can use Offline Sorter to:

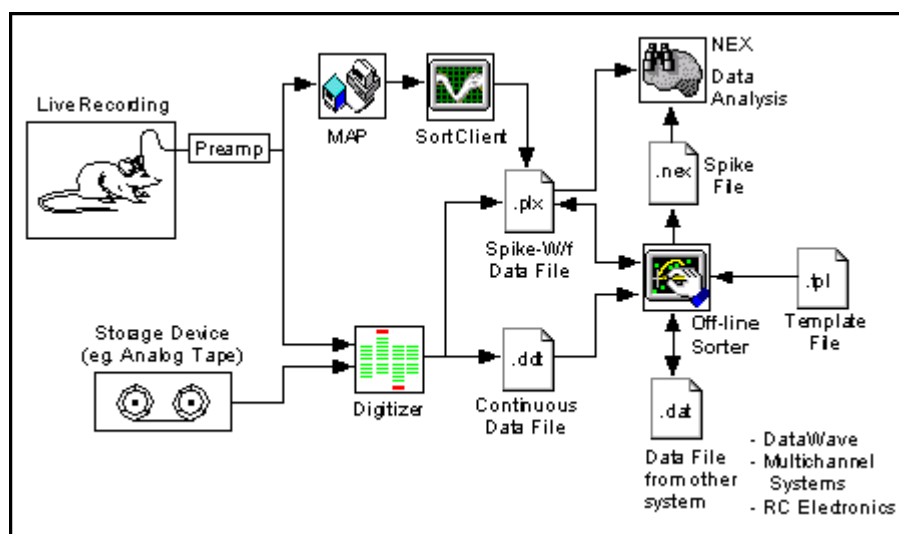
- Extract and sort action potential waveforms (spikes) from continuously recorded data, such as the data generated by Plexon programs like Recorder (DDT file) or other data acquisition systems that include:
  - Panasonic MED
  - RC Electronics
  - Data Translation

Offline Sorter detects spike waveforms within a continuous file using a threshold-crossing method. You can apply thresholding for waveform detection to the:

- Raw voltage signal amplitude
- Energy of the signal
- Non-linear energy of the signal
- Re-sort action potential waveforms from previously sorted spike/waveform data generated either by the Plexon Multichannel Acquisition Processor (MAP) data acquisition and Sort Client system (PLX file), or some other data acquisition system from the following list:
  - Bionic Technologies
  - CED Spike 2
  - DataWave
  - Instrutech
  - MultiChannel Systems
  - Neuralynx

The result of the sorting process is a set of timestamps that correspond to the spike times of the sorted units. Offline Sorter can save these spike trains in the form of a NEX datafile. If you use Offline Sorter to open a Spike/Waveform data file (e.g. a Plexon PLX file, or a data file from the DataWave or Bionic Technologies systems), you can save the resorted waveforms and spike times in a file of the same format, or in a new PLX file. You can then analyze the resulting spike

files (PLX, NEX, etc.) with NeuroExplorer (NEX). A diagram of the typical analysis pathways and files related to the Offline Sorter follows.



## Offline Sorter Inputs

Offline Sorter can **Open** or **Import** data files recorded from Plexon data acquisition hardware (PLX or DDT files), and files from various other data-acquisition hardware vendors. In general, these data files contain at least three types of data:

- **Continuously-recorded data:** With this type of data, an analog signal is digitized at a certain rate and every digitized sample is stored in the data file. Examples of this type of file are Plexon DDT files and “continuous” recording channels within a Plexon PLX file.
- **Spike Waveform data:** It is generally desirable to record the waveforms resulting from true neural firings at higher frequencies (e.g. 40 kHz). However, if you store the digitized samples between the neural firings, this can lead to large data files. Therefore, neural data acquisition systems generally employ a thresholding or some other “triggering” mechanism to record only the digitized data in a narrow time window around what is likely to be a neural firing. These waveform segments are generally called “spikes”.
- **Event data:** Many data acquisition systems have the ability to record timestamped markers in the data stream when some external event happens. While these external event markers are somewhat tangential to Offline Sorter’s main purpose of sorting spikes, Offline Sorter can display them and you can use the event markers to select time periods of interest in the data.

Offline Sorter can also read in batch command files to direct its processing and files of sorting parameters that you can apply to data files. For more information see [“Batch File Processing” on page 5-39](#).

## Sorting Techniques

Basic to multi-neuron study is the ability to resolve the spike waveforms associated with the multiple individual neurons, which often are present on single or multiple electrodes. For reviews of spike sorting techniques, see Wheeler and Heetderks, 1982, Schmidt, 1984, Lewicki, 1998, Wheeler, 1999.

For many data acquisition systems, spike detection is done on-line, and Offline Sorter can read in and use the already-detected spike waveforms. For continuously-recorded data files, Offline Sorter can perform spike detection by using a threshold-crossing algorithm. In addition to using the raw waveform amplitude for thresholding, Offline Sorter can also use calculated waveform energy, or a calculated nonlinear energy.

When spike waveforms are present, Offline Sorter helps you to separate the waveforms collected from single or multiple electrodes into distinct waveform groups or classes that are associated with individual neurons. Each class of waveforms is called a unit or a cluster.

Offline Sorter currently offers the following spike-sorting algorithms, grouped into categories:

- Manual Methods:
  - Waveform Crossing
  - Contours
- Semi-Automatic Methods:
  - Template
  - K-Means
  - Standard E-M (Expectation Maximization)
- Automatic Methods:
  - Valley-Seeking
  - T-Distribution E-M

With Manual spike sorting methods, you have full control of the sorting process, and you can manually indicate which waveforms should be assigned to each unit. With the Waveform Crossing method, you manually select the units by using the actual waveforms. You specify the unit by crossing a bundle of waveforms with the mouse pointer.

All sorting methods work in so-called “feature space”, except for the Waveform Crossing and Template sorting methods. For more information on template methods, see [“Template Files” on page 5-33](#). Instead of using the entire raw waveform, sorting methods that work in feature space characterize the essence of the waveform by using a few calculated features. You can define the feature space by selecting features from among several for each axis, which are described under



---

[“Features Available for Sorting” on page 5-9](#). The features available for assignment to the axes in feature space include:

- Projections onto principal components
- Waveform heights at chosen times (“slices”)
- Peak height, valley height, peak-valley difference, widths
- Timestamp of the waveform
- For stereotrode and tetrode data, per-electrode and ratio-between-electrodes features

Thus the waveform can be represented by the point  $(x, y)$  in a 2D feature space, or  $(x, y, z)$  in a 3D feature space, where you can choose the  $x$ ,  $y$  and  $z$  axes from the available features.

In the 3D view of the feature space points, you can navigate the eyepoint through the 3D space to get a better idea of what the data set looks like and to find a vantage point where clusters are best separated. After you achieve the desired perspective in 3D space, you can apply the sorting methods.

Using the Contours method, you specify a cluster by drawing an arbitrary shape in feature space (usually around a visible cluster). Offline Sorter assigns all waveforms inside the arbitrary shape to that cluster. When you apply the Contours method from the 3D view, it operates in a 2D projection of the 3D feature space that is a standard perspective projection based on the chosen viewpoint. That is, the 3D feature space is first transformed to a 2D feature space (“screen space”) by using a standard perspective projection, then the contour sorting algorithm runs in screen space. You can also run the other sorting algorithms in projected 2D space, or you can run them in true 3D space.

The Semi-automatic sorting methods generally require you to specify cluster centers (and therefore implicitly indicate the number of clusters), and then the algorithm assigns waveforms to the clusters.

The Template method requires you to pick an existing waveform to serve as the template for a unit. Offline Sorter then adds other waveforms to the unit based on their similarity (in the least-squares difference sense) to the template unit, with a tolerance you can define.

The K-Means method is a well-known iterative algorithm that assigns each waveform to one of the cluster centers you specify, based on Euclidean distance in feature space. Then it recomputes the cluster centers, and it repeats the process until no more waveforms change units.

The Standard E-M (Std E-M) method is a variation of the Expectation Maximization algorithm. The Expectation Maximization algorithms in general fit a mixture of Gaussians to the point densities in feature space by varying the normal distri-

bution parameters (means, covariances) to maximize a likelihood function. This algorithm uses the center points you specify to start the search.

The Semi-automatic sorting methods are also capable of using an existing set of sorted units as a starting point, in what is referred to as “Continue” sorting. The algorithms can run using the centroids or template waveforms of existing clusters as the starting points, instead of requiring you to manually specify them.

The Automatic sorting methods are capable of automatically finding the optimal number of clusters, and of assigning waveforms to the clusters. Thus they require no intervention from you to arrive at the initial clustering of the data, and they can operate on completely unsorted data.

The Valley-Seeking automatic method uses a valley-seeking algorithm that it applies to inter-point distances to automatically determine the number of clusters and the cluster memberships. The T-Distribution E-M (T-Dist E-M) method is another variation of the Expectation Maximization algorithm that fits a mixture of T-Distributions instead of Gaussians to the point densities in feature space. This algorithm is also capable of adjusting the number of clusters as it runs (by removing unfavorable clusters) to arrive at the optimum number of clusters.

Offline Sorter provides many mechanisms for inspecting and manually adjusting the sorting results; you can add or remove waveforms from clusters or you can invalidate them completely.

With Offline Sorter you can also resort the waveforms by using any of the clustering methods. After you define the units, you can save the contours, templates and principal components to a TPL file and you can use them to sort waveforms in other data files.

For more information on the sorting algorithms, see [“Details of the Sorting Algorithms” on page 5-13](#).

**References:**

Lewicki, M.S., *A review of methods for spike sorting: the detection and classification of neural action potentials*, Network: Comput. Neural Syst., 9, R53-R78, 1998.

Schmidt, E.M., *Computer separation of multi-unit neuroelectric data: a review*, J. Neurosci. Meth., 12, 95-111, 1984.

Wheeler, B.C., *Automatic Discrimination of Single Units*, *Methods for Neural Ensemble Recordings*, ed. by M. Nicolelis, CRC Press, Boca Raton, 61-77, 1999.

Wheeler, B.C., and Heetderks, W.J., *A comparison of techniques for classification of multiple neural signals*, IEEE Trans. Biomed. Eng., 29, 752-759, 1982.

---

## Offline Sorter Outputs

Offline Sorter provides a variety of ways to view and save the results of spike-sorting. The Offline Sorter **File** menu includes the following actions.

### Printing Results

You can print the 2D and 3D Cluster Views, as well as the Waveform Display and the various grid-based views. The [“Sorting Summary View” on page 3-46](#) provides a convenient summary of the sorted units on a channel, and you can print or export it to a Microsoft PowerPoint\* or an Adobe PDF\* file. For more information, see [“Printing” on page 5-44](#).

### Saving Results

For PLX, Datawave, Bionic Technology NEV, and CED Spike-2 data files, Offline Sorter can save the unit assignments for each spike back into an output file of the same type. Offline Sorter preserves all records from the original file in the output file, except that it modifies the unit assignments for spike records to reflect the sorting that you have performed in Offline Sorter. For more information, see [“Saving Timestamps and Waveforms” on page 4-63](#).

#### Save As New PLX File

For any of the data formats that Offline Sorter can read, it can write the spike data out as a PLX file. With this feature, Offline Sorter can effectively translate any format data file into a Plexon PLX file. For more information, see [“Saving Timestamps and Waveforms” on page 4-63](#).

#### Save As NEX File

Offline Sorter can export the timestamps of sorted waveforms into a NEX file, for use with Neuroexplorer. For more information, see [“Saving Timestamps in a NEX File” on page 4-62](#).

### Exporting Per-Waveform Data

To accommodate later analyses of the data, Offline Sorter can write out selected information about each waveform into either a text or a MATLAB<sup>®</sup> file, or into an Excel<sup>®</sup> spreadsheet. For a given waveform, you can export any of the calculated feature values, along with the unit assignment, timestamp, raw waveform voltage values, etc. For more information, see [“Exporting Information” on page 4-65](#).

### Exporting Per-Channel Data

To accommodate later analyses of the data, Offline Sorter can write out selected information about each channel into either a text or a MATLAB file, or into an Excel spreadsheet. For a given channel, you can export any of the sorting quality statistics, and also the unit counts, sorting criteria, etc. For more information, see [“Exporting Information” on page 4-65](#).

### Exporting Per-Unit Data

To accommodate later analyses of the data, Offline Sorter can write out selected sorting information about each sorted unit into either a text or a MATLAB file, or into an Excel spreadsheet. For a given unit, you can export items like cluster center positions, cluster sigmas, ISI percentages, etc. For more information, see [“Exporting Information” on page 4-65](#).

### Reproducing Results

Offline Sorter provides several methods to apply or re-apply sorting criteria to data.

#### Export Sort Info to .TPL

You can write the sorting criterion used in sorting channels to a TPL file. The TPL file remembers information such as the features used for sorting, the sorting method, and any contours, templates, or parameters that control or affect the sorting. For more information, see [“Saving Templates” on page 5-33](#).

#### Sort Using .TPL file

Offline Sorter can read the TPL file and use the information to resort the current file. For more information, see [“Sort Using Saved Templates” on page 5-34](#).

#### Sort Using .NEX file

Offline Sorter can read a previously-saved .NEX file and apply the sorting embodied in the .NEX file to the currently loaded file. This only makes sense if the .NEX file was saved from the same file as is currently loaded. For more information, see [“Re-sorting from a NEX File” on page 4-63](#).

#### Update .EXP file with Sort Info

Offline Sorter can read a Plexon Sort Client experiment file (EXP file), and replaces the unit templates with the unit templates calculated from sorted units in the current file. For more information, see [“Save Templates to a Sort Client EXP File” on page 5-35](#).

#### Execute Batch Command File

Offline Sorter has a batch-processing facility that can perform sorting actions based on simple commands read from a text file. This facility is useful for automatically sorting large files, which can be time-consuming. For more information, see [“Batch File Processing” on page 5-39](#).

## Chapter 3

# Screen Elements and Displays

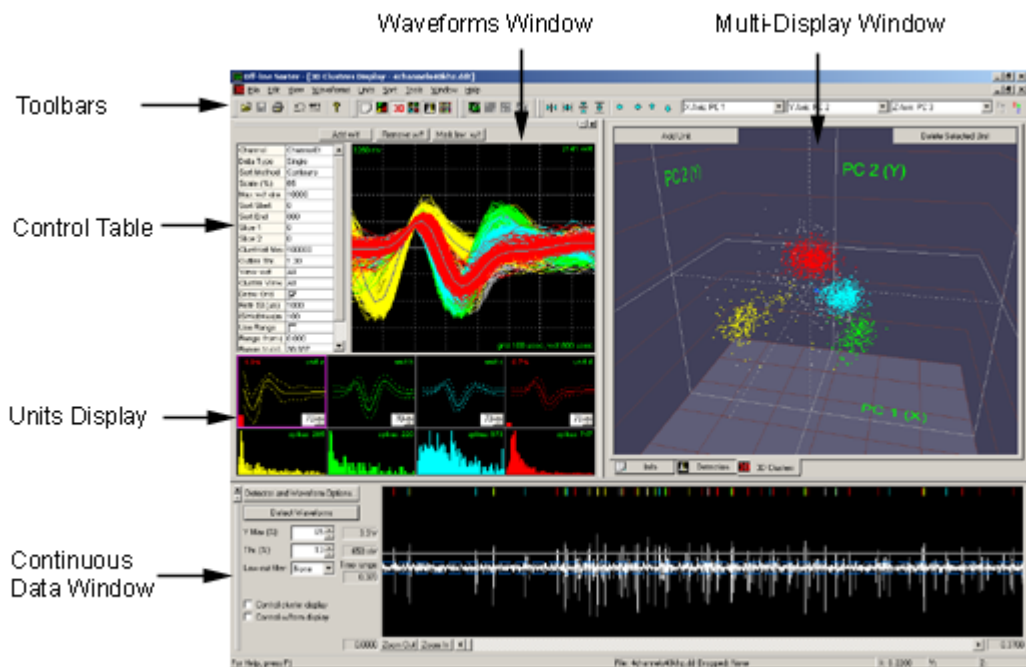
---

Screen Elements Overview .....	3-2
Control Table .....	3-5
Waveform Window .....	3-9
Units Display .....	3-11
Multi-Display Window .....	3-13
Continuous Data Window .....	3-54

---

## Screen Elements Overview

Offline Sorter contains the following screen elements.



### Toolbars

Frequently-used menu items have equivalent buttons on the toolbars. Move the mouse pointer over a toolbar button to open a tooltip window that describes the function of that button. You can move the toolbars and dock them to the top or bottom of the screen. You can rearrange and customize the contents of the toolbars by using the **Tools | Customize Toolbars** menu item.



#### HINT

#### Keyboard Bindings

You can conveniently and quickly perform common operations by binding a keyboard shortcut to any function available from the menus or toolbars. Use **Tools | Keyboard Shortcuts** menu item.

### Control Table

Use the Control Table to select a specific channel from an opened data file and to adjust sorting parameters and viewing options. See [“Control Table” on page 3-5](#).

### Waveform Window

The Waveform Window shows the waveforms for the selected channel, which are all superimposed. You can use the mouse to add and remove waveforms from units in this display. For more information, see [“Waveform Window” on page 3-9](#).

---

## Units Display

The Units Display shows the templates (average waveforms) for each sorted unit, plus the remaining unsorted waveforms. Click in this view for the fastest way to select the current unit. The graphs at the bottom of the view display the histograms of interspike intervals.

**Note:** The Control Table, the Waveform Window, and the Units Display are grouped together into the **Current Channel Window**, which can be hidden or shown using the **View | Current Channel Window** menu command. It can also be detached, floated, or docked to the other side of the main window frame.

## Continuous Data Window

There are two versions of the Continuous Data Window. The version shown earlier is visible if you open a file with continuously recorded signals. You can use the display to specify an optional digital filter, to select the threshold for waveform extraction, and to select individual waveforms for inspection. See [“Continuous Data Window” on page 3-54](#).

If you open a file that contains only spike waveforms, the Continuous Data Window shows the sorted and unsorted waveforms in their temporal sequence.



## Multi-Display Window

The Multi-Display Window can display various views. If you maximize the window, you can see each view by selecting tabs at the bottom of the display. When not maximized, you can change the views in the Multi-Display Window to be tiled, cascaded, or raised to the foreground by using menu items available in the **Window** menu. Also, each view has a check-marked option under the **View** menu that shows/hides it.

The **Info View** (the tab labeled **Info**) itself contains sub-tabs at the bottom of the view to select different panes of information.

- The **Channels** sub-tab contains counts of waveforms and units for each channel.
- The **Chan Params** sub-tab contains channel-by-channel information, such as thresholds, gains, and filtering.
- The **Chan Map** sub-tab shows the association of spike and continuous channels, and it allows changing channel types and the data type.

- The **Sort Details** sub-tab shows the details about the sorting performed on each channel, like what features and sorting methods were used.
- The **Global** sub-tab contains the global waveform and sorting parameters and file information.
- The **Events** sub-tab lists the external events present in the data file. It allows events to be displayed with custom colors on the continuous data views.
- The **Stats** sub-tab contains multi-variate ANOVA tests on the separation of unit clusters and other sorting quality-related statistical information.
- The **PCA** sub-tab displays the eigenvectors and values of the principal component analysis on the waveforms that is used to generate the principal component displays.
- The **Templates** sub-tab displays the mean template for each sorted unit.
- The **Scan** sub-tab is an exploration tool that can automatically resort a channel numerous times while stepping the values of parameters. This allows one to get a feeling for the optimal values of the sorting parameters.



## HINT

### Keystroke Binding

You can select the individual sub-tabs under the **Info** tab on the **View** - **Info View Tabs** sub-menu. You can bind keystrokes to these sub-menu items that quickly take you to a sub-tab with a single keypress.

The **2D Cluster View** shows the waveforms as points in 2D feature space. You can select each axis from any of the [“Features Available for Sorting” on page 5-9](#). You can use this display to select the clusters of waveforms. For more information, see [“Cluster Views” on page 3-19](#).

The **3D Cluster View** (shown in the screenshot above) shows the waveforms as points in 3D feature space. You can select each axis from any of the [“Features Available for Sorting” on page 5-9](#). You can use this display to select the clusters of waveforms. For more information, see [“Cluster Views” on page 3-19](#).

The **Waveform Inspection View** compares individual waveforms selected in the Continuous Data Window with the average waveform (template) of each sorted unit. For more information, see [“Waveform Inspection View” on page 3-37](#).

The **Waveform Detection View** is only available for files with continuously recorded signals (e.g. DDT files). It displays histograms of detected peak heights that enables you to set a spike detection threshold. For more information, see [“Waveform Detection View” on page 3-38](#).

The **Feature vs. Feature View** displays miniature 2D Cluster view windows for various combinations of features. You can use it to quickly identify features that are the most useful in the sorting of units. For more information, see [“Feature vs. Feature View” on page 3-42](#).



The **Sorting Summary View** displays a concise, printable summary of the sorting on a channel. You can also export the Sorting Summary View to Power-Point\*. For more information, see [“Sorting Summary View” on page 3-46](#).

The **PCA Results View** graphically displays the information that can be found on the **PCA** sub-tab of the **Info View**. You can plot each of the PCA eigenvectors, along with a bar graph of the eigenvalue magnitude. For more information, see [“PCA Results View” on page 3-50](#).

The **3D Phase Space View** shows the waveforms plotted in 3D phase space. This view provides you with an interesting new way to visualize the waveforms. For more information, see [“3D Phase Space View” on page 3-35](#).

The **ISI Histograms View** shows the same ISI Histograms that are plotted in the Unit Display (see below), but you can print the view and it offers several other display options, such as log scaling of the axes. For more information, see [“ISI Histograms View” on page 3-41](#).

The **Cross-Correlograms View** shows a matrix of cross-correlation plots between all of the units in the current channel. For more information, see [“Cross-Correlograms View” on page 3-51](#).

The **Rasters View** shows Raster Plots of the spike and event times for all channels, and the continuous data for the selected channel. For more information, see [“Rasters View” on page 3-52](#).

## Control Table

Channel	None
Data Type	Single
Sort Method	Contours
Zoom	1.0
Max w/f draw	100000
Sort Start	0
Sort End	800
Slice 1 (μs)	0
Slice 2 (μs)	0
Slice 3 (μs)	0
Slice 4 (μs)	0
Outlier Thr.	1.30
Parzen Mult.	1.00
View w/f	All
Cluster View	As above
Contin. View	As above
Draw Grid	<input type="checkbox"/>
Draw Thr.	<input checked="" type="checkbox"/>
Refr ISI (μs)	1000
ISIHistMax(ms)	100
Use Range	<input type="checkbox"/>
Range from (s)	0.000
Range to (s)	0.000
Invalid w/f	Hide

Use the Control Table to set the following parameters:

**Channel** – is the currently active channel. Click to show a drop-list of all the available channels. You can also select the channel by double-clicking it in the Info View. For more information, see [“Info View” on page 3-13](#).

**Data Type** – is the electrode recording mode. **Single** indicates that each channel is viewed separately. **Stereotrode** indicates that channels are arranged in groups of two. **Tetrode** indicates that channels are arranged in groups of four. **Data Type** automatically selects the appropriate mode for Plexon PLX, Datawave UFF, and Neuralynx files. If you change **Data Type**, you force Offline Sorter to interpret the file as if it contained the selected data type. For more information, see [“Stereotrode and Tetrode Data” on page 5-4](#).

**Sort Method** – is the cluster selection method. Offline Sorter provides the following methods:

- Contours ([page 4-29](#))
- W/f Crossing ([page 4-36](#))
- Templates ([page 4-41](#))
- K-Means ([page 4-42](#))
- Standard E-M ([page 4-45](#))
- Valley-Seeking ([page 4-46](#))
- T-Distribution E-M ([page 4-49](#))

When you select one of these options, you set the current sorting mode, but you do not actually initiate a sorting operation.

**Zoom** – is the vertical zoom factor in the Waveform Window and the Continuous Data Windows. You can use the up or down arrows to increase or decrease the factor by steps, but you can also enter any floating point value by typing the numbers. By default, the up-down arrows increase or decrease the factor by increments of 0.1 when the current zoom factor is small, and by increments of 1.0 when the zoom factor is higher. You can adjust the value at which the 1.0 increment applies by typing the value in the **Zoom control steps by 1.0 starting at value:** box on the **Display** tab in the **Tools | Options** window.

**Max w/f draw** – Maximum-waveform-draw is the maximum number of waveforms drawn in the Waveform Window and the Cluster views. For example, if there are 30000 waveforms in the active channel, and **Max w/f draw** is set to 10000, then only every 3rd waveform is drawn. Change the **Max w/f draw** setting from the large default value only to improve drawing speed on slower computers when you are viewing channels with many waveforms. When the number of waveforms in a channel is greater than the **Max w/f draw** setting, the Waveform Window shows the waveform count section highlighted in red, and there is

---

a divided-by indicator that shows the fraction of waveforms that are being displayed; e.g., for the preceding example, the divided-by indicator is 3.



### CAUTION

#### Misleading Sort

It can be misleading to sort units on a channel where **Max w/f draw** is less than the number of waveforms. If Offline Sorter takes too long to redraw the display, it is generally safer for you to use the deferred drawing feature. See [“Deferring Waveform Drawing” on page 5-46](#). In general, if the waveform count on a channel exceeds the **Max w/f draw** setting, do not select the **Use only currently visible waveforms** checkbox in the **Tools | Options General** tab; that selection can result in valid waveforms being unassigned when you select units by using the waveform crossing or contour sorting methods.

**Sort Start** and **Sort End** – are the lower-bound and upper-bound of the time interval (See [“Adjusting the Sort Range” on page 4-25](#).) of the waveform that Offline Sorter uses to calculate principal components and feature values and in [“Details of the Sorting Algorithms” on page 5-13](#). The default value for Sort Start is zero. The default value for the Sort End is the total waveform length. Offline Sorter only uses the time points of the waveforms between Sort Start and Sort End in the template sorting algorithm, in the recalculation of the principal component scores, and in the calculation of feature vectors. For options in principal component calculations, see [“Adjusting the Sort Range” on page 4-25](#).

**Slice 1, Slice 2, Slice 3, Slice 4** – are the time points in the waveform (slices in time) that Offline Sorter uses to calculate the Slice feature. A vertical line in the Waveform View represents the Slice time. You can view the voltage of the waveforms at the Slice time point on any axis of the Cluster views and you can use it in the sorting algorithms.

**Outlier Thr.** – Outlier Threshold can be used for several purposes. The value in this field is in units of standard deviations (sigmas). When Offline Sorter renders ellipses in the cluster views, this field controls where Offline Sorter draws the lines or surfaces. For the [Semi-Automatic Clustering using K-Means](#) sorting algorithm, and for the **Tools | Remove Outliers** menu selection, this field controls the removal of outliers. The K-Means clustering algorithm assigns every waveform to a cluster. To eliminate the outliers, Offline Sorter calculates the standard deviation  $\sigma$  of the distance from all points in the cluster to its center, for each cluster. Then, if the distance from a given point to the cluster center is more than  $\sigma \times \text{Outlier\_Threshold}$ , Offline Sorter removes the point from the cluster.

**Parzen Mult., D.O.F. Mult., or Beta** – Only one of these values appears, which depends upon the sorting method you have currently selected. All the values control parameters in the sorting algorithms, as follows. The [Automatic Clustering using Valley-Seeking Method](#) algorithm works by counting neighbors that are within the Parzen radius times this **Parzen Mult.** multiplier around the point in feature space. If you decrease the value, it generally results in more numerous,

smaller clusters. The [Automatic Clustering using T-Distribution E-M Method](#) has a setting for the number of degrees of freedom (DOF) for the fitting procedures. Use of the theoretically-predicted value of this parameter results in poor sorting. The **D.O.F Mult.** value multiplies the theoretical DOF setting. Smaller values of this multiplier generally result in slow convergence to a larger number of clusters, larger values result in rapid convergence to a smaller number of clusters. The [Semi-Automatic Clustering using Standard E-M](#) algorithm depends upon an “annealing rate” parameter **Beta**. When **Beta** is 1, a completely standard E-M algorithm runs. When the value is other than one, a variant of the standard E-M algorithm runs.

**View w/f** – View Waveforms specifies which waveforms Offline Sorter draws in the Waveform Window (either **All** waveforms, **Selected Unit(s)**, **Unsorted**, or only the **Selected Waveform**). **Selected Unit(s)** enables you to make a quick selection of units by clicking in the Units Display.

**Cluster View** – specifies which waveforms appear in the Cluster views (either **All** waveforms, or **Same as above**—which displays the waveforms specified in the **View w/f** parameter).

**Contin. View** – Continuous View specifies which waveforms are shown in the Continuous Data View (either **All** waveforms, or **Same as above**—which displays the waveforms specified in the **View w/f** parameter).

**Draw Grid** – draws a grid in the Waveform Window.

**Draw Thr.** – Draw Threshold draws a horizontal dashed white line in the Waveform Window to show the position of the threshold for the channel. For more information, see [“Re-thresholding Spike Data”](#) on page 5-38.

**Refractory Interspike Interval** – is the minimum valid interspike interval. If there are intervals smaller than the refractory interval, Offline Sorter displays a red bar to the left of the template display in the Units Display that corresponds to the proportion of spikes with ISIs less than the refractory interval (See [“Units Display”](#) on page 3-11.). Lines between spikes with ISIs less than the refractory interval are also drawn in the Cluster views when that option is enabled. The ISI is specified in microseconds. For more information, see [“Cluster Views”](#) on page 3-19.

**ISI Hist Max** – Interspike Interval (ISI) Histogram Maximum is the maximum value of the X axis in the Interspike Interval Histogram in the [Units Display](#), in milliseconds.

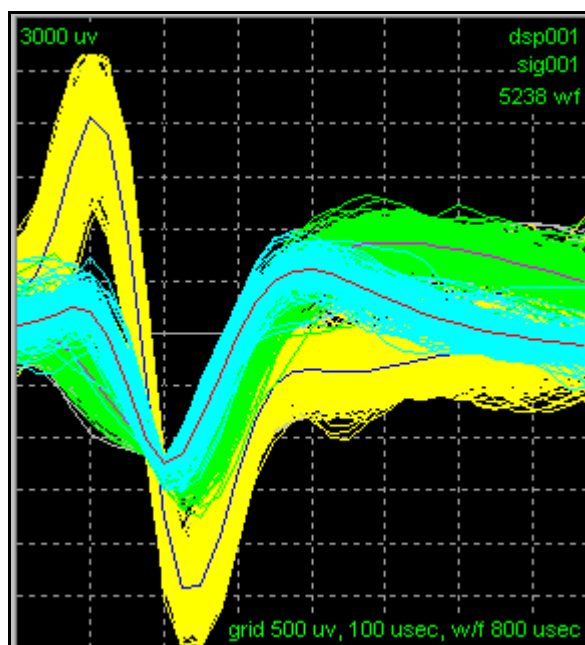
**Use Range** – is an option that displays the waveforms for the specified time range (**Range From** and **Range To** parameters that follow). By default, the program displays all the waveforms from a selected channel. Use this option to display only a subset of waveforms based on a time range. You can also change the Time Range by using the **Waveforms | Select Time Range** menu command.

---

**Invalid w/f** – Invalid Waveforms is an option to **Show** or **Hide** the waveforms marked as invalid in the Waveform and Cluster views. Offline Sorter retains this value between invocations.

## Waveform Window

The Waveform Window shows the superimposed waveforms aligned, by default, at the threshold crossing times. Offline Sorter draws the waveforms assigned to units in the colors associated with those units. Offline Sorter also draws the template for each unit in a contrasting color.



You can align waveforms (See [“Waveform Alignment” on page 4-22.](#)) at their minimum or maximum values by using the **Tools | Align Waveforms** or the **Waveforms | Align** menu commands.

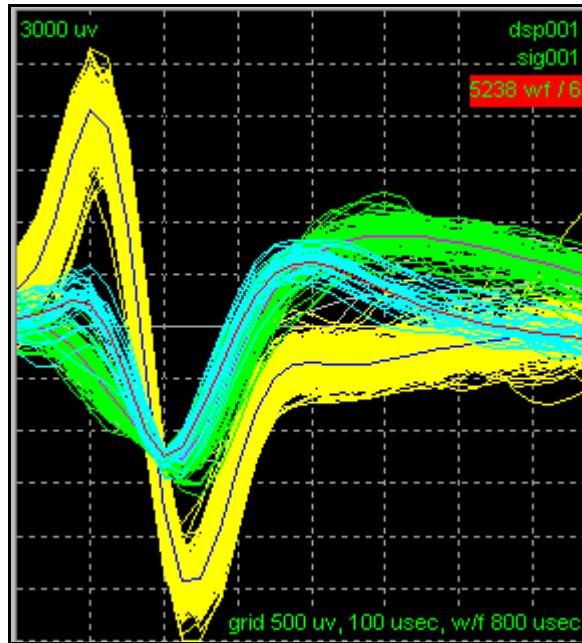
You can create unit clusters in this view by using the mouse pointer to cross waveform bundles; see [“Manual Clustering using the Waveform Crossing Method” on page 4-36.](#) You can also remove specific waveforms from a unit (i.e. “change them to unsorted”) in the Waveform View. See [“Cleaning the Selection” on page 4-54.](#)

To control the temporal range and type of waveforms shown in the Waveform Window, you can use the scroll bar at the bottom of the [Continuous Data Window](#) (when the **Control waveform display** check box is checked) and the **View w/f** parameter in the [Control Table](#).

The **Draw Grid** option in the [Control Table](#) determines if a reference grid appears in the Waveform Window.

**Note:** The Waveform Window can be **refreshed/redrawn** by **right-clicking** in the window.

If the number of waveforms in a channel exceeds the **Max w/f draw** value in the [Control Table](#), the Waveform Window and the Cluster Displays draw only every  $N$ th waveform, where the divisor  $N$  is calculated to bring the number of drawn waveforms under the **Max w/f draw** limit. This ensures that drawing occurs with an acceptable speed on slower computer systems. When this “decimation” is in effect, Offline Sorter highlights the wf count text in red.



In the preceding screen shot, the **Max w/f draw** is set to 1000. The divisor  $N$  in this case is 6 as shown in the red highlighted section, which indicates that only every 6th waveform is being drawn.

When the **Control waveform display** checkbox in the Continuous Data Window is checked, decimation is disabled and all waveforms in the selected time interval are displayed.



### CAUTION

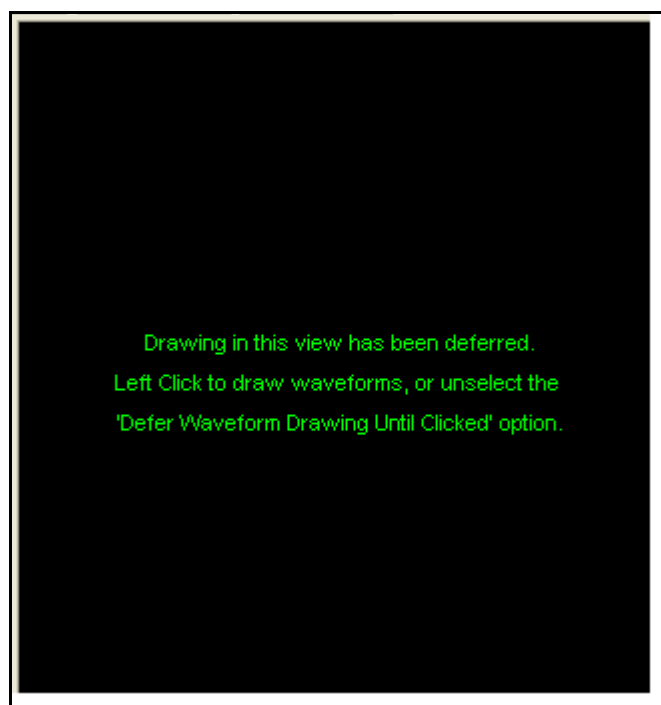
#### Sorting During Decimation

If you perform sorting while decimation is occurring, it can result in lost information. If possible, increase the **Max w/f draw** setting so that decimation does not occur. The **Use only currently visible waveforms** setting in the **Tools | Options | General** menu dialog controls whether waveforms that are not being drawn are included in waveform crossing and contour sorting.

As an alternative to using the **Max w/f draw** feature to speed up the display when viewing large files, you can configure the Waveform Window to only redraw “on

---

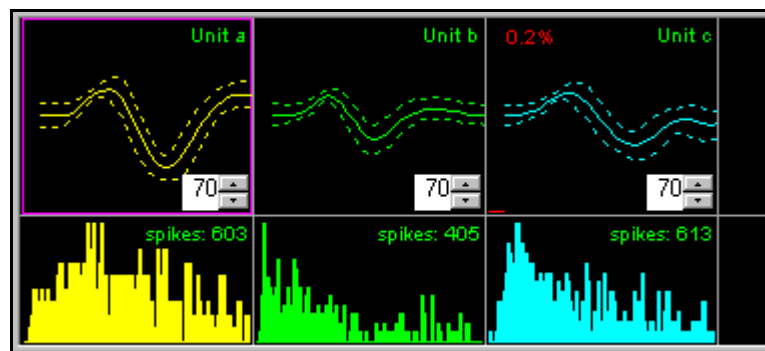
demand” by using the **Defer Waveform Drawing Until Clicked** checkbox in the **Tools | Options | Display** dialog box. When you check this item, the waveform window displays the following view instead of drawing the waveforms.



Left-clicking in the Waveform Window causes all the waveforms to be redrawn. After that, any operation that would normally cause the Waveform Window to be drawn results in the above text instead. The **Defer Waveform Drawing Until Clicked** option also affects drawing in the Continuous Data Views.

## Units Display

Offline sorter provides information about each of the currently-defined units in the Unit Display along with a display for the waveforms that are currently unsorted.



The top row of graphs of this display shows the template (average waveform) for the unit (solid line) as well as  $\pm 3$  standard deviations from the template (dashed



lines). You can use the edit controls with the spin buttons to adjust the fit tolerances for Template sorting.

You can set the default template Fit Tolerance (70 in the preceding screenshot) by using the **Sort** tab on the **Tools | Options** menu panel. After setting a new value for the Fit Tolerance, you can elect to apply the new value to existing units, or to only use the new value when new units are defined. For additional information on templates, see [“Template Algorithm” on page 5-14](#).

The bottom row of graphs of this display shows the **Interspike Interval Histograms**. The X axis of each histogram runs from zero to **ISIHistMax**. The graphs also show the total number of spikes for each unit. You can display larger printable versions of these Interspike Interval Histograms in the [“ISI Histograms View” on page 3-41](#).

The Unit Display shows the percentage of spikes with Interspike Intervals less than the **Refractory Interspike Interval** in **red**, which also includes a **red bar** in the upper portion of the display. You can change the value of **ISIHistMax** and **Refractory Interspike Interval** in the [Control Table](#). You can also set the **Refractory Interspike Interval** value by using the **Refractory Period** tab on the **Tools | Options** menu command.

## Selecting Units

You can select units in the Units Display by left-clicking on the corresponding template. Offline Sorter draws a solid highlighted box and a colored background on the box corresponding to the currently selected unit. For purposes of selecting units, the Unsorted “unit” behaves just like the actual sorted units. When the **View w/f** setting in the Control Table is **Sel.Unit(s)**, selecting units by clicking in the Units Display is a convenient means to quickly see the differences and similarities between units.

Offline Sorter uses the concept of selecting a single primary unit and multiple secondary units. The primary unit is the unit that is affected by operations such as **Delete Selected Unit**, **Add w/f**, and **Remove w/f**. Offline Sorter shows the multiple secondary units with a dashed box highlighting the unit template, but with the default background color. When the **View w/f** setting in the Control Table is set to **Sel.Units(s)**, Offline Sorter also displays the waveforms for the multiple selections in the Waveform Window, Cluster views, and Continuous Data View. This is a convenient way to compare and contrast units. Also, certain operations like **Units | Combine Units** and **Swap Units** use multiple unit selections.

To select multiple units, follow the Windows standard for selecting multiple items:

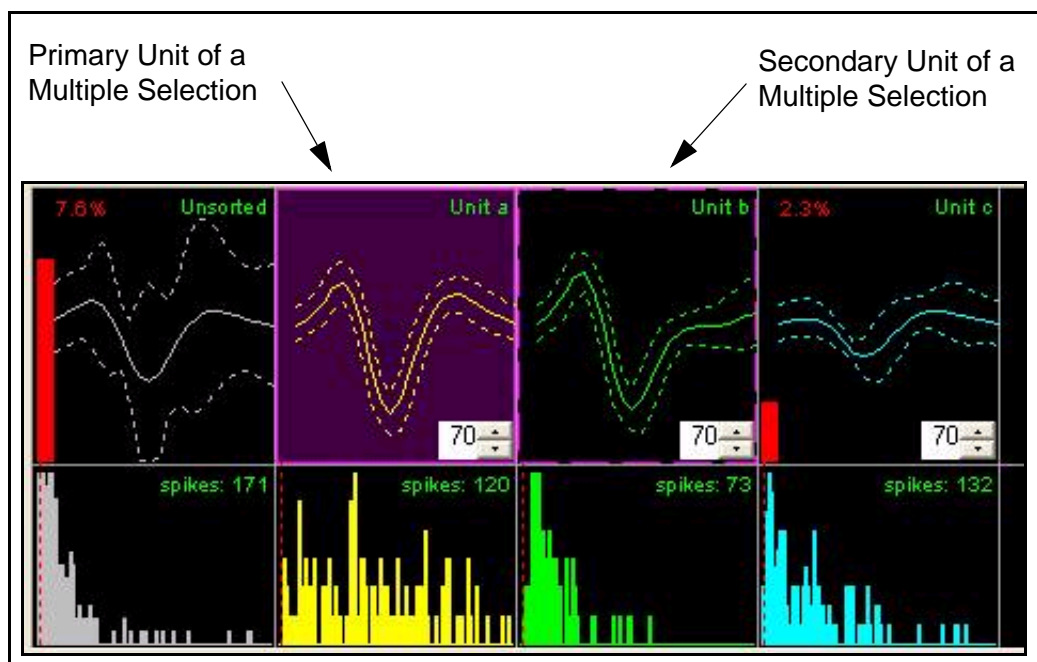
- To select a unit, click the unit.
- To select non-adjacent units, press and hold down **CTRL**, and then click each unit.



- To select adjacent units, click the first unit, press and hold down **SHIFT**, and then click the last unit.
- To cancel an existing multiple selection, click any unit.

You can also select units by using the **Units | Select Unit(s)** dialog box.

To set the colors assigned to each of the sorted units, select the **Colors** tab of the **Tools | Options** dialog box.




## Multi-Display Window

### Info View

The **Info** view displays information in a spread sheet format. It has several sub-tabs to show different kinds of information.

To display the Info view

- Click the Info toolbar button , or,

select the **View | File Info View** menu command.

	Name	Type	Count	Unsorted	Unit a	Unit b	Unit c	Unit d	Unit e	Unit f	Unit g	Unit h
1	sig001	Spike	214730	64388	128480	2005	19857	-	-	-	-	-
2	sig002	Spike	2673	11	2662	-	-	-	-	-	-	-
3	sig003	Spike	14362	2966	11396	-	-	-	-	-	-	-
4	sig004	Spike	7984	6	7978	-	-	-	-	-	-	-
5	sig005	Spike	16304	1415	8173	4435	1306	975	-	-	-	-
6	sig006	Spike	8219	5	8214	-	-	-	-	-	-	-
7	sig007	Spike	41351	11987	17853	9060	2451	-	-	-	-	-
8	sig008	Spike	2918	3	2915	-	-	-	-	-	-	-
9	sig009	Spike	1387	5	1382	-	-	-	-	-	-	-
10	sig010	Spike	14475	3104	5130	6241	-	-	-	-	-	-
11	sig011	Spike	6147	7	6140	-	-	-	-	-	-	-
12	sig012	Spike	2640	6	2634	-	-	-	-	-	-	-
>>13	sig013	Spike	8420	4	4097	3968	351	-	-	-	-	-
14	sig014	Spike	2454	5	2449	-	-	-	-	-	-	-
15	sig015	Spike	2669	6	2663	-	-	-	-	-	-	-
16	sig016	Spike	866	5	861	-	-	-	-	-	-	-
17	AD01	Slow Cont	3591121	0	-	-	-	-	-	-	-	-
18	AD17	Slow Cont	3591121	0	-	-	-	-	-	-	-	-
19	AD18	Slow Cont	3591121	0	-	-	-	-	-	-	-	-

Channels Chan Params Sort Details Global Events Stats PCA Templates Scan Chan Map

Info

The Info view contains several spreadsheet pages: **Channels**, **Chan Params**, **Sort Details**, **Global**, **Events**, **Stats**, **PCA**, **Templates**, **Scan**, and **Chan Map**.

### Channels Tab

The **Channels** tab is initially displayed when a file is first opened.



#### HINT

#### Viewing Individual Channels

You can select individual channels for viewing or sorting, or both, by double-clicking the channel # on the far left-hand side of the page.

The **Channels** tab offers a quick overview of the contents and the sorting status of a file.

The **Channels** tab contains a row for each channel in the file. The columns that display in the **Channels** tab depend upon the file type that is open.

For data files with timestamped spikes (e.g. PLX files, as pictured earlier), the **Channels** tab contains the following columns:

- **Name** – refers to the channel name. You can rename channels by clicking in this column and typing a new name.
- **Type** – includes either **Spikes** for time-stamped spike data, or **Slow Cont** or **Spike Cont** for continuously digitized data.
- **Count** – refers to the number of spikes for spike channels, or the number of data points for continuous channels.
- **Invalidated** – shows the number of invalidated waveforms in the channel.

**Note:** By default, this column is very narrow and it can appear to be blank. To see the column information, you must expand the column by dragging the dividing lines on the header bar for the column.

- **Unsorted** – shows the number of spikes that have not been assigned to a unit.
- **Unit a....z** – shows the number of spikes assigned to each unit.

For data files with continuously recorded signals (e.g. DDT files), the **Info** view also displays a count of data points in the channel.

## Chan Params Tab

The **Chan Params** (channel parameters) tab contains more detailed, per-channel information about the settings in effect when the data was acquired.

	Name	Type	SIG ch	Freq	Gain	Filter	Threshold	Zoom
1	sig001	Spike	1	40000	32000	On	-262 (-11.99 $\mu$ V)	1.0
2	sig002	Spike	2	40000	32000	Off	-221 (-10.12 $\mu$ V)	1.0
3	sig003	Spike	3	40000	1000	Off	-7 (-10.25 $\mu$ V)	3.0
4	sig004	Spike	4	40000	1000	Off	-7 (-10.25 $\mu$ V)	3.0
5	sig005	Spike	5	40000	1000	Off	-7 (-10.25 $\mu$ V)	4.0
6	sig006	Spike	6	40000	1000	Off	-7 (-10.25 $\mu$ V)	3.0
7	sig007	Spike	7	40000	1000	Off	-6 (-8.789 $\mu$ V)	4.0
8	sig008	Spike	8	40000	1000	Off	-8 (-11.72 $\mu$ V)	5.0
9	sig009	Spike	9	40000	1000	Off	-7 (-10.25 $\mu$ V)	5.0
10	sig010	Spike	10	40000	1000	Off	-6 (-8.789 $\mu$ V)	5.0
11	sig011	Spike	11	40000	1000	Off	-6 (-8.789 $\mu$ V)	5.0
12	sig012	Spike	12	40000	1000	Off	-7 (-10.25 $\mu$ V)	5.0
>>13	sig013	Spike	13	40000	1000	Off	-6 (-8.789 $\mu$ V)	5.0
14	sig014	Spike	14	40000	1000	Off	-7 (-10.25 $\mu$ V)	5.0
15	sig015	Spike	15	40000	1000	Off	-7 (-10.25 $\mu$ V)	5.0
16	sig016	Spike	16	40000	1000	Off	-7 (-10.25 $\mu$ V)	5.0
17	AD01	Slow Cont	1	1000	10000	N/A	0 (0 $\mu$ V)	33.0
18	AD17	Slow Cont	17	1000	1000	N/A	0 (0 $\mu$ V)	1.0
19	AD18	Slow Cont	18	1000	1000	N/A	0 (0 $\mu$ V)	1.0

Channels
**Chan Params**
Sort Details
Global
Events
Stats
PCA
Tel

Info

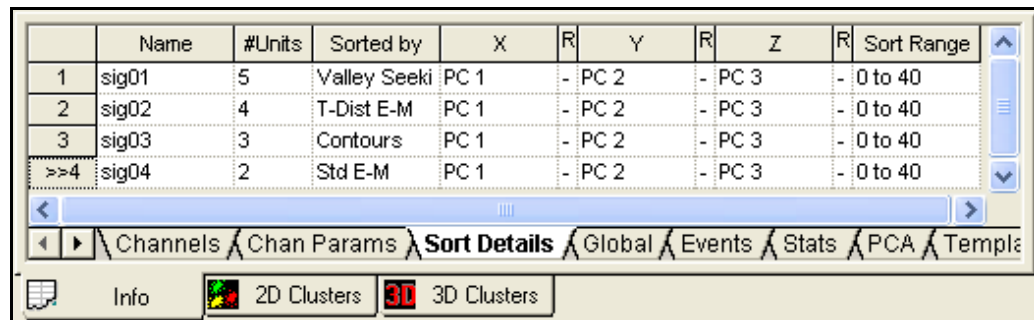
The **Chan Params** tab can contain different items depending on the file type and the data in the file.

- **Name** – refers to the channel name. You can rename channels by clicking in this column and typing a new name.
- **Type** – includes either **Spikes** for timestamped spike data, or **Slow Cont** or **Spike Cont** for continuously digitized data.
- **SIG Chan** – is for PLX files only and it includes the channel number on the SIG board that is associated with this channel.
- **Filter** – is for PLX files only and it indicates if the filter is switched on for the channel.
- **Gain** – indicates the overall gain setting for this channel. For Plexon DDT files, you can change the gain by typing a new value in this column.

- **Threshold** – shows the threshold value in A/D counts and the associated voltage that was used to collect the waveforms. For continuously-recorded data files, you can change the threshold for the currently selected channel by typing a new value.
- **Zoom** – refers to the scaling factor that Offline Sorter uses to display the waveforms in the Waveform Window and cluster display. You can change the scaling factor for the currently selected channel by typing a new value.

### Sort Details Tab

The **Sort Details** tab shows information about the channels that have been sorted during this session of Offline Sorter.



	Name	#Units	Sorted by	X	R	Y	R	Z	R	Sort Range
1	sig01	5	Valley Seeki	PC 1	-	PC 2	-	PC 3	-	0 to 40
2	sig02	4	T-Dist E-M	PC 1	-	PC 2	-	PC 3	-	0 to 40
3	sig03	3	Contours	PC 1	-	PC 2	-	PC 3	-	0 to 40
>>4	sig04	2	Std E-M	PC 1	-	PC 2	-	PC 3	-	0 to 40

Channels / Chan Params / **Sort Details** / Global / Events / Stats / PCA / Templates

Info 2D Clusters 3D Clusters

The **X**, **Y**, and **Z** columns show which feature is selected for each dimension of feature space.

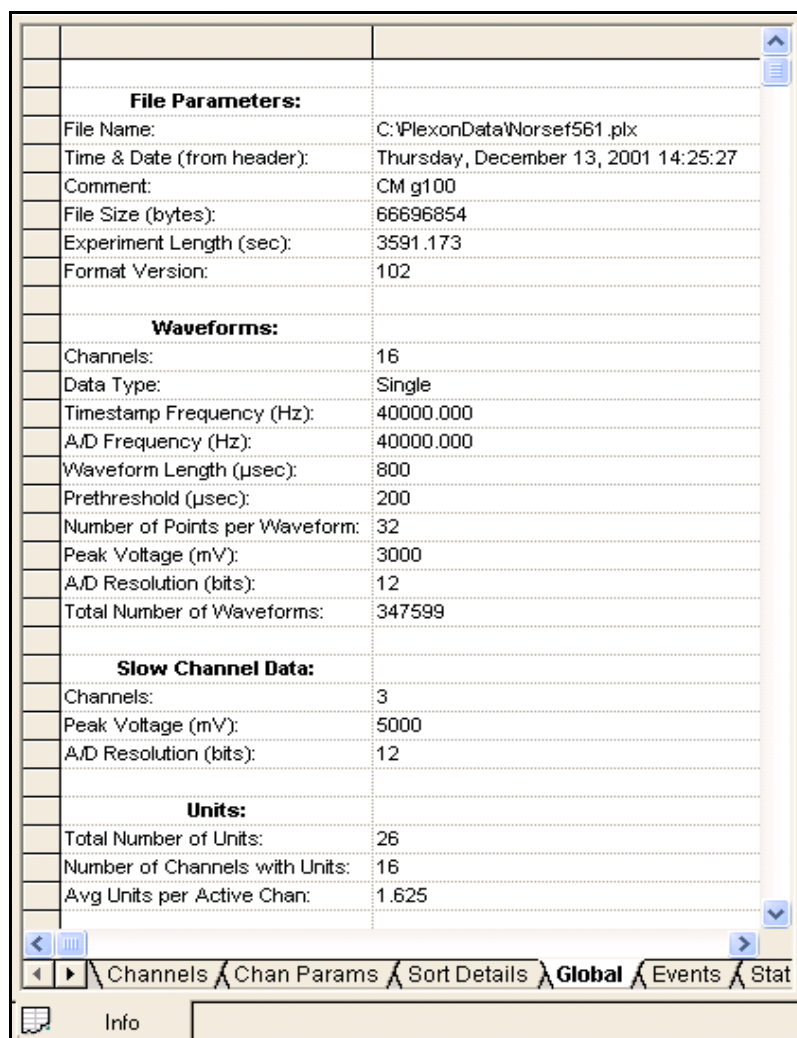
The **R** columns show the range covered by the feature values - you can expand the column width by dragging the column header to see these values.

The **Sort Range** column indicates how much of the waveform width (in clock ticks) was used in the sorting.

The contents of this view is essentially what is written out to a TPL file. For more information, see [“Saving Templates” on page 5-33](#).

## Global Tab

The **Global** tab contains various items of interest for the currently-loaded file, in a two-column spreadsheet format.



<b>File Parameters:</b>	
File Name:	C:\PlexonData\Norsef561.plx
Time & Date (from header):	Thursday, December 13, 2001 14:25:27
Comment:	CM g100
File Size (bytes):	66696854
Experiment Length (sec):	3591.173
Format Version:	102
<b>Waveforms:</b>	
Channels:	16
Data Type:	Single
Timestamp Frequency (Hz):	40000.000
A/D Frequency (Hz):	40000.000
Waveform Length (µsec):	800
Prethreshold (µsec):	200
Number of Points per Waveform:	32
Peak Voltage (mV):	3000
A/D Resolution (bits):	12
Total Number of Waveforms:	347599
<b>Slow Channel Data:</b>	
Channels:	3
Peak Voltage (mV):	5000
A/D Resolution (bits):	12
<b>Units:</b>	
Total Number of Units:	26
Number of Channels with Units:	16
Avg Units per Active Chan:	1.625

Navigation: < > Channels Chan Params Sort Details **Global** Events Stat

Info

### File Parameters:

- **File Name** – shows the full path to the currently-loaded file on disk.
- **Time & Date (from header)** – shows the time and date that the data was collected, as the data acquisition software wrote the information into the file header. If the time and date is unavailable from the file header, the time & date of the file itself appears.
- **Comment** – lists the comment field in the file header, if present.
- **File Size (bytes)** – indicates the size of the currently-loaded file in bytes.
- **Experiment Length (sec)** – is the total amount of data in seconds that is in the current file.

- **Format Version** – shows (for some file formats) the version of the file format currently loaded.

#### Waveforms:

- **Channels** – shows the total number of spike channels in the file.
- **Data Type** – indicates whether the data type is Single, Stereotrode, or Tet-rode.
- **Timestamp Frequency (Hz)** – shows the frequency in Hz of the timestamp clock.
- **A/D Frequency (Hz)** – shows the frequency in Hz of the analog-to-digital (A/D) conversion while the data was collected.
- **Waveform Length (μsec)** – indicates the number of clock ticks (at the analog-to-digital frequency) of data for each spike waveform.
- **Prethreshold (μsec)** – indicates the number of clock ticks collected prior to the triggering event (threshold crossing).
- **Number of Points per Waveform** – indicates the number of samples taken per spike waveform (waveform length \* A/D frequency).
- **Peak Voltage (mV)** – shows the full-scale voltage to produce the maximum analog-to-digital converter count for spike channels.
- **A/D Resolution (bits)** – is the resolution of the analog-to-digital (A/D) converter for the spike channels.
- **Total Number of Waveforms** – indicates the total spikes (waveforms) in the file.

#### Slow Channel Data:

- **Channels** – indicates the total number of slow continuous channels in the file.
- **A/D Frequency (Hz)** – shows the frequency at which the slow channels have been digitized (this can vary from channel to channel).
- **Peak Voltage (mV)** – shows the full-scale voltage to produce the maximum analog-to-digital converter count for the slow channels.
- **A/D Resolution (bits)** – is the resolution of the analog-to-digital converter for the slow channels.

#### Units:

- **Total Number of Units** – the total number of units on all channels for the current file.
- **Number of Channels with Units** – indicates the number of channels that have >0 units.

- 
- **Avg Units per Active Chan** – is the value in **Total Number of Units** divided by the value in **Number of Channels with Units**.

### Events Tab

The **Events** tab displays the events that are present in the data file. You can select individual events to show in the Continuous Data views and the Rasters views. For more information, see [“Examining Events” on page 5-23](#).

### Stats Tab

The **Stats** tab contains the results of statistical tests on the separation of selected unit clusters in 2D and 3D feature space for the current channel. The page is blank if a channel is not loaded. For more information, see [“Info View Statistics” on page 4-59](#).

### PCA Tab

The **PCA** tab displays the eigenvectors and eigenvalues for the waveforms in the current channel. The first 8 eigenvalues from the PCA appear at the top of the display. The eigenvector corresponding to each of those eigenvalues appears in a column directly underneath the eigenvalue. This is a grid-based view that numerically shows the same information displayed in graphical form in the [“PCA Results View” on page 3-50](#).

### Templates Tab

The **Templates** tab displays a numerical representation of the template for each unit. You can copy and paste these templates into other analysis programs.

### Scan Tab

The **Scan** tab is a data exploration tool that enables you to do multiple automatic and semi-automatic sortings on a channel while stepping the value of a sorting parameter over a range of values. This enables you to get a feeling for how the sorting parameters affect the sorting quality. For more information, see [“Parameter Scans” on page 5-36](#).

### Chan Map Tab

The **Chan Map** tab displays information about the associations between channels. For each channel, the tab displays the Name, Channel Type, and any associated channels (See [“Channel Types and Associated Channels” on page 5-2](#)). You can also use the Chan Map view to change the Data Type, the Channel Type, or to associate or disassociate channels.

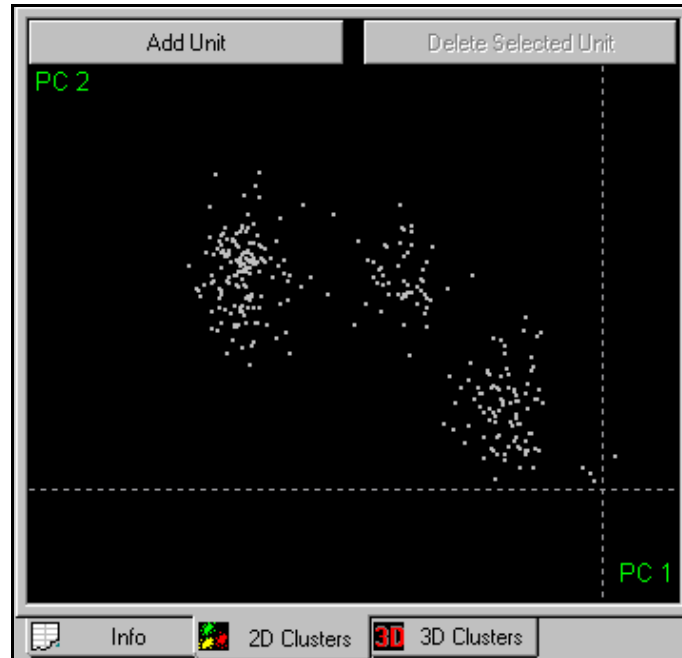
## Cluster Views

Although they are treated as separate views in Offline Sorter, the 2D Cluster view and the 3D Cluster view share many features in common. Both cluster views show the waveforms in feature space. The 2D Cluster view shows the waveforms as points in a 2-dimensional feature space, and the 3D Cluster view displays

points in a 3-dimensional feature space. Both the 2D and 3D Cluster views can be active at the same time.

To display the 2D Cluster View

- Click the 2D Clusters toolbar button  , or, select the **View | 2D Cluster View** menu command.

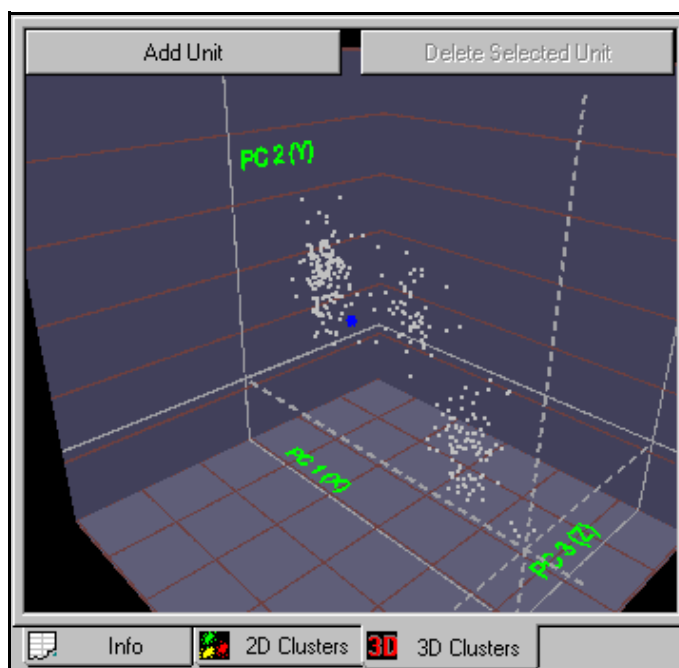




---

To display the 3D Cluster View

- Press 3D Clusters toolbar button , or, select the **View | 3D Cluster View** menu command.



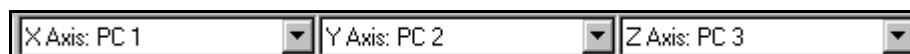
When you initially open a channel, you can create either the 2D Cluster view, the 3D Cluster view, or both. You can control this via the **Initial Display** radio buttons under the **Display** tab of the **Tools | Options** dialog box.

Features common to both the 2D and 3D views are discussed in the [Cluster View Common Functionality](#) section that follows. Features specific to either view are discussed in “[2D Cluster View Details](#)” on page 3-28 and in “[3D Cluster View Details](#)” on page 3-31.

### Cluster View Common Functionality

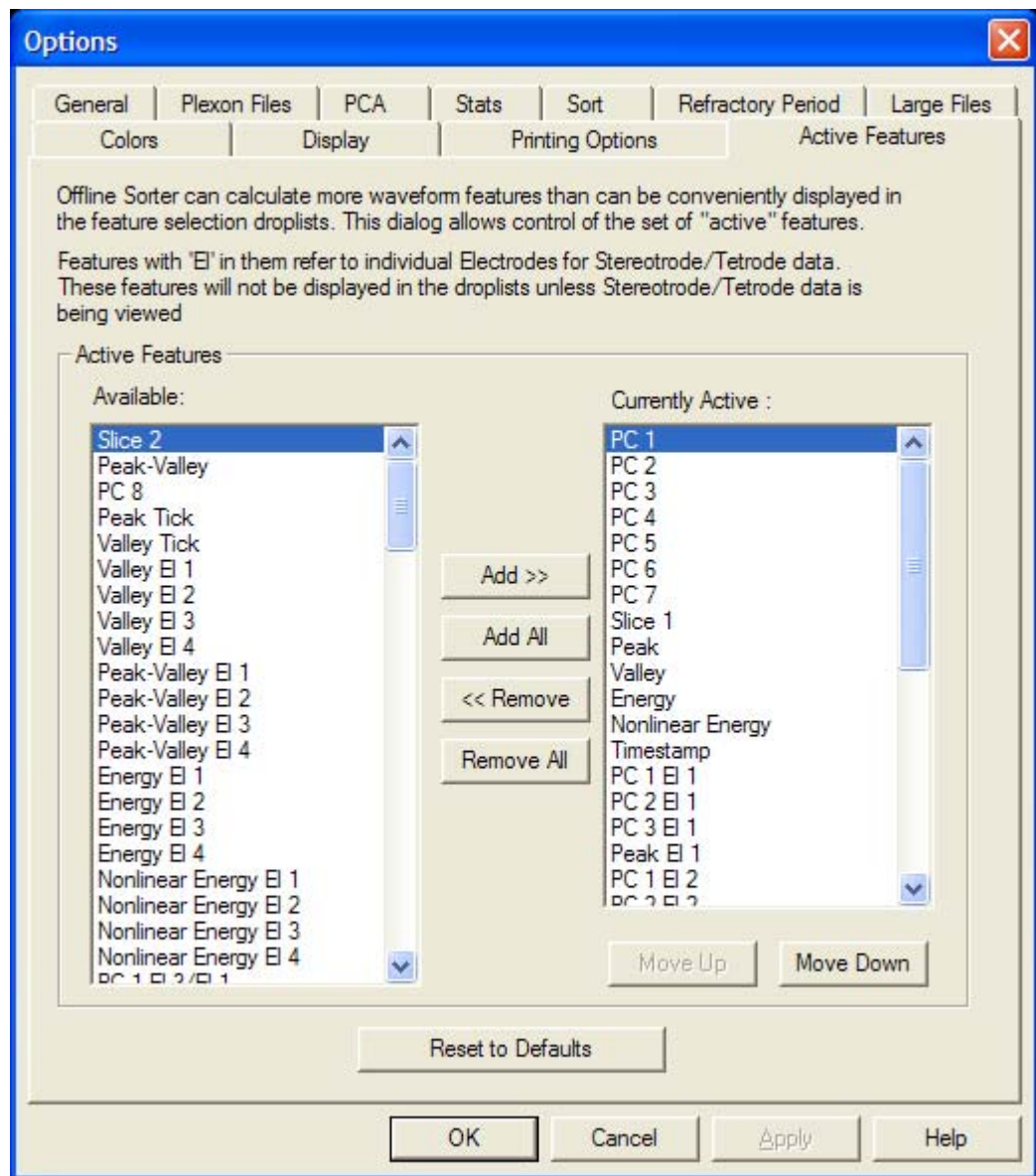
The following features are common to both the 2D Cluster view and the 3D Cluster view:

**Axis Feature Selection.** The variables shown on the **X Axis** and **Y Axis** (for the 2D Cluster view) and the **Z Axis** (for the 3D Cluster view) can be selected from among any of the “[Features Available for Sorting](#)” on page 5-9 by using the **X Axis**, **Y Axis**, and **Z Axis Drop-Down Feature Lists** on the **Toolbar**.



When the 2D Cluster view is active, only the X Axis and Y Axis drop-down feature lists are enabled. When the 3D Cluster view is active, the Z Axis drop-down feature list is also enabled.

Only the currently “active” features show up in the drop-down feature lists. The currently active features in Offline Sorter are a subset of the “[Features Available for Sorting](#)” on page 5-9. To control which features are active, use the **Active Features** tab in the **Tools | Options** dialog box.



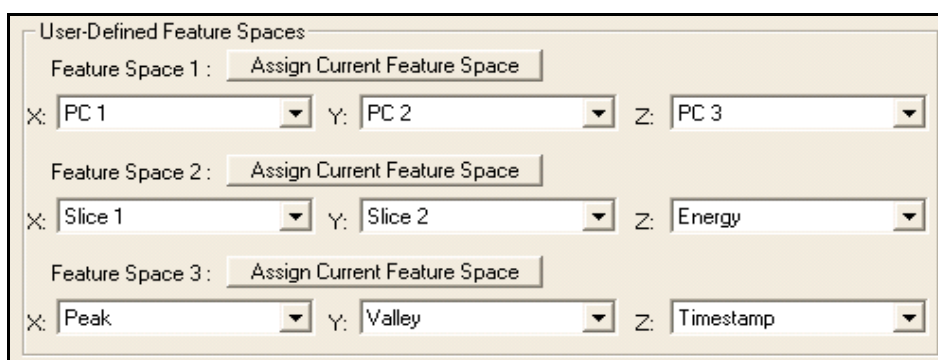
The currently active features appear in the right-hand list box, and the currently inactive (and thus available for activation) features appear in the left-hand list box. Select one or more features in either list box, then press the **Add >>** or **<< Remove** buttons to transfer features between list boxes. The order in which the active features show up in the drop-down feature lists is the same as the order in the **Currently Active** list box. Re-order the items in this list box by selecting one or more features and pressing the **Move Up** or **Move Down** buttons. Pressing the **Reset to Defaults** button reverts the set of active features to the as-shipped defaults.

---

Some of the features are relevant only for stereotrode or tetrode data; these features, although they can be active, do not show up in the drop-down feature lists when single-electrode data is being viewed in Offline Sorter.

**User-Defined Feature Spaces.** Offline Sorter can remember up to three Feature Space definitions that you commonly use, and it enables you to quickly put those pre-defined Feature Spaces into effect. For example, if you find yourself constantly switching between a {PC1, PC2, PC3} feature space and a {Slice 1, Peak, Valley} feature space, you can define Feature Space 1 to be {PC1, PC2, PC3} and Feature Space 2 to be {Slice 1, Peak, Valley} by using the **User-Defined Feature Spaces** section of the **Feature Calc** tab of the **Tools | Option** window. Then, you can use the toolbar buttons corresponding to each of the three user-defined feature spaces or the menu items under **View | Feature Space Definitions** (which you can bind to keystrokes for faster selection) to flip between the feature spaces, without having to use the toolbar feature drop lists.

User-Defined Feature Spaces include three dimensions of a feature space. If you use or display a 2D feature space, only the first two dimensions are used.



To define a user-defined feature space, use the drop-lists to select the feature for each dimension. Or, if the Cluster View is already showing the desired feature space, click **Assign Current Feature Space**.

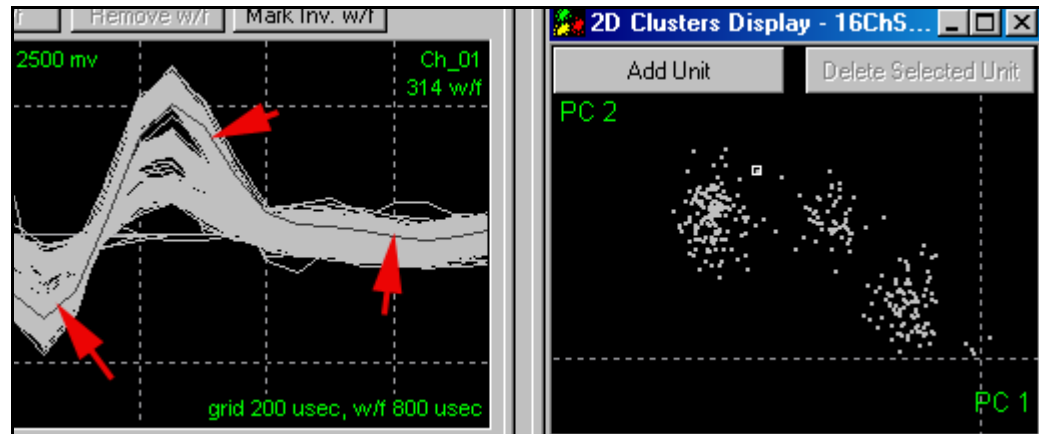
**Waveform Highlighting.** The Cluster views enable you to view the waveform shape corresponding to any point in the display.

To view the waveform

- 1 Left-click in the 2D Clusters view.
- 2 Left-click while holding down the ALT key in the 3D Cluster view.

The pointer become crosshairs, the nearest point in the cluster view is highlighted with a box, and the corresponding waveform is shown in the Waveform Window to the left (red arrows). If the waveform is in the time range being displayed, the

selected waveform is also highlighted with a white box in the Continuous Data Window.



If you move the mouse pointer while you press the left button, the Waveform Window and Continuous Data Window select the waveform currently closest to the mouse pointer.

On the bottom status bar of the application, Offline Sorter displays the exact X, Y, and Z (if applicable) feature coordinates for the highlighted waveform

X: 606.9104 Y: -2395.08 Z: -2395.08

as well as the timestamp associated with the highlighted waveform.

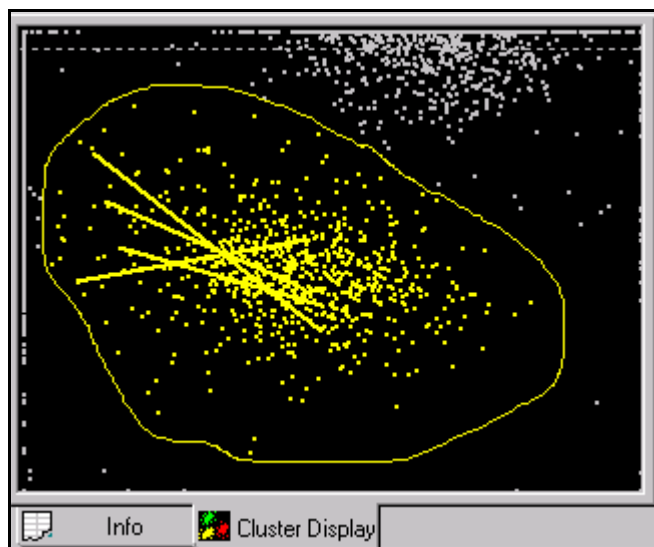
**Controlling Which Waveforms Are Displayed.** If you select the **Control cluster display** checkbox, you can use the scroll bar at the bottom of the [Continuous Data Window](#) to control the temporal range of the waveforms shown in the Cluster view. You can use the **Cluster View** parameter in the [Control Table](#) to restrict which waveform types are displayed in the Cluster view (either **All**, or only those shown in the Waveform Window).

You can select unit clusters in feature space by using either the manual Contours method (See [“Units Display” on page 3-11.](#)) or other sorting methods. You can add or remove specific waveform points from a unit cluster by selecting the **Units | Add (Remove) Waveforms from the Selected Unit** menu commands. You can also use the Waveform Window to add or remove waveforms by crossing them. Also, you can select and add or remove individual waveforms from a unit in the Continuous Data Display by using the right-click menu.

**Short ISI Spike Pairs Display.** The Cluster Display shows the waveforms assigned to the same unit with interspike intervals less than the Refractory ISI

---



(inter-spike interval) by drawing a line between the two corresponding dots in feature space.







You can turn on or turn off the drawing of lines between spikes with small ISIs by using either **Show Short ISI Lines** in the context menu or by using the **Tools | Options** menu command and selecting the **Refractory Period** tab. You can set the Refractory ISI either in the **Tools | Options | Refractory Period** tab or in the [Control Table](#) using the **Refr ISI** edit box.


**Cluster Positioning.** You can expand or contract the cluster image in feature space along the

x ( ,  ) and

y ( ,  ) axes or move it

left , right , up , or down 

by using the corresponding toolbar buttons, or by using selections from the context menu. Also, you can right-click and move the mouse up-and-down to zoom the display in and out. If you left-click and hold the `SHIFT` key down, you can pan the display by moving the mouse.

The zoom reset function (  in the toolbar, or **Reset Range and Zoom** in the context menu) resets the zoom and pan to the original range, which is calculated so as to include all the points.

**Buttons, Hints and Statistics Display.** The buttons displayed on top and the operational hints shown at the bottom of the Cluster views can be *on* (following left) or *off* (following right); use the **Show Buttons** and **Show Hints** menu com-

mands in the right-click menu (see later). These functions are also available via the **View | Cluster View Options** menu.



When you select the **Show Statistics** menu command, the hint box shows the cluster quality metrics described in [“Sorting Quality Statistics” on page 5-19](#). When the statistics are being shown, and the MANOVA p-value is greater than the **Probability threshold for significance** setting in the **Stats** tab of the **Options** dialog box, the text appears in red to indicate that the clustering can be invalid.

If the **Show Cluster Quality** menu command is selected, a box at the top of the display shows the current value of a selected cluster quality metric, along with the percentage change of that metric that resulted from the last sorting-related change. If the last change improved the cluster quality, the box appears in green and if it degraded the cluster quality, the box appears in red. This can be very useful when you are deciding which unit to assign “borderline” waveforms to. Any of the available cluster quality metrics described in [“Sorting Quality Statistics” on page 5-19](#) can be used for this display—you can choose the metric in the **Stats** tab of the **Tools | Options** dialog box.

If the **Show Ellipses** (or **Show Ellipsoids** on the 3D Cluster view) menu command is selected, ellipses/ellipsoids for each unit are shown on the display. The ellipses are calculated by finding the standard deviations along the principal axes. The ellipse or ellipsoid is scaled by the number of standard deviations specified by the **Outlier Threshold** in the Control Table.

**Colors.** To set the colors assigned to each unit in the Cluster views (and Units Display), select **Tools | Options** menu command and select the **Colors** tab.

**Point Size.** To set the point size in the Cluster views, select the **Display** tab under the **Tools | Options** dialog box.

**Automatic Cluster Redraw.** To update the Cluster view immediately when the **Slice 1**, **Slice 2**, **Slice 3**, or **Slice 4** values change, select the **Redraw Cluster Display Immediately when Slice Position Changes** in the **Display** tab under the **Tools | Options** dialog box.

---

**Zooming and Panning.** Both of the Cluster views support similar mechanisms to zoom and pan in the other display. Hold down the right mouse button and move the mouse vertically to zoom the display in and out. To pan the display, hold down the `SHIFT` key, then hold down the left mouse button, and move the mouse in the direction that you want to pan the display. The pointer becomes a hand-shaped symbol when zooming or panning occurs.

**Note:** The displays can be zoomed and panned using toolbar buttons or context menu entries. Both zooming and panning can be reset by using the **Reset Zoom** toolbar button, or by using the **Reset Range and Zoom** context menu option. This sets the display to default such that all the points are visible in the field of view.

**Context Menus.** Both of the Cluster views support context menus (commonly known as right-click menus), although the menu items available in each differ slightly. To bring up the context menu, click the right mouse button anywhere within the Cluster view, or, hold down the `ALT` key and then click the right mouse button.

The menu items common to both the 2D and 3D views are:

- **Add Unit** – when using Contour sorting, creates a new unit. This is the same as pressing the **Add Unit** button.
- **Delete Selected Unit** – removes the unit currently highlighted in the Units Display. This is the same as pressing the **Delete Selected Unit** button.
- **Delete All Units** – removes all units currently defined for the channel.

**Note:** The above options are also available from the **Units** submenu on the main menu bar.

- **Show Buttons** – shows the buttons as described in [“Buttons, Hints and Statistics Display” on page 3-25](#).
- **Show Hints** – shows text hints as described in [“Buttons, Hints and Statistics Display” on page 3-25](#).
- **Show Statistics** – shows the Sort Quality Statistics in the Hints area. When Hints are being shown, they use the same area of the screen.
- **Show Cluster Quality** – shows the selected Sort Quality metric changes at the top of the display, as described earlier.
- **Show Ellipses (Ellipsoids)** – shows the ellipses for units (ellipsoids for the 3D Cluster view) as described earlier.
- **X Range, Y Range** – expands, contracts, or shifts the viewing range as described in Zooming and Panning earlier.
- **Reset Range And Zoom** – resets the zooming and panning for the view to values calculated so that all the points are visible.

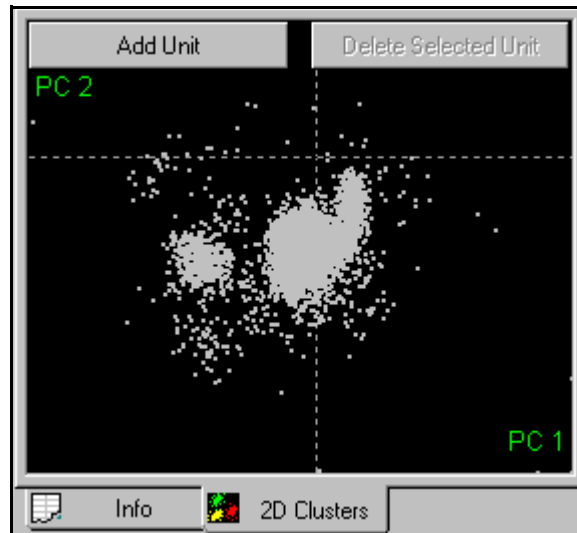


## 2D Cluster View Details

**Showing Dots or Dot Frequencies.** In response to the selection you make with the right-click menu, the 2D Cluster view can display the feature space points as dots, dot frequencies, or both. Displaying as dot frequencies builds a 2-dimensional histogram of the points, and maps the histogram bin contents to a gray-scale.

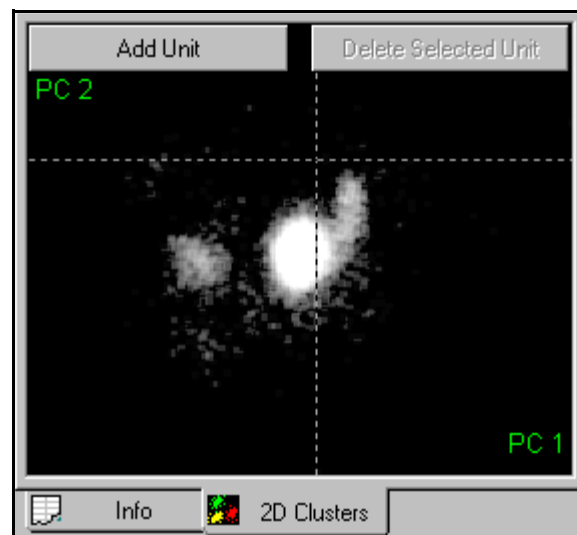
To display the feature space points as dots (the default)

- Select **Show Dots Only** on the **2D Clusters** right-click menu.



To display the feature space points as dot frequencies

- Select **Show Dot Frequencies** on the **2D Clusters** right-click menu.

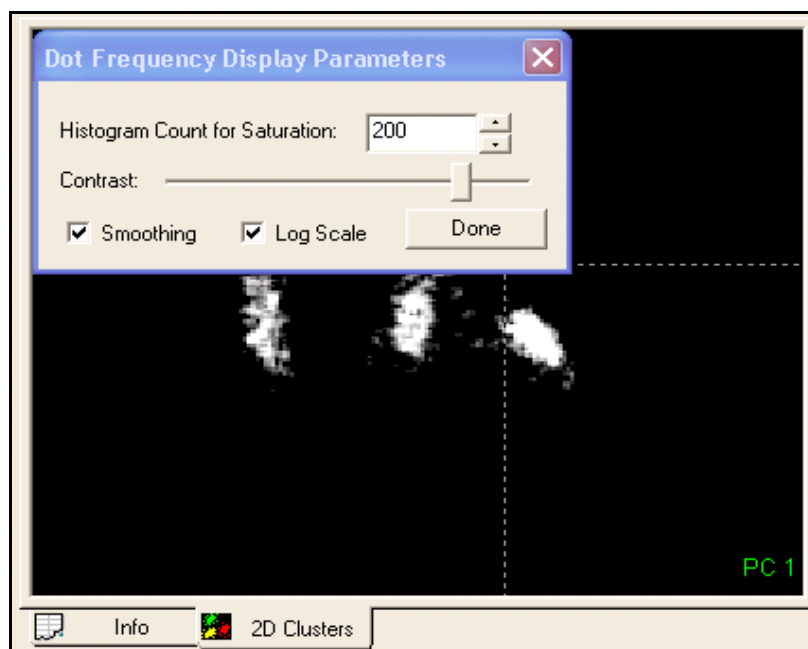


The dot frequencies are shown in linear or log scale. When dot frequencies are display, Offline Sorter enables the **Control Dot Frequency Display** right-click menu command. Selecting this displays a floating dialog that can be used to set



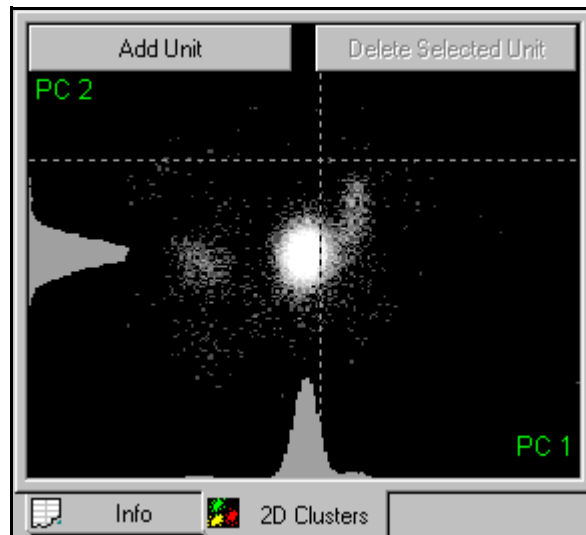
---

the contrast, enable/disable smoothing, and select between linear and log scaling for the dot frequency color assignments. The **Histogram Count for Saturation** parameter controls the mapping between the dot density and the grey level scale. If **Histogram Count for Saturation** = 10, for instance, then all cells with more than 10 dots are shown in white.



Both dots and dot-frequency density can be viewed simultaneously in the 2D Cluster view by selecting the **Show Dots and Frequencies** right-click menu command.

**Showing Histograms.** The 2D Cluster view can display histograms of the point distributions in each dimension by selecting the **Show Histograms** right-click menu command. The histograms are displayed along the margins.



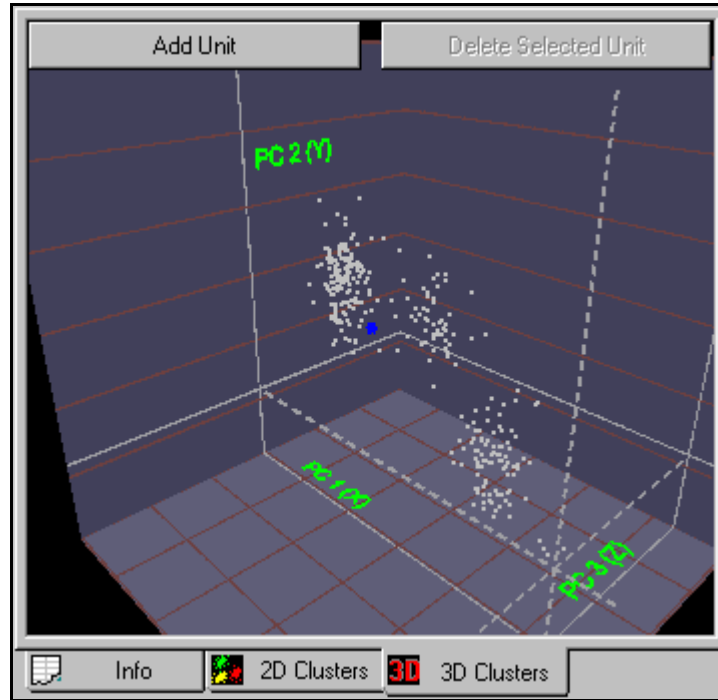
**Showing Ellipses.** The 2D Cluster view can display standard deviation ellipses around clusters by selecting the **Show Ellipses** right-click menu command, or by pressing the **E** key while the 2D Cluster view has input focus. The ellipses are oriented along the principal axes of the cluster of points, with the lengths of the major and minor axes for the ellipse being the standard deviation (sigma) of the points along each axis scaled by the value of the **Outlier Threshold** in the Control Panel. The ellipses in the following illustration show the 4 sigma boundaries around each unit. Offline Sorter draws the ellipses with dashed lines, in the color associated with the unit.



---

## 3D Cluster View Details

The 3D Cluster view allows changing the eyepoint in 3D feature space via mouse and keyboard actions in order to better visualize the data, and to find the viewing angle where cluster separation is maximized.



For the 3D Cluster view, the +Z axis is initially into the screen. All the points are contained inside a box with grid lines to provide a frame of reference. The dashed grey lines represent the X, Y, and Z axis, and they intersect at (0,0,0) in feature space. The solid grey lines show the projection of the X, Y, and Z axes onto the reference box. The axes projections onto the box are labeled with the feature name represented by that axis.

Components in the 3D display such as the box wall, the box top and bottom, and the grid lines on the box can be shown or hidden using the **Display** tab under the **Tools | Options** dialog. Their colors can be set using **Colors** tab.

The grid lines provide only cues—they do not correspond to actual values of features. The number of grid lines drawn can be controlled in the **Tools | Options Display** tab.

**Positioning the Eyepoint.** You can position the eyepoint with the following procedure.

To rotate the eyepoint

- Click and hold the **left mouse button**. The pointer changes to a hand. Mouse movement in the left-right direction then rotates the eyepoint around the y axis, and movement in the up-down direction rotates it about the x axis (pitch the view up-down).

To move the eyepoint nearer or farther away

- Click and hold the **right mouse button**. The pointer changes to a hand. Moving the mouse up moves the eyepoint farther away, and moving it down moves the eyepoint nearer.



## HINT

### Manipulating the Eyepoint

To get accustomed to manipulating the eyepoint, think of the hand grabbing onto and turning the point clouds.

By default, the eyepoint motion is constrained to not allow “upside down” eyepoints, which can be confusing. You can relax this constraint by unchecking the **Constrain view rotation so that +Y axis is always upwards** checkbox in the **Tools | Options | Display** tab.

To animate the eyepoint rotation

- The 3D Cluster view supports a mode of operation that enables the eyepoint for the 3D view to be continuously animated. You can control this by using the **Animate the 3D rotation after left mouse release** option in the **Tools | Options Display** tab. When you enable this option, the eyepoint continues to move in whatever direction it was going when the mouse button is released. This is as if the left mouse button “throws” the eyepoint into an orbit around the look-at point. You can quickly toggle this option on or off by pressing the **A** key while the 3D Cluster view has focus.

**The Look-At Point (the point of rotation in 3D).** Wherever the eyepoint is positioned, the “direction of gaze” is always adjusted so that the eye is looking at a certain point in space. This Look-At point is marked on the 3D Cluster view with a small blue (by default) axis icon.



The display of the Look-At icon can be turned on and off using the **3D Look-At Icon** checkbox in the **Tools | Options Display** tab. The color of the Look-At icon can be changed in the **Tools | Options Colors** tab.

The look-at point can be changed in two ways:

- Manually positioning
- Positioning it to the center of gravity of clusters

---

To manually position the Look-At point

- 1 Hold down the **Shift key**, then click and hold the **left mouse button**.
- 2 The pointer changes to a hand. Mouse movement then translates the eyepoint in screen space. That is, moving the mouse in the left-right direction translates the Look-At point in the screen x direction, and moving the mouse up-down translates it in the screen y direction.
- 3 To move the Look-At point nearer or farther from the eye, hold down the **SHIFT** key, then click and hold the right mouse button. The pointer changes to a hand. Moving the mouse up and down translates the Look-At point in the screen z direction.



#### HINT

#### Moving the Look-at Point

To get accustomed to manually moving the Look-at Point, think of the hand grabbing and pulling the enclosing 3D grid box.

To position the Look-At point at cluster centers

- Press the **C** or **L** key to position the Look-At point at the center of gravity (CG) of the unit clusters, and make a selection by cycling through each assigned unit cluster with a key press. The center of gravity of all points, and the center of gravity of all unassigned points are also included in the cycle. For example, if a channel contains two 2 sorted units, pressing the **C** or **L** keys moves the Look-At point as follows:

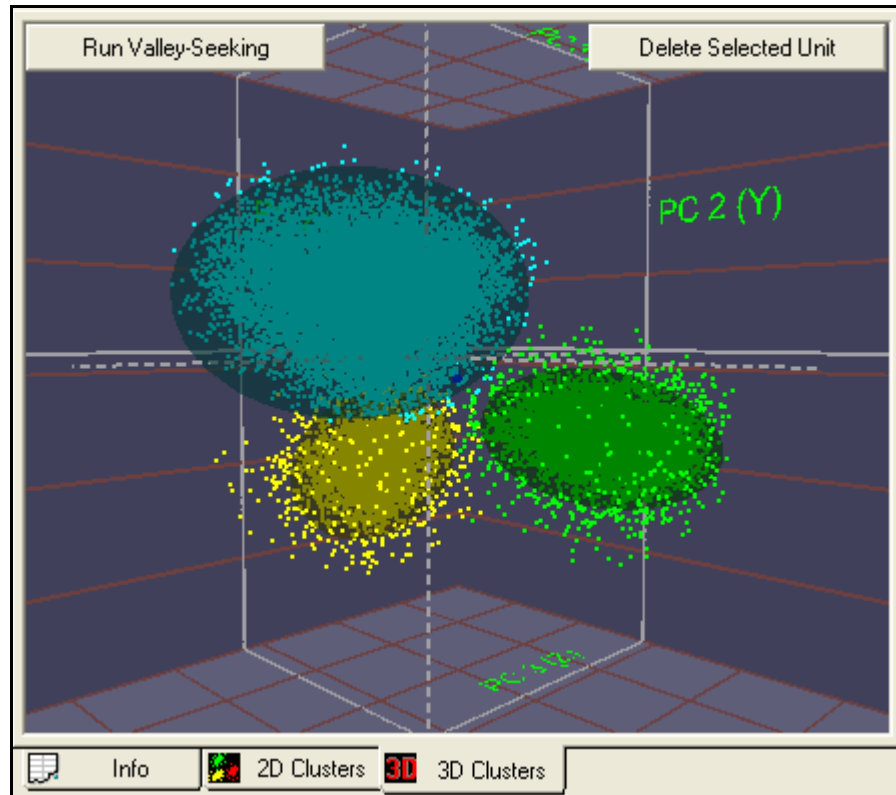
CG of All points -> CG of Unit a points -> CG of Unit b points -> CG of all unassigned points

#### Resetting the Eyepoint and Look-At points

- Pressing the left and right mouse buttons simultaneously resets the eyepoint and Look-At points to their defaults (looking directly down the Z axis, with the Look-At point at the center of gravity of all points).

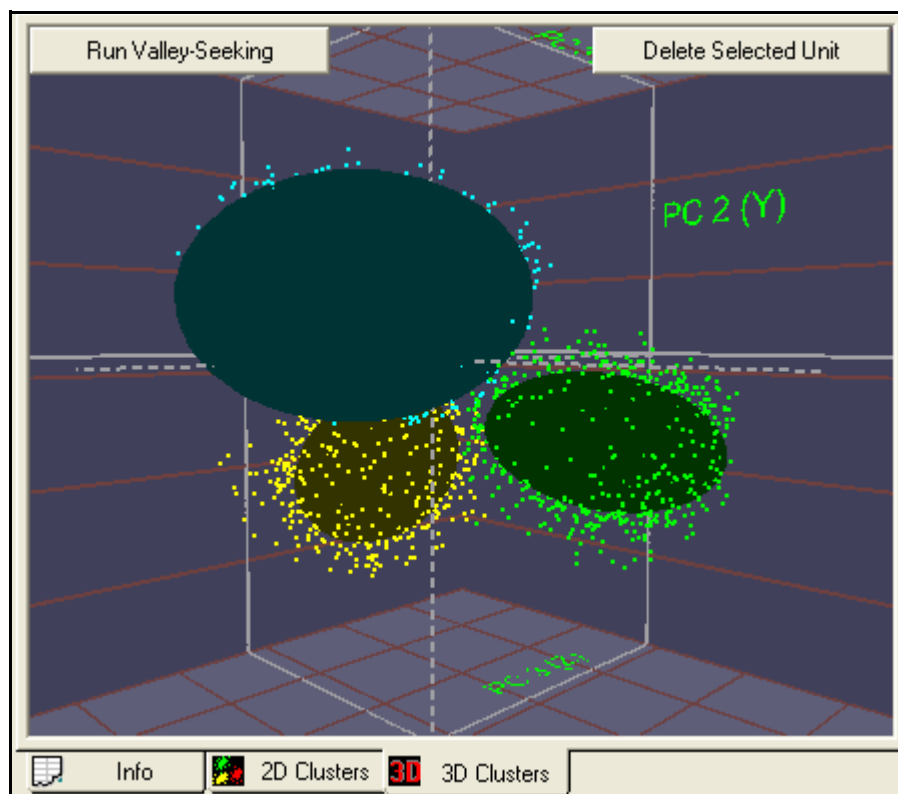
**Showing Ellipsoids.** The 3D Cluster view can display ellipsoids around clusters by selecting the **Show Ellipsoids** right-click menu command, or by pressing the **E** key while the 3D Cluster view has input focus. The ellipsoids are oriented along the principal axes of the cloud of points, with the lengths of the ellipsoid axes being the standard deviation (sigma) of the points along each axis scaled by the value of the Outlier Threshold in the Control Panel. For example, the ellipses shown below show the 3 sigma boundaries around each unit, because the Outlier Threshold was set to 3. Offline Sorter draws the ellipses semi-transparently, in the

color associated with the unit. Observe that you can see the points both inside and outside the ellipsoid.



You can change the transparency of the ellipsoid by pressing the  $\uparrow$  and  $\downarrow$  keys (for fade-in, fade-out) while the 3D Cluster view has input focus. The following

screenshot is the same as the previous screenshot, except the  $\perp$  key has been repeatedly pressed until the ellipsoids became totally opaque.




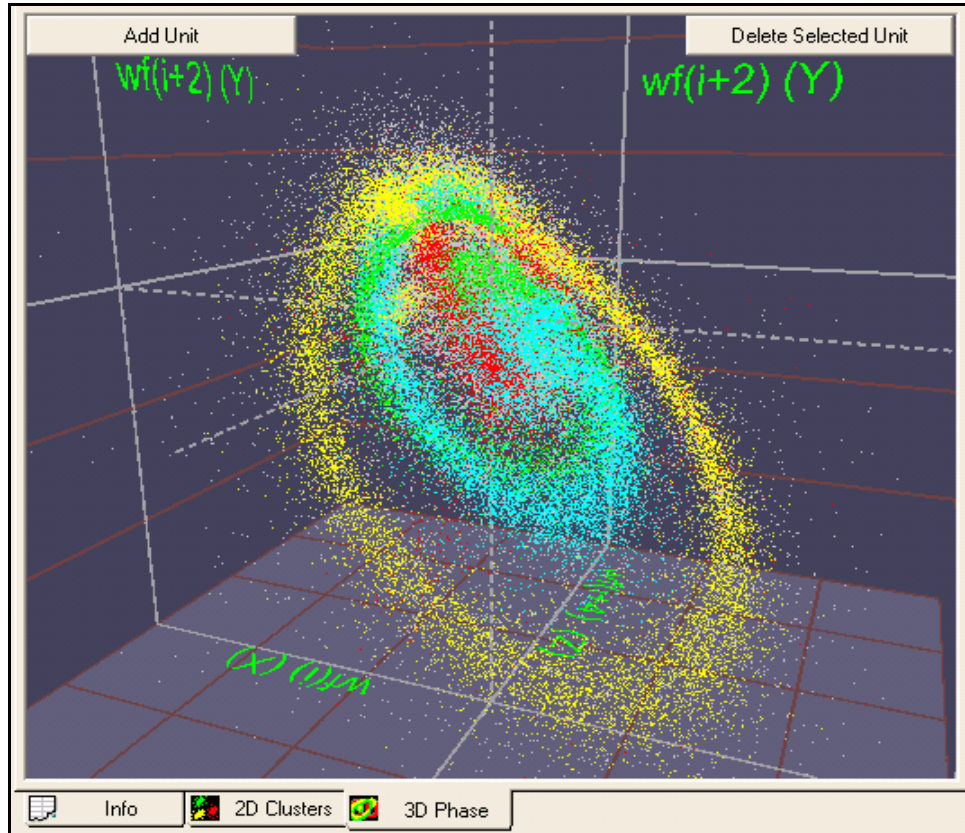
**Note:** If enabling the ellipsoid display does not cause the ellipsoids to appear, press the  $\perp$  key repeatedly, as it is possible that the ellipsoids are there but are completely transparent.

### 3D Phase Space View

The 3D Phase Space View is an experimental view that provides you with an interesting new way to view the waveforms. The 3D Phase Space View is similar to the 3D Cluster View, but instead of showing the waveforms as points in feature space, it shows them as lines (or a sequence of points) in phase space. The line for a waveform is referred to as the waveform's "trajectory" in phase space. You can apply most of the operations from the 3D Cluster View to the 3D Phase Space View, such as rotating, zooming, moving the Look-At point, selecting waveforms, and even defining contours for contour sorting.

To display the 3D Phase Space View

- Select the **View | 3D Phase Space View** menu command, or  
press the  button on the toolbar.



In phase space, each waveform is a series of points in space connected by a line. By default, only the points at the vertices of the trajectory are shown in the 3D Phase Space View; you can toggle the display by pressing the **D** key to show the trajectory lines, but this often results in a very slow and hard-to-decipher display. To display the entire trajectory associated with the point nearest the click, with a box drawn around the starting point, select a point by using **ALT + left-click**.

You can plot the 3D Phase Space View waveforms in two different kinds of phase space; a space formed by “embedding” the waveform, or a derivative-based phase space; you can toggle the display between the two different kinds of phase spaces by using the **E** key. Offline Sorter forms the embedding phase space by taking triplets of waveform sample values at points on the waveform separated by a “Lag”. For example, for waveform sample values  $wf(i)$ , the phase space coordinates are  $\{wf(i), wf(i+1*Lag), wf(i+2*Lag)\}$ . By default, the Lag parameter is 2; you can increase or decrease by pressing the **S** or **T** keys. The derivative phase space plots the waveform ADC values ( $wf(i)$ ) on the X dimension, the first derivative ( $dwf(i)/dt$ ) of the waveform on the Y dimension, and the second derivative ( $d2wf(i)/dt2$ ) on the Z dimension.



---


Although you can do contour sorting by circling points on the 3D Phase Space View, it is not necessarily a superior way to go about sorting units.

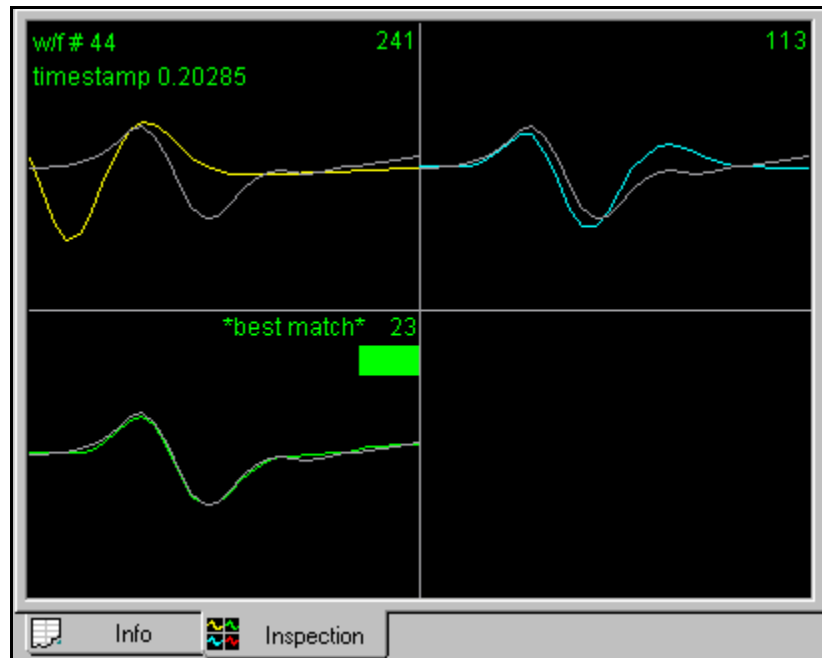
## Waveform Inspection View

The Waveform Inspection View is used for comparing individual waveforms with the template (average waveform) for each sorted unit.

To display the Waveform Inspection View

- Select the **View | Waveform Inspection View** menu command, or

press the  button on the toolbar.



The Waveform Inspection Window shows the average waveform (template) of each sorted unit in the appropriate unit color. The individual waveforms, which you select either in the [Continuous Data Window](#) (left-click on waveform) or in either of the [Cluster Views](#) (left-click on waveform's dot in feature space), are displayed in grey and positioned over each template for comparison.

### Display Features:

- **colored waveforms** – template (average waveform) for each sorted unit
- **grey waveform** – waveform selected in the Continuous Data Window or Cluster views
- **w/f #** – number of the selected waveform in the temporal sequence of waveforms on the channel
- **timestamp** – time of occurrence of the selected waveform in seconds

- **dotted vertical lines** – sort start and sort stop times (if present); see [“Adjusting the Sort Range”](#) on page 4-25.
- **# in upper right-hand corner of each plot** – sum of squared difference between the selected waveform and the template
- **\*best match\***- displayed on the unit with the best match (least sum of squared difference) between the unit template and the selected waveform
- **Colored bar** – displayed on the unit assignment of the selected unit (green bar above)

#### Waveform Detection View

The Waveform Detection View is only available when you open a channel containing continuously digitized data. The view displays distributions of quantities (e.g. peak heights, energy, or non-linear energy) that can be used for thresholding the data for spike detection.

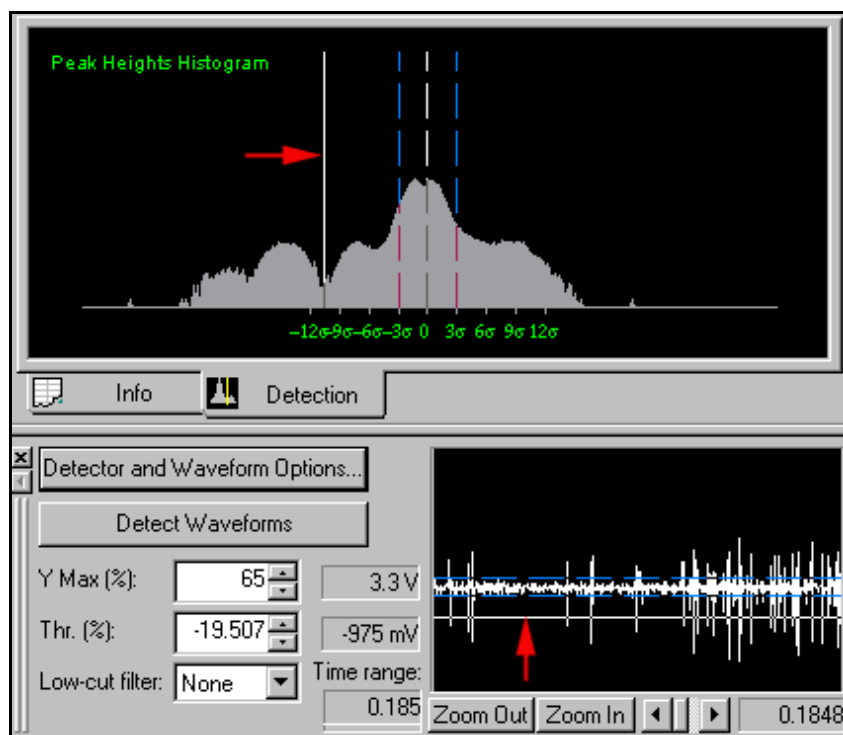
To display the Waveform Detection View

- Select the **View | Waveform Detection View** menu command, or,

press the  button on the toolbar.

By default, the Waveform Detection View is selected when you open a continuous data channel. The Waveform Detection View shows a histogram of the logarithm of the waveform peak heights. The dashed white line indicates zero, and

dashed blue lines indicate the 3 sigma noise level. The solid white line (marked with a red arrow below) indicates the selected threshold.



## Adjusting the Threshold for Waveform Detection

The threshold can be changed by clicking and dragging the solid white line, either on the Waveform Detection View or on the Continuous Data Window. Thresholding can be done on either positive or negative amplitudes. Alternately, you can set a threshold by entering a value in the **Thr. (%)** edit box in the Continuous Data Window, or by clicking on the associated spin control.

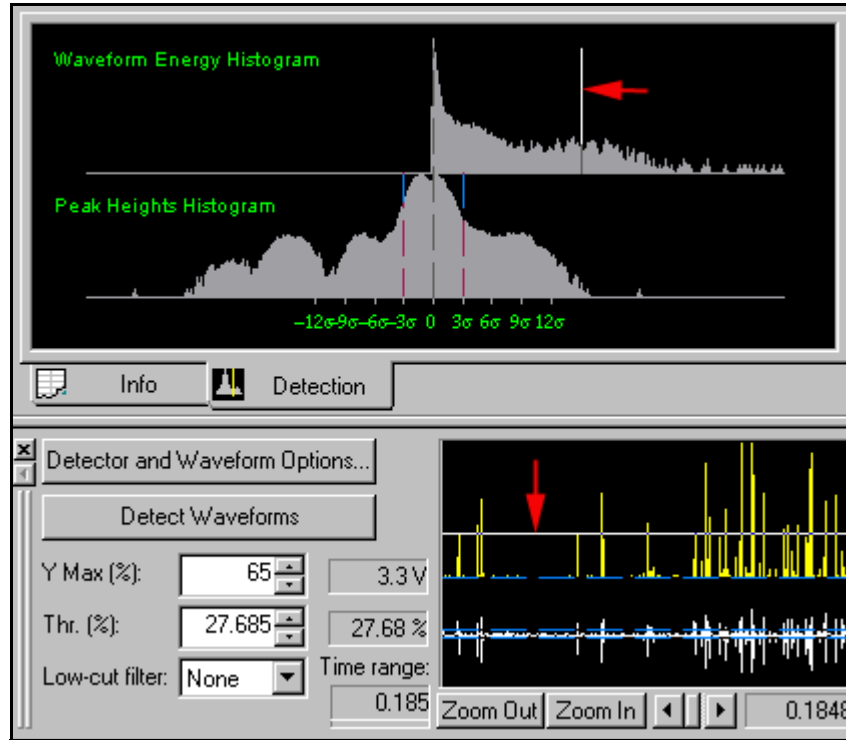
## Changing the Detector

The waveform detection algorithm can use thresholds on quantities other than raw waveform amplitude. As explained in [“Specifying the Waveform Detection Options” on page 4-6](#), you can use the **Detector and Waveform Options** menu dialog to select the signal threshold to be the:

- Raw Voltage Signal (the default), or
- Energy of the Signal, or
- Nonlinear Energy of the Signal

When you select a signal other than **Raw Voltage** amplitude for use in spike detection, the Waveform Detection View and the Continuous Data Window show both the selected parameter (top display) and the **Raw Voltage** (bottom display).

Shown below are the Waveform Detection View and the Continuous Data Window when the **Energy of the Signal** threshold has been selected.



The top histogram in the Waveform Detection View above shows the distribution of the logarithm of the signal energy, and the top trace in the Continuous Data Window shows the waveform energy. The thresholding is set in the top histogram or trace.

Offline Sorter calculates the Energy of the Signal by taking a moving-window average of the square of the voltage amplitude. The width of the window used for the averaging is controlled in the **Detector and Waveform Options** dialog. Non-linear Energy is akin to a derivative of the signal in that it is zero if the voltage is constant and maximum if the waveform is rapidly varying.

Mathematically, the energy and nonlinear energy of the signal at time point  $i$  is given by:

$$\text{Energy}(i) = \left( \frac{1}{\text{width}} \right) \sum v(j) v(j)$$

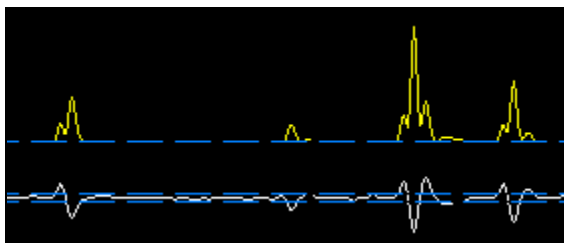
$$\text{Nonlinear Energy}(i) = \left( \frac{1}{\text{width}} \right) \sum [v(j) v(j) - v(j-1) v(j+1)]$$

where  $\Sigma$  is the sum over the energy window centered around  $i$  (i.e., the sum over  $j$  where

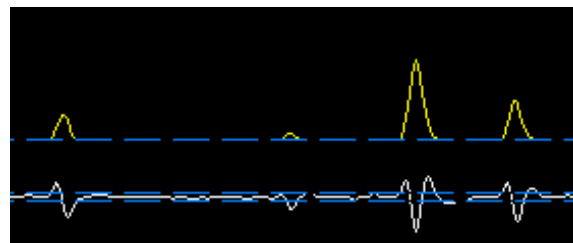
$$j = i - \frac{\text{width}}{2} \dots i + \frac{\text{width}}{2}$$

It is instructive to look closely at these calculated quantities for typical waveforms. The top yellow trace is the calculated quantity, and the bottom trace is the raw waveform amplitude.

**Energy of the Signal**



**Nonlinear Energy of the Signal**



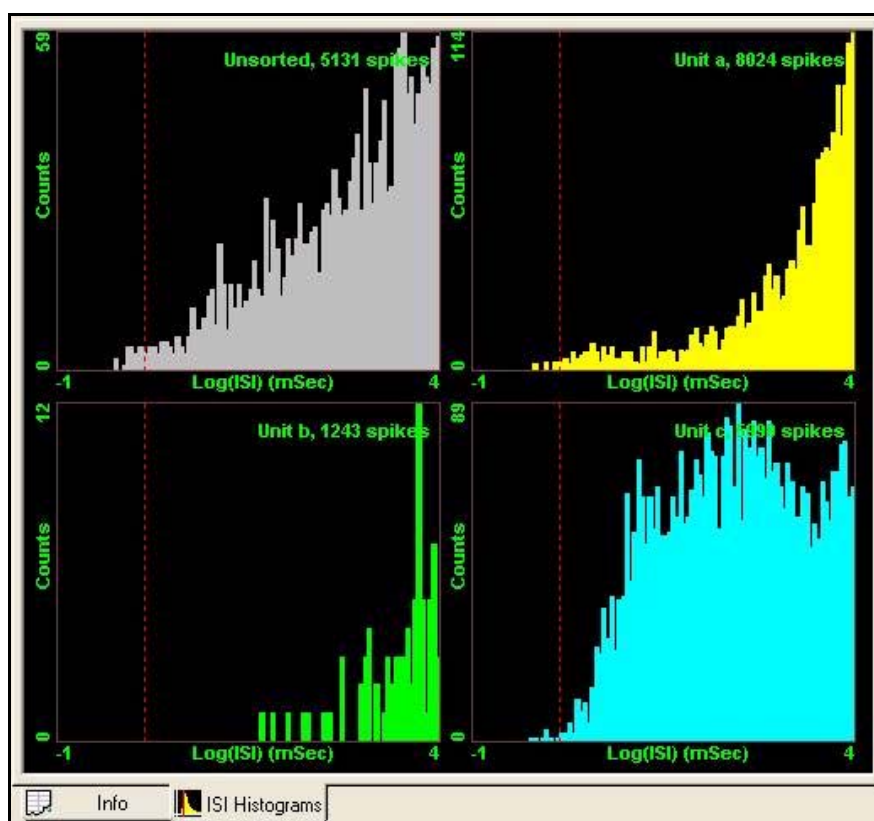
### ISI Histograms View

The ISI Histograms View provides a larger, printable version of the Inter-spike Interval Histograms that appear in the Cluster Display. The range for the ISI Histograms is 0 (zero) to the ISIHistMax(ms) Control Panel setting.

To display the ISI Histograms view

- Select the **View | ISI Histograms View** menu command, or,

press the  button on the toolbar.



You have several options for displaying the ISI Histograms from the right-click context menu. You can configure the view to show the ISI Histograms from all sorted units, from all sorted units plus an ISI Histogram for the unsorted waveforms, or only for the currently-selected unit. Additionally, you can display the Time axis or the Counts axis, or both, using a log scale.

### Feature vs. Feature View

You can use the Feature vs. Feature view to quickly view the feature space point distributions for many different features. This can give you a quick impression of which features might be most useful for sorting. Offline Sorter divides the screen into boxes, and each box contains a 2D Cluster view. Much of the functionality from the 2D Cluster view is available in the Feature vs. Feature view, including the point selection, zooming, and contour sorting functionality.

To display the Feature vs. Feature view

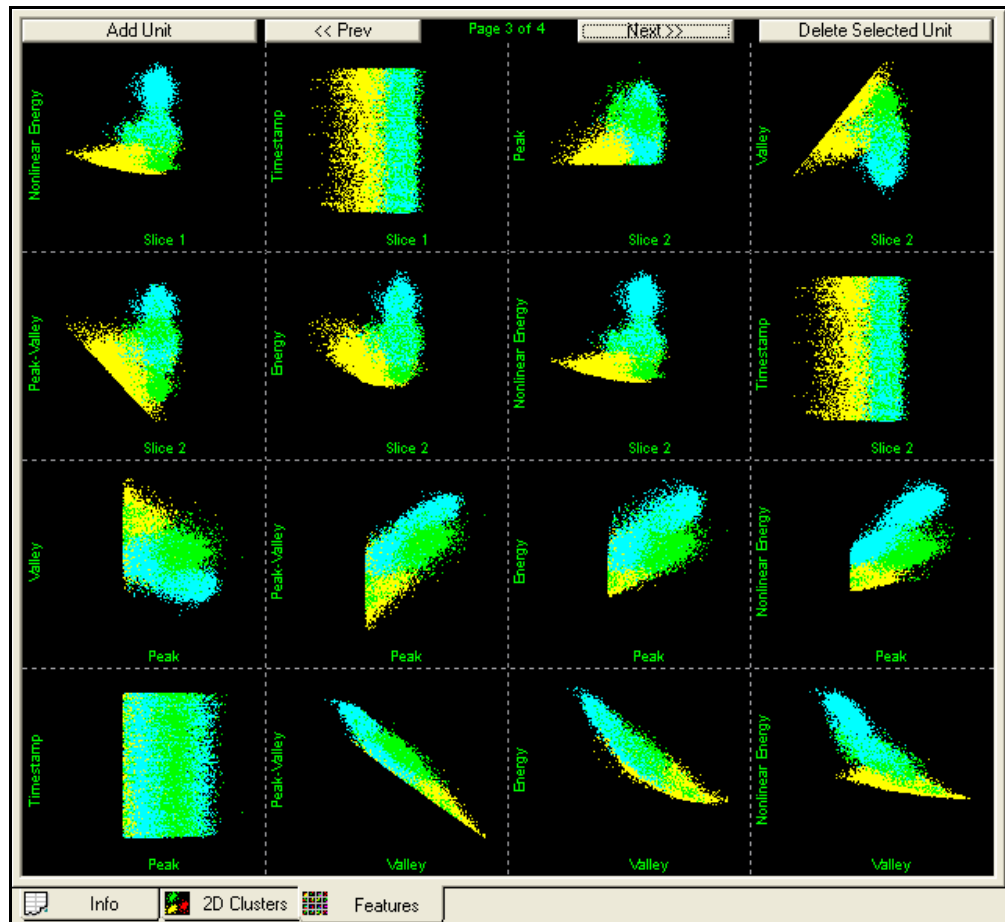
- Select the **View | Feature vs. Feature View** menu command, or,

press the  button on the toolbar.

The Feature vs. Feature view can operate in one of several modes, which differ in the combinations of features that are shown in the plots. You can select different display combinations via the context menu that is available by holding down the ALT key and right-clicking in the view. The modes are:

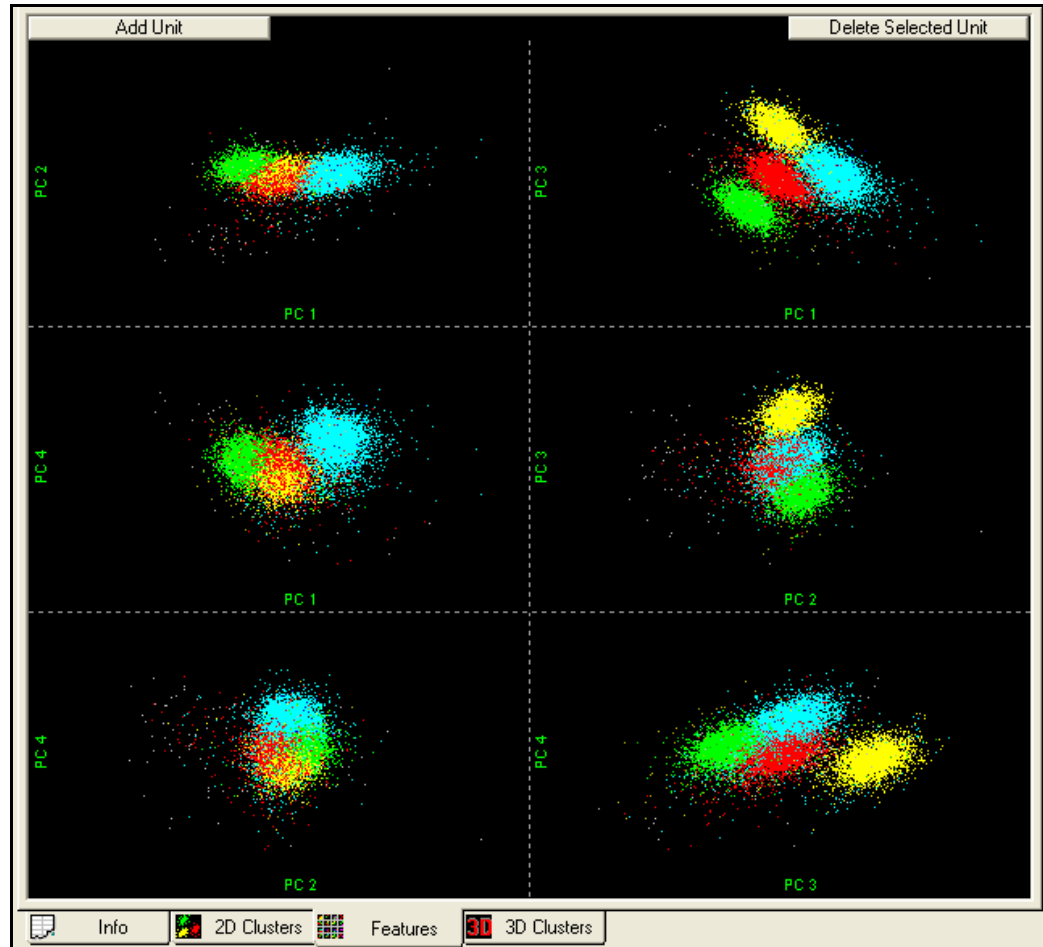
- **Show All Active Feature Combinations:** shows every permutation of the currently Active Features. This is the default mode for non-Tetrode Data Types. Generally the number of permutations is large enough that several pages are required to show all of the plots. To step through the pages, use the

**Next** >> and << **Prev** buttons on the top of the screen, or the N (Next) or P (Previous) keys.



- **Show PC1-4 Combinations:** shows the 6 permutations of the first 4 Principal Component projections; PC1 vs. PC2, PC1 vs. PC3, PC1 vs. PC4, PC2 vs.

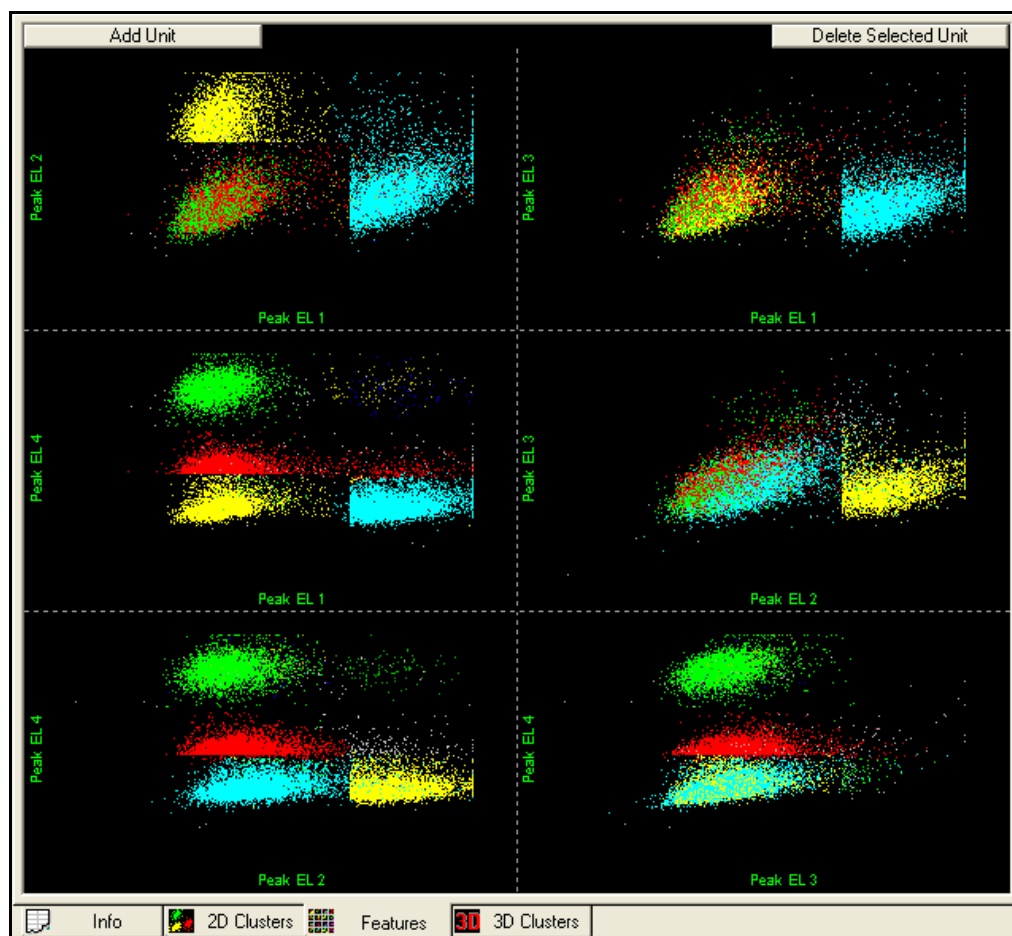
PC3, PC2 vs. PC4, and PC3 vs. PC4. In this mode, press the N or P keys to go to the next or previous channel (as in the 2D Cluster view).



- **Show Tetraode Electrode Combinations:** shows the 6 permutations of 4 tetraode electrodes, using a tetraode per-electrode feature value (Per-Electrode Peak Height by default). For example, this mode shows plots of Peak Height for Electrode 1 vs. Peak Height for Electrode 2, Peak Height for Electrode 1 vs. Peak Height for Electrode 3, ..., Peak Height for Electrode 3 vs. Peak Height for Electrode 4. This is the default mode for Tetraode Data Types. In



this mode, press the **N** or **P** keys to go to the next or previous channel (as in the 2D Cluster view).



To select which Per-Electrode feature is used for the Tetraode electrode display mode, select that feature for the X Axis in the 2D Cluster view using the feature droplists in the toolbar. For example, if you select PC 1 EL 1 for the X Axis of the 2D Cluster view, the Feature vs. Feature View operating in Tetraode electrode display mode shows the plots PC1 EL 1 vs. PC1 EL 2, PC1 EL 1 vs. PC1 EL 3, ... , PC1 EL 3 vs. PC1 EL 4. If you select a non-Per-Electrode feature for the X Axis in the 2D Cluster view, the tetraode display uses Per-Electrode Peak Height by default.

## Common Functionality

In general the plots in the Feature vs. Feature view behave the same as the 2D Cluster view plot, responding to the **Add/Remove Waveforms/Units** buttons and menu items, the **View W/f** selection, and the zooming/panning toolbar buttons in the same way. For example, the following mouse functions from the 2D Cluster view are available:

- **Left-Click** selects the nearest waveform point, highlighting that point on all displays and in the waveform view.

- **Control-Left-Click** while in Contour sorting mode adds a new unit.
- **Shift-Left-Click** pans the display.
- **Right-Click** and hold zooms the display in and out.
- **Alt-Right-Click** displays the context menu.

Additionally, **Double-click** any sub-box to bring up the full-size 2D Cluster view with the corresponding features on the X and Y axes.

## Sorting Summary View

The Sorting Summary view can be used to quickly get an overview of the results of sorting a channel. The Sorting Summary view is intended to be printable so that it can serve as a permanent record of the sorting. The contents of the Sorting Summary view can also be exported to PowerPoint. Much of the information on the Sorting Summary view can be found in other views or windows in Offline Sorter; the value of the Sorting Summary view is that it gathers the most important sorting information together in one view.

To display the Sorting Summary view

- Select the **View | Sorting Summary View** menu command, or

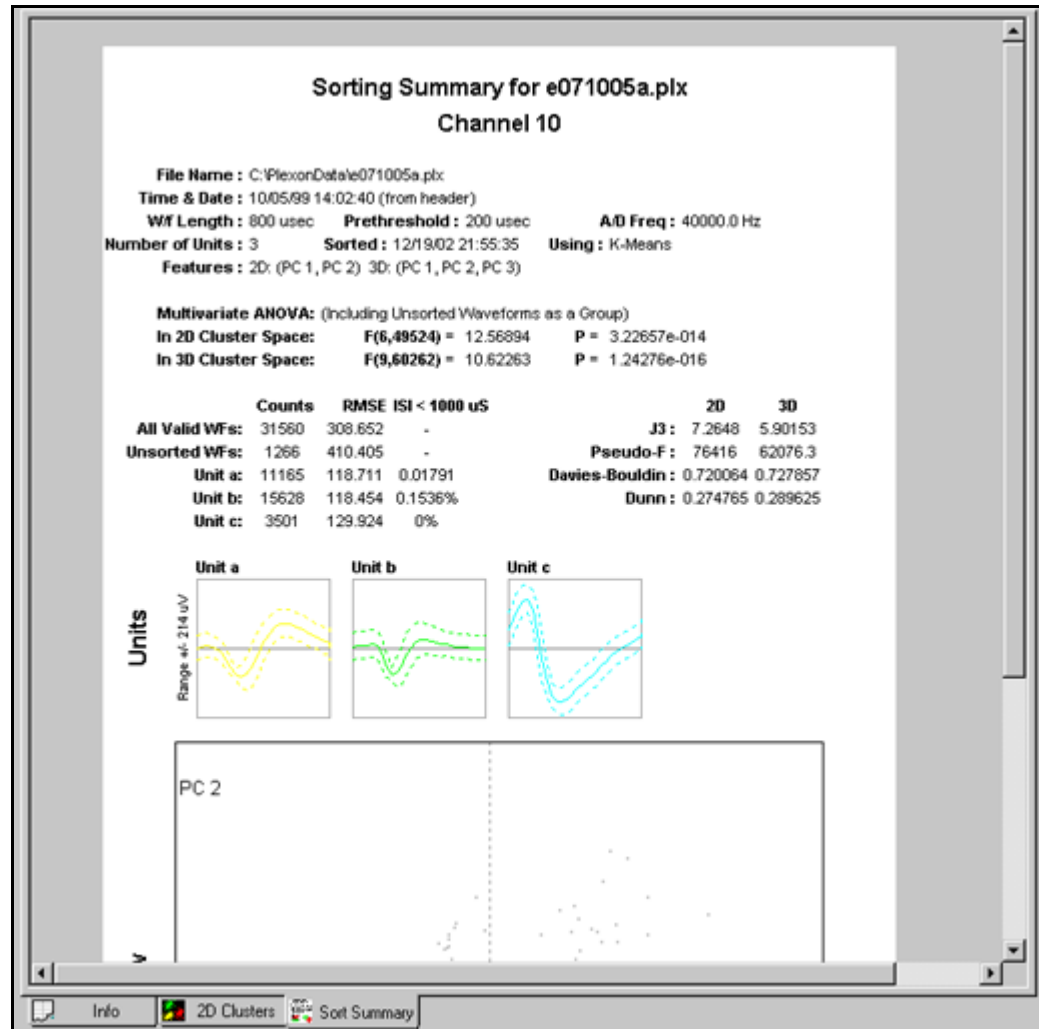
press the  button on the toolbar.



### HINT

#### Is the Sorting Summary Button Missing?

If this button does not appear on your toolbar, select **Tools | Customize Toolbars**, and select the line labelled **View** in the list box in the Toolbars tab. Do not uncheck the checkbox—just select the line. Then press the **Reset** button.



The title on the top contains the channel number, and optionally the file name. To change whether the file name shows up in the title, select the **Show Filename in Title** entry in the right-click menu on the Sorting Summary view.

The other entries on the form include:

- **Filename:** the full path to the file
- **Time & Date:** the time and date that the data file was collected. Whenever possible, this time and date is obtained from the header records in the data file, so **(from header)** appears on this line. If the time and date cannot be obtained from the file header, it is taken from the time/date on the file itself, in which case **(from file)** appears on this line.
- **W/f Length:** the total length of the waveform, in microseconds
- **Prethreshold:** the time between the start of the waveform and the threshold crossing, in microseconds
- **A/D Freq:** the digitization rate, in Hertz

- **Number Of Units:** the number of units found on this channel
- **Sorted:** the time & date that sorting was done. Currently this is always the current date.
- **Using:** the name of the algorithm used for the sorting
- **Features:** the features selected to define the features space. All the sorting quality statistics depend on the choice of features.
- **Multivariate ANOVA:** the string in parenthesis indicates whether or not unsorted waveforms are treated as a unit in the ANOVA and the other statistics calculations. The next lines contain the ANOVA F and p-values, as explained in [“Sorting Quality Statistics” on page 5-19](#).

The next section contains a table with a row per unit, plus entries for all waveforms and for unsorted waveforms. The columns are:

- **Counts:** The total number of waveforms (in the **All Valid WFs** row), the number of unsorted (unassigned) waveforms (in the **Unsorted WFs** row), and the number of waveforms in each unit.
- **RMSE:** the summed root mean squared for all waveforms in the category. This is a sample-by-sample summation of the deviation from the average waveform, summed across all waveforms. This sum is then divided by the total number of ticks, then the square root is taken.

$$\text{RMSE} = \sqrt{(\sum \sum (v(w, t) - \text{avg}(t))^2) / (N_w N_t)}$$

where

the sums run over  $N_w$  waveforms  $w$  and  $N_t$  ticks  $t$  in each waveform

$v(w, t)$  is the waveform amplitude at tick  $t$  for waveform  $w$

$\text{avg}(t)$  is the average amplitude of all waveforms at tick  $t$

Thus, this number is smaller for compact clusters. The Plexon MAP system calculates a similar summed squared difference for the real-time online template sorting unit assignment algorithm.

- **ISI Percentage:** lists the percentage of waveforms that occur within the Refractory ISI period (as set by from the Control Table) of each other. Large percentages here for a unit can indicate that errors have been made in the sorting of that unit.

The next section to the right displays the 2D and 3D versions of each of the J3, Pseudo-F, Davies-Bouldin, and Dunn cluster validity metrics (See [“Sorting Quality Statistics” on page 5-19](#)). The feature space used for this sorting quality statistics displayed under **Features** above.

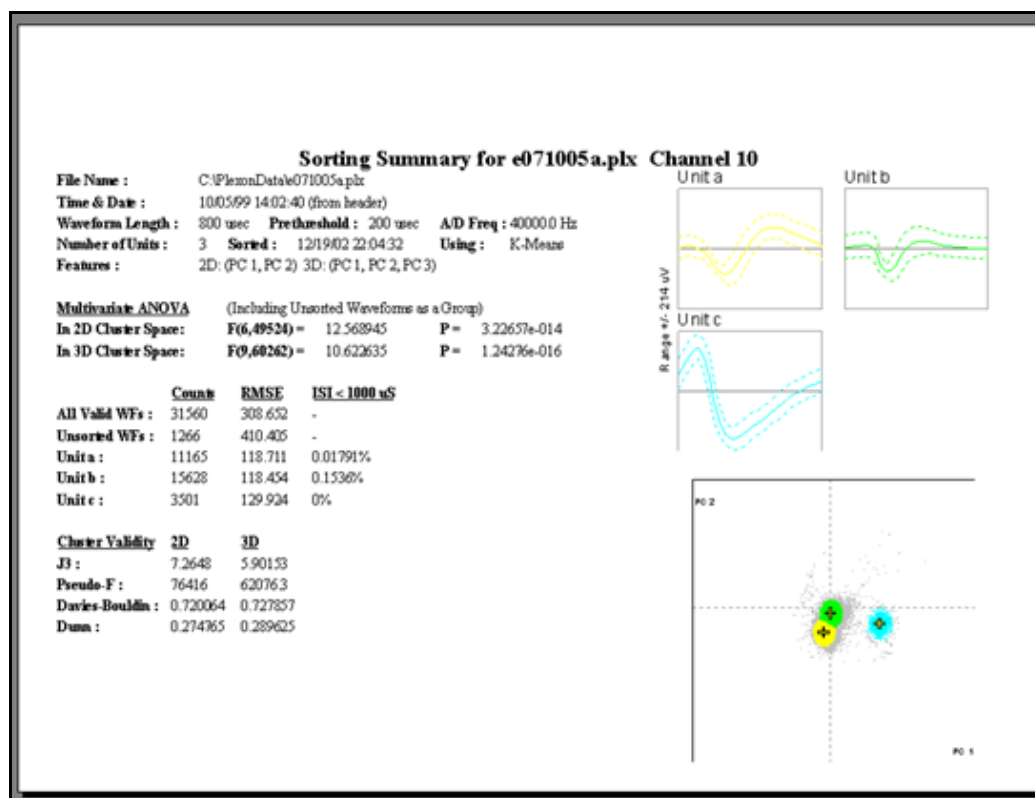
The next section displays the template (average) waveform for each of the units on the channel, as described in the “Units Display” on page 3-11. The dashed line represents the 3-sigma deviation of the waveforms around the average.

**Note:** The colors displayed for the units are controlled by the printing colors, not the display colors. You can set the printing colors in the **Printing Options** tab of the **Tools | Options** dialog.

The bottom section is the 2D Cluster view. Again the colors displayed are controlled by the printing colors, not the display colors.

## Output to PowerPoint

You can send the Sorting Summary view to PowerPoint by selecting **Send To PowerPoint** in the right-click menu in the Sorting Summary view. For this feature to be available, PowerPoint must be installed on the computer.



When the **Send To PowerPoint** menu item is selected, PowerPoint opens (if it is not already open), and a new slide is added to the presentation. Sending another Sorting Summary view to PowerPoint results in another slide being added to the same presentation. PowerPoint does not save the presentation to disk—select **Save** from the PowerPoint menu to save your presentation.

### Outputting for All Channels

The **File – Print All Sort Summaries** and **File – Send All Sort Summaries to PowerPoint** iterate through all channels and print or export the Sort Summary for each channel to a single PowerPoint file, or to a single print job.

### PCA Results View

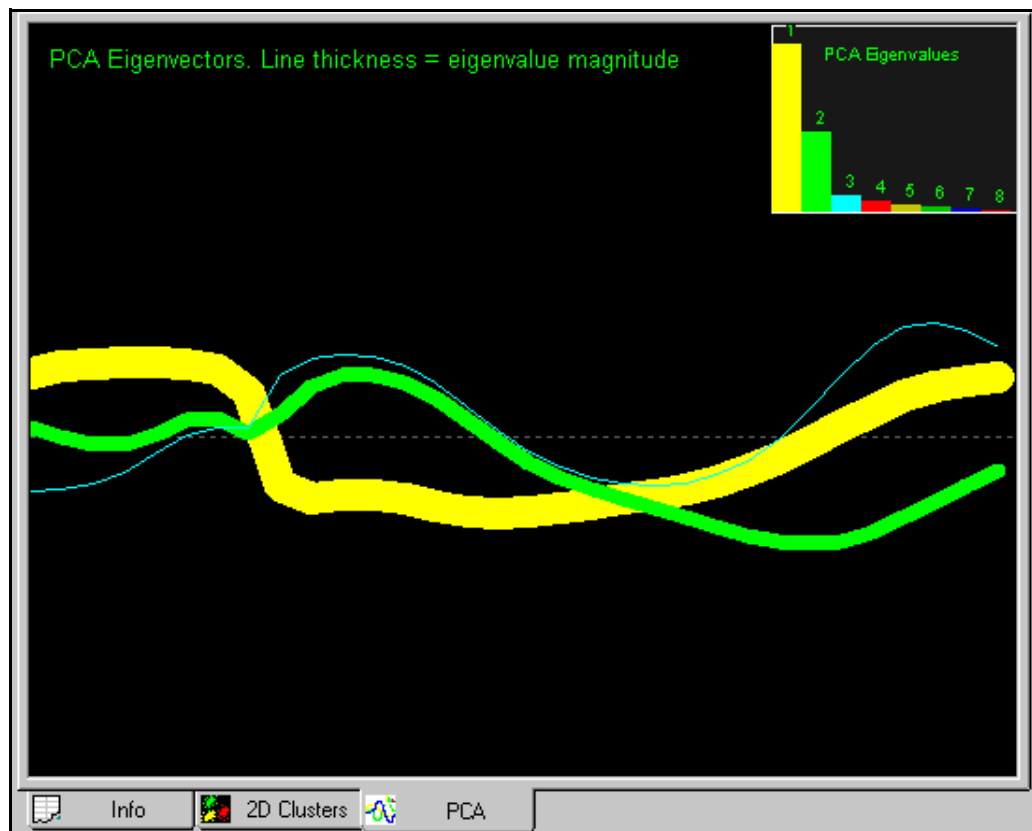
The PCA Results View graphically displays the results of the Principal Components Analysis (PCA) that runs on the waveforms. The calculated principle component eigenvectors form a basis onto which waveforms can be projected in order to get the PC1-PC8 feature values. A bar graph of the eigenvalues associated with each eigenvector is also displayed. The same PCA information is displayed in table form in the PCA tab of the [“Info View”](#) on page 3-13.

To display the PCA Results View,

- Select the **View | PCA Results View** menu command, or,

press the  button on the toolbar.

**Note:** If this button does not appear on your toolbar, select **Tools | Customize Toolbars...**, and select the line labelled **View** in the list box in the **Toolbars** tab. Do not uncheck the checkbox—just select the line. Then click the **Reset** button.



---

By default, Offline Sorter plots the first 3 principal component eigenvectors in the main portion of the view. Each different color corresponds to an eigenvector/eigenvalue pair. Offline Sorter calculates a maximum of 8 principal components eigenvalues/eigenvectors. You can display all 8 eigenvectors by selecting **Show All 8 Eigenvectors** from the right-click menu.

The width of the line used to plot the eigenvector is proportional to the magnitude of the corresponding eigenvalue. A bar chart graph of the eigenvalue magnitudes is shown in the upper right corner by default. You can move the graph to other corners by selecting **Move Eigenvalue Graph** from the right-click menu. The eigenvalue bars are scaled for display in the graph such that the first eigenvalue takes up 90% of the range. This graph is sometimes referred to as a “Scree Plot”.

The magnitude of an eigenvalue for a principal component eigenvector, relative to the total of all eigenvalues summed, represents the fraction of the total variance between waveforms that has been captured in that PCA dimension. In the above screenshot, it is clear that the first two principal components capture the vast majority of the variance. So, there is likely to be very little discriminatory power in the higher principal components PC3-PC8. The detailed shape of the PCA eigenvectors are not always straightforward to interpret. Generally, the first principal component (the yellow line) is largest in those sections of the waveform where the most variance is. The shapes of the higher principal component eigenvectors are constrained by the fact that all the eigenvectors must be orthogonal to each other, which makes their interpretation difficult. Note though that they approach zero at the trigger point (e.g. at a time value equal to the prethreshold into the waveform). All waveforms have an amplitude equal to the trigger threshold value at this point in the waveform, with essentially no variance. Thus, all the eigenvectors are zero at that point.

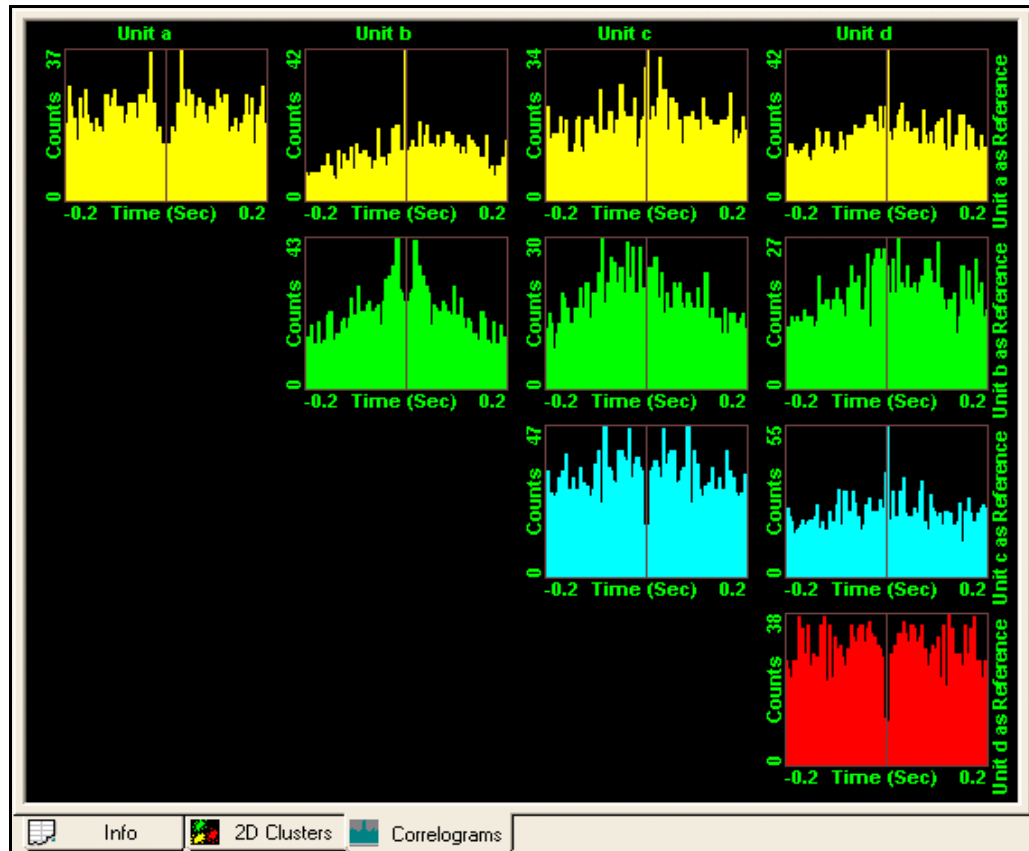
## Cross-Correlograms View

The Cross-Correlograms View displays a half-matrix of cross-correlograms between all units (and optionally the unsorted waveforms as well) in a channel. The firing of each unit is used as a reference time (time zero) to form a correlogram of a different unit’s firings, using an adjustable timescale. The cross-correlograms are arranged on a grid where the reference unit increases along the vertical axis and the plotted unit increases along the horizontal axis, with the redundant (mirror-image) correlograms omitted.

To display the Cross-Correlograms View,

- Select the **View | Cross-Correlograms View** menu command, or,

press the  button on the toolbar.



The parameters for the Cross-Correlograms are set by using the right-click context menu. The view can be configured to show the cross-correlograms formed by permuting either: all sorted units, all sorted units plus the unsorted waveforms, or just the currently selected units. The labeling can be enabled or disabled, which is useful if there are large numbers of units. Also, whether or not the reference unit firing is counted in the correlogram can be selected. The timescale and bin widths used to build the correlograms can be set using the 'Binning' section of the dialog that is displayed when 'Properties...' is selected from the context menu.

### Rasters View

The Rasters View displays color-coded tick marks that correspond to the times of unit firings, which is similar to the Continuous Data Views, but the Rasters View does so for all channels simultaneously.



---

To display the Rasters View,

- Select the **View | Rasters View** menu command, or,

press the  button on the toolbar.



In addition, the Rasters View can show the times when Events occur. The Events can appear either as tick marks on separate raster lines, or as lines overlaid across all rasters.

The Rasters View can only display the timestamps for channels that have been loaded into memory. When a channel is loaded, its timestamps are available for display. By default, the Rasters View only displays the timestamps for the channels that have been loaded. To load all timestamps for all channels into memory, select **Load All Rasters** from the right-click context menu. For .PLX files, you can automatically load the timestamps for all channels when the file loads by checking the **Automatically Show Rasters View and Pre-Load Timestamps** checkbox in the **Display** tab of the **Tools - Options** dialog. However, if you check this option, Offline Sorter uses more system memory and this slightly increases the time it takes to load .PLX files.

The Rasters View displays the same time range as the time range that appears in the Continuous Data Views, and the Rasters View responds to the **Zoom In** button, the **Zoom Out** button, and the time-range scroll bar on the Continuous Data Views. By default, the Rasters View displays all channels that have spikes. You

can change the number of channels displayed vertically by selecting **Zoom In To Show Fewer Channels** or **Zoom Out To Show More Channels** from the context menu, or by pressing the  $\uparrow$  or  $\downarrow$  keys. To scroll the display vertically, move the mouse wheel or left-click and drag the mouse in the label area on the left. You can also scroll the display vertically by using the `PAGE UP` and `PAGE DOWN` or `HOME` and `END` keys. You can also grab the dividing line between the labels and the tick marks and move it by left-clicking and dragging the line.

The Rasters View reflects the **Invalid w/f** setting in the Control Grid, however it does reflect the **View w/f** setting; the Rasters View always shows the tick marks for all units, and you can optionally show the unsorted waveforms by checking **Show Spike Marks For Unsorted** in the context menu. By default, each channel appears on a separate raster; however, if you uncheck **Show All Units for a Channel on One Raster** in the context menu, you can instead show a line for each sorted unit (and possibly also the unsorted units, which are denoted by a “U” postfix, e.g. “sig002U” is the unsorted waveforms on the “sig002” channel).

By default, any Events that you select to show, by clicking **Show** in the **Events** sub-tab of **Info View**, are plotted each on its own raster lines, in the color you select in the **Events** sub-tab. You can show the rasters for Events either on the top (default) or on the bottom of the spike timelines by checking or unchecking **Show Events on Top-most Rasters** the context menu. Alternatively, select **Show Events Across All Rasters** in the right-click context menu to plot the Events as vertical lines that cross all the raster lines. Also, you can select **Show All Events** in the context menu, which is equivalent to clicking **Show** for all Events.

## Continuous Data Window

The Continuous Data Window shows the data, either continuously digitized or waveform segments, in its proper temporal sequence.

This window can be used to:

- specify an optional digital filter and select a threshold and extract waveform segments for continuously recorded data
- control the temporal range of waveforms shown in the Waveform Window and Cluster views
- select individual waveforms (left-mouse click on waveform) for viewing in the Waveform Window and Waveform Inspection Window

---

There are two versions of the Continuous Data Window. One is for continuously-recorded data and the other is for spike waveform data.

### For Continuously-Recorded Data

If you open a channel with continuously recorded signals, the following version of the Continuous Data Window appears.



The Continuous Data Window consists of the following components:

- **Continuous stripchart display** – The vertical ticks above the waveform segments that exceed the threshold represent timestamps, which are color-coded to represent the assigned unit. You can select individual extracted waveform segments by left mouse clicking the waveform. When you select a waveform, it is highlighted in the other displays.
- **Blue Dotted Horizontal Lines** – These lines represent  $\pm 3$  standard deviations of noise. The noise is calculated by setting the Threshold to  $\pm 3$  standard deviations of the original signal and then removing waveform segments 250 msec before and 750 msec after the threshold crossings.
- **Detect Waveforms** – Click this button to extract waveform segments around threshold crossings
- **Detector and Waveform Options** – This item displays a dialog box for defining the Total Length, Prethreshold Period, Dead Time for the waveforms extracted, Signal to be thresholded (Raw, Energy, Nonlinear Energy), and the Energy Window Width.
- **Zoom** – This item represents the zoom factor for the voltage display, which is the same as the **Zoom** parameter in the Control Table. The voltage from zero to peak appears to the right of the **Zoom** control.
- **Thr. (%)** – This is the voltage threshold control for waveform extraction; the threshold voltage level is shown to the right of the **Threshold** control and it is represented by a solid horizontal line in the window display. You can adjust the threshold by dragging the horizontal line up or down with the mouse.
- **Low-cut filter** – This item selects the cutoff frequency for the 4th order low-cut Butterworth IIR filters

- **Control cluster display** – This item restricts waveforms described in the “Cluster Views” on page 3-19 to those contained in the **Scroll range** setting of the Continuous Data Window.
- **Control w/form display** – This item restricts waveforms shown in the [Waveform Window](#) to those contained in the scroll range of the Continuous Data Window.
- **Center Selected** – If checked, when you select a waveform using the [Cluster Views](#) it causes the time range displayed to change such that the selected waveform is shown in the center of the window.
- **Scroll bar, Zoom Out, Zoom In** – Use these controls to change the limits of the scroll range for the Continuous Data Window.
- **Scroll range** – The scroll window length and start time appear to the left of the scroll bar; the stop time appears to the right.

### For Spike Waveform Data

If you open the channel that contains extracted spike waveforms, the Continuous Data Display window has the following layout:



The Data Display window consists of the following components:

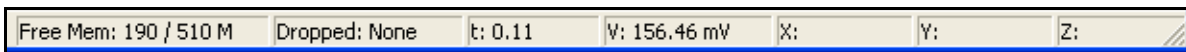
- **Pictured Waveform Segments** – The display shows waveform segments positioned at their corresponding timestamp. Colors represent the assigned unit. You can select individual waveforms for viewing in the Waveform Window and in the Waveform Inspection Window by clicking on the waveform with the left mouse button.
- **Control cluster display** – This item restricts waveforms shown in [Cluster Views](#) to those contained in the **Scroll range** setting of the Continuous Data Window.
- **Control waveform display** – This item restricts waveforms shown in the [Waveform Window](#) to those contained in the scroll range of the Data Display Window.
- **Show w/f within scroll range** – This item shows waveforms in the Data Display window that are contained within the scroll range [ScrollRangeMin, ScrollRangeMax].

- 
- **Show w/f from 0 to scroll pos.** – This item shows the cumulative display of waveforms in the Data Display window, from zero to the scroll position [0, ScrollRangeMax].
  - **Scroll Range** – This item displays the time range of the Data Display window in seconds.
  - **W/f Zoom** – This item controls the horizontal zoom (scaling) of the waveform segments.
  - **Y Zoom** – This item controls the vertical zoom for Continuous Data Display window. This item has the same effect as the zoom control in the control table.
  - **Center Selected** – If checked, when you select a waveform using the [Cluster Views](#) it causes the time range display to change such that the selected waveform is shown in the center of the window.
  - **Scroll bar, Zoom Out, Zoom In** – These controls change the limits of the scroll range of the Data Display Window.
  - **Scroll range** – The scroll window length and the start time appear to the left of the scroll bar; the stop time appears to the right.

### Other Features of the Continuous Data Windows

Right-clicking on a waveform brings up a menu that enables you to assign that particular waveform to the current unit, remove it from it's currently assigned unit, mark it as being invalid, or to use it as a template to define a new unit using the Templates sorting method (See [“Semi-Automatic Clustering using Templates”](#) on page 4-41.).

For both the waveform and continuous data displays, the status bar on the bottom of the screen shows the time and the voltage corresponding to the current mouse position within the display.





## Chapter 4

# Spike Sorting Step-by-Step

---

Procedure Summary.....	4-2
Waveform Acquisition From Continuously Recorded Data..	4-3
Waveform Acquisition From Spike/Waveform Data.....	4-10
Preparation For Sorting.....	4-12
Spike Sorting.....	4-27
Unit Cleaning.....	4-50
Examining Sorting Statistics.....	4-58
Saving the Sorting Results .....	4-62

---

## Procedure Summary

The spike sorting process with Offline Sorter involves the following general procedures:

- Waveform Extraction (can be omitted, depending on type of data files)
- Preparation for Sorting
- Unit (Cluster) Selection
- Waveform Inspection and Unit Cleaning
- Saving Sorted Spike Times

Offline Sorter is capable of sorting action potential waveforms (spikes) in either continuously digitized data, or in previously sorted spike/waveform data files (re-sorting).

For sorting spikes in Continuously Recorded data

- 1 Open the continuously digitized data file. See [“Opening the Digitized Data Files” on page 4-3](#).
- 2 For each of the channels that contain spike data:
  - Select the data channel. See [“Selecting the Continuous Data Channel” on page 4-4](#).
  - Specify the low-cut digital filter if needed. See [“Selecting the Low-Cut Digital Filter” on page 4-6](#).
  - Specify the waveform detection options. See [“Specifying the Waveform Detection Options” on page 4-6](#).
  - Set the threshold and extract the waveforms. See [“Extracting the Thresholded Waveforms” on page 4-9](#).
  - Go to [Step 3 on page 4-2](#).

For Re-sorting Spikes from Previously Sorted Waveform/spike Data files

- 1 Open the spike/waveform data file. See [“Opening the Spike/Waveform Data File” on page 4-10](#). Go to [Step 3 on page 4-2](#).
- 2 Convert Data Type if necessary. See [“Stereotrode and Tetrode Data” on page 5-4](#).
- 3 For each channel with spike waveforms:
  - Inspect the data. See [“Data Inspection” on page 4-12](#).
  - Remove artifact waveforms. See [“Marking Artifact Waveforms as Invalid using the Waveform Window” on page 4-16](#).
  - If necessary, select intervals of valid data. See [“Interval Selection” on page 5-27](#).




- 
- Check Waveform Alignment if necessary. See [“Waveform Alignment” on page 4-22](#).
  - Adjust Sort Width if necessary. See [“Adjusting the Sort Range” on page 4-25](#).
  - Select a Feature Space in which to perform sorting. See [“Selecting Features to Use For Sorting” on page 4-26](#).
  - Select units using one of the supported spike sorting methods. See [“Spike Sorting” on page 4-27](#).
- 4 For each sorted unit:
    - Clean the selection if necessary. See [“Cleaning the Selection” on page 4-54](#).
    - Inspect individual waveforms if necessary. See [“Waveform Inspection” on page 5-29](#).
  - 5 Save the timestamps as a NEX data file. See [“Saving Timestamps in a NEX File” on page 4-62](#), or save the timestamps and waveforms in a data file. See [“Saving Timestamps and Waveforms” on page 4-63](#).

## Waveform Acquisition From Continuously Recorded Data

### Opening the Digitized Data Files

Offline Sorter can sort waveforms from files containing continuously digitized data, such as that generated by the Plexon Recorder program (DDT file), other vendor’s files, or from other binary files that contain multiplexed A/D values recorded from several channels.

To open a continuously digitized DDT data file

- press File Open toolbar button , or, select **File | Open** menu command.

To open a continuously digitized data file from another vendor’s hardware

- Select **File | Import** menu command and select the vendor.

To open an arbitrary binary file with continuously digitized data (header plus multiplexed A/D values stored as 2-byte signed integers)

- 1 Select **File | Import | Binary file with Continuously Digitized Data** menu command.
- 2 Specify the file name, number of channels, data offset, digitization frequency, A/D conversion resolution, and maximum voltage in the dialog box.

**Note:** The binary file is expected to conform to the following specifications:

- The data file can contain a fixed-size header block at the beginning of the file. The size of this header block must be known

and entered into the dialog so that this part of the file can be skipped over. Zero is a valid size for the header.

- Once the continuously digitized data starts after the header, it cannot have any kind of “block structure” to it—it must contain just A/D values arranged in sweeps; each sweep contains a single A/D value from each channel.
- Each A/D value is stored in a signed 16-bit word, regardless of the A/D conversion resolution. If you check the **Swap Bytes** checkbox, Offline Sorter interprets the 16-bit words in “big-endian” (MSB first) format instead of the default Intel-standard “little-endian” format.

### Selecting the Continuous Data Channel

Since loading the data channel from a large file can require considerable time, Offline Sorter does not open any channel when you initially load the data file.

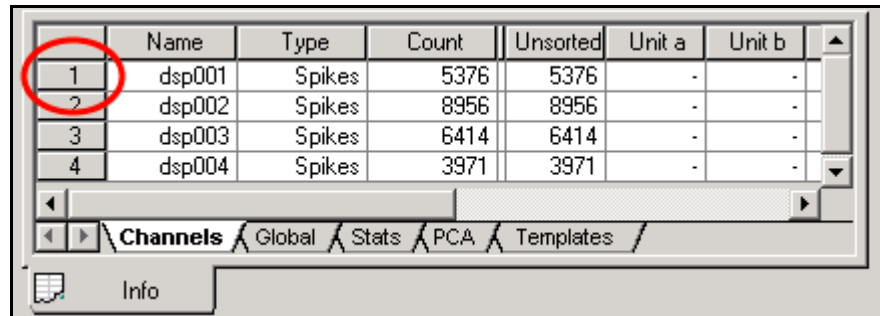
To select a particular data channel

- 1 Click on the **Channel** drop list box in the **Control Table** (shows **None** in the illustration that follows).
- 2 Select the channel from the drop list.

Channel	None
Data Type	Single
Sort Method	Contours
Scale (%)	100
Max w/f draw	10000
Sort Start	0
Sort End	800
Slice 1	0
Slice 2	0
ClustHist Max	100000
Outlier Thr.	1.30
View w/f	All
Cluster View	All
Draw Grid	<input checked="" type="checkbox"/>
Refr ISI (us)	1000
ISIHistMax(ms)	100
Use Range	<input type="checkbox"/>
Range from (s)	0.000
Range to (s)	100.000
Invalid w/f	Show

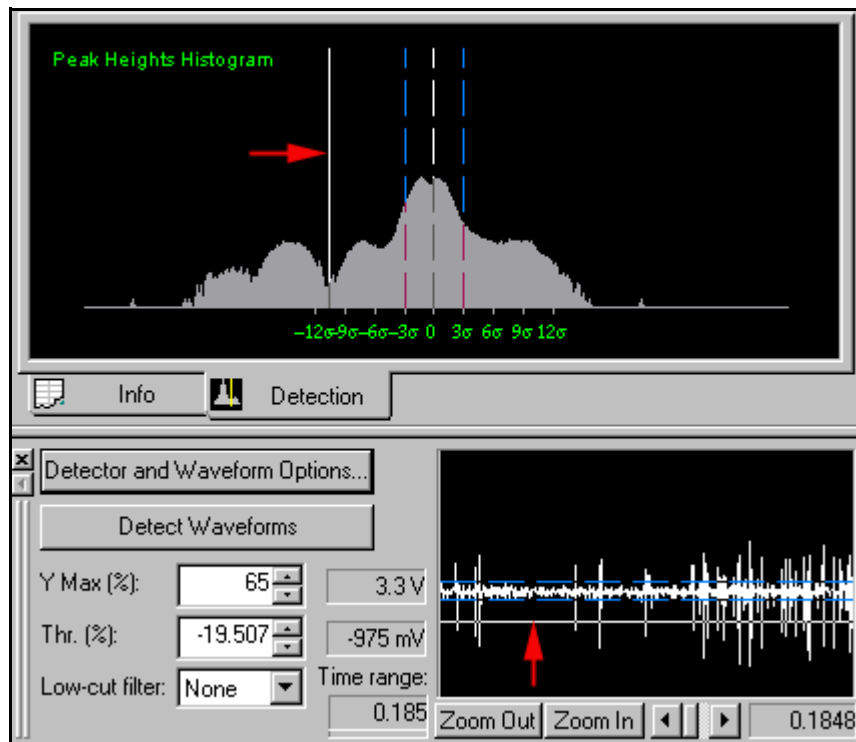
or

In the **Channels** sub-tab of the Info View, double-click on the row of the desired channel.



	Name	Type	Count	Unsorted	Unit a	Unit b
1	dsp001	Spikes	5376	5376	-	-
2	dsp002	Spikes	8956	8956	-	-
3	dsp003	Spikes	6414	6414	-	-
4	dsp004	Spikes	3971	3971	-	-

The Continuous Data Window then shows the channel's data, and the Waveform Detection View appears in the Multi-Display Window.



- Offline Sorter maintains a history of previously-opened channels. To revisit a previously-opened channel, click **View | Channel Back** or click **View | Channel Forward**. You can also use the following toolbar buttons:

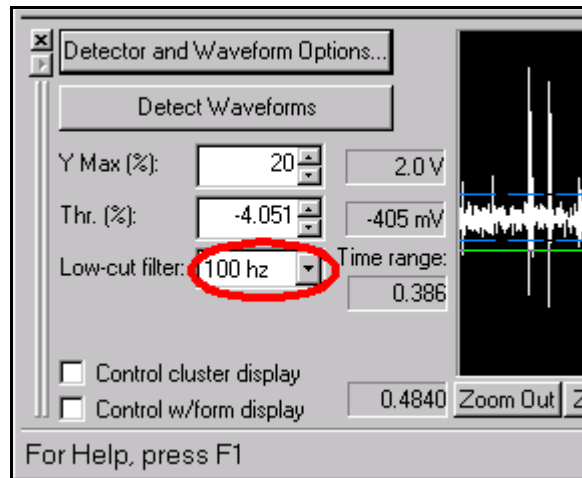


The default setting is the Raw Signal detector, in which case the Waveform Detection view shows a histogram of the logarithm of the Peak Heights. The white lines on the displays (red arrows) indicate the current threshold setting.

## Selecting the Low-Cut Digital Filter

Offline Sorter can filter the signal from an active channel by using the low-cut digital filters. The filters available in the program are the 4th-order Butterworth IIR filters.

To select a filter, click the arrow of the **Low-cut filter** menu, which is highlighted in the following illustration, and then select the cut-off frequency value from the drop-down list.



**Note:** When you change the filter it causes the histograms and the waveform displays to be redrawn. In the **Display** tab of the **View | Options** window, if you click **Auto Zoom**, the zoom also adjusts.

## Specifying the Waveform Detection Options

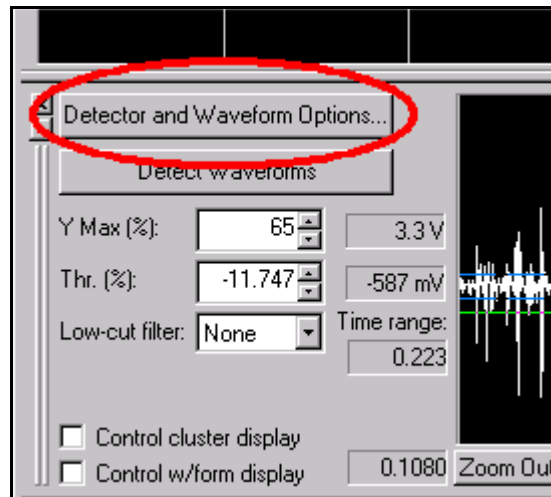
The **Total Waveform length** and **Prethreshold Period** parameters specify the time range of the waveform segments extracted from the continuous record.

The **Dead Time** is the minimum time between extracted waveform segments. For example, if the DeadTime = 1 msec and the recorded signal crosses the threshold at times  $t = 5000.0$  msec and  $t = 5000.8$  msec, then only one waveform segment aligned around the  $t = 5000.0$  msec crossing is detected and extracted. If the DeadTime = 0.5 msec however, then two waveform segments are extracted: one aligned around the  $t = 5000.0$  msec crossing and the other aligned around the  $t = 5000.8$  msec crossing.

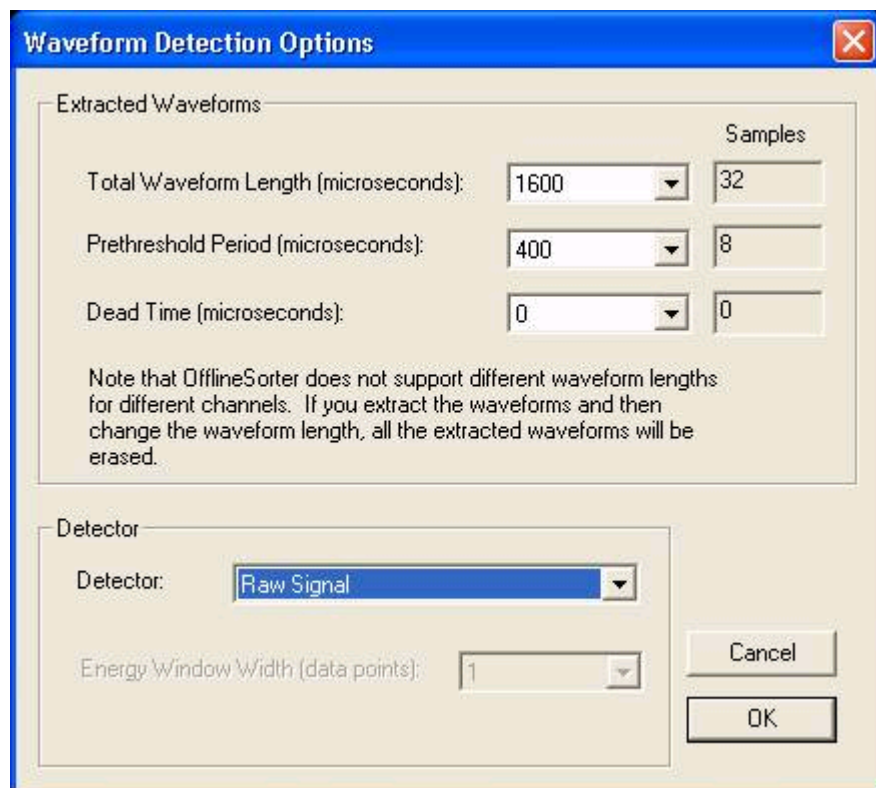
The **Detector** settings specify the quantity used in the threshold crossing algorithm for spike detection (See [“Waveform Detection View”](#) on page 3-38.). By default Offline Sorter uses the **Raw Signal** amplitude. If you select either **Energy of the Signal** or **Nonlinear Energy of the Signal** you can specify the Energy Window Width.

To set the waveform detection parameters

- Select the **Detector and Waveform Options** button in the Continuous Data Window.



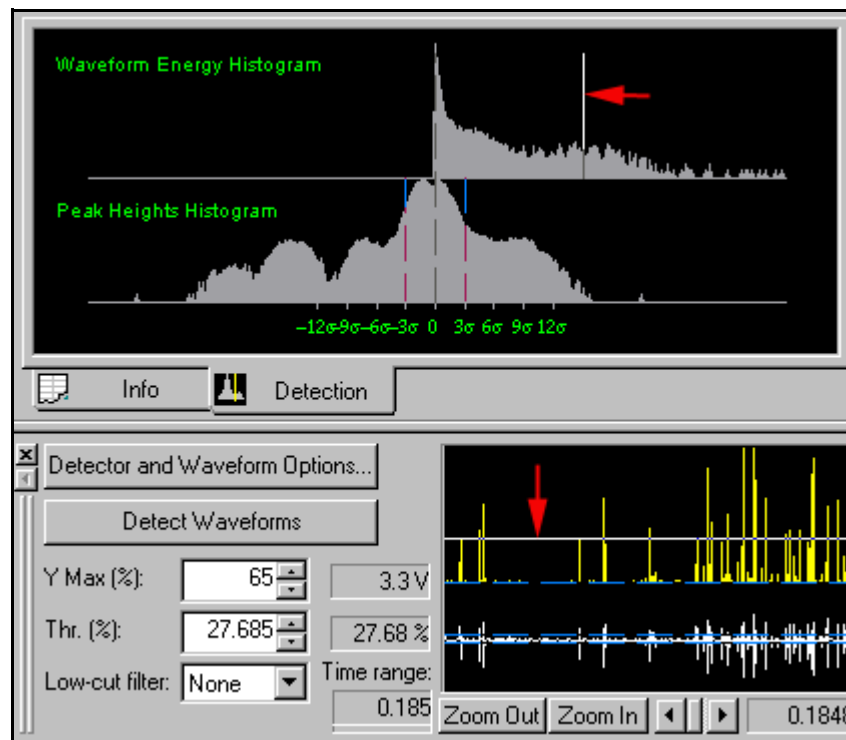
The **Waveform Detection Options** window opens.



The **Waveform Options** window contains the following items:

- **Total Waveform Length** – This item specifies the total length, in  $\mu\text{sec}$  and number of samples, of the thresholded waveforms that are extracted from the continuous data.
- **Prethreshold** – This item specifies the time, in  $\mu\text{sec}$  and number of samples, prior the threshold crossing included in the extracted waveforms.
- **Dead Time** – This item specifies the minimum time, in  $\mu\text{sec}$  and number of samples, between extracted waveform segments.
- **Detector** – The item indicates quantity used for thresholding, either **Raw Signal**, **Energy of the Signal**, or **Nonlinear Energy of the Signal**.
- **Energy Window Width** – This item specifies the width of the moving window over which the quantity is averaged (for energy and nonlinear energy detector options only).

If energy or nonlinear energy is selected as the signal to be thresholded in the **Detector** options menu, the quantity is displayed in the Continuous Data window and the histogram of the quantity (log axis) is displayed in the Waveform Detection view.



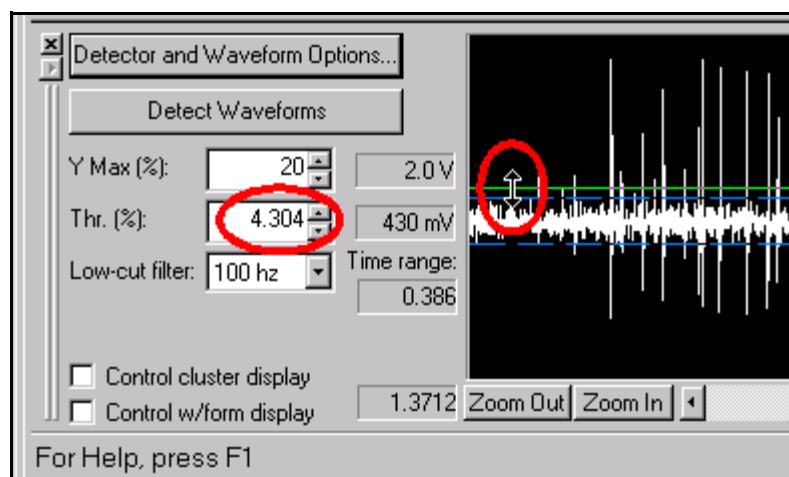
You can then adjust the threshold (red arrows) to different energy/nonlinear energy levels or histogram values.

## Extracting the Thresholded Waveforms

Offline sorter uses the following algorithm to extract the waveforms from continuously recorded signals:

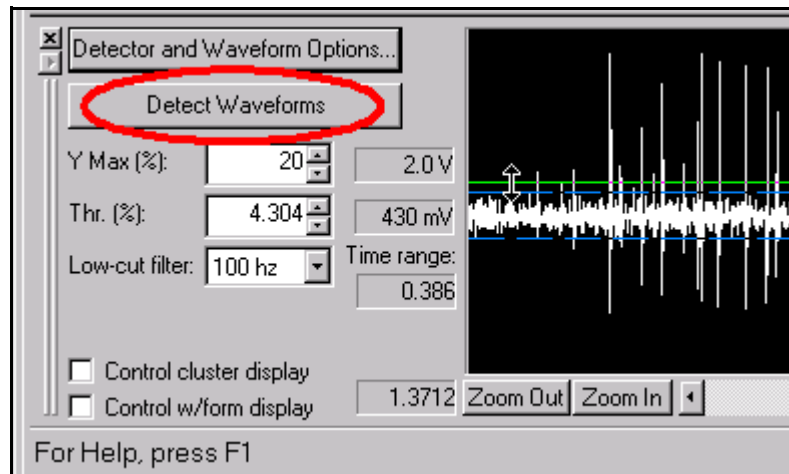
- If the threshold is *positive or zero*, the program looks for the moment when the signal crosses the threshold in the positive direction:  
  
find the first  $t$  such that  $\text{signal}[t - dt] < \text{threshold}$  and  $\text{signal}[t] \geq \text{threshold}$
- If the threshold is *negative*, the program looks for the moment when the signal crosses the threshold in the negative direction:  
  
find the first  $t$  such that  $\text{signal}[t - dt] > \text{threshold}$  and  $\text{signal}[t] \leq \text{threshold}$
- The time  $t$  then becomes the timestamp of the waveform. The program stores the specified number of data points before  $t$  and the number of points after  $t$  in the extracted waveform that is used later for sorting (See [“Specifying the Waveform Detection Options” on page 4-6](#) to see how to set the waveform length and the number of prethreshold points.).
- When the current waveform is extracted (that is, the specified number of data points after the threshold crossing is stored), the program moves ahead in time by the amount specified by the **Dead Time** parameter—see [“Specifying the Waveform Detection Options” on page 4-6](#)—and then it starts looking for the next waveform. This means that waveforms that cross the threshold less than the dead time after a detected spike are not detected.

To set the threshold, use the threshold edit control in the Continuous Data window or drag the horizontal threshold line with the mouse.



Alternately, to set the threshold, you can drag the vertical threshold line on the histogram in the Waveform Detection view to the left or right.

Click **Detect Waveforms** to extract the waveforms using the specified detector, total waveform length, prethreshold period, dead time, and threshold level.



## Waveform Acquisition From Spike/Waveform Data

### Opening the Spike/Waveform Data File

Offline Sorter can read native data files containing spike timestamps and spike waveform segments created by the Plexon MAP system (PLX) as well as other data acquisition systems (Bionic Technologies, CED Spike 2, DataWave, Instrutech, MultiChannel Systems, Panasonic MED, Neuralynx, RC Electronics).

To open a data file

- Click the **File Open** toolbar button,



, or  
select **File | Open** menu command.

To open a spike/waveform file from another vendor's hardware

- Select **File | Import** menu command and select the vendor.

Once the data file is opened, the waveforms are viewable and ready for sorting.

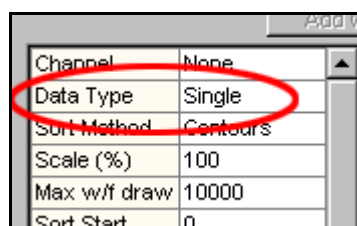
### Stereotrode and Tetrode Data Type Selection

Offline Sorter can sort files with signals from single electrodes, stereotrodes, or tetrodes (see also [“Stereotrode and Tetrode Data”](#) on page 5-4).



---

The **Data Type (Single/Stereotrode/Tetrode)** is displayed in the Control Table.



A screenshot of a software window titled 'Offline Sorter'. It contains a 'Control Table' with several rows. The 'Data Type' row is highlighted with a red oval and shows the value 'Single'. Other rows include 'Channel' (None), 'Sort Method' (Contours), 'Scale (%)' (100), 'Max w/f draw' (10000), and 'Sort Start' (0).

Channel	None
Data Type	Single
Sort Method	Contours
Scale (%)	100
Max w/f draw	10000
Sort Start	0

For Plexon PLX files, DataWave UFF files, and Neuralynx files, Offline Sorter automatically detects stereotrode or tetrode data when the file is opened, and sets the **Data Type** accordingly. For all other file types, Offline Sorter initially assumes single-channel data.

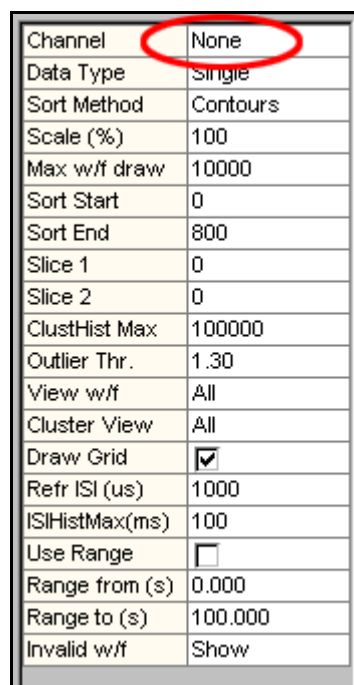
If the data type displayed in the Control Table is not appropriate for the data file opened, Offline Sorter can sometimes convert the data type. For a description of this procedure, see [“Converting Data Type” on page 5-5](#).

### Selecting the Data Channel

Since loading the data channel from a large file may require considerable time, Offline Sorter does not open any channel when you load the data file.

To select a particular data channel

- 1 Click on the **Channel** drop list box in the Control Table (shows **None** in the illustration that follows).
- 2 Select the channel from the drop list.

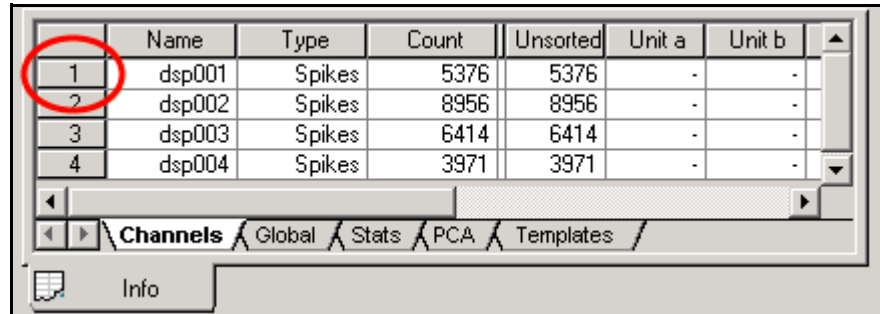


A screenshot of the 'Offline Sorter' window, showing a more complete 'Control Table'. The 'Channel' row is highlighted with a red oval and shows the value 'None'. Other rows include 'Data Type' (Single), 'Sort Method' (Contours), 'Scale (%)' (100), 'Max w/f draw' (10000), 'Sort Start' (0), 'Sort End' (800), 'Slice 1' (0), 'Slice 2' (0), 'ClustHist Max' (100000), 'Outlier Thr.' (1.30), 'View w/f' (All), 'Cluster View' (All), 'Draw Grid' (checked), 'Refr ISI (us)' (1000), 'ISIHistMax(ms)' (100), 'Use Range' (unchecked), 'Range from (s)' (0.000), 'Range to (s)' (100.000), and 'Invalid w/f' (Show).

Channel	None
Data Type	Single
Sort Method	Contours
Scale (%)	100
Max w/f draw	10000
Sort Start	0
Sort End	800
Slice 1	0
Slice 2	0
ClustHist Max	100000
Outlier Thr.	1.30
View w/f	All
Cluster View	All
Draw Grid	<input checked="" type="checkbox"/>
Refr ISI (us)	1000
ISIHistMax(ms)	100
Use Range	<input type="checkbox"/>
Range from (s)	0.000
Range to (s)	100.000
Invalid w/f	Show

or

In the **Channels** sub-tab of the Info view, double-click anywhere on the row of the desired channel.



The main views in Offline Sorter then show the channel's data.

- Offline Sorter maintains a history of previously-opened channels. To revisit a previously-opened channel, click **View | Channel Back** or click **View | Channel Forward**. You can also use the following toolbar buttons:



## Preparation For Sorting

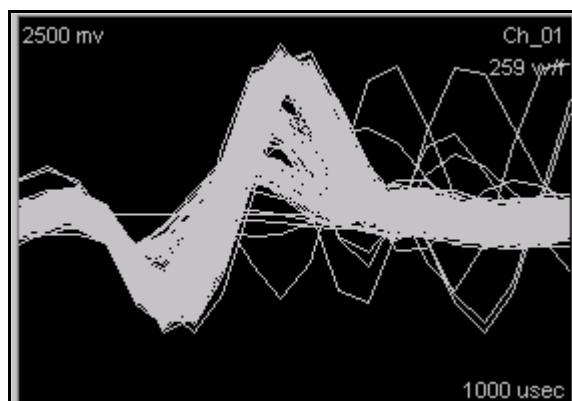
### Data Inspection

After the segmented waveforms (spikes) are acquired, they are displayed in the Waveform Window and the Cluster views. Offline Sorter allows the user to quickly get an overview of the data.

---

## Waveform Window

The Waveform Window shows the superimposed waveforms aligned at the threshold crossing times:



### HINT

#### Controlling Which Waveforms Are Displayed

The scroll bar at the bottom of the [Continuous Data Window](#) (use the **Control w/form display** check box) and the **View w/f** parameter in the [Control Table](#) can be used to control which waveforms are displayed in the Waveform Window. If you limit the waveforms displayed in this way it can often reveal structure that is hidden when all waveforms are drawn.

By default, Offline Sorter automatically determines the zoom factor for the Waveform Window and the Continuous Data View. You can disable the zoom factor by unchecking the **Auto Zoom** checkbox in the **Tools | Options | Display** dialog box. The auto-zoom calculation finds the peak amplitude present in any spike waveform, and sets the zoom to the largest integer value that displays the waveforms without clipping them. You can perform the Auto-zoom procedure at any time by selecting the **Waveforms | Auto-Zoom Waveform Displays** menu item.

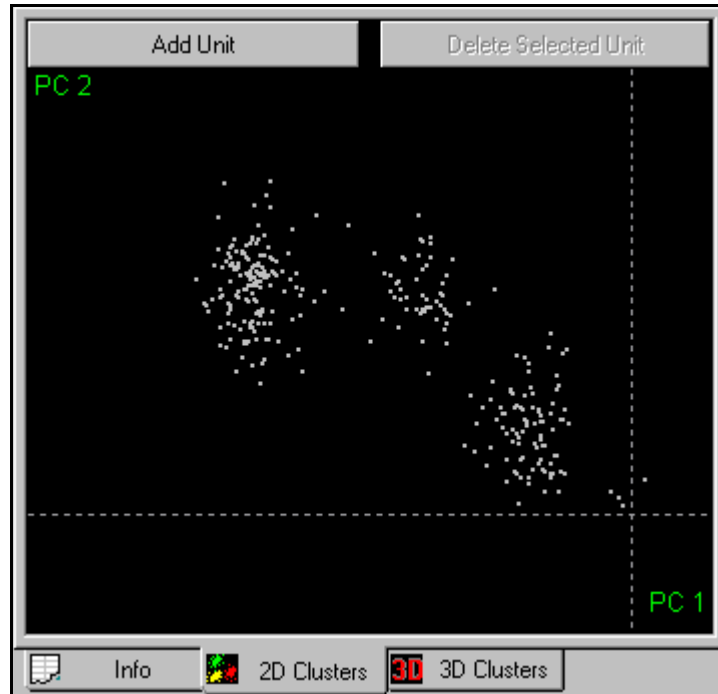
Offline Sorter reads spike waveforms from a file or detects them from continuous data, and stores them in memory. You can invert the waveforms stored in memory by using **Tools | Invert Waveforms** on the main menu. After an **Invert Waveforms** operation is done, all principal component and feature calculations use the inverted version of the waveform. If you select **Export to New .PLX**, Offline Sorter writes the inverted version of the waveforms to the output file. However, if you do a normal **Save** or **Save As** operation, Offline Sorter does not save the inverted waveforms.

## Cluster Views

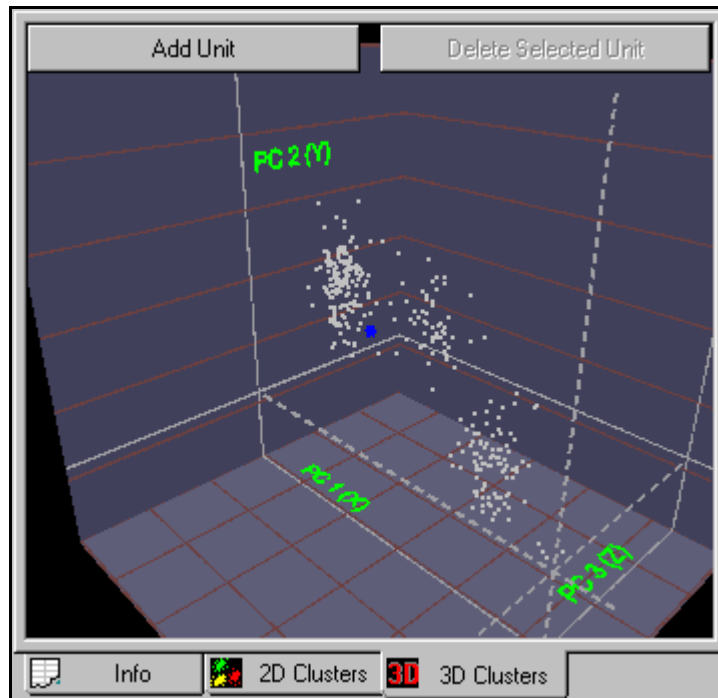
The “[Cluster Views](#)” on [page 3-19](#) show the waveforms as points in feature space. You can choose to use either the 2D Cluster view or the 3D Cluster view, or both. The variables shown on the X, Y and Z axes can be chosen from a list of active features. For more information, see the “Axis Feature Selection” section

under “[Cluster View Common Functionality](#)” on page 3-21. By default, the cluster views show the PC1, PC2, and PC3 features on the X, Y, and Z axes.

## 2D Cluster View



## 3D Cluster View



---

## Continuous Data Window

The Continuous Data window shows the waveforms in their temporal sequence. This is useful for determining if the data comes in “bursts” or if there are gaps in the data. Also, the times of occurrences of external events can be superimposed on the Continuous Data window (See [“Examining Events” on page 5-23.](#)), which enables you to see the spike activity surrounding those events.



## Invalidating Waveforms

Offline Sorter contains a mechanism to mark waveforms as being “invalid”, so that they are not considered in subsequent operations such as spike sorting. There are several reasons to do this:

- 1 Waveforms with shapes uncharacteristic of neuronal action potentials, such as those caused by stimulus or movement artifacts, can be marked as invalid and removed before spike sorting is performed. You can remove these waveforms by following the procedure in the [“Marking Artifact Waveforms as Invalid using the Waveform Window” on page 4-16.](#) In general, artifact waveforms are best invalidated, because they can significantly influence the principal components and the display ranges for the other views (invalidated waveforms are not used when auto-scaling and auto-zooming operations are performed by Offline Sorter).
- 2 Waveforms can be just “noise” that satisfied the threshold but that did not represent action potentials. The easiest way to invalidate these noise waveforms is to use the Cluster views, as outlined in the [“Marking Noise Waveforms as Invalid Using the Cluster Views” on page 4-17.](#) This step is optional; you can also just leave these noise waveforms as unsorted.
- 3 Another reason to mark waveforms as invalid is that they are simply not of interest. For example, if the experimental animal was not exhibiting “interesting” behavior during certain time intervals, the waveforms collected from those intervals can be invalidated simply to get them out of the way. Offline Sorter contains a tool that enables you to select waveforms within some interval in time around an Event (or between pairs of Events) in the data file. For more details, see [“Interval Selection” on page 5-27.](#)

**Note:** You can mark waveforms as invalid in either waveform space or in feature space.



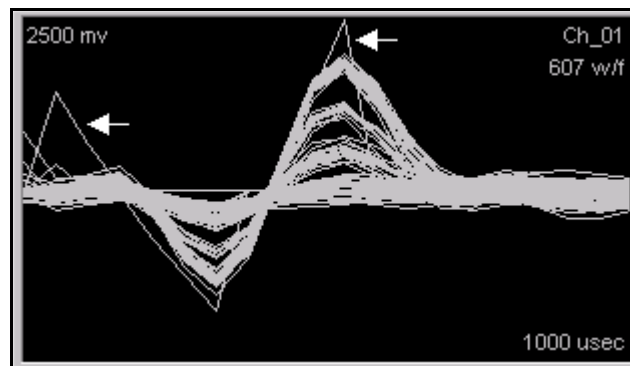
### CAUTION

#### Invalidated waveforms not marked

Invalidated waveforms may not be marked as such in the output file for some file formats (notably PLX). When the file is read in again, the invalidated waveforms will be valid again. However, the **File | Export to New .PLX** feature offers an option to not write invalidated waveforms to the output file.

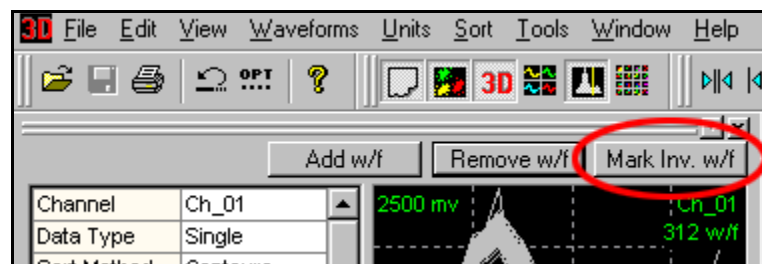
### Marking Artifact Waveforms as Invalid using the Waveform Window

In the Waveform Window shown below, two artifacts (arrows) can be clearly identified due to their sharp waveform shapes.




To mark the artifact waveforms as invalid and remove them from the display

- 1 Click **Mark Invalid w/f**.

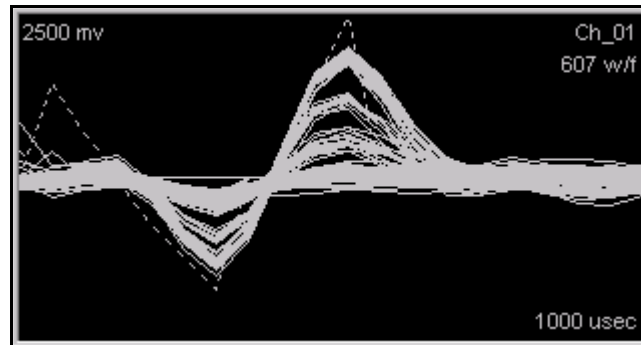


- 2 Move the pointer to the **Waveform Window**.

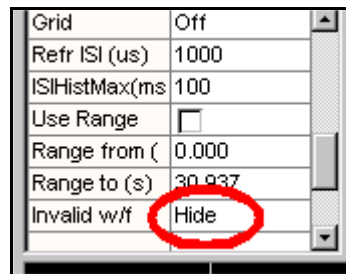
The pointer changes to a “drawing hand” .

- 3 Click the left mouse button and drag a line across the artifact waveforms.

*The invalidated waveforms appear as dashed lines.*



The invalid waveforms are ineligible for unit selection in any of the spike sorting methods. You can remove them from the Waveform window and the Cluster views by setting the **Invalid w/f** parameter in the Control Table to **Hide**.



You can undo the invalidation operation by using the undo mechanism, or by selecting the **Waveforms | Mark All Waveforms As Valid** menu command.

### Marking Noise Waveforms as Invalid Using the Cluster Views

Artifact waveforms generally appear as outlying points on the Cluster views. You can mark the points as invalid by using a procedure similar to that described in this section.

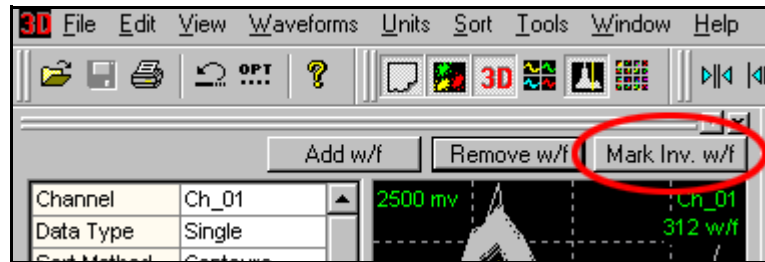
Another reason to mark waveforms as invalid is because they are due to “noise” that does not represent a valid neural action potential. This section describes how to invalidate noise waveforms when the 2D Cluster view is showing PC1 vs. PC2. You can use the same procedure in the 3D Cluster view and in any feature space.

The principal components of the extracted waveforms represent linear combinations (linearly independent projections) of the waveforms which account for maximal variance in the waveform data. Maximal variance implies maximal separation of the noise cluster from the unit clusters and the unit clusters from one another. For low amplitude noise, the noise cluster is centered close to the origin (0,0) in principal component space.


If the noise cluster is localized near zero and isolated, it can be removed from the principal component calculations so that the new principal components maximally separate the remaining unit clusters from one another. You can then select the maximally separated unit clusters by using the contour method.

To invalidate the noise waveforms near zero

- 1 Click **Mark Invalid w/f**.

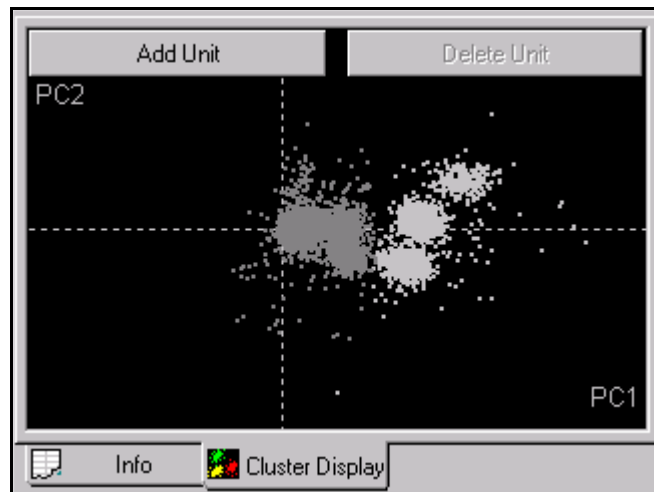


- 2 Move the mouse pointer to the Cluster view.

The pointer changes to a “drawing hand” .

- 3 Click the left mouse button and drag a line around the invalid “noise” waveforms near zero.

In the following 2D Cluster view, the noise cluster near the origin is invalidated (gray).



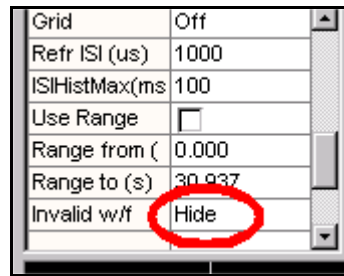
**Note:** You can also use the 3D Cluster view for the following procedure.



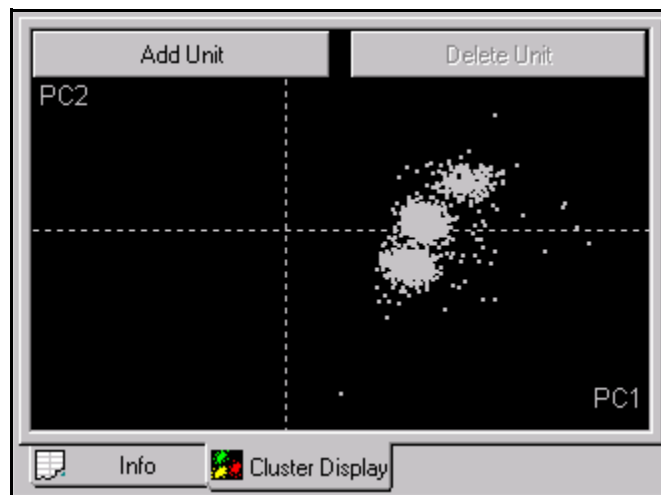
---

To remove the invalid waveforms from the display

- Set the **Invalid w/f** parameter in the Control Table to **Hide**.

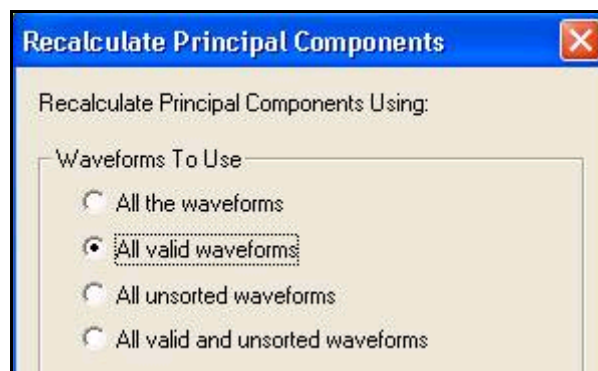


*The invalid waveforms disappear from the display.*

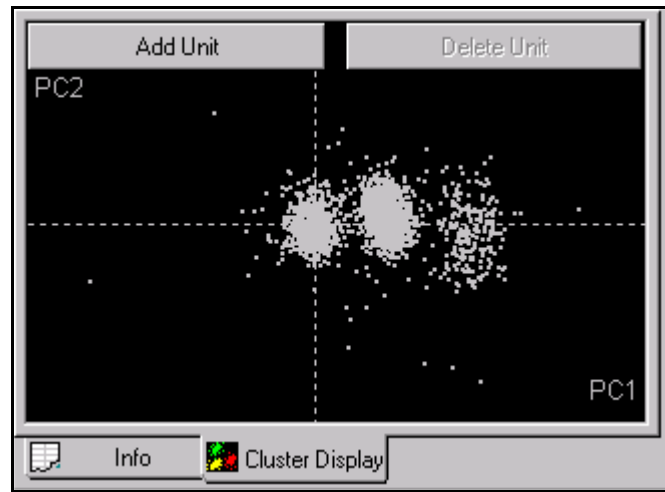


To recalculate the principal components using only the valid unit cluster data:

- Select the **Tools | Recalculate the Principal Components** menu command and rerun the principal component analysis using **All valid waveforms**.



The resulting 2D Cluster view can resemble the following one.



The three unit clusters (white dots) are now slightly better separated from one another than they were when the noise cluster was included in the PCA calculations.

**Note:** You can undo the invalidation operation by selecting the **Waveforms | Mark All Waveforms As Valid** menu command.

### Invalidating All Except

The preceding invalidation procedures designated the waveforms to be *invalidated*. You can “invert” that process to designate the waveforms to be *retained* by using the **Waveforms | Invalidate All Except Selected Waveforms** setting. When you select this setting, all waveforms that are not selected are invalidated. This selection applies in both the Waveform View (cross the waveforms to be *retained*) and in the Cluster views (circle the waveforms that are to be *retained*).

### Automatically Invalidating Artifacts

Offline Sorter includes a tool that automatically finds and invalidates artifacts that occur simultaneously (or nearly so) across several channels. The Invalidate Artifacts tool only uses spike times. The tool does not consider waveform shapes or amplitudes.

To use the Invalidate Artifacts tool, you must set two criteria:

- 1 a time width that ranges downstream from the first artifact threshold crossing point on any channel
- 2 a presence criterion that sets the number of channels on which a spike must be simultaneously present within the time range, to be identified as an artifact

When the tool determines that an artifact is present at time  $T$  on a sufficient number of channels, the tool invalidates all waveforms that are present on all channels between the times  $T_{\min}$  and  $T_{\max}$ , where  $T_{\min}$  and  $T_{\max}$  are the minimum and

---

maximum time stamps of the spikes in the artifact. The tool runs on all channels, not just the currently loaded channel.

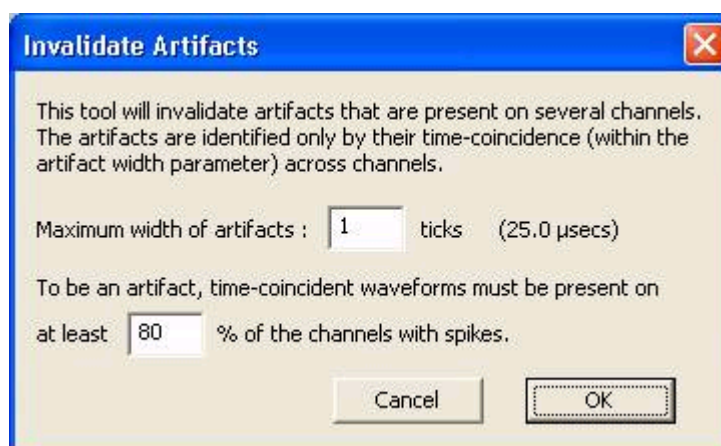
To be identified as artifacts, *simultaneous* spikes must be present on at least a critical number of channels. You set the presence criterion in percent multiplied by the number of channels with spikes. The Invalidate Artifacts tool completely ignores any channels that have 0 spikes in them and assumes they are inactive or disabled channels.

Because artifact waveforms can cross different thresholds at slightly different times, different threshold settings on different channels effect the timestamps of artifacts. The time width setting provides a way to bracket timestamps so that artifacts are still recognized with different thresholds. You can set the channel presence criterion to take into account the fraction of channels that contain high-quality, analyzable data, as opposed to occasional noise triggers.

Using automatic artifact invalidation

- 1 Click **File | Invalidate Cross-Channel Artifacts**.

*The **Invalidate Artifacts** window opens.*



- 2 In the **Maximum width of artifacts** box, type the maximum time width in ticks (units of 25.0  $\mu$ s) of the artifacts to be invalidated.

**Note:** For example, if the timestamps on two channels differ by this number of ticks or less, the tool considers the spikes on the two channels to be simultaneous. If you enter 0 (zero), the tool only invalidates artifacts with identical timestamps.

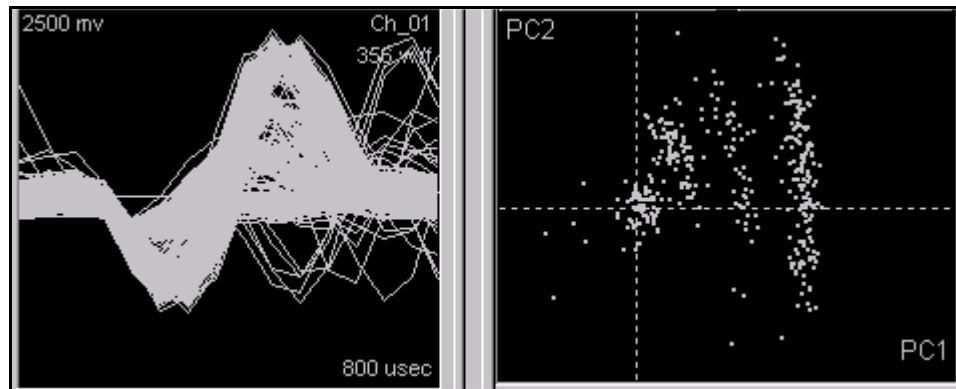
- 3 In the **% of the channels with spikes** box, type the channel presence criteria in percent. For example, in the illustration in Step 1, the tool invalidates

artifacts of the specified width and with simultaneous timestamps that are present on 80% of the channels with spikes.

- 4 To start the tool, click **OK**.

### Waveform Alignment

After detection and extraction, the waveforms are initially aligned at the point of threshold crossing in the Waveform Window (below left). The principal components are calculated using the threshold-aligned waveforms. You can select the projections onto the principal components in the 2D Cluster view, as follows in the right-side illustration.

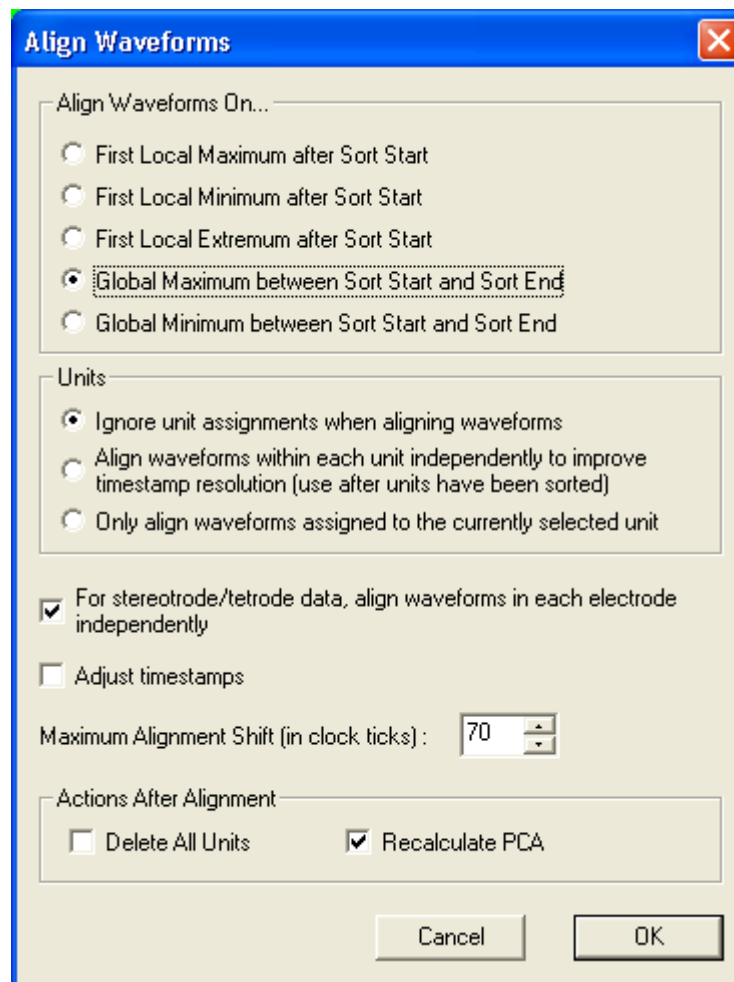


The variance within waveform groups can frequently be reduced, and the separation between waveform groups enhanced, by aligning the waveforms at their minimum or maximum values.

To set the alignment criteria

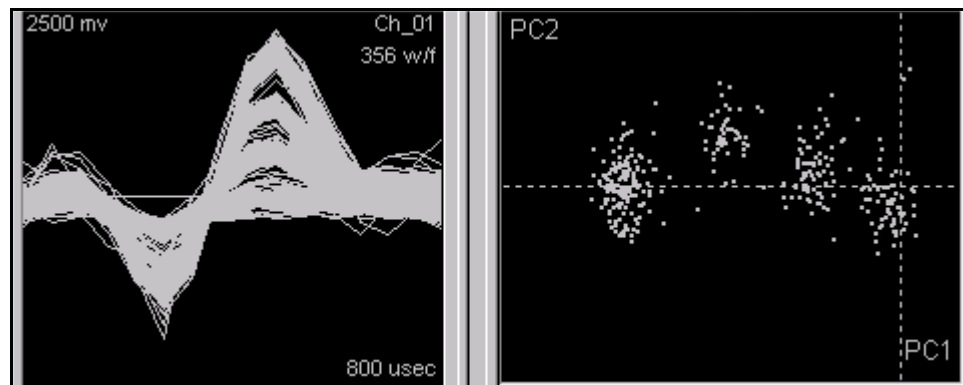
- 1 Select the **Waveforms | Align** menu command,  
or,  
select the **Tools | Align Waveforms** menu command.

- 2 Then select the desired alignment method.



Under most circumstances, aligning around the **Global Maximum** or **Global Minimum** provides the best alignment of the waveforms. Optionally, all units can be deleted (the **Delete All Units** checkbox) and the PCA recalculated (**Recalculated PCA** checkbox) when the waveforms are realigned.

After alignment around the global minimum (as follows), note how the waveforms are more tightly clustered in waveform space and principal component space.



By default, unit assignments are ignored when performing alignment (the **Ignore unit assignments when aligning waveforms** radio button is checked). You can check the **Align waveforms within each unit independently** radio button after the units have been sorted. This actually performs N independent alignment operations, where N is the number of units. For each alignment, only the waveforms in a given unit are considered when choosing the alignment center for the operation. This has the effect of adjusting timestamps to correct for triggering jitter.



### CAUTION

#### View reports with caution

Use the **Align waveforms within each unit independently** setting with caution! By its nature, it appears to “tighten up” the clusters. However, this is a bit misleading because it was done with a-priori knowledge of the clusters. After performing this operation, view the probabilities and F-statistics reports with caution.

The **Only align waveforms assigned to the currently selected unit** option is similar to the previous option in that it only performs the alignment on the current primary selected unit.

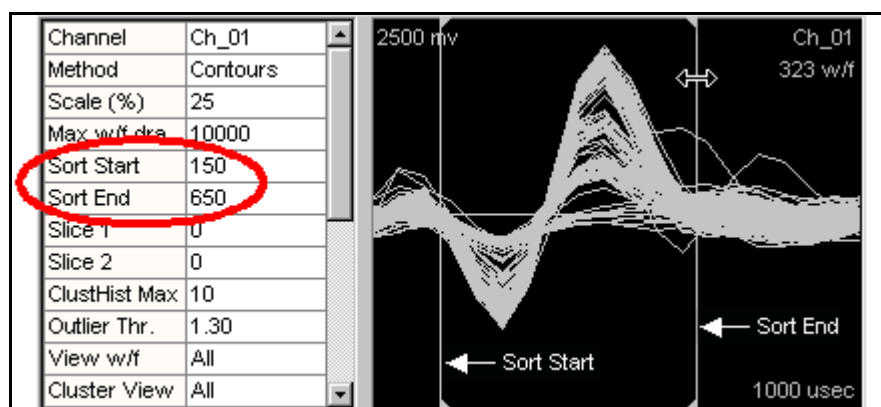
The **For stereotrode/tetrode data, align waveforms in each electrode independently** checkbox is only relevant when working with stereotrode/tetrode data. When unchecked, the concatenated waveform (See “[Stereotrode and Tetrode Data](#)” on page 5-4.) is aligned as a whole—the min or max of the entire concatenated waveform is found and the concatenated waveform is shifted as a whole. When checked, the alignment is instead performed per-electrode. For example, for tetrode data, four independent alignments are performed, one for each electrode. Note that the waveform in each electrode can be shifted a different amount.

The **Adjust Timestamps** checkbox controls whether or not the timestamp of the waveform is shifted by the amount the waveform is shifted.

The **Maximum Alignment Shift** adjustment enables you to put a cap on the number of clock ticks that any waveform is shifted. It is seldom desirable to shift a waveform by more than a few clock ticks, because zeros are added onto the end of the waveform to backfill the shifted amount. Waveforms that would be shifted by more than the specified number of clock ticks are not shifted at all.

## Adjusting the Sort Range

Often the extreme tail and the earliest samples of the waveforms are not useful in separating units from each other or separating the units from noise. To specify the range of samples within the waveform that is used to calculate the Principal Components (See [“Principal Components”](#) on page 5-11.) and used in feature calculation and template sorting, adjust the **Sort Start** and **Sort End** parameters in the Control Table; see the following illustration to locate the parameters. When they are visible, you can adjust the **Sort Start** and **Sort End** times by dragging the corresponding vertical line with the mouse; see the double-ended arrow in the following illustration.

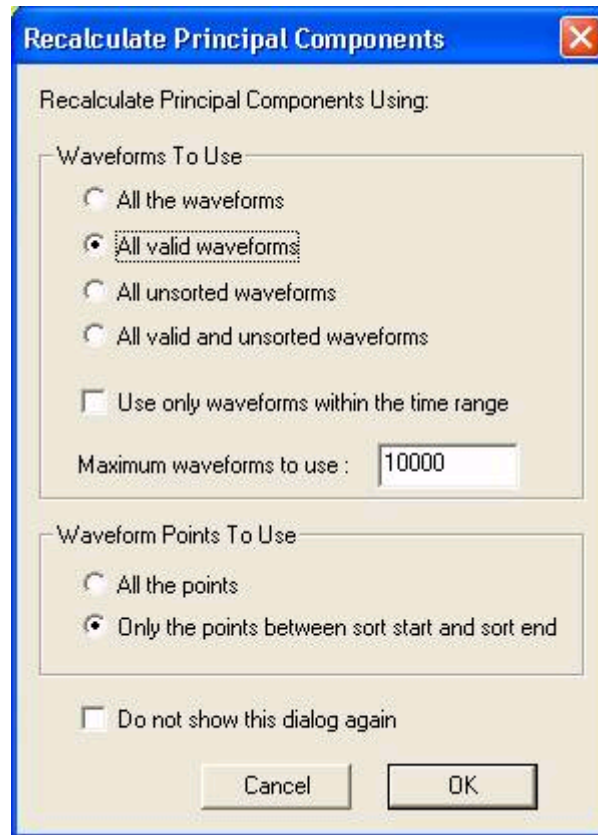


The **Sort Start** and **Sort End** time point is the lower and upper bound, respectively, of the time interval of the waveform used to calculate principal components, feature values, and for template sorting. The **Sort Start** and **Sort End** times are represented by vertical lines in the Waveform view; see the arrows in the preceding illustration. Small arrows on the top and bottom of the lines indicate whether it's the start or end. In the previous example, the total waveform length is 1000 msec, the **Sort Start** time is 150 msec, and the **Sort End** time is 650 msec.

**Note:** For stereotrode and tetrode data, the sort range is specified per-electrode, not for the concatenated waveform.

To recalculate the principal components using the time range between the **Sort Start** and **Sort End** time points, select the **Tools | Recalculate Principal Com-**

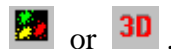
**ponents** menu command and recalculate using **All the points** and **Only the points specified by sort start and sort end**.



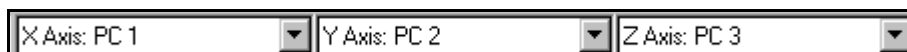
### Selecting Features to Use For Sorting

For those sorting methods that use feature space (i.e. all methods except waveform crossing and templates), you must first decide which feature space to work in. You do this by selecting particular features for each of the 2 (for the 2D Cluster view) or 3 (for the 3D Cluster view) axes in feature space (See [“Features Available for Sorting” on page 5-9.](#)). Offline Sorter does all spike sorting and calculates all statistics in the feature space you select, so your objective is to find a combination of features that gives maximal cluster separation. The description of the [“Feature vs. Feature View” on page 3-42](#) can be useful in selecting which features to use.

First, select either the **2D Cluster View** or the **3D Cluster View** using the **View** menu command or the toolbar buttons



Then, select features for the X, Y, (and Z) axes from the drop-down lists on the toolbar:





---

**Note:** Only the “active” features appear in the drop-down lists, which are a subset of all features. Use the **Active Features** tab of the **Tools | Options** dialog box to select which features are active. For more information, see [“Cluster View Common Functionality” on page 3-21](#).

When switching channels within the currently-opened data file, the selected features can either be “remembered” on a per-channel basis, or the same feature space can be used across all channels. You can control this by the **Remember selected features for each channel** checkbox in the **Tools | Options | Feature Calc** dialog box.

If you use the sorting algorithms, remember that they work differently in 2D and 3D feature space. The feature space that the algorithms use depends on which view you have active when the sorting operation initiates. See [“Feature Space Used for Sorting” on page 5-13](#).

## Spike Sorting

Finally, the waveforms can now be sorted into units. Offline Sorter offers several methods to accomplish this.

### Manual, Semi-Automatic, and Automatic Clustering Methods

Offline Sorter offers manual, semi-automatic, and automatic methods to identifying units or clusters. On a given channel, you can use any or all of these methods.

**Manual Clustering Methods.** With these methods, you can manually specify the unit assignment for each waveform.

- [“Manual Clustering using the Waveform Crossing Method” on page 4-36](#)
  - Waveforms are manually selected as belonging to a unit or cluster by crossing a bundle of waveforms in the Waveform window with the mouse.
- [“Manual Clustering Using Contours” on page 4-29](#)
  - Waveforms are manually selected in feature space in the Cluster views by drawing an arbitrary shape around a visible cluster.

**Semi-Automatic Clustering Methods.** With these methods, the user specifies a number of units, and “starting” cluster centers, and the algorithms then assign waveforms to the units.

- [“Semi-Automatic Clustering using Templates” on page 4-41](#)
  - The user selects an individual waveform to serve as the template for the unit, waveforms are then assigned to clusters using a template matching algorithm.
- [“Semi-Automatic Clustering using K-Means” on page 4-42](#)
  - The user selects tentative cluster centers in feature space in the Cluster views and then runs the K-Means algorithm which assigns points to clusters automatically.

- [“Semi-Automatic Clustering using Standard E-M ” on page 4-45](#)
  - The user selects tentative cluster centers in feature space in the Cluster views and then runs the Standard E-M algorithm which assigns points to clusters automatically.

**Automatic Clustering Methods.** With these methods, the number of clusters is decided algorithmically, and the waveforms are assigned to the clusters automatically.

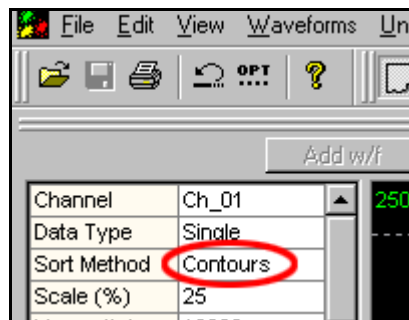
- [“Automatic Clustering using Valley-Seeking Method” on page 4-46](#)
  - The number of clusters and cluster membership is automatically determined via a valley-seeking algorithm operating on nearest-neighbor distances in feature space.
- [“Automatic Clustering using T-Distribution E-M Method” on page 4-49](#)
  - The number of clusters and cluster membership is automatically determined via an E-M algorithm based on the T-Distribution.

For the sorting methods that work in feature space, any of the active features can be assigned to the 2 or 3 axes of feature space and used in the sorting algorithms. See [“Selecting Features to Use For Sorting” on page 4-26](#). In the examples that follow, the feature space axes are generally left to the defaults values (PC1, PC2, and PC3 if using the 3D Cluster view).

You can re-sort the waveforms after the units have been selected by using the manual or automatic clustering procedures.

### Selecting the Sorting Method

This can be done either by selecting the **Sort Method** parameter in the Control Table,



or, you can also use the **Sort | Change Sort Method** menu command to select the method, or you can use the corresponding buttons on the Toolbar. In general, you use only *one* of the sort methods that are described in the following sections.

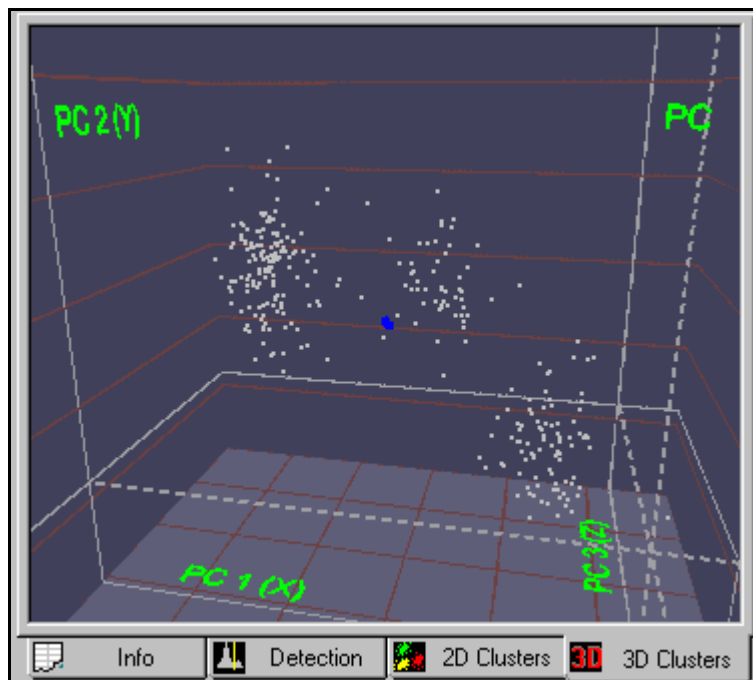
---

## Manual Clustering Using Contours

Select the Contour method of sorting by selecting **Contours** in the **Sort Method** drop list in the Control Table, or by selecting the **Sort | Change Sort Method | Use Contours to Define Clusters** menu item.

For details on how to use the Cluster views, see [“Cluster Views” on page 3-19](#).

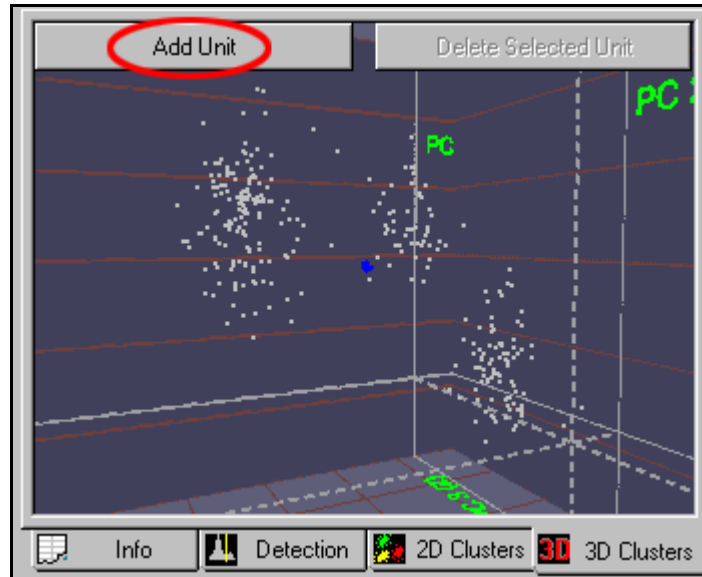
If the 3D Cluster view is being used, orient the eyepoint as described in [“3D Cluster View Details” on page 3-31](#) so as to maximally separate the clusters.



To add a unit using the Contour sorting method

**Note:** Adding a unit is to identify a subset of waveforms that are presumably generated by the same neuron.

- 1 In the Cluster view, click **Add Unit**,




or, select the **Units | Add Unit** menu command,

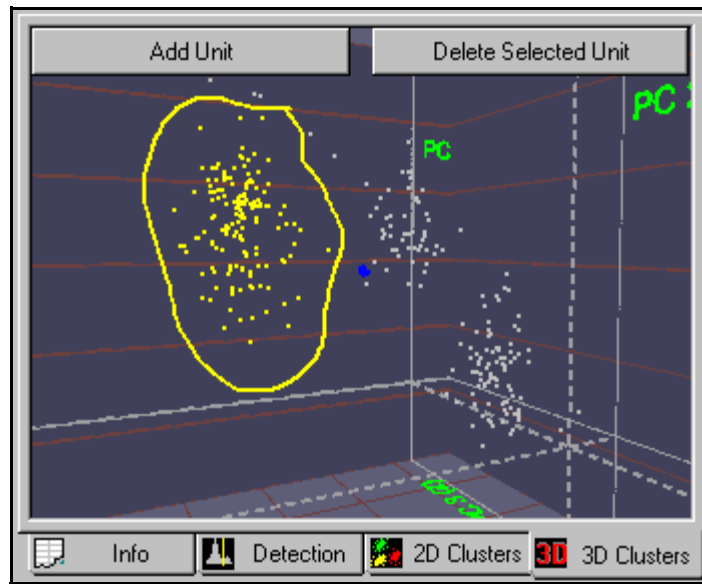
or, hold down the **ALT** key and right-click in the Cluster view and select the **Add Unit** menu command,

or, hold down the **CTRL** key.

- 2 Move the mouse pointer to the Cluster view.

The pointer changes to a “drawing hand”  .

- 3 Click the left mouse button and drag a line around the cluster.



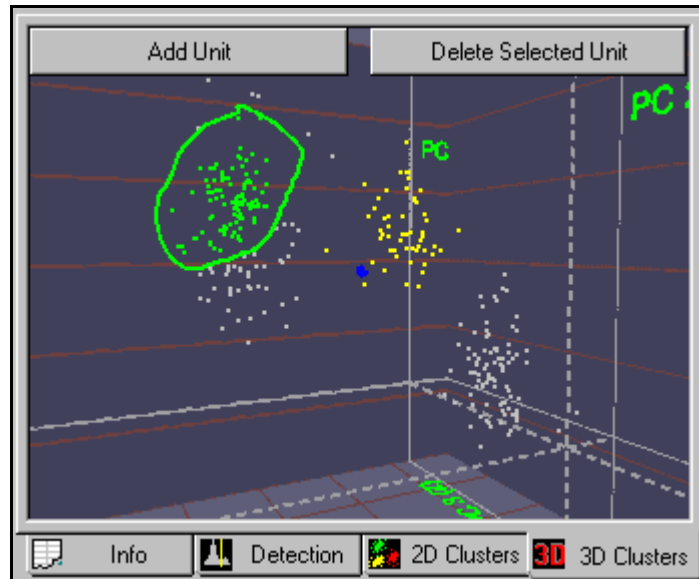
**Note:** In the 3D Cluster view, contours are drawn in “screen space”—the projection of the 3D feature space points onto the 2D screen. So the contour sorting operation is performed on a 2D projection of the 3D feature space. Up to 26 units can be sorted for each data channel.

### Adjusting the Sorting

After any sorting (not just sorting performed using the Contours method), Offline Sorter can be used to adjust the sorting by adding or removing individual or groups of waveforms from clusters, or deleting units entirely.

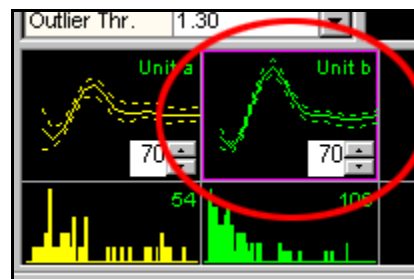
**Adding Waveforms to Selected Units in the Cluster Views.** You can add waveforms to existing units. The following illustration shows an existing unit,

which is outlined in green. Use the procedure that follows to add additional waveforms to the unit.

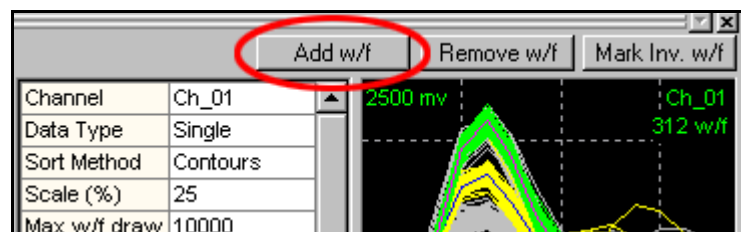


To Add Waveforms to a Unit

- 1 Select the unit by clicking it's template in the Units Display.




- 2 Click **Add w/f** above the Waveform window,

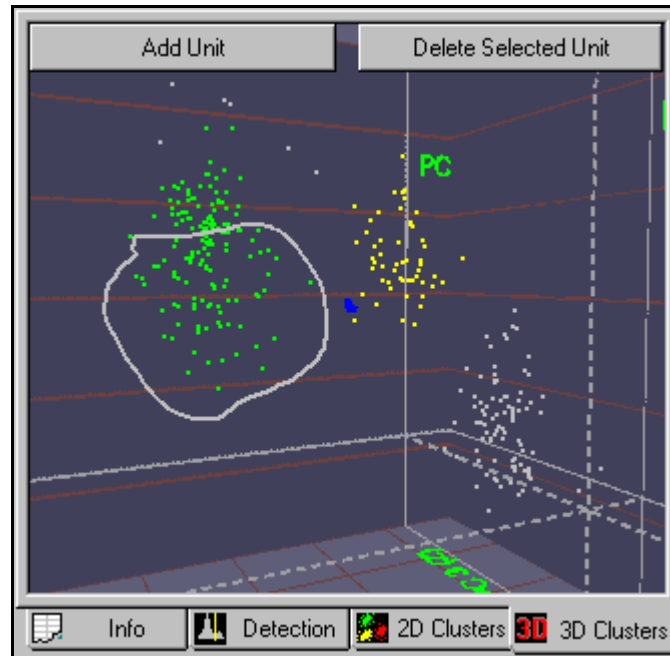


or, Select the **Units | Add Waveforms to the Selected Unit** menu command.

- 3 Move the mouse pointer to the Cluster view.

The pointer changes to a "drawing hand" .

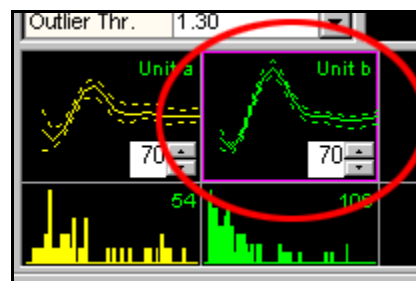
- 4 Click the left mouse button and drag a line around the waveforms to add.  
*The circled waveforms are added to selected unit.*



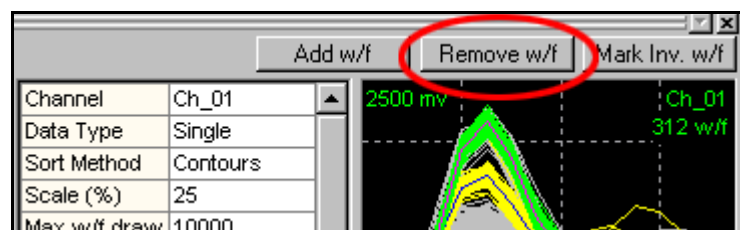
**Removing Waveforms from Selected Units In The Cluster Views.** You can also trim waveforms away from existing units.

To Remove Waveforms from a Unit

- 1 Select the unit by clicking it's template in the Units Display.




- 2 Press the **Remove w/f** button above the Waveform window,



or, select the **Units | Remove Waveforms from the Selected Unit** menu command.

- 3 Move the mouse pointer to the Cluster view.

*The pointer changes to a “drawing hand”* .

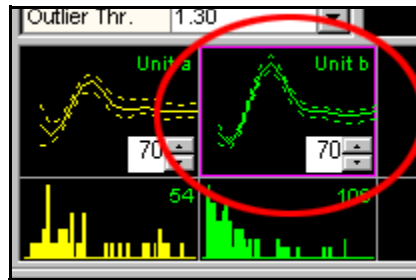
- 4 Click the left mouse button and drag a line around the waveforms to remove.

*The circled waveforms are removed from the selected unit.*


**Retaining Waveforms In Selected Units In The Cluster Views.** Offline Sorter also includes a reciprocal function that retains the selected waveforms and removes all the unselected waveforms. This function is similar to previous removal function, except that all waveforms NOT circled are removed from the cluster.

To Remove All Unselected Waveforms from a Unit

- 1 Select the unit by clicking it's template in the Units Display.



- 2 Select the **Units | Remove All Except Selected Waveforms from the Selected Unit** menu item, or click the corresponding toolbar button.
- 3 Move the mouse pointer to the Cluster view.

*The pointer changes to a “drawing hand”* .

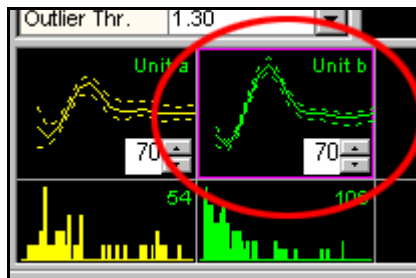
- 4 Click the left mouse button and drag a line around the waveforms to retain.

*The uncircled waveforms are removed from the selected unit.*

**Deleting Units.** You can remove units entirely.

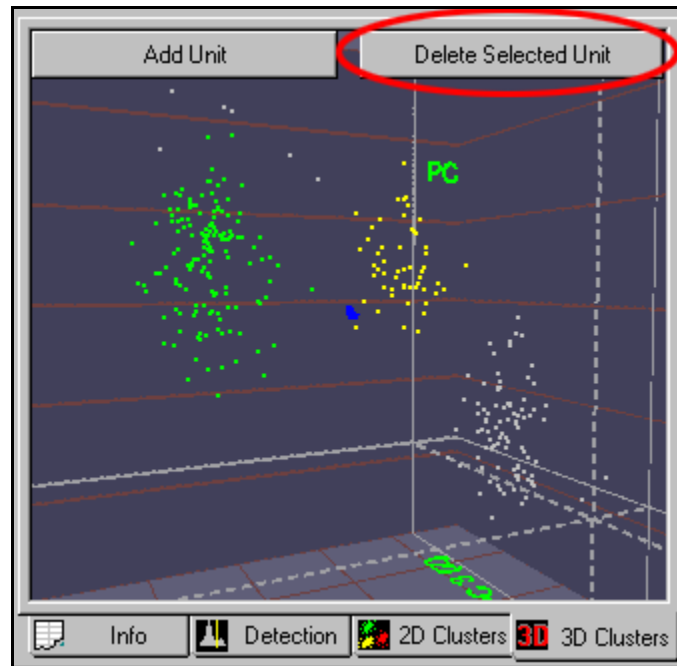
To Delete a Unit

- 1 Select the unit by clicking it's template in the Units Display.





- 
- 2 Click **Delete Selected Unit** in the Cluster view,

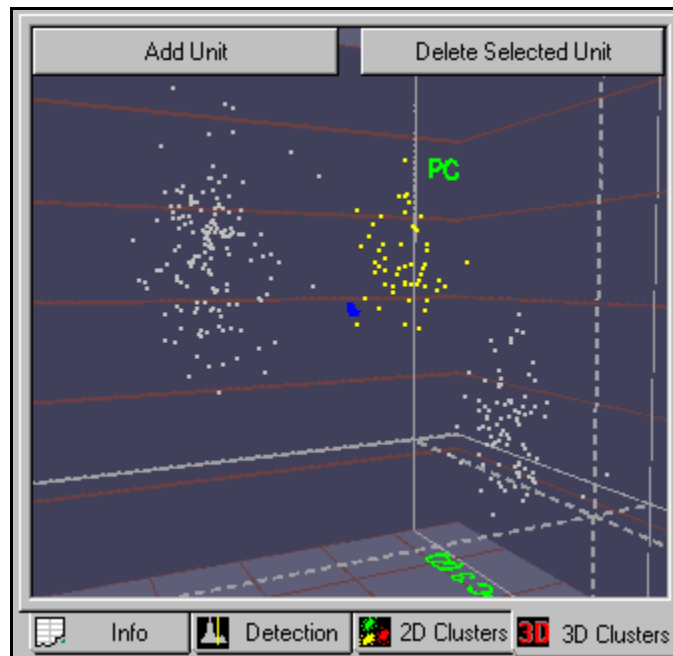


or, select the **Units | Delete Selected Unit** menu command,

or, right-click in the Cluster view and select the **Delete Selected Unit** menu command. (Use **ALT**+right-click in the 3D Cluster view).

*The selected unit is deleted.*

**Note:** Other units can shift down after the unit is deleted. For example, if Unit a and Unit b existed prior to deletion, when you delete Unit a, Unit b becomes Unit a.

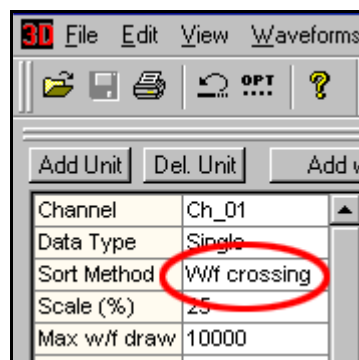


To delete all the sorted units

- Select the **Units | Delete All Units** menu command,  
or, hold down the ALT key and right-click in the Cluster view and select the **Delete All Units** menu command.

## Manual Clustering using the Waveform Crossing Method

Set the **Sort Method** parameter in the Control Table to **W/f crossing**:



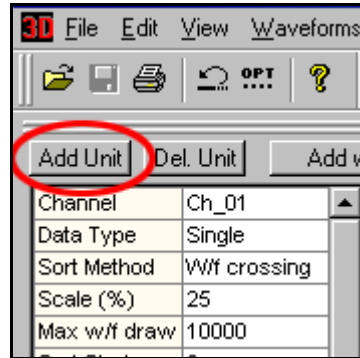
Or, select the **Sort | Change Sort Method | Cross Waveforms** in the **Waveform Display** menu item.

---


To add a unit

That is, to identify a subset of waveforms that comprise a unit.

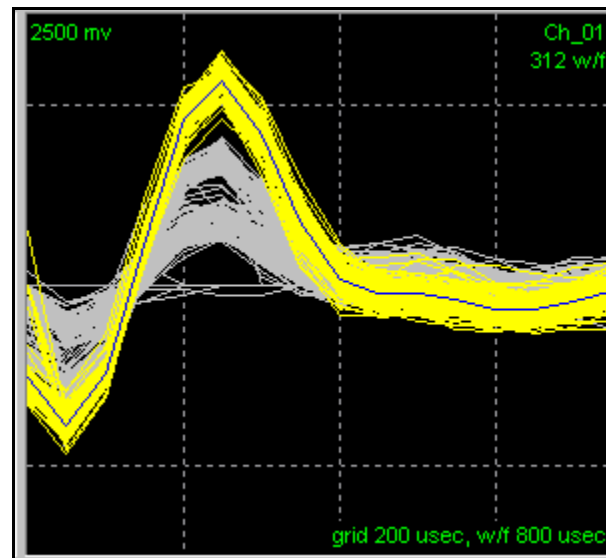
- 1 Click **Add Unit**.



Move the mouse pointer to the Waveform window.

The pointer changes to a “drawing hand” .

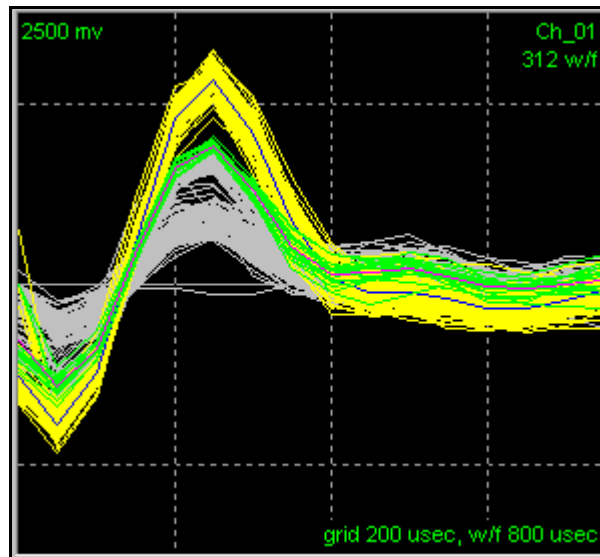
- 2 Click the left mouse button and drag a line across the bundle of waveforms.



**Note:** Up to 26 units can be sorted for each data channel.

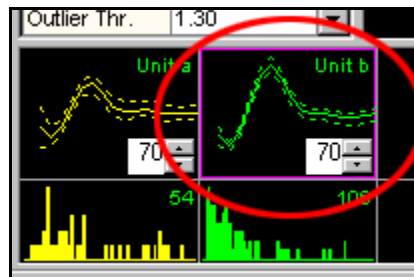
## Adjusting the Sorting

**Adding Waveforms to Selected Units in the Waveform Window.** Waveforms can be added to existing units by crossing waveforms in the Waveform window.

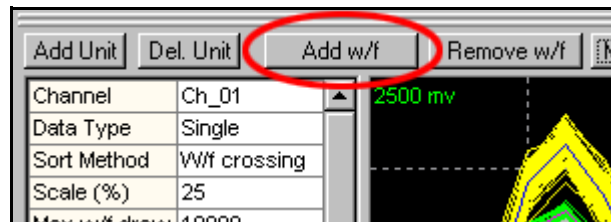


To Add Waveforms to a unit (green unit above)

- 1 Select the unit by clicking it's template in the Units Display.




- 2 Click **Add w/f** above the Waveform window.

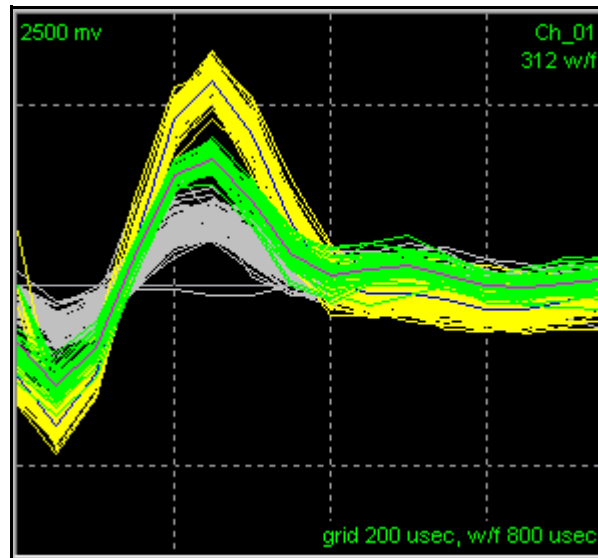


or, Select the **Units | Add Waveforms to the Selected Unit** menu command.

- 3 Move the mouse pointer into the Waveform window.

The pointer changes to a "drawing hand" .

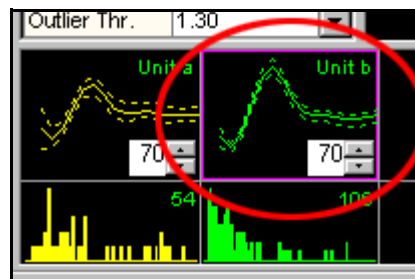
- Click the left mouse button and drag a line across the waveforms to add.  
*The crossed waveforms are added to selected unit.*



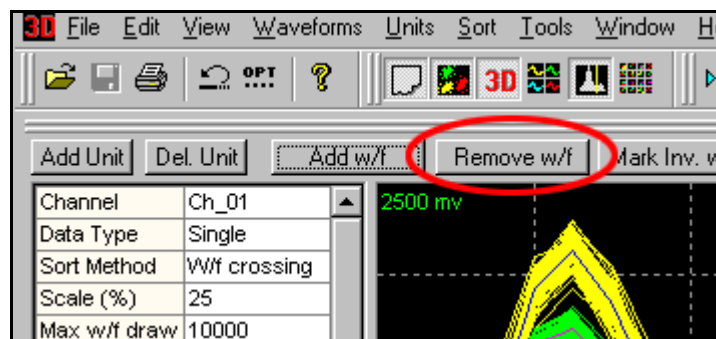
**Removing Waveforms from Selected Units in the Waveform Window.** You can trim waveforms away from existing units.

To Remove Waveforms from a Unit

- Select the unit by clicking its template in the Units Display.




- Click **Remove w/f** above the Waveform window,



or, select the **Units | Remove Waveforms from the Selected Unit** menu command.

- 3 Move the mouse pointer into the Waveform window.

The pointer changes to a “drawing hand” .

- 4 Click the left mouse button and drag a line across the waveforms to remove.

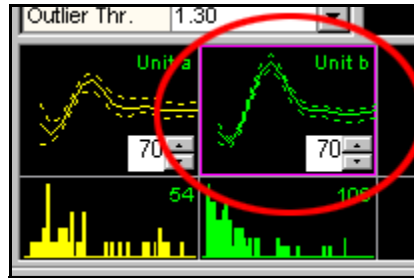
The crossed waveforms are removed from the selected unit.

For methods to display only the selected unit while removing waveforms, see [“Cleaning the Selection” on page 4-54](#).

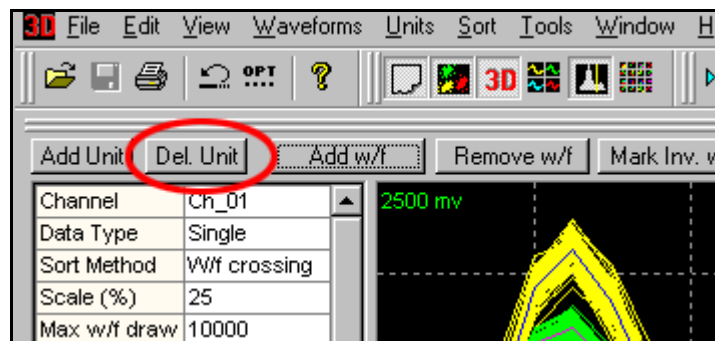
**Deleting Units.** Use the following procedure to delete a unit.

To Delete a Unit

- 1 Select the unit by clicking it's template in the Units Display.



- 2 Click **Delete Unit** above the Control Table,



or, select the **Units | Delete Selected Unit** menu command.

The selected unit is deleted.

To delete all the sorted units

- Select the **Units | Delete All Units** menu command,  
or, hold down the ALT key and right-click in the Cluster view and select the **Delete All Units** menu command.

---

## Semi-Automatic Clustering using Templates

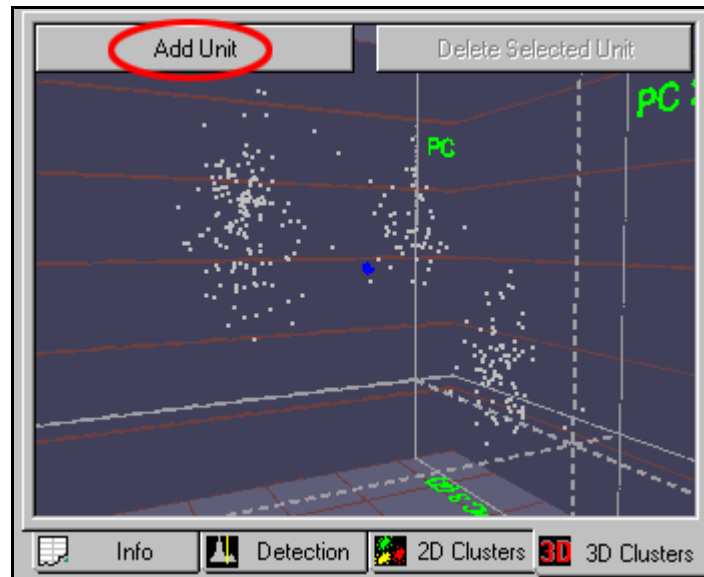
Select the templates method of sorting by selecting **Templates** in the **Sort Method** drop list in the Control Table, or by selecting the **Sort | Change Sort Method | Use Templates From Designated Waveforms** menu item.

The Templates algorithm works in waveform space, not in feature space. The goal of the algorithm is to select an existing waveform to serve as the template waveform for a new unit. However, you can use the Cluster views (showing feature space) to select the template waveform.

To add a unit using the Template sorting method

The goal is to select a waveform to be used as a unit template.


- 1 Press **Add Unit** button in the Cluster view,



or, select the **Units | Add Unit** menu command,

or, hold down the **ALT** key and right-click in the Cluster view and select the **Add Unit** menu command.

- 2 Move the mouse pointer to the Cluster view.

The pointer changes to a “drawing hand”  and the display is in a mode that selects the waveform closest to the pointer position.

**Note:** For a discussion of waveform selection, see [“Cluster View Common Functionality” on page 3-21](#). The selected waveform appears in the Waveform display. If you select **Sel.Wave** in the **View w/f** entry in the Control Table, it only shows the selected waveform, which can make it easier to select a template.

- 3 When you select the desired waveform (shown with a box around it), click the left mouse button.

*A new unit is added, and Template sorting is carried out using all of the templates.*

To add a unit using the Template sorting method using the Continuous Data View

Alternately, you can use the bottom Continuous Data View to select a template waveform.

- 1 In the Continuous Data View, right-click and move the mouse to select the desired waveform.
- 2 Release the right mouse button
- 3 Select **Use Waveform As Template For New Unit And Sort** from the context menu.

### Semi-Automatic Clustering using K-Means

Select the K-Means method of sorting by selecting **K-Means** in the **Sort Method** drop list in the Control Table, or by selecting the **Sort | Change Sort Method | Use K-Means with Cluster Centroids** menu item.

The K-Means algorithm assigns waveforms to a pre-determined number of clusters (units) using an iterative algorithm that operates on the waveform points in feature space. For more details, see [“K-Means Algorithm” on page 5-15](#). The K-Means algorithm operates differently depending on which view it is initiated from and some options settings; see [“Feature Space Used for Sorting” on page 5-13](#).

There are two steps involved:

- 1 [“Selecting Initial Cluster Centroids” on page 4-42](#)
- 2 [“Running the K-Means Clustering Algorithm” on page 4-44](#)

#### Selecting Initial Cluster Centroids

For the K-Means algorithm to work, you need to specify:

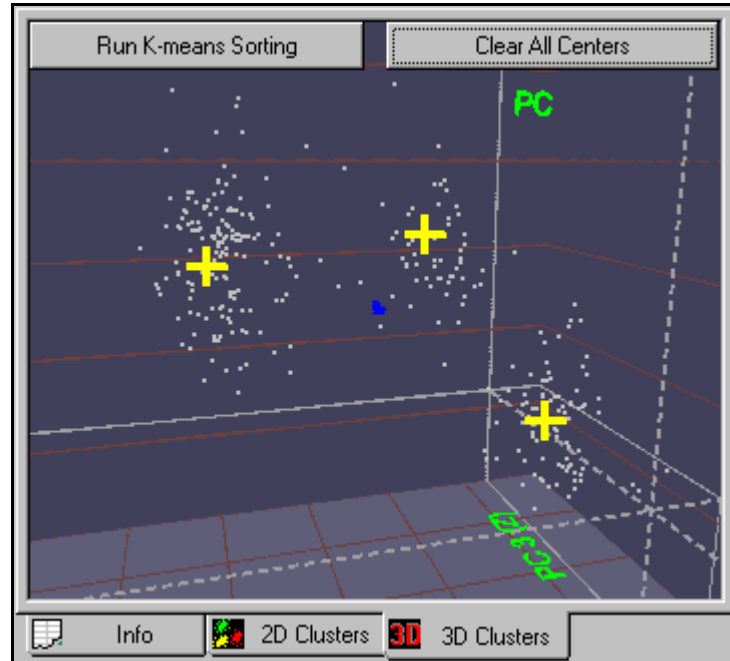
- how many clusters there are
- the approximate cluster centroids



---

To define a cluster center

- 1 Move mouse pointer to the Cluster view.
- 2 Point to the center of the cluster.
- 3 Double-click the left mouse button, or hold down **CTRL** and single-click.



**Note:** You cannot move or drag the centroids. To adjust the centroids positions, you need to remove all the centroids first and then select the new centroids.

**Note:** When using the 3D Cluster view, take care not to rotate the 3D viewpoint while you select cluster centroids. Any rotation or movement of the eyepoint performs a "Clear All Centers" operation so that the cluster centers must be re-selected.

To remove the centroids

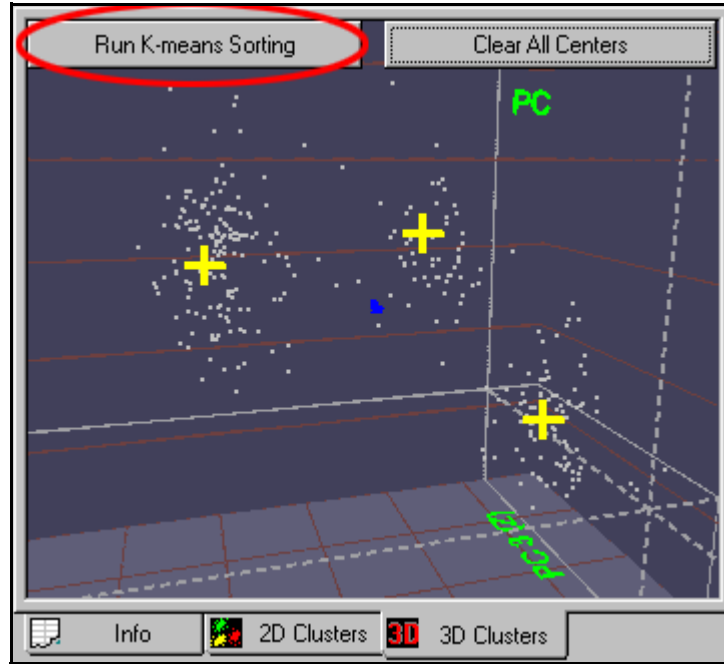
- 1 Click the **Clear All Centers** button,  
or
- 2 Select **Remove All Cluster Centers** in the floating menu.

### Running the K-Means Clustering Algorithm

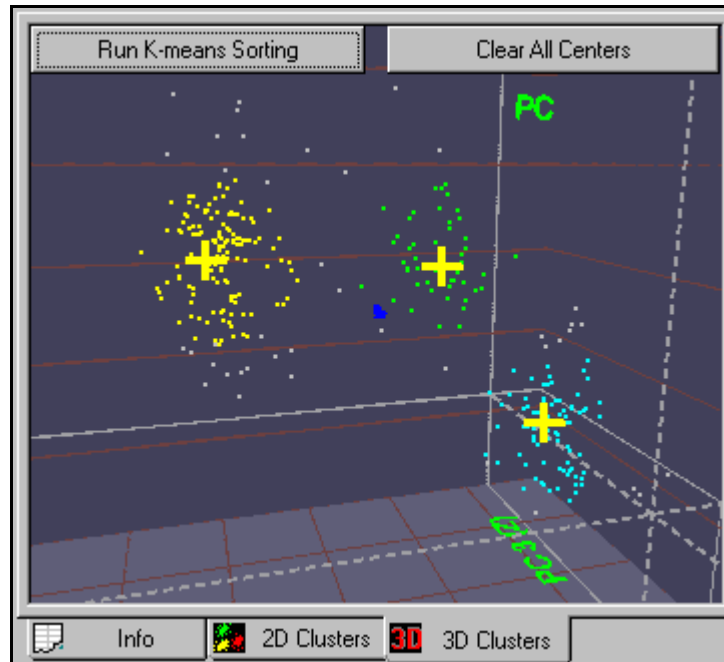
After you have specified the cluster centers, start the algorithm.

To run the K-Means clustering algorithm

- Click **Run K-means Sorting** in the Cluster view.

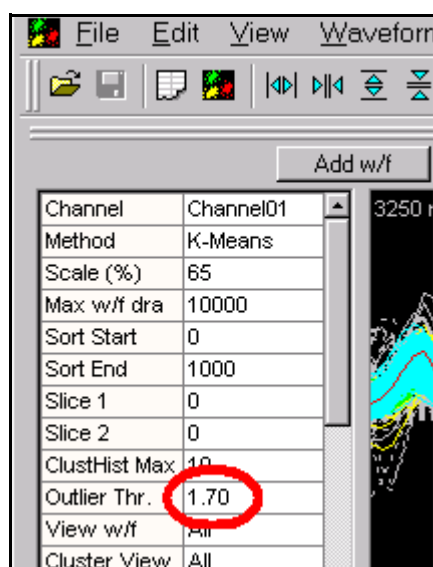


*When the algorithm completes, the points are sorted.*



---

As part of its operation, the K-Means algorithm removes the outliers from the clusters. This is affected by the **Outlier Thr.** parameter, which you can adjust by using the Control Table:



To sort using the new Outlier Threshold, you must rerun the K-Means sorting algorithm.

You can “clean up” the clustering results after the K-Means sorting is complete by adding or removing waveforms from each unit:

- either in the Waveform window:
  - See [“Adding Waveforms to Selected Units in the Waveform Window”](#) on page 4-38 and
  - [“Cleaning the Selection”](#) on page 4-54.
- or in the Cluster views:
  - See [“Adding Waveforms to Selected Units in the Cluster Views”](#) on page 4-31 and
  - See [“Removing Waveforms from Selected Units In The Cluster Views”](#) on page 4-33.

### Semi-Automatic Clustering using Standard E-M

Select the Standard E-M method of sorting by selecting **Std E-M** in the **Sort Method** drop list in the Control Table, or by selecting the **Sort | Change Sort Method | Use Standard E-M Sorting Method** menu item.

The Standard E-M algorithm assigns waveforms to a pre-determined number of clusters (units) using an iterative algorithm operating on the waveform points in feature space. For more details, see [“Standard E-M Algorithm”](#) on page 5-16. The

Standard E-M algorithm operates differently depending on which view you initiate it from and which options you set; see [“Feature Space Used for Sorting” on page 5-13](#).

There are two steps involved:

- 1 Selecting initial cluster centers. This is done in exactly the same fashion as for the K-Means algorithm described above. See [“Selecting Initial Cluster Centroids” on page 4-42](#) for a step-by-step description.
- 2 [“Running the Standard E-M Clustering Algorithm” on page 4-46](#)

### Running the Standard E-M Clustering Algorithm

After the cluster centers have been specified, start the algorithm.

To run the Standard E-M clustering algorithm

- Click the **Run E-M** button in the Cluster view.

The **Run E-M** button changes to **Cancel** after the sorting starts. The **Cancel** button stops the algorithm after the next iteration completes. If it is not cancelled, the algorithm runs until convergence, or until a specified maximum number of iterations has completed. This **Max Iterations** parameter is set in the **E-M Sorting** section of the **Sort** tab under the **Tools | Options** dialog box.

There is a parameter in the Control Table labelled **Beta** that appears when the Standard E-M sorting method is selected. This parameter controls the convergence properties of the algorithm; for details, see [“Standard E-M Algorithm” on page 5-16](#).

The clustering results can be “cleaned-up” after the sorting is completed by adding or removing waveforms from each unit:

- either in the Waveform window:
  - See [“Adding Waveforms to Selected Units in the Waveform Window” on page 4-38](#) and
  - [“Cleaning the Selection” on page 4-54](#).
- or in the Cluster views:
  - See [“Adding Waveforms to Selected Units in the Cluster Views” on page 4-31](#) and
  - [“Removing Waveforms from Selected Units In The Cluster Views” on page 4-33](#).

### Automatic Clustering using Valley-Seeking Method

Select the Valley Seeking method of sorting by selecting **Valley Seek** in the **Sort Method** drop list in the Control Table, or by selecting the **Sort | Change Sort Method | Use Valley-Seeking Sorting Method** menu item.

---

The Valley-Seeking algorithm uses the inter-point distances in feature space to assign the waveforms to an optimal number of clusters. For more details, see [“Valley Seeking Algorithm” on page 5-16](#). The Valley-Seeking algorithm operates differently depending on which view you initiate it from and which options you have set, see [“Feature Space Used for Sorting” on page 5-13](#).

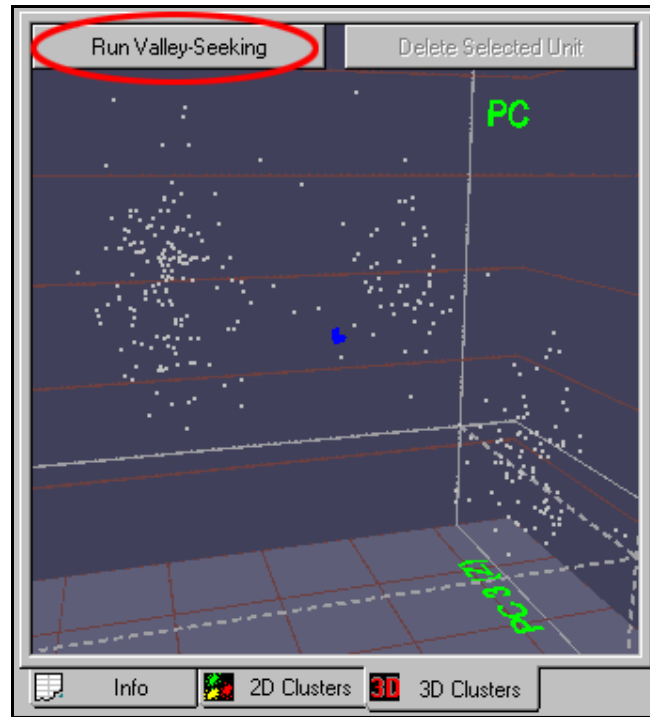
**Note:** The Valley-Seeking algorithm scales as  $N^2$ , where  $N$  is the number of points. So, for large number of points (more than a few thousand), the algorithm can use unreasonable amounts of memory and take a long time to complete.

Offline Sorter can limit the number of points that are used as input to the Valley-Seeking algorithm. The **Sort** tab on the **Tools | Options** dialog box contains the **Limit Number of Waveforms used in Valley-Seeking sorting** checkbox that, when checked, limits the number of waveforms sorted to be the number in the associated edit box. Offline Sorter then uses every  $n$ th waveform in the valley seeking calculation, where  $n$  is the smallest integer such that the total number of waveforms used is less than the specified maximum. After the valley seeking algorithm sorts the waveforms, you can optionally operate on the unsorted waveforms (those that were skipped) by using the Assign Unsorted Waveforms tool. This is the same tool that is available from the **Tools | Assign Unsorted to Closest Unit** menu selection. See [“Assigning Unsorted Points to Closest Unit” on page 5-39](#). You can use the Assign Unsorted Waveforms tool to assign the waveforms to one of the clusters the valley seeking algorithm creates.

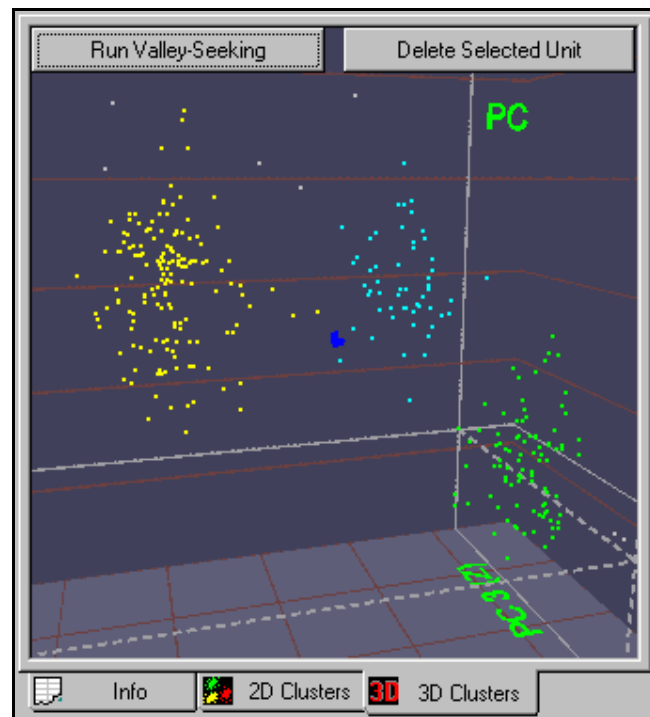
The outlier threshold specified in the Control Table also affects how these skipped points are assigned to units. Finally, you can use the **Tools | Remove Outliers** menu selection to optionally run following the sort, in order to remove outliers using the outlier threshold value. For more information, see [“Automatic Outlier Removal Tool” on page 4-54](#).

To run the Valley-Seeking clustering program

- Click **Run Valley-Seeking** in the Cluster view, or



select the **Sort | Sort Current Channel | Run Valley-Seeking Sorting** menu command.



---

*The algorithm attempts to determine the correct number of clusters and assign the waveforms to those clusters.*

**Note:** The **Run Valley-Seeking** button changes to **Cancel** after sorting starts. Click **Cancel** to stop the algorithm after the next iteration completes. If you do not cancel it, the algorithm runs until convergence, or until a maximum number of iterations completes.

There is a parameter in the Control Table labelled **Parzen Mult.** that appears when you select the Valley Seeking sorting method. This parameter controls the behavior of the algorithm; see [“Valley Seeking Algorithm” on page 5-16](#) for more details.

You can “clean up” the automatic clustering results after sorting completes by adding or removing waveforms from each unit:

- either in the Waveform window
  - See [“Adding Waveforms to Selected Units in the Waveform Window” on page 4-38](#) and
  - [“Cleaning the Selection” on page 4-54.](#)
- or in the Cluster views
  - See [“Adding Waveforms to Selected Units in the Cluster Views” on page 4-31](#) and
  - [“Removing Waveforms from Selected Units In The Cluster Views” on page 4-33.](#)

## Automatic Clustering using T-Distribution E-M Method

Select the T-Distribution E-M method of sorting by selecting **T-Dist E-M** in the **Sort Method** drop list in the Control Table, or by selecting the **Sort | Change Sort Method | Use T-Distribution E-M Sorting Method** menu item.

The T-Distribution E-M algorithm is an iterative procedure that operates on the points in feature space to assign the waveforms to an optimal number of clusters. For more details, see [“T-Distribution E-M Algorithm” on page 5-17](#). The T-Distribution E-M algorithm operates differently depending on which view you initiate it from and what options you set, see [“Feature Space Used for Sorting” on page 5-13](#).

To run the T-Distribution E-M clustering algorithm

- Click the **Run T-Dist E-M** button in the Cluster view, or  
select **Sort | Sort Current Channel | Run T-Dist E-M Sorting**.

The **Run T-Dist E-M** button changes to **Cancel** after sorting starts. Click **Cancel** to stop the algorithm after the next iteration completes. If it is not cancelled, the algorithm run until convergence, or until a specified maximum number of iterations completes. You can set the **Max Iterations** parameter in the **E-M Sorting** section of the Sort page under the **Tools | Options** dialog box.

There is a parameter in the Control Table labelled **D.O.F. Mult.** that appears when the T-Distribution E-M sorting method is selected. This parameter controls the convergence properties of the algorithm; for details see [“T-Distribution E-M Algorithm” on page 5-17](#).

The first step of the T-Distribution sorting algorithm is actually to do a K-Means sort using a preset number of “seed” cluster centers. The E-M algorithm can then combine clusters together as the algorithm progresses to arrive at the optimal number of final clusters. You can set the number of seed clusters used initially in the **E-M Sorting** section of the **Sort** page under the **Tools | Options** dialog box. The T-Distribution sorting algorithm never produces more clusters than the initial number of seed clusters.

You can “clean up” the clustering results after the sorting is completed by adding or removing waveforms from each unit:

- either in the Waveform window
  - [“Adding Waveforms to Selected Units in the Waveform Window” on page 4-38](#) and
  - [“Cleaning the Selection” on page 4-54](#).
- or in the Cluster views
  - See [“Adding Waveforms to Selected Units in the Cluster Views” on page 4-31](#) and
  - [“Removing Waveforms from Selected Units In The Cluster Views” on page 4-33](#).

## Unit Cleaning

### Controlling Waveforms Displayed

To inspect the sorting more closely, it is sometimes desirable to select which waveforms are displayed in the windows. There are two main ways of limiting the waveforms that display: units and time interval.

### Controlling waveforms viewed by Units

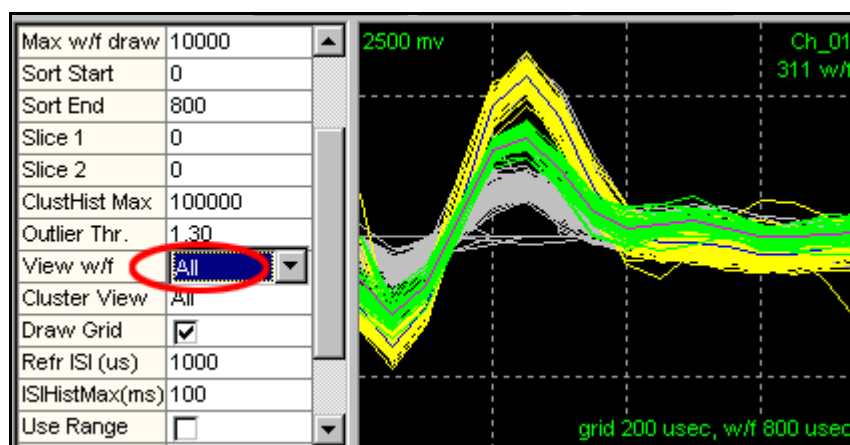
Offline Sorter can display the waveforms from only selected unit(s) in the Waveform window and in the Cluster views.

**Controlling waveforms in the Waveform Window by Units.** You can change the type of waveforms viewed in the Waveform window by using the **View w/f** parameter in the Control Table, or by using the **Waveforms | View Only Currently Selected Unit(s)** menu entry.



To view waveforms from all units in the Waveform Window

- Set the **View w/f** parameter in the Control Table to **All**.



or, unselect (make so that it is not checked) the **Waveforms | View Only Currently Selected Unit(s)** menu entry.

To view waveforms from specific units in the Waveform Window

- 1 Set the **View w/f** parameter in the Control Table to **Sel.Unit(s)**, or select (make so that it is checked) the **Waveforms | Use Only Currently Selected Unit(s)** menu entry.
- 2 Select the unit by clicking on the unit's template in the Unit's Display.



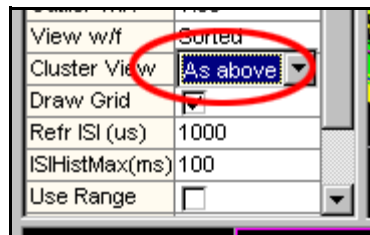
**Note:** To see the waveforms that are currently unsorted, select the unsorted "unit" in the display.

**Note:** You can select multiple units as described in [“Units Display” on page 3-11](#). Holding down the **CTL** key and left-clicking on a unit in the Units Display selects the clicked unit, but leaves the previously-selected unit(s) in the multiply selected state (denoted by a dashed box drawn around the unit). Holding down the **SHIFT** key and left-clicking a unit selects that unit and keeps the previously selected unit(s) in the multiple-select state, and also multiple-selects all units between the previous and the clicked unit.

If you select **Sel. Wave** from the **View w/f** drop list, Offline Sorter displays only the waveform currently selected in the Continuous Data window, or the Cluster view.

If you select **Unsorted** from the **View w/f** drop list, Offline Sorter displays only the unsorted waveforms, regardless of what unit you select. This selection can help you to add unsorted waveforms into an existing unit; you can view only unsorted waveforms, yet still have the “target” unit selected. If you choose this selection, use one of the following waveform selection options: **Use Only Visible Waveforms** or **Use Only Unsorted Waveforms**. For more information, see [“Waveform Selection Options” on page 4-53](#).

**Controlling waveforms in the Cluster Views by Units.** By default, the type of waveforms shown in the Cluster views matches those shown in the Waveform window because the **Cluster View** parameter in the **Control Table** defaults to **As above**:



However, you can override the default by selecting **All** in the Control Table, which forces all waveforms to be shown in the Cluster views, regardless of what is showing in the Waveform window.

**Controlling waveforms in the Continuous Data Views by Units.** By default, the type of waveforms shown in the Continuous Data view matches those shown in the Waveform window because the **Contin.View** parameter in the **Control Table** defaults to **As above**. However, you can override the default by selecting **All** in the Control Table, which forces all waveforms to be shown in the Cluster views, regardless of what is showing in the Waveform window.

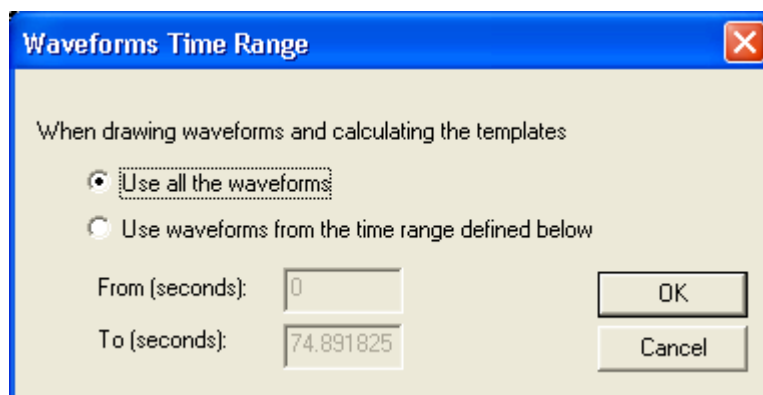
### Controlling waveforms viewed by Time

**Using the Continuous Data Views.** The **Control cluster display** and **Control w/form display** checkboxes in the lower left portion of the [Continuous Data Window](#) detailed on [page 3-54](#), which you can use to restrict the waveforms shown in the Cluster view and Waveform window, respectively, to those wave-

---

forms contained in the scroll region of the Continuous Data view. You can use this feature to display “animations” of how things progressed during the course of the file—for details, see [“Playing a Movie of the data” on page 5-22](#).

**By selecting a Time Range.** The **Waveforms | Select Time Range** menu entry brings up a dialog box that enables you to enter a time range that is relative to the beginning of the data file. All display, selection, and sorting operations only consider the waveforms within this time range. The PCA calculation can optionally only consider waveforms within this time range.



Also, the Control Table item **Use Range** enables you to limit the waveforms to the time range, which you can also set by using the **Range From (s)** and **Range To (s)** entries in the Control Table.

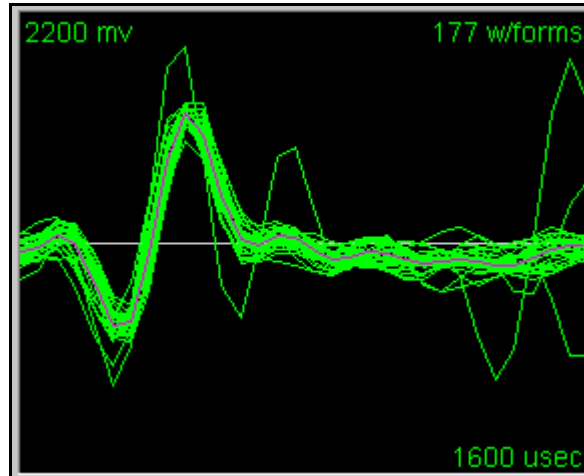
## Waveform Selection Options

Use the **Operations Involving Waveforms** section in the **Tools | Options | General** dialog to determine which waveforms are included in waveform operations, such as sorting, selecting waveforms, adding waveforms, removing waveforms, etc. There are three options: either use all waveforms, or use only the currently visible waveforms, or use only the unsorted waveforms. If you click the **Use Only Currently Visible Waveforms** checkbox, Offline Sorter restricts operations to affect only those waveforms that are currently visible in the Waveform window or Cluster views (as appropriate). For methods for selecting which waveforms are made visible, see [“Controlling Waveforms Displayed” on page 4-50](#). If you click the **Use Only Unsorted Waveforms** checkbox, Offline Sorter restricts operations to affect only those waveforms that are currently unsorted. You must use either the **Use Only Currently Visible Waveforms** option or the **Use Only Unsorted Waveforms** option; you cannot use both options together.

You can also use the **Waveforms | Use Only Currently Visible Waveforms** or **Waveforms | Use Only Unsorted Waveforms** menu options, or the corresponding toolbar buttons; they have the same effect as the checkboxes.

## Cleaning the Selection

Offline Sorter also provides the tools to remove waveforms from a selected unit, that is, to label the waveforms as unsorted. For example, you may want to remove some of the outlying waveforms from this unit:



### Automatic Outlier Removal Tool

The **Tools | Remove Outliers** menu selection provides a tool that automatically removes outliers from clusters, based on the **Outlier Thr.** setting in the Control Table. This outlier removal works by considering the location of the points corresponding to the waveforms in the currently selected feature space. If the 3D Cluster view appears when the tool is selected, then 3D feature space is used, otherwise 2D feature space is used.

The outlier removal tool automatically does the following for each unit:

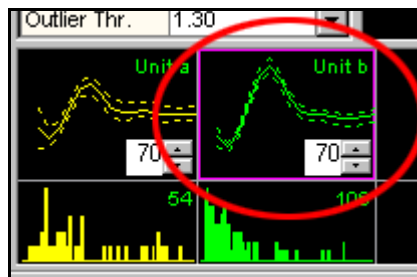
- Calculates the centroid and sigmas along the principal axis in the 2D or 3D feature space.
- For each point, calculates the Mahalanobis distance between the point and the centroid. The units of Mahalanobis distance are standard deviations.
- If this distance is greater than the **Outlier Thr.** setting, it sets the waveform point to unsorted.

### Manual Waveform Removal

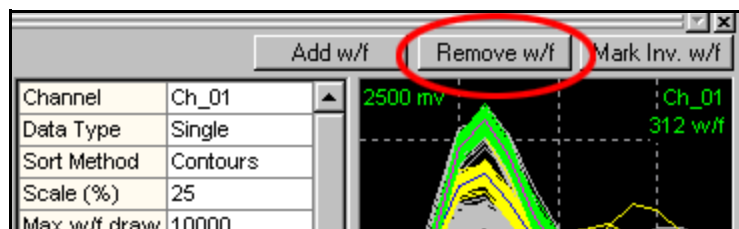
You can also manually remove any waveform (not just outliers) from a unit. This feature is easier to use if the Waveform view shows only the waveforms of a particular unit, not all the waveforms for the channel. To select which waveforms are shown in the Waveform view, use the **View w/f** parameter in the Control Table.

To remove some of the waveforms from a particular unit


- 1 In the Control Table, change **View w/f** parameter value to **Sel. Unit(s)** (Selected Units).
- 2 To select the unit, click at the unit's template display.



- 3 Click **Remove w/f**.

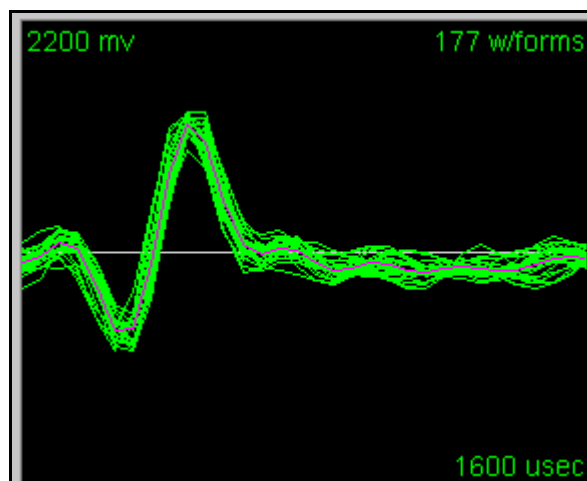


- 4 Move the mouse pointer to the Waveform view.

The pointer changes to a "drawing hand" .

- 5 Click the left mouse button and drag a straight line across the waveforms you want to exclude from the unit.

After several iterations of clicking **Remove w/f** and drawing a line, the unit can appear as follows:



## Adding or Removing Waveforms using other views

**Note:** Waveforms can be added or removed from the selected unit by circling them in the cluster views:

- after you press the **Add w/f** or **Remove w/f** buttons, or
- by selecting the **Units | Add Waveforms to the Selected Unit** or
- the **Units | Remove Waveforms from the Selected Unit**

- 1 After selecting one of these commands, move the pointer to the Cluster view.

**Note:** The pointer changes to a blue drawing hand



for adding waveforms and a red drawing hand



for removing waveforms.

- 2 Click the left mouse button and drag a line around the units to be added or removed.

**Note:** Waveforms can be un-sorted or added to the current view from the continuous data window. Right-click on a waveform to select it, and, from the pop-up menu, select **Unsort Waveform** or **Add Waveform to Selected Unit**.

## Combining Units

Sometimes what initially appeared to be multiple units is really a single unit. The **Units | Combine Units** menu selection enables you to select several units and combine them into a single unit. If you select multiple units to combine in the Units Display, these units also appear as selected units in the **Combine Units** dialog box.

## Swapping Units

For comparing units across files, you can control which unit gets which unit label. You can use the **Units | Swap Units** menu selection to select two units to be swapped.

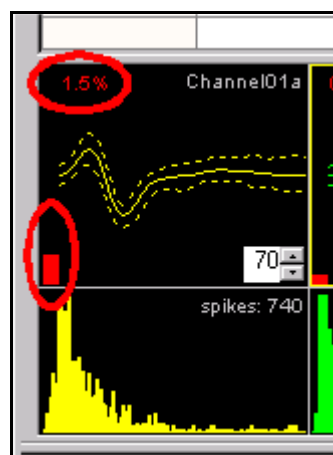
## Invalidating Units

Occasionally, an entire unit's worth of waveforms can be "noise" or other undesirable artifacts. You can use the **Units | Invalidate All Waveforms in Selected Unit** or **Waveforms | Invalidate All Waveforms in Selected Unit** menu selection to invalidate all waveforms in the primary selected unit (See ["Invalidating Waveforms" on page 4-15.](#)).

## Removing Spikes with Short Inter-spike Intervals

Most neurons have absolute refractory periods after generating an action potential. Effective isolation of a single unit within a cluster should therefore have relatively few spikes with inter-spike intervals (ISI) less than the **Refractory Interspike Interval**. You can set the value for the Refractory ISI in the Control Table or through the **Refractory Period** tab on the **Tools | Options** dialog box.

You can use the Units Display to inspect the interspike interval-related aspects of the current units. The top row of graphs of this display shows the template (average waveform) for the unit (solid line) as well as  $\pm 3$  standard deviations from the template (dash lines). You can use the edit controls with spin buttons to adjust the tolerances for Template sorting.



The bottom row of graphs of this display shows the interspike interval histograms. The X axis of each histogram is from zero to **ISIHistMax**—a value you can also set in the Control Table. The total number of spikes for each unit is also shown.

The percentage of spikes with interspike intervals less than the **Refractory Interspike Interval** appears in **red** and a **red bar** indicator also appears in the upper portion of the Units Display. You can change the value of **ISIHistMax** and **Refractory Interspike Interval** in the Control Table. You can also set the Refractory Interspike Interval value by using the **Tools | Options | Refractory Period** tab.

Offline Sorter provides a means for visualizing which spikes in a cluster have inter-spike intervals less than the refractory period and for removing them.

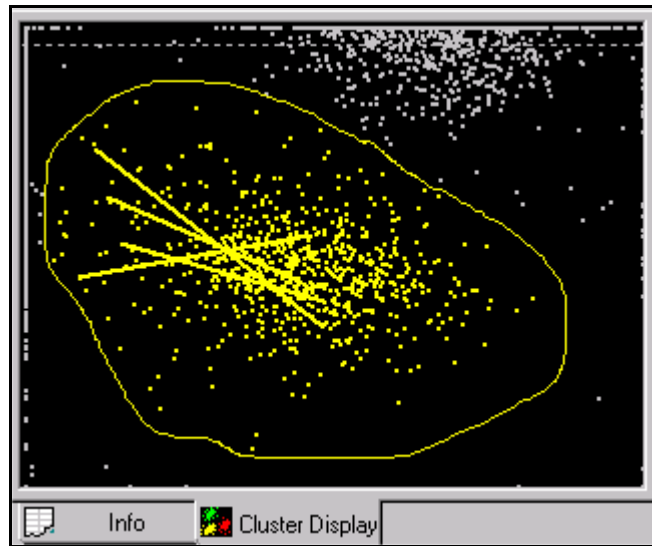
To Draw Lines in the Cluster View between spike pairs with interspike intervals less than the Refractory ISI

- 1 Select **Draw lines showing detected small interspike intervals** in the **Refractory Period** tab on the **Tools | Options** dialog box.
- 2 After zooming in to the cluster in question using the zoom

(, , , ) and move

(, , , ) control buttons on the toolbar,

the Cluster view might then look like this:



### Automatic Short ISI Spike Removal

Offline Sorter provides a tool to automatically remove waveforms with ISIs below the refractory period. Selecting **Tools | Remove Short ISI Waveforms** from the main menu removes waveforms from the clusters according to the following 2D algorithm:

For each unit:

- 1 Calculate the center of gravity of points in that unit.
  - *This gives a point in 2D feature space that is near the center of the cluster.*
  - For each pair of waveforms with ISIs below the Refractory Period, calculate which member of the pair is furthest from the center of gravity point for the unit.
- 2 Remove that waveform from the unit.

**Note:** After this tool runs, short ISI waveform pairs can still exist if a particular waveform was within the refractory period of more than one other waveform. If needed, re-run the tool to get rid of all short ISI waveforms.

It is also possible to manually remove spikes with short ISIs from units. See [“Manual Short ISI Spike Removal” on page 5-25](#)

## Examining Sorting Statistics

There are two methods to examine sorting statistics. Each one is described in detail in this section.



## Info View Statistics

The **Stats** sub-tab under the **Info** tab contains more detailed statistical information about the sorting performed on a channel. The [Sorting Summary View](#) detailed on [page 3-46](#) provides a convenient place to look to see a more concise summary of the most important sorting quality statistics.

Sorting Statistics for Channel 1											
Features : 2D: (PC 1, PC 2) 3D: (PC 1, PC 2, PC 3)											
Multivariate ANOVA: (Not Including Unsorted Waveforms as a Group)											
In 2D Cluster Space:		F(4,54724) = 11.67906		p = 1.75465e-009							
In 3D Cluster Space:		F(6,54722) = 8.104329		p = 8.8966e-009							
		2D		3D							
J3 :		3.84931		2.6387							
Pseudo-F :		189415		129844							
Davies-Bouldin :		0.306595		0.372886							
Dunn :		1.73042		1.38495							
Per-Unit Statistics											
Counts		ISI < 100		Center (2D)		Sigmas (2D)		Center (3D)		Sigmas (3D)	
Unit a:		7366	0.3530	-568.5524, 554.481		324.8189, 226.2428		-568.5524, 554.4814, -218.42		345.5869, 226.2481, 163.0215	
Unit b:		37570	0.1544	-262.5755, -1160.0		280.8709, 218.1188		-262.5755, -1160.0490, -510.2		292.0075, 239.2599, 193.5820	
Unit c:		53482	0.2412	870.6649, -892.622		324.7222, 202.5226		870.6649, -892.6222, -501.00		345.0548, 271.4262, 179.9306	
Normality Statistics											
Counts		Skew(2D)		Kurtosis(2D)		Skew(3D)		Kurtosis(3D)			
All Valid WFs:		98418	-1.1101, 17.4594		-1.1775, 4.4132		-1.1101, 17.4594, -0.1175		-1.1775, 4.4132, 0.1527		
Unsorted WFs:		0	0.0000, 0.0000		0.0000, 0.0000		0.0000, 0.0000, 0.0000		0.0000, 0.0000, 0.0000		
Unit a:		7366	-0.3921, -0.2196		1.6444, 0.0334		-0.3921, -0.2196, -0.3578		1.6444, 0.0334, 0.0735		
Unit b:		37570	-0.3803, 0.4460		0.0153, 1.3004		-0.3803, 0.4460, 0.4680		0.0153, 1.3004, 0.2264		
Unit c:		53482	-0.4105, -0.2450		-0.0598, 0.2803		-0.4105, -0.2450, 0.3551		-0.0598, 0.2803, 0.2417		
Pairwise Sort Quality											
Channels Chan Params Sort Details Global Events Stats PCA Templates Scan											

### Top Half of the Window

In the top half of the window, the **Features** line shows which features were used for each dimension of the 2D and 3D feature spaces. All sorting statistics are dependent on the feature space.

The **Multivariate ANOVA** section shows the p-value and the F Statistics (with the Hypothesis and Error degrees of freedom in parenthesis) from a 2D and 3D Multivariate Analysis of Variance (ANOVA) on the sorted units in the currently displayed in 2D and 3D feature space:

- **In 2D Cluster Space:** (first row) Parameters from Multivariate ANOVA test on clusters in 2D feature space
- **In 3D Cluster Space:** (second row) Parameters from Multivariate ANOVA test on clusters in 3D feature space

You can display the p-value, along with all the other sorting quality statistics, in the Cluster views by selecting **Show Statistics** in the right-click menus in the

Cluster views, or by selecting **Show Statistics** using the **View | Cluster View Options** menu command.

You can specify a threshold p-value in the **Stats** tab of the **Tools | Options** dialog box.

The **p-value** is displayed in **red text** if it is not significant with respect to the threshold probability (i.e. greater than the threshold p-value).

### Bottom Half of the Window

The bottom half of the window, lists the values of the [Sorting Quality Statistics](#), which are detailed on [page 5-19](#), for both the 2D and 3D feature spaces.

**Per-Unit Statistics Section.** The next section is a table that present the means and standard deviations of the multivariate normal distributions for the waveforms of each unit, for both the 2D and 3D feature spaces. The percentage of waveforms in each unit that do not meet the minimum ISI (interspike interval) period is also displayed.

**Normality Statistics Section.** The next section is a table that calculates the Skew and Kurtosis characteristics of the multivariate normal distribution for groups of waveforms (all valid waveforms, the unsorted waveforms, and the waveforms for each Unit), for both the 2D and 3D feature spaces. Large values of Skew and/or Kurtosis indicate that the data is not well-fit by a multivariate normal distribution. Some analyses have underlying assumptions of normally-distributed data (e.g. the Multivariate ANOVA analysis).

**Pairwise Sort Quality Section.** This section contains a number of tables, for the p-values and each of the [Sorting Quality Statistics](#), which are detailed on [page 5-19](#), for both 2D and 3D feature spaces. Each table has the units listed across both the rows and the columns; the value at the intersection of the *i*th row and the *j*th column represents the pairwise statistic calculated for Unit *i* and Unit *j*. The value is calculated as if only those two units existed in the channel, and only the waveforms assigned to those units existed. All other waveforms (in other units, or unsorted) are ignored.

**Note:** In each table, the “matrix” formed by the values is symmetric ( $a(i, j) = a(j, i)$ ), and the diagonals (that correspond to the pairwise statistics calculated between some unit and itself) are meaningless and are set to 0.

The values in the tables in some sense indicate how well-sorted each pair of units is. A poor sort quality value between two units can indicate that the two units need to be combined into a single unit. Take care with the interpretation of these pairwise results, as they are a post-hoc test, and the probability is high that somewhere in the tables there are values that indicate exceptionally well-sorted units.

The calculation of these values can take a noticeable amount of time to complete. Because of this, the calculation is optional, and it is disabled by default. To enable it, check the **Calculate unit pairwise statistics after each change** checkbox in the **Tools | Options | Stats** dialog box.

---

## Statistics Feedback While Sorting

On the 2D and 3D Cluster views, there are mechanisms you can use to get immediate feedback on the sorting quality while you perform sorting. It is sometimes useful for you to know immediately if an action (e.g. removing outliers from a cluster) helped or hurt the sorting quality. If the action hurt the sort quality, you can immediately undo it by using the **Edit | Undo** menu selection.

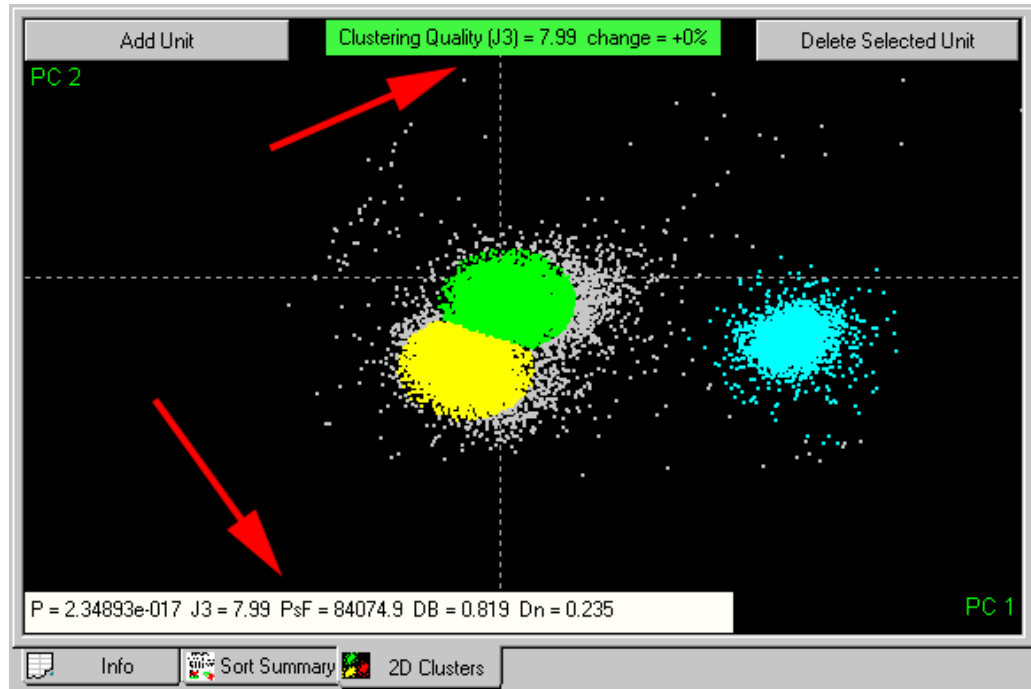
There are two similar ways to display the sort quality statistics on the Cluster views. The descriptions follow.

You can display the absolute values of all [Sorting Quality Statistics](#), which are detailed on [page 5-19](#), in a white bar along the bottom of the view by selecting **Show Statistics** from the right-click menu. The text is red if the p-value is below the **Probability Threshold for significance** in the **Tools | Options | Stats** tab. The abbreviations appear in this line are:

- p: the p-value from the multivariate ANOVA
- J3: the J3 statistic
- PsF: the Pseudo-F statistic
- DB: the Davies-Bouldin statistic
- Dn: the Dunn statistic

There are several available Sorting Quality Statistics that you can choose for display at the top of the view in a green or red bar by selecting **Show Cluster Quality** from the right-click menu. You can choose the Sorting Quality Statistics to display in the top bar from the **Sorting Quality Statistic to use on the Cluster View Displays** droplist on the **Tools | Options | Stats** tab. The value of that statistic appears, along with the percentage change from its previous value that the last

sorting-related action caused. If the action made the clustering better according to the statistic, the bar is green; if the sorting quality decreased, the bar is red.



## Saving the Sorting Results

### Saving Timestamps in a NEX File

The main result of the sorting process is a set of spike trains (arrays of timestamps) corresponding to the sorted units. You can analyze these spike trains with NeuroExplorer (NEX), which is a software package that is separately available from Plexon.

To save the timestamps and other data, such as continuous analog variables, in a NEX data file

- 1 Select **File | Export to NEX** menu command.
- 2 Select the file name for the saved file
- 3 Select a time range of the timestamps to save to the file, or accept the default setting where all timestamps are saved to the file.

You can optionally export a NEX waveform variable for each unit's template. The template waveforms have timestamp 0.

Also you can optionally write out the timestamps for the unsorted waveforms to the NEX file, as if it were a sorted unit. The NEX variable names that hold the unsorted waveform timestamps have a 'U' appended to denote Unsorted waveforms.

---

## Re-sorting from a NEX File

Offline Sorter can also resort a file by using the unit assignments that are saved in a NEX file. This function is only useful if you apply it to the data file originally used to produce the NEX file. You can use this feature to facilitate retaining only the original, raw data file on disk, rather than retaining the raw data file plus possibly several versions of the data file sorted slightly differently.

To apply the sorting from a NEX file to the currently loaded file, select **File | Sort The File Using Saved .NEX**. The operation works by matching timestamps from the Nex file to the waveform timestamps in the currently loaded file; there is also an option that you can use to control what happens to the waveforms that cannot be matched to timestamp from the NEX file. You can also multiply the timestamps from the Nex file by a factor to take into account the ‘timestamp compression’ that can happen due to Nex’s 31-bit timestamp limit.

## Saving Timestamps and Waveforms

If you use **Offline Sorter** to resort spikes from a file with previously sorted waveform/spike data (e.g. a Plexon PLX data file, Datawave, Bionic Technology NEV, and CED Spike-2), the resorted waveforms and timestamps can be saved into another file of the same format. All records in the original file are preserved in the output file, except that the unit assignments for spike records are modified to reflect the sorting that has been performed in Offline Sorter.

To save the waveforms and re-sorted timestamps

- Select **File | Save As** menu command.

**Note:** You cannot overwrite the original file with this operation.

## Overwrite/Backup Control on File-Saves

When files are saved using the **File | Save** menu option or the toolbar **Save** button, Offline Sorter provides control over saving files in terms of whether the data files are overwritten or backups are kept. You can control this behavior using the **File-Save Behavior** drop-down list in the **General** tab of the **Tools | Options** dialog box.

The options available are:

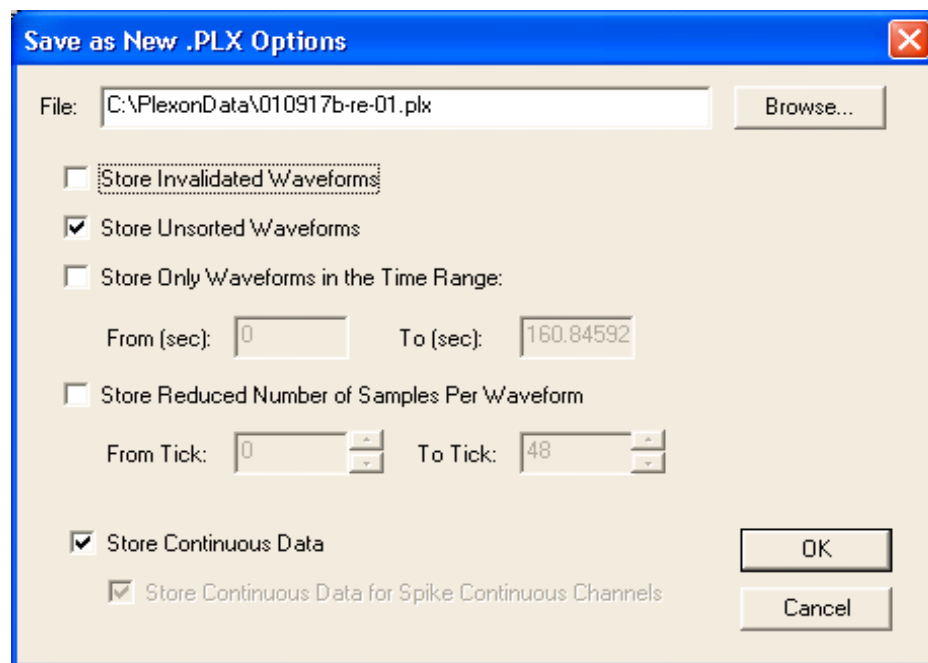
- **Preserve Original, prompt for new filename (defaults to filename-0n)** – This is the behavior available in older versions of Offline Sorter. The original data file is never altered, but Offline Sorter prompts you for a new file name to write the sorting results into. During the prompt, a default file name appears that is constructed by appending numbers to the original filename until an unused filename has been made. The latest sorting results are in the filename-0n file with the largest *n*.
- **Overwrite Original, backup to filename-0n before each save** – With each save, Offline Sorter always overwrites the original file with the latest sorting results, but it saves the previous version of the file to a backup file. Offline

Sorter constructs the backup file name by appending numbers to the original file name.

- **Overwrite Original, backup to filename-0n on first save only** – This item is the same as above, except that Offline Sorter makes the backup only on the first save. That is, the first time that a file is saved, Offline Sorter saves a backup of the original to a backup file whose filename it constructs by appending numbers to the original file name. Offline Sorter overwrites the original file. Subsequent saves overwrite the original file, but they do not save backups. With this option, the filename-0n backup file contains the data originally read into Offline Sorter, no matter how many times the file is saved.
- **Overwrite Original, backup to filename-0n and update each save** – On each save, Offline Sorter overwrites the original file and the backup file. The backup file contents are always “one save behind” the original file.
- **Overwrite Original, do not backup** – This item just overwrites the original file on each save and doesn't make backups. Plexon does not recommend this option unless the you back up the data files elsewhere.

### Saving a New PLX File

Offline Sorter also supports a more flexible means to save timestamps and waveforms into a PLX file. No matter how the waveforms are read into Offline Sorter (from any supported file format, or from a DDT file after waveform extraction has been performed), you can save the waveforms into a PLX file. Select the **File | Export to New .PLX** menu command to bring up the following dialog box:



- **File** – Select the name of the output file. Click the **Browse** button to bring up a file browser.

- 
- **Store Invalidated Waveforms** – If you select this item, Offline Sorter still writes waveforms that have been invalidated to the output PLX file. If unselected (default), Offline Sorter does not write the invalidated waveforms to the file.
  - **Store Unsorted Waveforms** – If you select this item (the default), Offline Sorter writes out the waveforms that are not sorted. If unselected, Offline Sorter writes out only sorted waveforms.
  - **Store Only Waveforms in the Time Range** – Select this item to enable the specification of a time range for an output file. If selected, Offline Sorter does not write waveforms with timestamps outside the specified range to the output file. If this item is unselected (the default), Offline Sorter writes all waveforms out regardless of timestamp.
  - **Store Reduced Number of Samples Per Waveform** – The item enables you to store shortened waveforms in the output file. For example, if the sort range is reduced, only the relevant portion of each waveform can be saved.
  - **Store Continuous Data** – Select this item to store continuous data channels in the .plx file. If this item is not selected, continuous channels are not stored in the .plx file, although there are continuous channels in the current data.
  - **Store Original Data in Spike Continuous Data Channels** – This option is only available when a ddt file has been read. Offline Sorter writes out the spikes extracted from each continuous-data channel as spike channels. If this option is checked, Offline Sorter writes out the original ddt data to the output file as additional continuous channels. Offline Sorter writes out the continuous waveform data using a block size that is specified in the **File Import** tab in the **Tools | Options** dialog box. If you use a value larger than the default (120) it can cause compatibility problems with programs that read the saved .plx file.

## Exporting Information

Offline Sorter has powerful and flexible facilities to export information for further analysis. Offline Sorter can export information to the following destinations:

- Text files: you can choose the delimiter between values.
- A MATLAB file: The data appears as a matrix variable when you open the file in MATLAB. This option is only appears MATLAB is installed on the same computer as Offline Sorter.
- A Microsoft Excel spreadsheet: This option only appears if Microsoft Excel is installed on the same computer as Offline Sorter. During the information export, Excel opens and accepts the exported information into the first sheet on the spreadsheet. You can then save the spreadsheet from within Excel.

Offline Sorter can export three varieties of information:

- Per-Waveform information via the **File | Export Per-Waveform Data** menu selection: This item generates a line of information (a row) for each waveform that satisfies the export criteria. The kind of information items that can be exported per-waveform include the timestamp, unit assignment, and the values of any feature.
- Per-Channel information via the **File | Export Per-Channel Data** menu selection: This item generates a line of information for each channel in the file. The kind of information items that can be exported per-channel includes the numbers of waveforms and units, and the Sorting Quality Statistics.
- Per-Unit information via the **File | Export Per-Unit Data** menu selection: This item generates a line of information for each unit in the file. The kind of information items that can be exported per-channel include the numbers of waveforms, the cluster center positions and the sigmas and other normality statistics.

When exporting to a text file or to Excel, you can choose an option to export a row of column-header text to help in deciphering the exported data.



## Exporting Per-Waveform Data

Selecting the **File | Export Per-Waveform Data** menu selection displays the following window:

**Export Per-Waveform Data**

**Format**

- ☒ Text file
- ☐ Matlab file
- ☐ Excel

**Delimiter to use between values:**

Space

☐ Export a Row of Column Headers

**Which Channels**

- ☐ Current Channel only
- ☒ All Channels into one file
- ☐ All Channels, one file per channel

**What Waveforms to Export**

- ☐ Export Invalidated Waveforms
- ☒ Export Unsorted Waveforms
- ☐ Export Only Waveforms in the Time Range From (sec): 0 To (sec): 5.299275

**What to Export for Each Waveform**

**Available:**

- PC 1
- PC 2
- PC 3
- Slice 1
- Slice 2
- Peak
- Valley
- Peak-Valley
- Energy
- Nonlinear Energy
- Timestamp
- PC 4
- PC 5
- PC 6
- PC 7

**Chosen:**

- Channel
- Unit

Buttons: Add >>, Add All, << Remove, Remove All, Move Up, Move Down

☒ Append Raw Waveform A/D Data in units of ☒ mV ☐ ADC Counts

**What File(s) to Store the Exported Data in**

File: C:\PlexonData\NSSample.txt

Buttons: Cancel, Export!

To export information

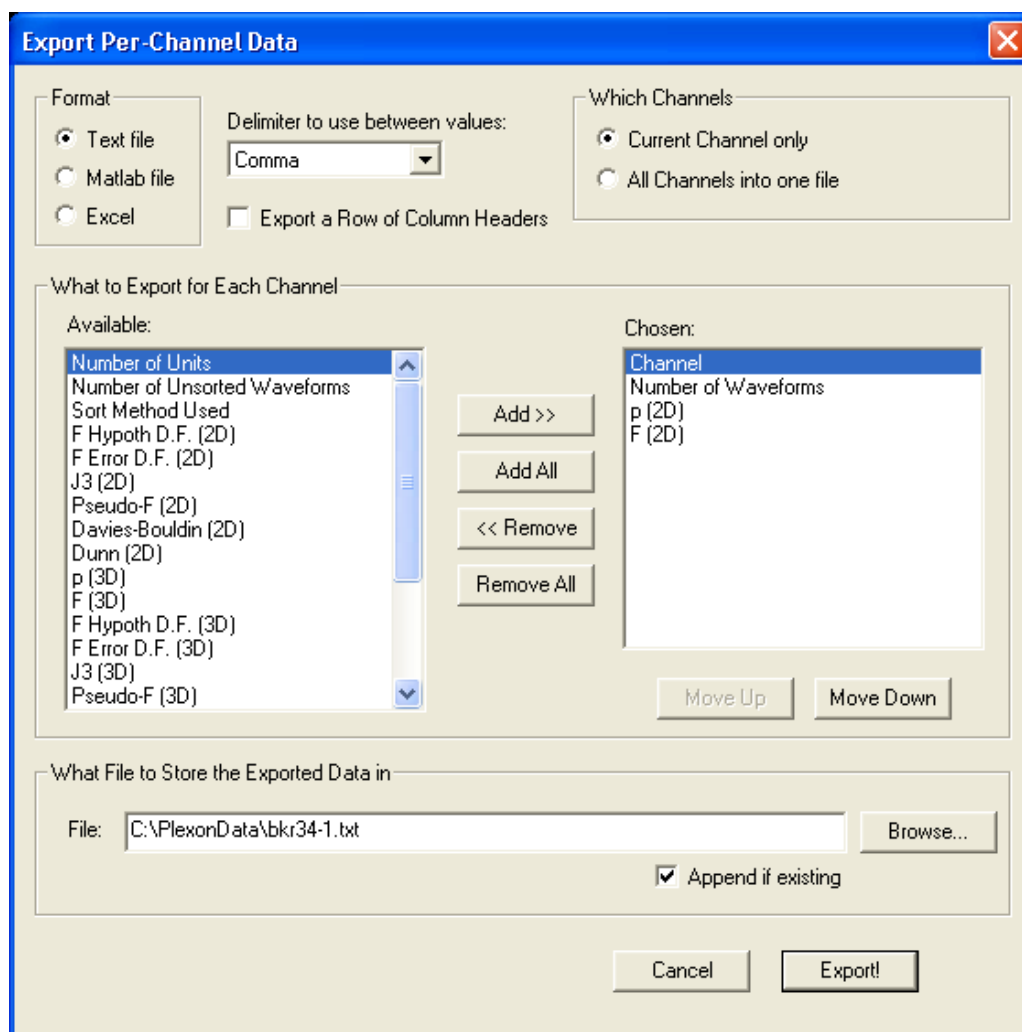
- 1 Select the **Format** of the export - either text file, MATLAB file (if MATLAB is available on the computer) or Microsoft Excel (if available on the computer).
- 2 For text export, choose which delimiter to use between values. Delimiter choices include comma, comma and space, space, or semicolon.
- 3 For Excel and text export, check the **Export a Row of Column Headers** to get a line of header text that labels each of the columns. This makes the exported file easier to read.
- 4 Select **Which Channels** to export. If a current channel has been selected in Offline Sorter when the **File | Export Per-Waveform Data** menu selection

was selected, the default is the current channel only. Also, all channels can be exported to a single file. For text and MATLAB exports, another option is to do all channels, but generate one exported file per channel. In this case, a channel number is appended to the filename.

- 5 Select **What Waveforms to Export**. By default, Offline Sorter exports all valid waveforms. Check the **Export Invalidated Waveforms** to also include any waveforms that have been marked invalid in the exported information. Uncheck **Export Unsorted Waveforms** to cause any waveforms that have not been assigned to a unit to be ignored during the export. You can also select a time range for exported waveforms.
- 6 Select **What to Export for Each Waveform**. This enables you to select what items appear in each column for each exported row of data. The **Available** list box on the left shows all the items that can be exported, and the **Chosen** list box on the right shows those that have been selected for export. Click on one or more items to highlight them and move items back and forth between the list boxes by then pressing the **Add>>** or **<<Remove** buttons. Use other convenience buttons to add all items, remove all items, or reset the chosen items to the “factory defaults”. The order that the items appear in the **Chosen** list box controls the order that they appear in the output file. You can change the order by selecting one or more items in the **Chosen** list box and pressing the **Move Up** or **Move Down** buttons. Offline Sorter remembers the chosen items, and their order, so that the next time you export information, the same items are chosen automatically. Besides the obvious Channel, Unit, and Timestamp items, you can export the value of any of the Features Available for Sorting for each waveform.
- 7 In addition to, or instead of, the exported feature values, Offline Sorter can additionally export the raw A/D values associated with each waveform. To do so, click the **Append Raw Waveform A/D Data** checkbox. The A/D values appear in additional data columns after the exported feature columns. To export only the waveform A/D values, deselect all features from the **Chosen** list box and click the **Append Raw Waveform A/D Data** checkbox. The number of appended columns is equal to the number of points per waveform. The waveform A/D values can be exported in units of microvolts (the default), or in raw ADC counts, by clicking the radio buttons near the checkbox.
- 8 Select **What File(s) to Store the Exported Data in** by typing a filename or selecting one using the **Browse** button. For exports involving multiple channels into multiple files, Offline Sorter appends a channel number to the filename entered.
- 9 Click the **Export!** button
- 10 For exports to Excel, save the spreadsheet using Microsoft Excel.

## Exporting Per-Channel Data

Selecting the **File | Export Per-Channel Data** menu selection displays the following window:



To perform an export

- 1 Select the **Format** for the exported information - either text file, MATLAB file (if MATLAB is available on the computer) or Microsoft Excel (if available on the computer).
- 2 For text export, choose a delimiter to use between values. Delimiter choices include comma, comma and space, space, or semicolon.
- 3 For Excel and text export, check **Export a Row of Column Headers** to get a header line of text that labels each of the columns. This makes reading the exported file easier.
- 4 Select **Which Channels** to export. If you have selected a current channel in Offline Sorter when the **File | Export Per-Channel Data** menu selection was

selected, the default is the **Current Channel only**. You can also export all channels to a single file.

- 5 Select **What to Export for Each Channel**. This enables you to select what items appear in each column for each exported row of data. The **Available** list box on the left shows all the items that can be exported, and the **Chosen** list box on the right shows those that you have selected for export. Click on one or more items to highlight them and move items back and forth between the list boxes by then pressing the **Add>>** or **<<Remove** buttons. Use other convenience buttons to add all items, remove all items, or reset the chosen items to the “factory defaults”. The order that the items appear in the **Chosen** list box controls the order that they appear in the output file. You can change the order by selecting one or more items in the **Chosen** list box and pressing the **Move Up** or **Move Down** buttons. Offline Sorter remembers the chosen items, and their order, so that the next time you export information, the same items are chosen automatically. The items available include the channel number, the number of units, waveforms, and unsorted waveforms in the channel, the sort method used to sort the channel, and the values of any of the Sorting Quality Statistics (both 2D and 3D). You can also export the features selected for each axis of feature space as a numerical code.
- 6 Select **What File to Store the Exported Data** in by typing a filename or selecting one by using the **Browse** button. To append the data onto the end of an existing file, click **Append if existing**.
- 7 Click **Export!**
- 8 For exports to Excel, save the spreadsheet using Microsoft Excel.

## Exporting Per-Unit Data

Selecting the **File | Export Per-UnitData** menu selection displays the following dialog box:

**Export Per-Unit Data**

**Format**

☐ Text file  
☐ Matlab file  
☒ Excel

Delimiter to use between values:  
Comma

☐ Export a Row of Column Headers

**Which Channels**

☒ Current Channel only  
☐ All Channels into one file

**What to Export for Each Unit**

Available:

- Number of Waveforms
- Sort Method Used
- Small ISI %
- RMSE
- Center X (2D)
- Center Y (2D)
- Ellipse Angle (2D)
- Sigma X (2D)
- Sigma Y (2D)
- Center X (3D)
- Center Y (3D)
- Center Z (3D)
- Sigma X (3D)
- Sigma Y (3D)
- Sigma Z (3D)

Chosen:

- Channel
- Unit Number

Add >> Add All << Remove Remove All Move Up Move Down

☒ Append Template Waveform Data in units of ☒ microVolts ☐ ADC Counts

**What File to Store the Exported Data in**

File: C:\PlexonData\UCSF\nb.13p1.06.xls Browse...

☒ Append if existing

Cancel Export!

To perform an export

- 1 Select the **Format** for the exported information - either text file, MATLAB file (if MATLAB is available on the computer) or Microsoft Excel (if available on the computer).
- 2 For text export, choose a delimiter to use between values. Delimiter choices include comma, comma and space, space, or semicolon.
- 3 For Excel and text export, check the **Export a Row of Column Headers** to get a header line of text that labels each of the columns. This makes reading the exported file easier.
- 4 Select **Which Channels** to export. If a current channel is selected in Offline Sorter when you select **File | Export Per-Unit Data**, the default is **Current**

**Channel only**, which means Offline Sorter exports all units on the current channel. You can also export all units on all channels to a single file.

- 5 Select **What to Export for Each Channel**. This enables you to select what items appear in each column for each exported row of data. The **Available** list box on the left shows all the items that can be exported, and the **Chosen** list box on the right shows those that you have selected for export. Click on one or more items to highlight them and move items back and forth between the list boxes by then pressing the **Add>>** or **<<Remove** buttons. Use other convenience buttons to add all items, remove all items, or reset the chosen items to the “factory defaults”. The order that the items appear in the **Chosen** list box controls the order that they appear in the output file. You can change the order by selecting one or more items in the **Chosen** list box and pressing the **Move Up** or **Move Down** buttons. Offline Sorter remembers the chosen items, and their order, so that the next time you export information, the same items are chosen automatically. The items available include the Channel number, unit number, number of waveforms, the values of the normality statistics (mean, sigma, skew and kurtosis) for both 2D and 3D. The numerical values are in units appropriate to the corresponding feature in the current feature space. You can also export the features selected for each axis of feature space as a numerical code.
- 6 In addition to, or instead of, the exported feature values, Offline Sorter can additionally export the raw A/D values associated with each waveform. To do so, click the **Append Raw Waveform A/D Data** checkbox. The A/D values appear in additional data columns after the exported feature columns. To export only the waveform A/D values, deselect all features from the **Chosen** list box and click the **Append Raw Waveform A/D Data** checkbox. The number of appended columns is equal to the number of points per waveform. The waveform A/D values can be exported in units of millivolts (the default), or in raw ADC counts, by clicking the radio buttons near the checkbox.
- 7 Select **What File to Store the Exported Data in** by typing a filename or selecting one using the **Browse** button. To append the data onto the end of an existing file, click **Append if existing**.
- 8 Click **Export!**
- 9 For exports to Excel, save the spreadsheet using Microsoft Excel.

## Chapter 5

# Reference Section

---

Channel Types and Associated Channels .....	5-2
Stereotrode and Tetrode Data.....	5-4
Features Available for Sorting .....	5-9
Details of the Sorting Algorithms .....	5-13
Sorting Quality Statistics .....	5-19
Playing a Movie of the data .....	5-22
Examining Events .....	5-23
Manual Short ISI Spike Removal .....	5-25
Interval Selection .....	5-27
Waveform Inspection .....	5-29
Template Files .....	5-33
Parameter Scans .....	5-36
Re-thresholding Spike Data .....	5-38
Assigning Unsorted Points to Closest Unit.....	5-39
Batch File Processing.....	5-39
Printing.....	5-44
Large File Tips .....	5-45

---

## Channel Types and Associated Channels

Offline Sorter designates each channel of data as having one of three Channel Types:

- Spike – a channel that contains waveform segments (spikes).
- Slow Continuous – a channel that contains continuous data digitized at slow, non-spike frequencies (e.g. less than 10 kHz). Data in these channels typically represents LFPs or other experimental data such as a joystick position. Offline Sorter, by default, does not run Waveform Extraction on Slow Continuous channels.
- Spike Continuous – a channel that contains continuous data digitized at fast, spike-appropriate frequencies (e.g. greater than 10 kHz). Furthermore, the continuous data is presumed to contain neural firings within it. A working definition of a Spike Continuous channel might be “channels for which it makes sense to perform Waveform Extraction”. A Spike Continuous channel can have *both* continuous data *and* spikes within it, which is the case after Waveform Extraction has been run on a Spike Continuous channel.

The Channel Type appears in the **Type** column in the **Channels** sub-tab of the File Info View. The Channel Type also appears in the Chan Map View.

If Offline Sorter loads a file type that can only contain continuous data, for example, a Plexon DDT file, it automatically assigns the channels as **Spike Continuous**, so that waveform extraction is possible. However, if Offline Sorter loads a file type that contains both spikes and continuous data, by default it does not automatically designate any continuous channels as **Spike Continuous** channels; this occurs because the **Allow designation of Spike Continuous channels at load time** checkbox in the **File Import** tab of the **Tools | Options** dialog defaults to unchecked. In this case, Offline Sorter designates all continuous channels as **Slow Continuous**. If the checkbox is checked though, Offline Sorter decides whether a continuous channel is **Spike Continuous** or **Slow Continuous** based on the channel's digitization frequency; if the digitization frequency is at least 10 kHz, the channel is assigned as **Spike Continuous**. In most cases, you can change the type of a continuous channel by using the **Chan Map** tab in the Info View.

Within a neural data file, each channel has a “raw” channel number that uniquely identifies that channel within the data file. However, when Offline Sorter reads in the data file, it refers to the channel by name (e.g. “sig001”). For example, the analog channel that uses channel identifier 2 in a PLX file can become Slow Cont channel “AD01” when Offline Sorter reads in the file. The Channel Map view shows which raw channel number (or numbers) is associated with each channel. For Spike Continuous channels, both a raw spike channel and a raw continuous data channel can be associated with a single Offline Sorter channel. For stereotrode and tetrode data, either two or four raw spike channels can be associated with a single Offline Sorter channel. In a future version of Offline Sorter, the Chan Map View will be able to change the mapping of raw channels to Offline Sorter channels.



---

**Note:** For continuous data channels in PLX files, there is a discrepancy between the channel identifiers in the data file (which are 0-based) and the channel names and numbers that the Plexon software displays (1-based channel numbers).

Starting with PLX file format version 104 (which is written by RASPUTIN Version 2.1 and above and Recorder version 2.3 and above), a PLX file can store “associations” between channels. You can configure SortClient to record *both* a spike channel *and* a continuous data channel *from the same physical electrode*. In this case, the PLX file can contain this association, that is, it records the fact that spike channel  $x$  and continuous data channel  $y$  are actually from the same source.

In Offline Sorter, only Spike Continuous channels can have associations, even though the PLX file format supports associations between spike channels and slow continuous channels. Therefore, by default, Offline Sorter ignores channel associations because the default setting of the **Allow designation of Spike Continuous channels at load time** checkbox (discussed previously) is unchecked. However, if this checkbox is checked, then when Offline Sorter loads a PLX file in which spike channel  $x$  and continuous channel  $y$  are associated, it shows the data as being from a single Spike Continuous channel  $x$ . The Spike Continuous channel contains both continuous data and waveform segments (spikes). The Waveform Extraction activity on the Spike Continuous channel then *replaces* the spikes on that channel with the new ones resulting from the extraction.

**Note:** With SortClient, you can associate a spike channel with a continuous channel that has a slow digitization rate (or any digitization rate that is not the same as the timestamp frequency). In this case, Offline Sorter does not automatically associate the spike and the continuous channels. Except under special conditions, Offline Sorter does not allow or ignores associations with any continuous data channel that does not have a digitization rate equal to the timestamp frequency.

The **Chan Map** sub-tab under the Info View allows inspection and in some cases changing both the Channel Type (e.g. changing a Slow Continuous into a Spike Continuous channel to enable Waveform Extraction), and the associations (e.g. associating a Spike channel with a Spike Continuous channel to indicate that both came from the same electrode) within a file. Saving the data using **Export to New .PLX** preserves these associations.

Suppose that spike channel  $x$  in the PLX file is really associated with continuous channel  $y$ .

Then:

For a Version 104 PLX File without associated channels (or a pre-Version 104 PLX File)	<ul style="list-style-type: none"> <li>• Offline Sorter displays the spike channel <math>x</math> as a channel with Channel Type “Spike” and a “Spike Continuous” channel <math>y</math>.</li> <li>• Performing Waveform Extraction on the Spike Continuous channel <math>y</math> does not affect the spikes in channel <math>x</math>.</li> <li>• If desired, you can use the Chan Map view to associate the channels.</li> </ul>
For a Version 104 PLX File with associated channels	<ul style="list-style-type: none"> <li>• Offline Sorter displays the associated spike channel <math>x</math> and continuous channel <math>y</math> as a single Spike Continuous channel <math>x</math>.</li> <li>• Performing Waveform Extraction on the Spike Continuous channel replaces the spikes originally in channel <math>x</math> in the PLX file.</li> <li>• If desired, you can use the Chan Map tab to disassociate the Spike Continuous channel; then, they appear as independent Spike and Spike Continuous channels.</li> </ul>

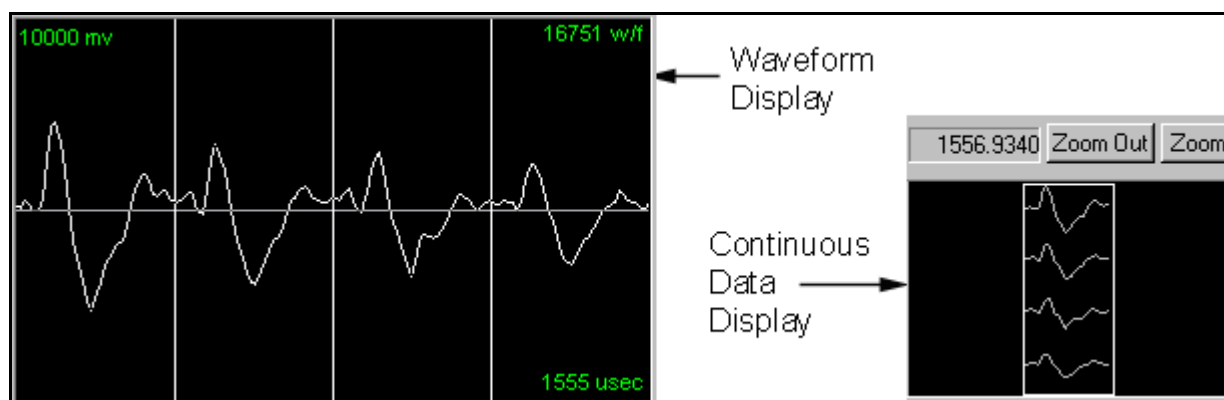
**Note:** There is some ambiguity inherent with the voltage scale for Spike Continuous channels. For PLX files, Spike channels generally have a maximum voltage at the ADC input of 3V, whereas for Spike Continuous channels the value is 5V (these values are recorded in the PLX file header). Thus, performing waveform extraction on a Spike Continuous channel can result in spikes that appear to have a voltage that is off by a factor of 5/3.

## Stereotrode and Tetraode Data

By default, Offline Sorter analyzes stereotrode and tetraode data by concatenating the individual waveforms from each electrode in the tetraode or stereotrode group into a single waveform. The resulting concatenated waveform is twice (for stereotrode) or four times (for tetraode) as long as each individual waveform.

The Waveform Display shows the concatenated waveform with vertical white lines separating the waveforms from each individual electrode channel within the

group. The Continuous Data Window represents the waveform group arranged vertically, indicating that they were acquired at the same point in time.



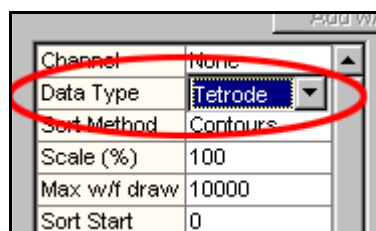
You can then apply all cluster display and sorting features to the concatenated waveforms. Offline Sorter includes sorting features that calculate values on the per-electrode waveforms instead of the concatenated waveforms (e.g. “Peak EL 1” is the peak value of the waveform on the first electrode only). Also, ratios can be formed from per-electrode features (e.g. “Peak EL 1 / EL 2” is the peak value on electrode 1 divided by the peak value on electrode 2).

**Note:** You can set the data type mode (**Single** electrode, **Stereotrode**, or **Tetrode**) by using the **Data Type** variable in the Control Table. By changing the data type, you force the file to be interpreted as containing data of that type. See the next section for details.

Stereotrode and tetrode data formats in Plexon PLX, DataWave UFF, and Neural-nx files are automatically detected when the file is opened.

## Converting Data Type

An advanced feature in Offline Sorter enables you to force-interpret spike data to be any data type. For example, data files that were acquired as single-channel data can be treated as if they were stereotrode or tetrode data. To convert the data type, use the **Data Type** drop-list in the Control Table,



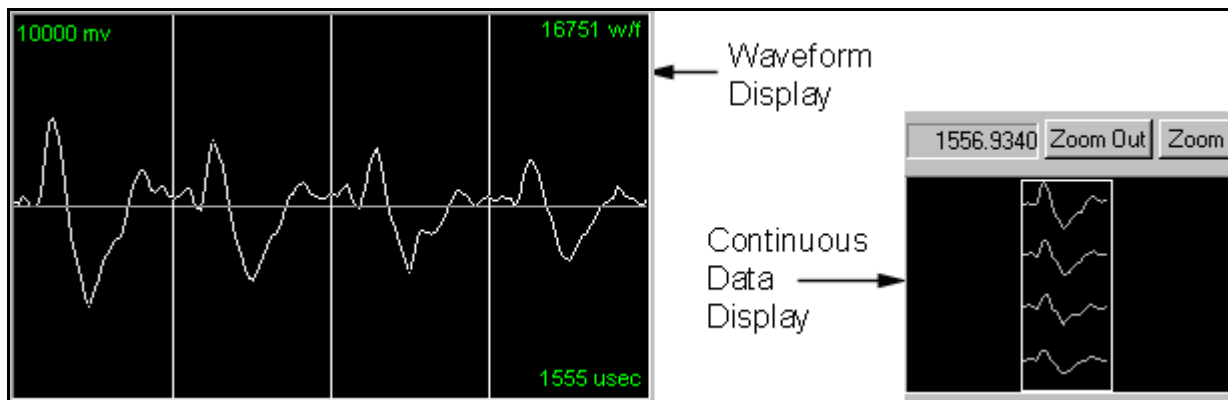
or, use the drop-down list in the **Chan Map** tab of the Info View.

When a different data type is selected, Offline Sorter performs a conversion algorithm on the data, which results in the data being re-loaded and all sorted unit assignments and parameters to be lost.

The algorithm used to convert data type is as follows, using the conversion of single-channel data into tetrode data as an example:

- 1 Group the channels into groups of four (two for stereotrode data). If the original single-channel data file has  $N$  channels, there will be  $N/4$  tetrode channels ( $N/2$  for stereotrode) after the conversion.
- 2 Using the first channel in each group, step through each spike in the channel and obtain its timestamp. Find the corresponding spikes with the same timestamp on the other three channels in the group. Failure to find a spike on the other channels with the same timestamp results in a “dropped waveform”. A count of dropped waveforms appears on the status bar when a converted channel is opened.
- 3 The waveforms for the four time-coincident spikes are concatenated together to form the “expanded waveform” for the single spike on tetrode channel. If each single channel contains 800 microsecond-long waveforms, the four spikes on the four single channels become a single spike on the tetrode channel with a length of  $4 \times 800 = 3200$  microseconds. Any dropped waveforms are represented as zeros in the corresponding position in the expanded waveform.

For tetrode data, the Waveform display shows the expanded waveform with vertical white bars separating the portions corresponding to each electrode. In the Continuous Data display, the waveforms for each electrode are stacked vertically to indicate their time-coincidence.



**Note:** At present, Offline Sorter cannot perform stereotrode- or tetrode-aware waveform extraction.

### Saving Converted Data

There are two ways of saving converted data, the normal **Save** function, and the **Save As New PLX** function.

The normal **Save** function saves a .PLX file with a data type that is the same as the original file. In this mode of saving a converted file, the waveforms from all channels in a channel group are tagged with the sorted unit number.

---

The **Save As New PLX** function records the data type with the file, and saves concatenated waveforms, and the saved file contains one “trodal” channel for each channel group.

Either way, the converted data type is stored in the PLX file, and Offline Sorter automatically displays the file with the correct “trodality” when it re-opens.

## Disabling Electrodes

With Stereotrode/Tetrode data, an electrode (or wire) can break or otherwise just produce noise or artifacts. You can choose to “ignore” individual electrodes within a stereotrode/tetrode by unchecking the **Use EL** checkboxes located above the Waveform Display. These checkboxes only appear if you load stereotrode/tetrode data into Offline Sorter.



When an electrode is disabled, a red box with the word 'Disabled' will appear in the space where that electrode is being displayed in the waveform display. Disabled electrodes do not contribute to PCA calculations or to calculations of any other Feature values for the waveforms.

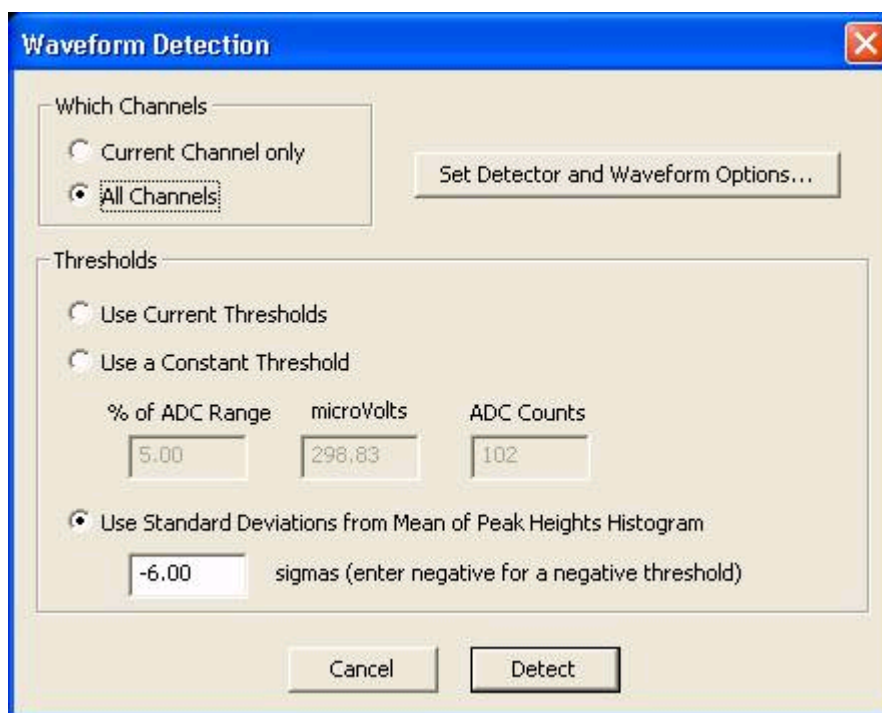
## Automatic Waveform Detection

Offline Sorter can also perform waveform detection on all the spike continuous channels in a file. Use the following procedure to set up waveform detection.

To initiate waveform detection

- 1 Click **Waveforms | Detect**.

The **Waveform Detection** window appears.



- 2 In the **Which Channels** area, click a channel selection. To perform waveform detection on the current channel, click **Current Channel only**. To perform waveform detection on all channels, click **All Channels**.
- 3 To open the **Waveform Detection Options** window, click **Set Detector and Waveform Options**. For more information on detector and waveform options, see [“Specifying the Waveform Detection Options” on page 4-6](#).
- 4 In the **Thresholds** area, click one of the following detection thresholds:
  - **Use Current Thresholds** – This setting uses the thresholds that are currently set for each channel. You can view and edit the threshold settings for each channel in the **Info | Chan Params** tab.
  - **Use a Constant Threshold** – Type a constant threshold value into of the following boxes:
    - % of ADC Range** – the percentage of the analog-to-digital converter (ADC) range indicated in the **Continuous Data** window (see [“Continuous Data Window” on page 3-54](#))
    - microVolts** – voltage threshold in  $\mu\text{V}$

---

**ADC Counts** – the threshold in raw analog-to-digital conversion counts

- **Use Standard Deviations from Mean of Peak Heights Histogram** – Type the number of standard deviations (sigmas) from the mean to set the threshold at that distance from the mean of the peak height histogram. A negative number sets the threshold to be less than the mean. A positive number sets the threshold to be greater than the mean.

5 To start the waveform detection process, click **Detect**.

## Features Available for Sorting

You can choose each dimension of the 2D or 3D feature space used in Offline Sorter from the list of available features. Any of the following features can be made “active”, so that they appear in the X axis, Y axis and Z axis feature drop-lists in the toolbar. See “Axis Selection” under “[Cluster View Common Functionality](#)” on page 3-21.

In the following,

$w(t) = [w(1), \dots, w(n)]$  is the waveform ( $n = \text{number\_of\_points\_in\_waveform}$ )

$p1(t) = [p1(1), \dots, p1(n)]$  is the first principal component vector

$p2(t) = [p2(1), \dots, p2(n)]$  is the second principal component vector

$p3(t) = [p3(1), \dots, p3(n)]$  is the third principal component vector

$\Sigma$  represents a sum over  $t$

The available features are:

- **PC1** – The waveform projection onto the first principal component

$$PC1 = \Sigma p1(t) * w(t)$$

- **PC2** – The waveform projection onto the second principal component

$$PC2 = \Sigma p2(t) * w(t)$$

- **PC3** – The waveform projection onto the third principal component

$$PC3 = \Sigma p3(t) * w(t)$$

- **PC4-PC8** – The waveform projections onto higher principal components

- **Slice 1** – The waveform height at a selected position in time. You can select the time position in the Control Table or the Waveform display.

$$\text{Slice 1} = w(i) : \text{the waveform voltage at time } t=i$$

- **Slice 2** – The waveform height at a selected position in time. You can select the position in the Control Table or the Waveform display.

$$\text{Slice 2} = w(j) : \text{the waveform voltage at time } t=j$$

- **Slice 3** – The waveform height at a selected position in time. You can select the position in the Control Table or the Waveform display.

$$\text{Slice 3} = w(k) : \text{the waveform voltage at time } t=k$$

- **Slice 4** – The waveform height at a selected position in time. You can select the position in the Control Table or the Waveform display.

$$\text{Slice 4} = w(1) \text{ :the waveform voltage at time } t=1$$

- **Peak** – The maximum voltage amplitude across the entire waveform length

$$\text{Peak} = \max(w(t))$$

- **Valley** – The minimum voltage amplitude across the entire waveform length

$$\text{Valley} = \min(w(t))$$

- **Peak-Valley** – The difference between the maximum and minimum waveform amplitude

$$\text{Peak-Valley} = \max(w(t)) - \min(w(t))$$

- **Peak, Valley FWHM** – The “Full Width at Half Maximum” for the peak or valley of the waveform, expressed in milliseconds. The exact points where the waveform crosses the half max value are obtained by interpolation.

- **Peak, Valley Tick** – The clock tick (as measured from the beginning of the captured waveform) at which the peak or valley occurs. No interpolation or waveform fitting is done, so these values are “quantized” to integer values.

- **Area** – The total integrated area underneath the waveform

$$\text{Area} = \sum \text{abs}(v(t))$$

- **Energy** – The waveform energy, as it is calculated for spike detection. See [“Waveform Detection View” on page 3-38](#).

$$\text{Energy} = \left(\frac{1}{\text{width}}\right) \sum v(t) v(t)$$

- **Non-Linear Energy** – The non-linear waveform energy, as it is calculated for spike detection. See [“Waveform Detection View” on page 3-38](#).

$$\text{Non-Linear Energy} = \left(\frac{1}{\text{width}}\right) \sum [v(t) v(t) - v(t-1) v(t+1)]$$

- **Timestamp** – The timestamp of the waveform in the file. Note: Selecting the timestamp as one of the dimensions in the Cluster view allows the viewing of the other feature(s) as a function of time. Discontinuities or drifts in time can indicate a change in the electrode position or the recording conditions.

- **Per-Electrode Features** – (only available for stereotrode or tetrode data)  
Most of the above features are available calculated per-electrode. For example, the feature “Peak EL 1” is the peak value of the waveform within the first electrode in a stereotrode (or tetrode). “Peak EL 2” is the peak value within the second electrode of the 'trode. For further discussion, see [“Stereotrode and Tetrode Data” on page 5-4](#).

- **Per-Electrode Ratio Features** – (only available for Stereotrode or Tetrode data) These features are formed by taking the ratio of Per-Electrode feature



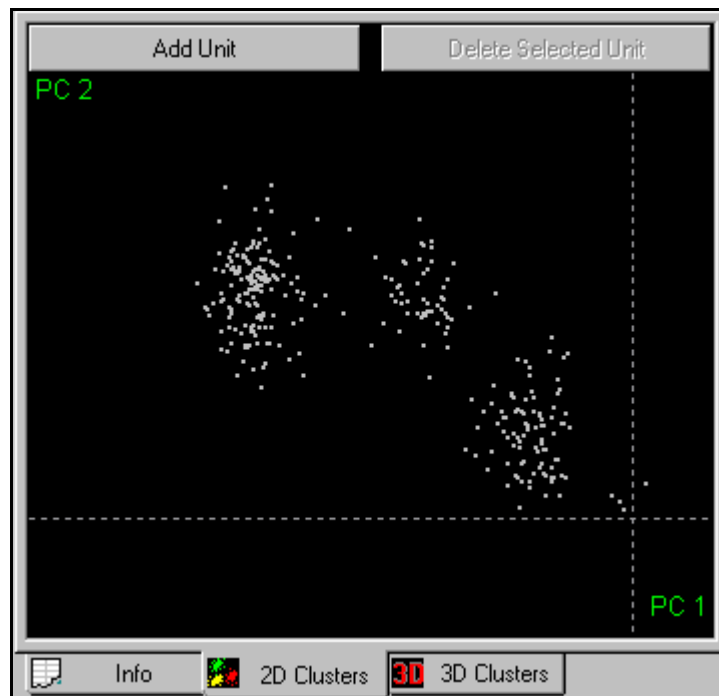
---

values. For example, the feature “Peak EL 1/EL 2” is “Peak EL 1” divided by “Peak EL 2”. Ratios with a denominator value of 0 are arbitrarily set to 0. For further discussion, see [“Stereotrode and Tetrode Data” on page 5-4](#).

## Principal Components

The projections of the first eight principal components can be plotted on any axis of the Cluster views. By default, the principal component vectors are calculated using all the valid waveforms of the current channel and all time points in the waveform between Sort Start and Sort End. For more information, see [“Adjusting the Sort Range” on page 4-25](#).

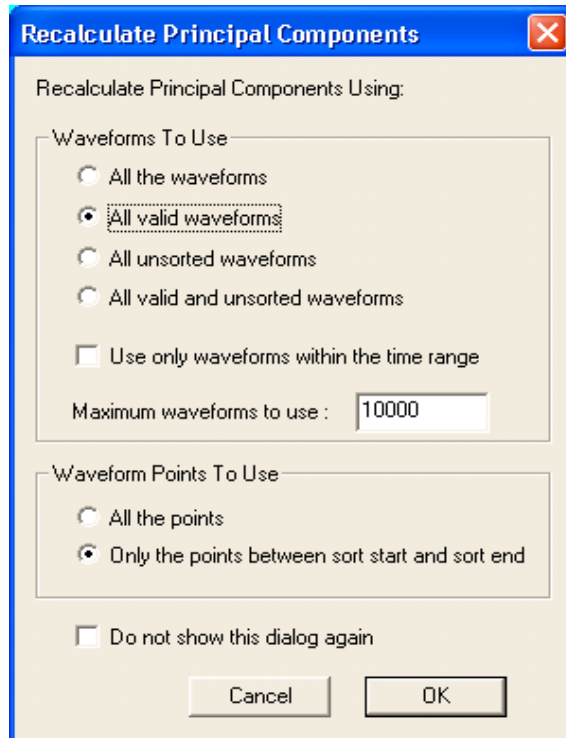
For example, they appear as follows when viewed in the 2D Clusters view.



The dashed lines are the X and Y axes, representing the zeroes of PC1 and PC2. The points representing the lowest-amplitude waveforms have principal component projections near zero (i.e. near the intersection of the two axes).

You can recalculate the principal components using subsets of waveforms (e.g. only valid waveforms, only unsorted waveforms, only waveforms within a time range) or using only the points specified by the **Sort Start** and **Sort End** parameters in the Control Table. The various options appear in the **Tools | Recalculate**

the **Principal Components** menu command, which displays the following window.



You can restrict the PCA Calculation to use only a certain number of waveforms. This is a useful feature for larger files, because the PCA calculation for channels containing large numbers of waveforms can be time-consuming. All of the options that are available in the **Recalculate Principal Components** window are also available under the **Feature Calc** tab of the **Tools | Options** window. Offline Sorter remembers all of the options that control the PCA calculation between invocations.

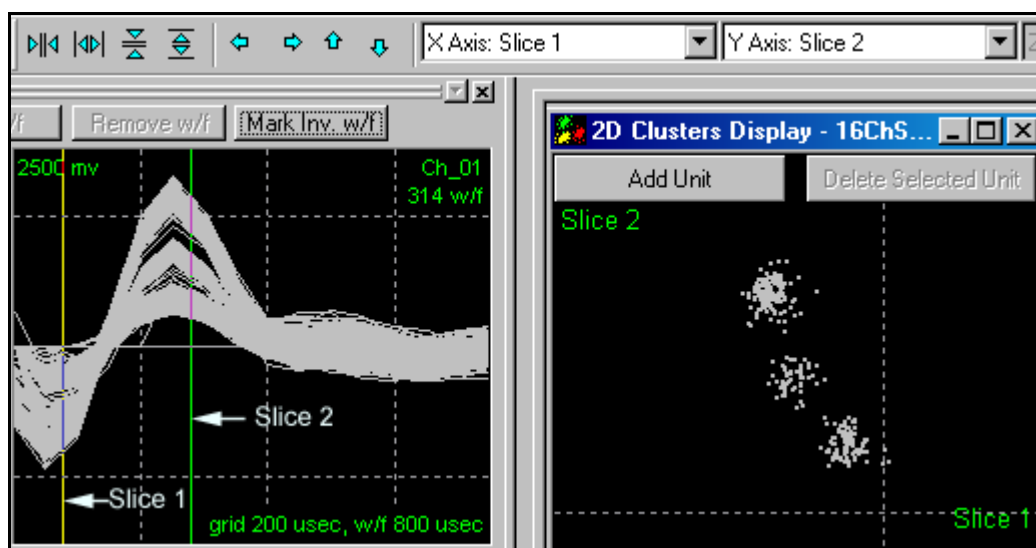
Click **Do not show this dialog again** to suppress the display of the **Principal Components** window, after you select or click **Recalculate the Principal Components**. To re-enable the display click to clear the **Do not show the Recalculate Principal Components dialog** checkbox on the **Feature Calc** tab of the **Tools | Options** window.

### Voltage at Specific Slices in Time

To view waveform voltages at specific points in time on the X, Y, or Z axis of the Cluster views:

- Select **Slice 1**, **Slice 2**, **Slice 3**, or **Slice 4** for the X, Y, or Z axis using the **Drop-Down List Boxes** on the **Toolbar**.
- Adjust the time values for **Slice 1**, **Slice 2**, **Slice 3**, or **Slice 4** in the Control Table. When visible, you can also drag the Slice 1, Slice 2, Slice 3, and Slice 4 lines into position with the mouse on the Waveform display.

- Right click the Cluster view and select **Redraw** in the pop-up menu.



- To update the Cluster view immediately when any of the slice n values are changed, select the **Redraw Cluster Display Immediately when Slice Position Changes** in the **Display** tab under the **Tools | Options** dialog box.
- By default, each channel “remembers” its slice positions independently. To have the same slice positions used across all channels, click **Use Same Slice Positions for all Channels** in the **Feature Calc** tab under the **Tools | Options** dialog box.

## Details of the Sorting Algorithms

### Feature Space Used for Sorting

The K-Means, Standard E-M, Valley Seeking, and T-Distribution E-M algorithms behave differently depending on how they are initiated and on the setting of the **Use Projected 2D space (Screen Space) when performing automatic sorting methods in 3D** checkbox on the **Tools | Options Sort** page. The available options are:

- If the sorting is initiated from the 2D Cluster view, the sorting is done in 2D feature space.
- If the sorting is initiated from the 3D Cluster view, and the **Use Projected 2D space (Screen Space) when performing automatic sorting methods in 3D** checkbox is checked, the sorting is done in 2D projected space based on the current 3D viewpoint. That is, the 3D points are first projected into a 2D screen space using a standard projection transformation, then the sorting algorithm is run using the point positions in this 2D screen space. The cluster centers are actually marked in this same 2D screen space. The sorting results depend on the viewpoint position and orientation in the 3D Cluster view. This was the default behavior prior to Offline Sorter version 2.2.

- If the sorting is initiated from the 3D Cluster View, and the **Use Projected 2D space (Screen Space) when performing automatic sorting methods in 3D** checkbox is not checked, then the sorting algorithm runs in true 3D feature space. That is, cluster centers and inter-point distances are calculated in true 3D space. The sorting results are independent of the viewpoint position and orientation in the 3D Cluster view. This is the default behavior.

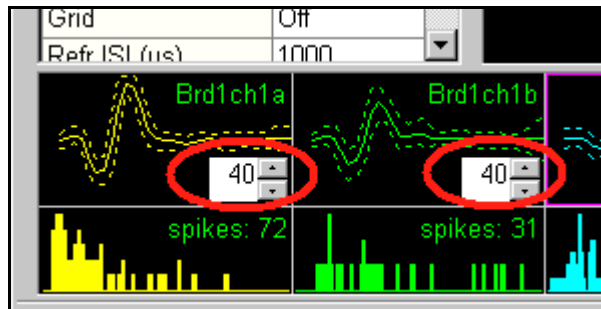
### Template Algorithm

The Template matching sorting algorithm is designed to mimic the online real-time sorting performed by the Plexon MAP data-acquisition hardware.

For each waveform, the algorithm:

- 1 calculates the sum-of-squares differences from this waveform to all the templates:
- 2 finds the template with the minimum sum-of-squares difference
- 3 if the minimum difference is less than the Fit Tolerance for this template, the waveform is assigned to this template's unit.

You can adjust the Fit Tolerance (measured in units of 0.01% of the maximum A/D value) in the Units view:



Let,

$$\text{waveform} = w = [w_1, \dots, w_N]$$

$$\text{template} = t = [t_1, \dots, t_N]$$

where

$w_i$  and  $t_i$  values are in AD units (typically between  $-2047$  and  $+2048$ )

$N$  = # digitized data points

**Note:** For Plexon MAP users: The MAP system has 12-bin A/D conversion ( $2^{12}=4096$ ). Therefore the AD values of waveform and template amplitudes are integers between  $-2047$  and  $+2048$ .

---

The difference measure ( $d_w$ ) between a waveform and a template is the square root of the mean sum of square error:

$$d_w = \sqrt{\frac{1}{M} \sum (w_i - t_i)^2}$$

where

$M$  is the sort width for the current channel in terms of the number of data points

$\sum$  represents a sum over  $i$  (the samples in the sort width)

For stereotrode and tetrode data, the sum is

$$S_u = \sqrt{\frac{1}{M} \sum \sum (w_i - t_i)^2}$$

where the first summation is over the electrodes.

The waveform is assigned to a template's unit if

$$\text{Template Fit Tolerance} < d_w$$

## K-Means Algorithm

The K-Means algorithm is described in any pattern recognition textbook. The algorithm runs in feature space, and requires starting positions for the clusters (and thus implicitly the number of clusters) to be specified before the algorithm begins.

Algorithm outline:

- 1 The user first sets initial cluster centers.
- 2 Each point is assigned to the cluster whose center is nearest, using a simple 2D or 3D Euclidean distance.
- 3 Cluster centers are then recalculated using the points assigned to each cluster, using a simple center-of-gravity method.

Steps 2 and 3 are repeated until the cluster centers no longer change.

When the K-Means algorithm converges, an outlier removal algorithm runs.

The standard deviation of the distances from each point to its cluster center computes. Any point that is more than a user-specified constant (OutlierThreshold) times this standard deviation away from its cluster center is removed from the cluster.

In Offline Sorter, you can run the K-Means algorithm on unsorted data after the user has specified cluster centers, or you can run it to re-sort data, using the cluster centers of the existing clusters as a starting point. To run K-Means with existing centers, click **Sort | Continue Sort Current Channel | Continue Sort**.

**References:**

Webb, Andrew, *Statistical Pattern Recognition* Second Edition, John Wiley and Sons 2002  
and many others...

**Standard E-M Algorithm**

The Expectation-Maximization algorithm is a widely-known iterative, parametric approach that fits a normal mixture model (a mixture of several Gaussians) to the available data points. It is so named because for each iteration the algorithm goes through two phases, an Expectation calculation followed by a Maximization phase. You can find the derivation of the equations for the algorithm in many sources, some of which are listed below.

The standard E-M algorithm requires the selection of the number of Gaussians to be fit (i.e. the number of units) and an initial starting assignment of points to clusters. In Offline Sorter, if a Continue Sort is being performed (See [“Continue Sorts” on page 5-19.](#)), the existing point assignments and number of clusters is used. If a normal Standard E-M sort is requested, the user is required to specify starting cluster centers, and a K-Means sort is used to get the initial assignment of points to clusters.

Offline Sorter actually implements a variant of the Standard E-M algorithm called REM-2, as discussed in the Sahani reference below. With this variant, a new parameter “beta” is introduced that influences the convergence of the E-M algorithm. In the Sahani reference, this beta parameter is systematically decreased as the algorithm progresses in an “annealing schedule”, which causes clusters to be “frozen out” as “phase transitions” are encountered at certain values of beta. When the beta value is set to 1.0, the REM-2 method becomes equivalent to the Standard E-M algorithm.

**References:**

Webb, Andrew, *Statistical Pattern Recognition* Second Edition, John Wiley & Sons 2002

Shahani, Maneesh *Latent Variable Models for Neural Data Analysis*, Doctoral Thesis, California Institute of Technology 1999

McLachlan and Peel, *Finite Mixture Models*, John Wiley & Sons 2000

**Valley Seeking Algorithm**

The valley seeking algorithm in Offline Sorter starts by calculating the number of neighbors that each point in feature space has. A neighbor is defined as a point that is within a certain critical distance of the target point.

The critical distance is given by

$$\text{Dist} = 0.25 * \text{sigma} * \text{Parzen Mult}$$

---

where sigma is the standard deviation of the distances of all points to the overall mean, and the Parzen Mult is a user-specified parameter so-named for its similarity to the Parzen density estimation kernel.

**Note:** Creating the neighbor counts inherently scales as  $N^2$ , where  $N$  is the number of points. This fact limits the number of points that can be handled by this algorithm to several thousand before computing resources are overwhelmed.

The neighbor count for each point is then put into a list and the list is sorted so that the points with the most neighbors are at the top of the list.

Next, a set of seed points are found that are “local maxima” in that they have more neighbors than any of their neighbors do. The search for seed points starts at the top of the list sorted according to the number of neighbors, so the seed points are points near the centers of the densest clusters. Points with less than a certain minimum number of neighbors cannot become seed points.

Then, initial clusters are formed from each seed points and all of the seed point’s neighbors. Also, if any of the seed point’s neighbors are themselves seed points, their initial clusters are combined.

At this stage, the clusters consist of a circle (or sphere in 3D) of points centered around each density maximum with a radius of roughly the critical neighbor distance, with the other points unassigned.

The unassigned points are then assigned to clusters using a modified Fukunaga algorithm. This is an iterative algorithm; at each step, unassigned points are assigned to the cluster that the majority of its neighbors belong to. This continues until points stop being assigned or a maximum number of iterations is completed.

**References:**

Fukunaga, Keinosuke *Introduction to Statistical Pattern Recognition*, Chapter 11, Academic Press 1972

## **T-Distribution E-M Algorithm**

A variant of the E-M algorithm by Shoham et al (see references that follow). This algorithm has two important features:

- 1) Instead of fitting a mixture of normal (Gaussian) distributions, T-Distributions are used instead. The T-Distribution is similar to a Gaussian, except that the tails are “fatter”. This characteristic can enable E-M algorithms based on T-Distributions to better handle outliers in the data set.
- 2) The algorithm has built into it a way to automatically decide on the number of clusters. This feature is agglomerative in nature, as the algorithm starts with a large number of clusters and combines them together as the algorithm proceeds, so as to minimize a penalized likelihood function.

As discussed in the Shoham reference, the algorithm does not perform well when using the theoretical value for the number of parameters per component. Offline Sorter introduces a multiplier on this theoretical degrees of freedom value that can be adjusted. Empirically, values of around 20-30 produce the most pleasing clusters. Smaller values generally lead to too many clusters, larger values lead to too few. The Scan feature in Offline Sorter can perform sweeps over this sorting parameter, so its affect on sorting performance can be investigated.

The algorithm must be started with an existing clustering of the data, generally with too many clusters defined, as the algorithm can agglomerate clusters together as it proceeds. The initial clusters can either be taken from the currently-defined units (See [“Continue Sorts” on page 5-19.](#)), or they can be generated automatically. To automatically generate the starting configuration, the K-Means algorithm is used. You can set the number of cluster centroids used in this initial sort by using the **Tools | Options | Sort** dialog in the **Number of seed clusters for T-Distribution algorithm** edit box. The initial cluster centers given to the K-Means algorithm are evenly spaced along a diagonal line through the bounding box in feature space that contains all of the points to be sorted.

#### References:

Shoham, S., Fellows, M., and Normann, R. *Robust, automatic spike sorting using mixtures of multivariate t-Distributions*, Journal of Neuroscience Methods, 127(2), 111-122 2003

Figueiredo, Mario and Jain, Anil *Unsupervised Learning of Finite Mixture Models*, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 25 pp.1150-1159, Sep 2003

### Running Automatic Sorting on All Channels

As a convenience, Offline Sorter can automatically run the Valley-Seeking or T-Distribution E-M sorting algorithms sequentially on all channels in the current file. You initiate this operation by selecting the desired sorting algorithm from the **Sort | Sort All Channels** menu. The current values of all parameters that affect the sorting algorithms (the Parzen Multiplier, D.O.F. Multiplier, Outlier Threshold, waveform limiting options, etc.) are used to sort all channels. This is completely equivalent to manually loading each channel in turn and running the sorting on it. You can cancel the sorting operation by pressing the **Cancel** button on the Cluster views or by selecting **Sort | Stop Sorting** menu entry.

The **When performing automatic sorting on all channels** option in the **Tools | Options | Sort** dialog controls whether 2D or 3D feature space is used when sorting all channels.

### Special Sorting

Normally, for the K-Means, Standard E-M, Valley-Seeking, and T-Distribution E-M sorting methods, initiating a Sort on a channel deletes all existing units and create new units according to the algorithm. However, other modes of operation are possible.



---

## Continue Sorts

Some of the sorting algorithms require a set of starting units to be defined. You can normally do this by specifying cluster centers, or by running an initial K-Means sort. However, if units already exist, those units can themselves be used as the starting point for sorting. This is called “Continue sorting” for lack of a better name. You can use the K-Means, Standard E-M, T-Distribution E-M, and (trivially) the Template sorting algorithms for a Continue sort.

To initiate a Continue sort, select the desired sorting algorithm from the **Sort | Continue Sort** menu. Alternately, you can initiate a Continue sort by holding down the **CTRL** key and pressing the button in the Cluster view that normally initiates the sort (i.e. the **Run K-Means Sorting**, **Run E-M**, **Run T-Dist E-M** buttons).

For Template sorting, a Continue sort is the default, in a manner of speaking. The nature of the Template sorting algorithm is such that all waveforms are sorted against all of the available units.

Also, you can do Continue sorts for all channels by selecting an algorithm from the **Sort | Continue Sort All Channels** menu.

## Partial Sorts

When the sorting algorithms are run when not all the waveforms are visible in the Cluster views (because the **Use Only Currently Visible Waveforms** option is in use and the waveforms being displayed is being limited by the **View w/f** setting or by the **Control Cluster Display** checkbox on the Continuous Data views), the behavior of the sorting algorithms changes. All existing units are preserved, and the sorting algorithm is only allowed to run on the currently visible waveforms. This generally results in some number of new units in addition to the originally present units. The final step in the sorting process is to remove any units that are empty (units with no waveforms in them).

You can also do Continue sorts when only some of the waveforms are visible.

## Sorting Quality Statistics

There is no universally agreed upon way to measure sorting quality, so Offline Sorter can calculate and display a number of sort quality metrics. It can also calculate pair-wise cluster separation statistics, and also other supporting statistical data such as measures of normality of distributions.

The following sections describe the sort quality metrics and other statistics available in Offline Sorter. You can view these values on the [Sorting Summary View](#), which is detailed on [page 3-46](#), or on the **Stats** sub-tab of the [Info View](#), which is detailed on [page 3-13](#). Also, you can choose one of the sorting quality metrics to be displayed on the top of the [Cluster Views](#), which are detailed on [page 3-19](#), where it can be used to give immediate feedback on the effectiveness of sorting decisions. See [“Statistics Feedback While Sorting” on page 4-61](#).

## MANOVA F and p-values

The degree to which the selected unit clusters are separated in the 2D and 3D cluster views is determined by a **Multivariate Analysis of Variance (MANOVA)** test. The Null hypothesis in the MANOVA is that all clusters actually come from the same underlying statistical distribution in 2D or 3D space. The p-value is the probability that an error is committed by rejecting the null hypothesis. Therefore, a small p-value indicates that each of the unit clusters has a statistically different location in 2D/3D space, and that the clusters are statistically well separated.

After clusters are selected in 2D or 3D cluster space, you can view the resulting F and p-values derived from the MANOVA test in two and three dimensions in the Stats Spreadsheet in the [Info View](#), which is detailed on [page 3-13](#), and in the cluster view; see [“Cluster View Common Functionality” on page 3-21](#).

### Some details on the MANOVA calculation

In the jargon of MANOVA, the MANOVA “number of groups” is our number of units, and the MANOVA “dependent variables” (DV) are our feature vectors (e.g. principal components). This analysis is that of a single “Factor” experiment, with the “Treatment” being the unit assignments.

The MANOVA F-statistic is calculated using a ratio of determinants to get Wilk's lambda, followed by Rao's method for converting Wilk's lambda into an F value. See “Applied Multivariate Statistics for the Social Sciences”, 2nd Edition by James Stevens, page 191 for more details.

An unweighted MANOVA is performed, in that all groups (regardless of the number of waveforms they contain) contribute equally to the F-statistic calculation, so as not to bias the MANOVA calculation toward rapidly firing units. See “Statistical Principles in Experimental Design”, 2nd Edition by B.J. Winer, page 232.

The conversion of the F-statistic to a p-value is done via a calculation of the incomplete beta function (See “Numerical Recipes in C”, section 6.2), assuming a single-tailed probability distribution.

### Including or not including the unsorted units in the ANOVA calculation:

The **Include unsorted waveforms as a unit in the statistics calculations** checkbox in the **Stats** tab in the **Tools | Options** dialog box controls whether the unsorted waveforms are treated as a unit in the statistics calculation, or simply ignored. The text in parenthesis on the **Multivariate ANOVA** line indicates the current setting for this option. For additional information, see [“Some details on the MANOVA calculation” on page 5-20](#).

**Note:** The multivariate ANOVA test requires a minimum of 2 clusters, one of which can be the unsorted waveforms—if they are included as a unit. Also, it does not make sense to run the ANOVA test if any unit is defined that contains zero waveforms (again, including the unsorted waveforms if they are included).

---

## Interpretation of the Multivariate ANOVA test

This section includes information on the interpretation of the Multivariate ANOVA tests. The test results appear in the lower half of the **Stats** sub-tab.

**Assumption:** The distributions of points is normally distributed in 2D or 3D cluster space.

**The Null hypothesis is:** All unit clusters actually have the same underlying distribution in 2D or 3D space and have the same mean (i.e. do not represent distinct spike waveforms).

**The p-value is:** The probability of rejecting the null hypothesis (that all clusters are actually the same) if it were actually true.

**Conclusion:** The smaller the p-value, the more confident one can be in rejecting the null hypothesis and concluding that the clusters are in fact distinct and represent different units.

## J3

J3 is a non-parametric measure of the quality of sorting (i.e. no normality assumptions are involved).

$J3 = J2/J1$ , where  $J1 = \sum \sum E(f(u, i) - m(u))$  is a measure of the average distance in feature space between points in a cluster ( $f$ ) from their center ( $m$ ).  $E(x)$  represents the Euclidean distance squared ( $\tilde{x} * x$ ). The summations are over units  $u$ , and over feature vectors (points) in each unit  $i$ .

$J2 = \sum N(u) E(m(u) - m)$  is a measure of the average distance between unit clusters. The summation is over units  $u$ ,  $N(u)$  is the number of points in unit  $u$ ,  $E(x)$  represents the Euclidean distance squared.  $m(u)$  is the cluster center for unit  $u$ , and  $m$  is the grand center of all points in all units.

So J3 takes on a maximum value for compact, well-separated clusters.

### References:

Wheeler, Bruce C., *Automatic Discrimination of Single Units in Methods for Neural Ensemble Recordings*, ed. by Nicolelis, M., CRC Press, Boca Raton, 1999

## Pseudo-F

The Pseudo-F statistic is closely related to the J3 statistic. It is essentially J3 that has been adjusted for the number of waveforms and the number of units.

$$\text{Pseudo-F} = (N - g) / (g - 1) J3$$

where  $N$  is the total number of waveforms, and  $g$  is the total number of units.

### Davies-Bouldin

The Davies-Bouldin validity metric is calculated as:

$$DB = \frac{1}{N} \sum_i \max_{j \neq i} \left[ \frac{(d'(i) + d'(j))}{d(i, j)} \right]$$

where  $d(i, j)$  is the Euclidean distance between the centroids of unit  $i$  and unit  $j$ , and  $d'(i)$  is the average distance of each point in unit  $i$  to the centroid of unit  $i$ . The sum is over all  $N$  units, and the max is taken over all units  $j \neq i$ .

Because the inter-cluster distance  $d(i, j)$  is in the denominator, the Davies-Bouldin validity statistic takes on small values for well-separated unit clusters. The Davies-Bouldin validity statistic is a non-parametric measure.

### Dunn

The Dunn validity metric is calculated as:

$$Dunn = \min_i \left[ \min_{j \neq i} \left[ \frac{d(i, j)}{\max_k (d'(k))} \right] \right]$$

where  $d(i, j)$  is the Euclidean distance between the centroids of unit  $i$  and unit  $j$ , and  $d'(k)$  is the average distance of each point in unit  $k$  to the centroid of unit  $k$ . The outermost min is taken over all units  $i$ , the inner min is taken over all  $j \neq i$ , and the max is taken over all units  $k$ .

The Dunn validity index assumes a maximum value for clusters that are compact and well-separated.

## Playing a Movie of the data

A useful technique with Offline Sorter is to “play back” the data in a file in time order. This is useful in determining whether waveform shapes are drifting as a function of time.

You can make Offline Sorter automatically advance time in the Continuous Data Views by using the operations under the **View | Animate** menu. The animation functions are completely equivalent to changing the time range by using the scroll bars on the continuous data views. The animation operations include:

**Play:** starts animating the playback of the file, but defaults advancing time at the normal rate of 1 second per second

**Pause:** pauses the animation; select it again to un-pause

**Stop:** stops the animation

**Slower:** slows down the rate of animation by a factor of  $\frac{1}{2}$

**Faster:** speeds up the rate of animation by a factor of 2

---

**Reset to 1X Speed:** resets to the normal animation rate (1 second per second)

**Rewind:** resets to the beginning of the file

**Reverse:** a toggle that makes the rate of time passage negative (e.g. time flows backwards)

**Loop:** when checked, the animation resets to the beginning of the file when the end of the file is reached

To make operation much easier, you can bind keystrokes to these keys by using the **Tools | Keyboard Shortcuts** menu item. Also, there is a toolbar that includes icons for these functions, but this toolbar is not displayed by default. To display this toolbar, select **Tools | Customize Toolbars** and click **Animation** in the **Toolbars** list.

You can increase the usefulness of the animation features by configuring the Waveform and Cluster Views to only show the waveforms within the current time range, thus what they display changes while the animation is in progress. You can accomplish this with the **Control Cluster Display** and **Control Waveform Display** checkboxes in the Continuous Data windows. If you click these checkboxes, the Cluster views or the Waveform window, or both, show only those waveforms that appear in the Continuous Data window. You can use the **Zoom In** and **Zoom Out** buttons on the Continuous Data window to control how many waveforms are visible in each “frame” of the animation.

Several commercially available screen capture packages are capable of making MPEG movie files of any on-screen activity, including the Offline Sorter displays.

## Examining Events

Although not strictly relevant to spike-sorting, Offline Sorter offers the ability to examine the events in data files. The **Events** sub-tab of the **Info View** lists the events present in the file by name and channel number, with a **Count** of the num-

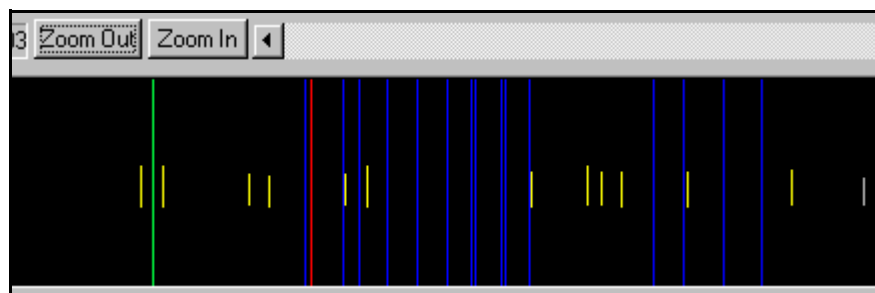
ber of times that events appears, along with a **Show** checkbox and a display color for each.

	Event Ch	Name	Count	Show	Display Color
1	3	FO	6	<input checked="" type="checkbox"/>	Dark Blue
2	4	VB	7	<input checked="" type="checkbox"/>	Teal
3	6	WN	7	<input checked="" type="checkbox"/>	Yellow
4	7	Lick	182	<input checked="" type="checkbox"/>	Blue
5	8	Water	10	<input type="checkbox"/>	Pink
6	258	Start	1	<input type="checkbox"/>	Red

Channels Chan Params Sort Details Global **Events**

Info 2D Clusters

When the **Show** checkbox is checked for an event, Offline Sorter plots each occurrence of that event as a line of the corresponding **Display Color** on the Continuous Data views and on the Rasters View, so that you can inspect the temporal relationships between events and nearby spikes. In the Continuous Data View, the full-height lines are the events, the short lines are spike waveforms.



You can quickly adjust the settings in the **Show** and **Display Color** columns by using **Set As Top Row** in the right-click context menu. The **Randomize Colors** entry in the context menu generates a random color for each different event.

Several of the data formats other than Plexon PLX files have an information type comparable to events (e.g. Datawave Markers). Wherever possible, these information records are translated into events when the data file is loaded.

Plexon PLX files also have the concept of strobed events. Strobed events have an event channel number of 257, but have an additional data fields that stores a data word. In Offline Sorter there is an option that controls the treatment of strobed events. If the **Create Separate Events for All Strobed Codes** option in the **Plexon Files** tab of the **Tools | Options** dialog box is unchecked, all strobed events simply show up as events with a channel of 257. When the **Create Separate Events for All Strobed Codes** option is checked, the strobed events are “expanded” into separate events for each possible data word value. The strobed events are given a name corresponding to the value of their associated data word,

---

the format of which you can control by using the **Name Generation Prefix** and the **Code Representation** options.

**Note:** These options also control how strobed events are exported into NEX files.

To accommodate downstream data analysis, Offline Sorter can create event occurrences when selected units fire. Click **Tools | Create Events at Unit Firing Times** to specify a Unit and an Event ID for the new event. You can display newly-created events in the Continuous Data Window or the Raster Views, and you can save the events to a PLX file by using the **Export to New .PLX** option.

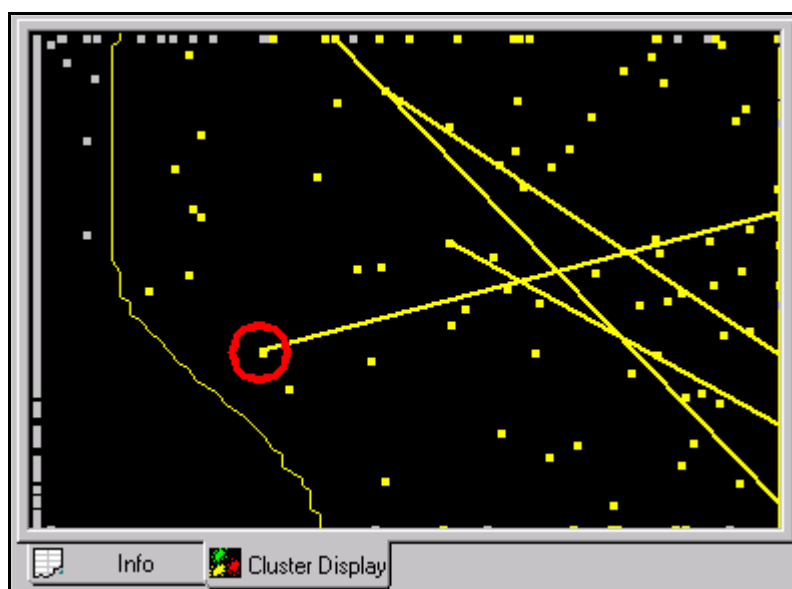
## Manual Short ISI Spike Removal

If a relatively small percentage of spikes have ISIs less than the refractory period, you can manually remove the spike in each pair which is further away from the cluster center.

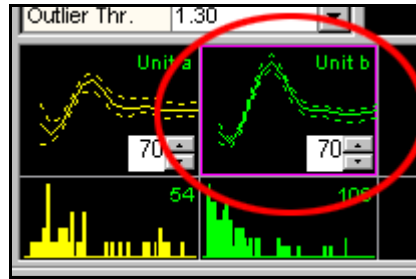
First, follow the procedure outlined in [“Removing Spikes with Short Inter-spike Intervals” on page 4-56](#) to get the lines drawn to connect pairs of spikes with short ISIs.


To remove a waveform from a selected unit

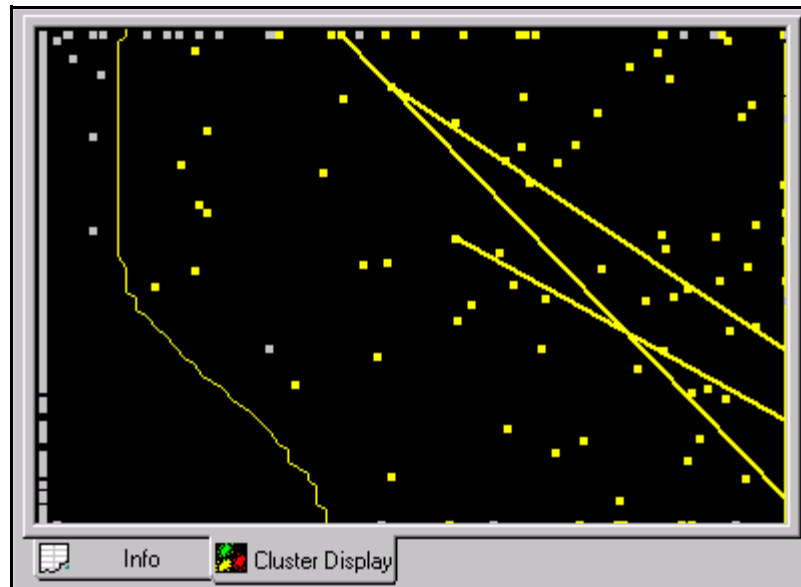
- 1 Zoom in to the waveform's dot (circled below) in the Cluster view using the toolbar control buttons.



- 2 Select the unit: click at the unit's template display.



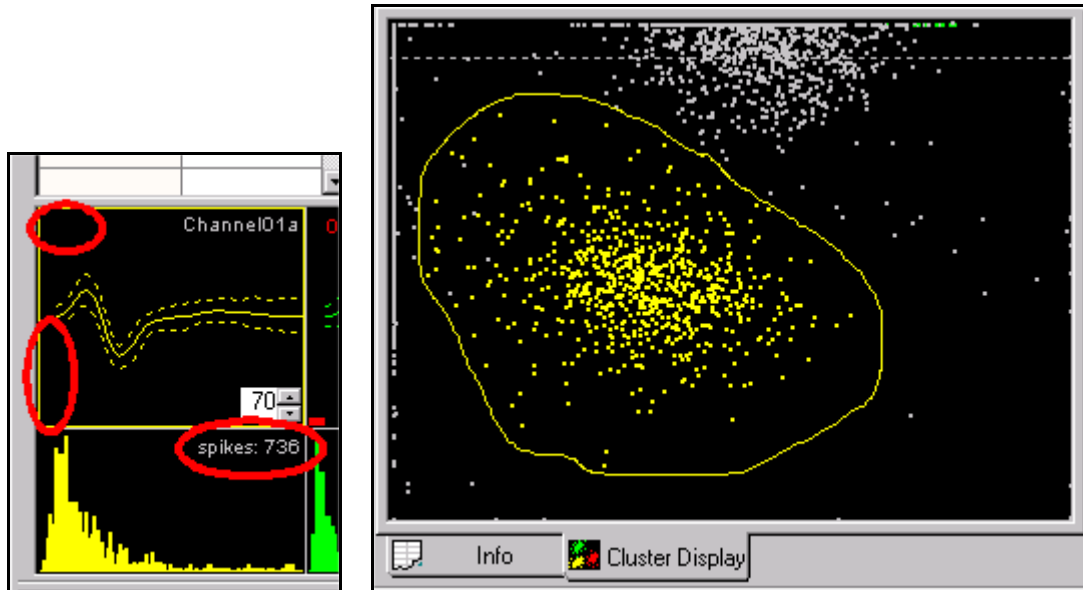
- 3 Select **Units | Remove Waveforms from the Selected Unit** menu command.
- 4 Move the mouse pointer to the Cluster view (note that the pointer changes to a “drawing hand”  ).
- 5 Press the left mouse button and draw a line around the waveform's dot in feature space while pressing the left button.
- 6 Release the mouse button.





---

You can repeat this procedure until all the waveforms furthest away from the cluster center with ISIs less than the refractory ISI have been removed:



If a waveform is removed by accident, you can add it to the selected cluster again within the Cluster view.

To add units

- 1 Select the unit: click the unit's template display.
- 2 Select **Units | Add Waveforms to the Selected Unit** menu command.
- 3 Move the mouse pointer to the Cluster view.

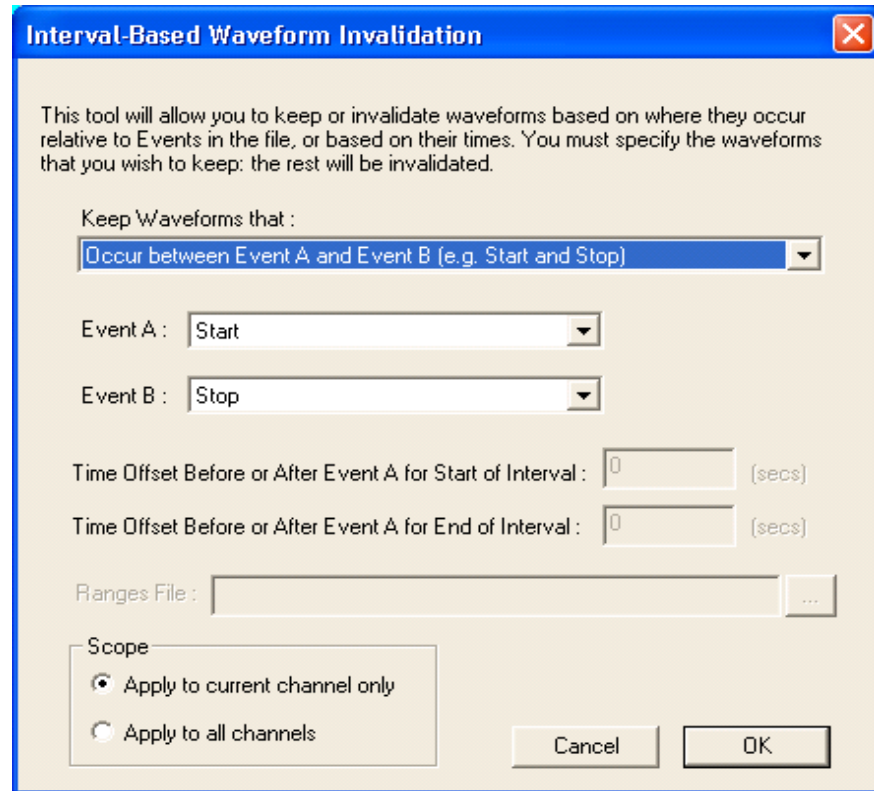
The pointer changes to a blue drawing hand  for adding waveforms.

- 4 Click the left mouse button and drag a line around the units to be added.

## Interval Selection

A valid reason to mark waveforms invalid is that they are simply not of interest. For example, if the experimental animal was not exhibiting “interesting” behavior during certain time intervals, you can invalidate the waveforms collected from that interval simply to get them out of the way. Offline Sorter contains a tool that enables the selection of waveforms within some interval in time around an event (or between events) in the data file. Or more accurately, it enables you to invalidate those waveforms that do not meet the selection criteria (those that are not in

the interval). Select **Tools | Interval Selection** from the main menu to display this window:



First, select the desired action in the **Keep Waveforms that** droplist. The rest of the controls on the dialog either enable or disable as is appropriate for the droplist selection. The **Keep Waveforms that** droplist selections are:

- **Occur between Event 1 and Event 2 (e.g. Start and Stop)** – This selection enables both the Event 1 and Event 2 droplists, and shows a list of all the event names that occur in the data file. Pick the Event 1 that marks the beginning of the desired waveforms, and pick the Event 2 that marks the end. The waveforms that are not between Event 1 and Event 2 are invalidated.
- **Occur within a time span around Event 1** – This selection enables the Event 1 droplist as well as the time offset edit boxes. Select the Event in the Event 1 droplist and enter the number of milliseconds before and after the event for which waveforms are to be kept.
- **Occur between a time offset from Event 1 and a time offset from Event 2** – This selection enables both the Event 1 and Event 2 droplists, as well as the offsets. This mode is like the first option, except the interval between the start and stop events extends by a time offset. The waveforms are kept if they occur between the time of Event 1 plus the first offset and the time of Event 2 plus the second offset.

- 
- **Occur within an arbitrary range of timestamps** – With this selection, no events need to be selected, only the beginning and end of a time interval (as measured from the beginning of the file - timestamp 0) needs to be specified. All waveforms that occur *outside* of this time range are invalidated.
  - **Occur outside of an arbitrary range of timestamps** – With this selection, no events need to be selected in this case, only the beginning and end of a time interval (as measured from the beginning of the file - timestamp 0) needs to be specified. All waveforms that occur *within* this time range are invalidated. This is the inverse of the above option.
  - **Occur outside of time ranges read in from a file** – This item selects and reads a text file that contains ranges of times to invalidate. Each line contains a channel number followed by the start and end times for the invalidation range, which are separated by a space, tab, or comma. The times are specified in seconds, from the beginning of the file. A special value of -1 for the channel number is interpreted as meaning “apply to all channels”.

You can choose to invalidate just the current channel, or all channels in the file. You can make this selection by using the radio buttons under Scope in the dialog box. If Offline Sorter reads the ranges in from a text file, it determines the channels affected based on the contents of the Ranges file.

As soon as the **OK** button is pressed, the filtering takes place. You can undo the operation by using the **Undo** mechanism, or by selecting the **Waveforms | Mark All Waveforms Valid** menu selection.


## Waveform Inspection

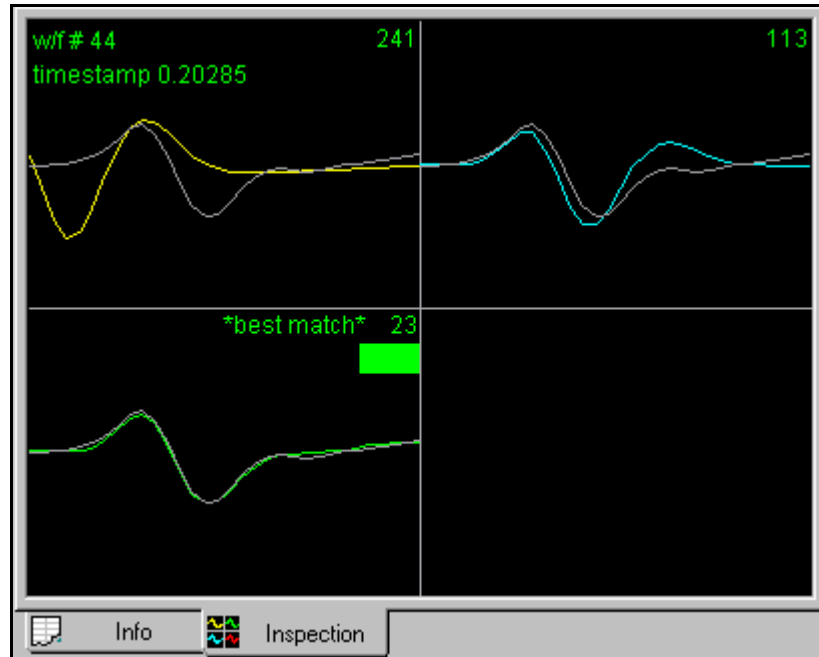
After units have been sorted, you can compare individual waveforms with the template (average waveform) of each of the sorted units in the Waveform Inspection window.

The waveform inspection process involves three steps:

- 1 [Activating the Waveform Inspection Window \(page 5-30\)](#)
- 2 Selecting a waveform, either by [Selecting Waveforms in the Continuous Data Window \(page 5-30\)](#) or by [Selecting Waveforms in the Cluster Views \(page 5-31\)](#).
- 3 [Comparing the Waveform with the Templates \(page 5-32\)](#)

## Activating the Waveform Inspection Window

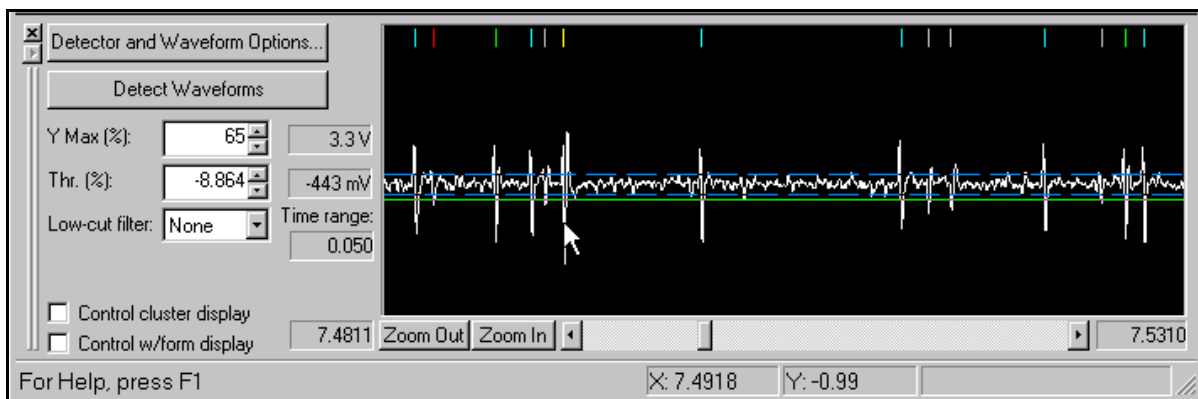
- Select the **View | Waveform Inspection Window** menu command, or, press the  button on the toolbar.



## Selecting Waveforms in the Continuous Data Window

For continuously digitized data files

- Use the **scroll bar** and **Zoom Out/Zoom In buttons** to locate the waveform to be inspected.  
**Note:** The **X** (time in seconds) and **Y** (amplitude in volts) coordinates of the mouse position is displayed in the status bar below the scroll bar.
- Left-click on the thresholded waveform** to select the waveform.



For data files with waveform segments (e.g. PLX files)

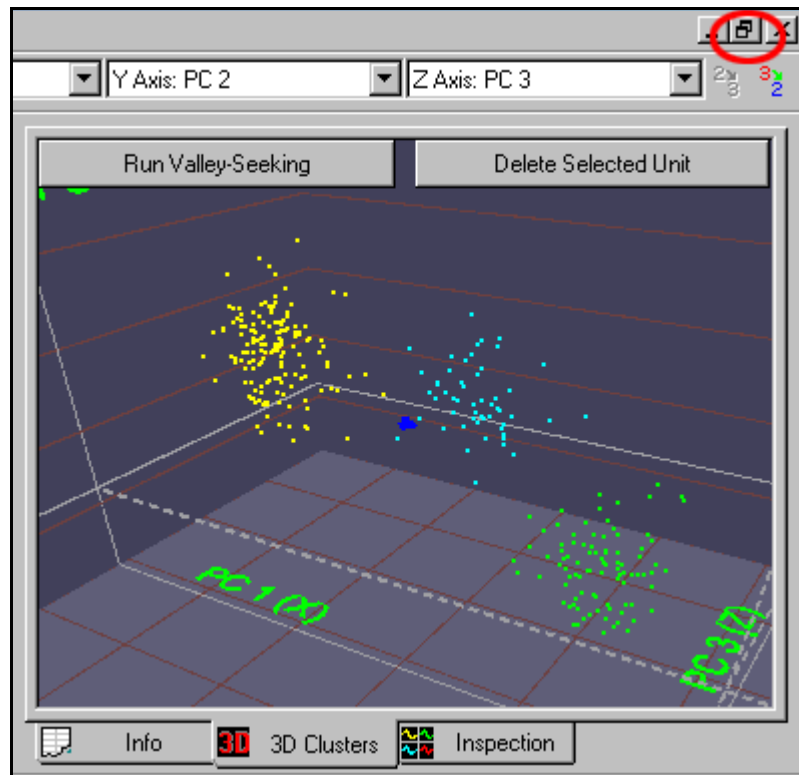
- 1 Use the **scroll bar** and **Zoom Out/Zoom In buttons** to locate the waveform to be inspected.
- 2 Adjust the **W/f zoom** to contract or expand the display of each waveform
- 3 **Left-click on the waveform** to select the waveform

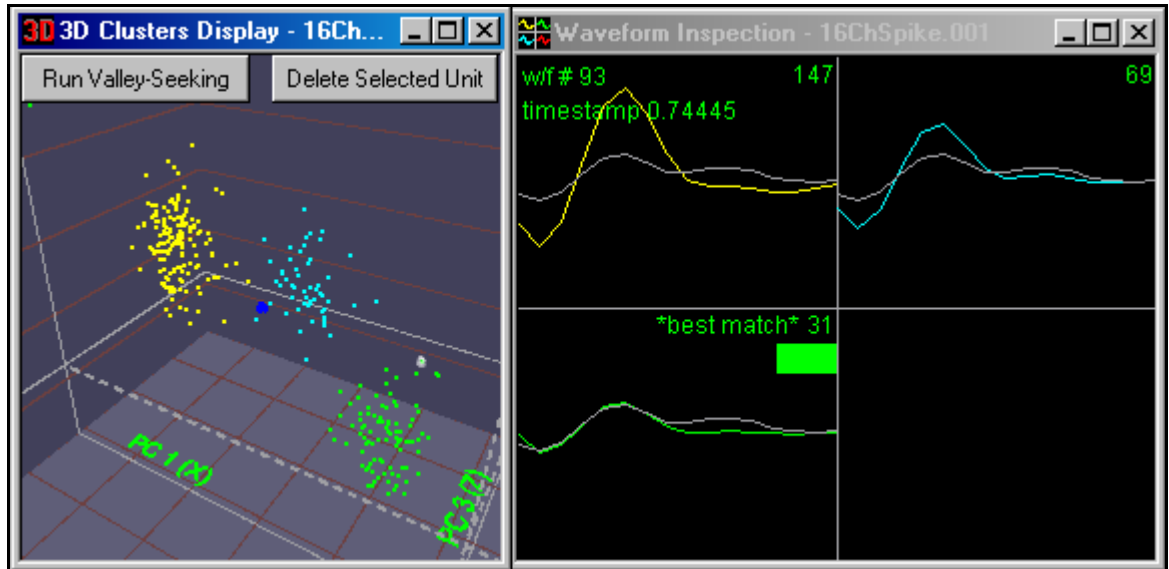


The selected waveform is then viewable in the Waveform Inspection Window.

Selecting Waveforms in the Cluster Views

- 1 Click the **restore window button** (circled below) in the Multi-Display window to view both the Cluster views and the Inspection window simultaneously:



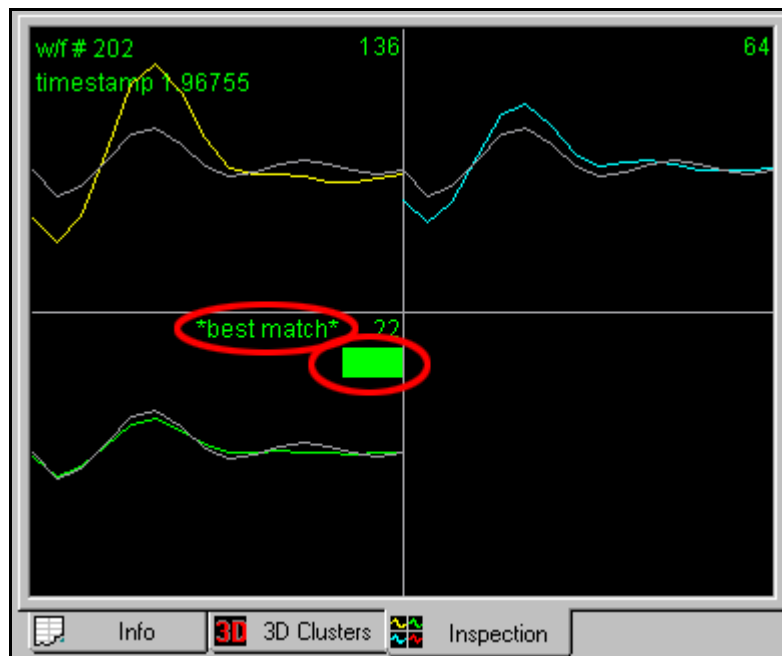


- 2 To select a waveform, **left-click on the waveform's dot** in the Cluster view (hold down the **ALT** key before left-clicking if the 3D Cluster view is being used)

The selected waveform is then viewable in the Waveform Inspection window, as shown above.

### Comparing the Waveform with the Templates

The waveform selected in the Continuous Data window or the Cluster views (drawn in white in the Waveform Inspection window) is plotted against the average waveform (template) of each sorted unit (drawn in color) in the Waveform Inspection window:



---

The degree to which the selected waveform matches each of the sorted unit templates is given by the sum of squared difference between the templates and the waveform, shown in the upper right of each plot.

The Unit with the minimum sum of squared difference with the selected waveform is denoted with the symbol **\*best match\*** in the upper portion of the box (circled in previous illustration).

The Unit to which the waveform is assigned is illustrated by a **colored bar** (circled in previous illustration).

In the earlier example, the selected Waveform (# 202 on this channel) occurred at time  $t = 1.96755$  seconds and is assigned to the green unit (Unit b).

**Note:** In most cases for correctly sorted units, the unit template (average waveform) that best matches the waveform (minimum sum-of-squared difference) should also be the unit to which the waveform is assigned. That is, the **\*best match\*** and the “colored bar” should be in the same template box.

## Template Files

### Saving Templates

As discussed above in the *Spike Sorting Step-by-Step* chapter, after you select units by using the manual or automatic clustering methods, a template can be used to resort the spikes. You can save the following items in a TPL text file:

- templates
- contours
- tolerances
- sort widths
- feature selections
- principal components
- thresholds, filters, and detection parameters for continuous data files

You can then use the saved information to perform sorting on other data files or export it to other analysis programs. Sorting using saved templates is useful for files recorded in succession with the same number of channels, gain on each channel, and global sorting parameters (waveform length and prethreshold time).

If you use Contour sorting to sort a channel, Offline Sorter saves the contour information in the TPL file. Offline Sorter also saves the sort method used for the channel in the TPL file, so you can resort some channels using templates and some with contours.

If waveform extraction has been performed on a channel with continuous data, Offline Sorter saves the extraction parameters in the TPL file. Likewise, if alignment was performed on the spikes, Offline Sorter saves the alignment parameters.

The **Sort Details** sub-tab of the **Info** tab shows what sorting has been performed on which channels using which sorting parameters, and it is a good indication of what will be written to the TPL file.

To save the information in a TPL file

- 1 Sort each channel using templates—see [“Template Algorithm” on page 5-14](#)—(adjust template tolerances, set the sort width, etc.) or using the Contours method.
- 2 Select **File | Save Sort Info to .tpl** menu command.

The TPL file contains information only for the channels that have been sorted.

**Note:** The only sorting methods that can be captured in a TPL file are Template and Contours. If a channel was sorted using another method (e.g. Valley-Seeking), you can approximately capture the results into a TPL file as a template sort. To do this, you need to adjust the template fit parameters so that the results of the template sort approximate the desired sorting.



### CAUTION

#### Uncaptured Operations

You can perform sequences of operations on a channel that cannot be captured in a TPL file. Thus, the same operations are not exactly duplicated when a saved TPL file is re-applied. Examples of operations that are not saved in TPL files include:

- any operation that invalidates waveforms (e.g. removing outliers)
- multiple applications of alignments or sorting operations. Only the last operation is saved to the TPL file.
- manually adding or removing waveforms from clusters.

## Sort Using Saved Templates

You can sort both continuous data files and spike/waveform data files using templates saved in TPL files. To sort a file using saved templates, first open the continuously digitized data file, or open the spike/waveform data file—see [“Opening the Digitized Data Files” on page 4-3](#).

To select the .tpl template file

- Select **File | Sort Using .TPL file** menu command, or, select **Sort | Sort Using .TPL file** menu command.

The templates, tolerances, contours, feature selections, extraction and alignment parameters, and sort widths that you saved for each channel in the TPL file are then used by Offline Sorter to sort the current data file.



---

For continuously digitized data files, threshold, low-cut filter, detector type (as described in [“Specifying the Waveform Detection Options” on page 4-6](#)) settings are also saved in the TPL file for each channel and applied to the new file to extract the waveforms for sorting.

When a channel is sorted using a saved TPL file, the order of operations is:

- 1 extract the waveforms using the saved parameters (for continuous data channels only)
- 2 perform alignment on the spikes (if indicated by the saved TPL file)
- 3 perform either template or contour sorting on the spikes

## Save Templates to a Sort Client EXP File

**Note:** This feature is only relevant for users of the Plexon MAP system.

The following scenario can occur: A Plexon MAP system collects data from an experimental animal, using the settings contained in a Sort Client experiment (EXP) file. Among other settings, the EXP file contains the templates for the units that the MAP system hardware sorts in real-time. After the data-taking session, Offline Sorter reads in the PLX file collected during the experiment and “cleans up” the real-time sorting done by the MAP system. This is sometimes necessary because units can appear or disappear and waveform characteristics can drift over the course of the experiment. Suppose that the same experimental animal is to be connected to the MAP system the next day for another session. It is preferable to use the cleaned-up unit templates in the MAP system instead of the original unit templates. With this feature of Offline Sorter, you can save the templates resulting from re-sorting done during a clean up session into a Sort Client EXP file for use by the MAP system.

The way in which this feature works is to read the original Sort Client EXP file, replace the unit templates in it with the new templates resulting from the sorting done within Offline Sorter, and re-write everything to a new EXP file. Offline Sorter saves all of the settings in the original EXP file into the new EXP file, except that it replaces the unit templates for some channels.

To use this feature, you must sort channels using the template method—see [“Details of the Sorting Algorithms” on page 5-13](#). Note: *only unit templates on those channels sorted with the template method are saved to the EXP file* -- Offline Sorter ignores all channels not sorted or those sorted using a different method. Then select **File | Update an .exp File with Sort Info** from the main menu. In the first file dialog that appears, select the original EXP used by Sort Client to collect the current PLX file. In the second dialog, select the name of a new EXP file (it must be different from the original). Offline Sorter then writes the EXP file. Then, load the new EXP file into Sort Client for the next data-taking session.

## Parameter Scans

Offline Sorter contains a tool to explore different sorting scenarios by scanning the values for parameters that affect the sorting over a range of values, re-running the sorting and recording the Sorting Quality Statistics at each step.

Select the **Scan** tab under the **Info** tab to display the Parameter Scan grid view:

**Parameter Scan for Channel 10**

Type: Scan Parzen Radius Multiplier for Valley-Seeking ☐ Use 3D Stats

Start: 1.000000 End: 2.000000 Step By: 0.200000 **Start Scan**

Scan Val	#Units	J3	Pseudo-F	D-B	Dunn	p
1	3	6.60147	69438.6	0.672273	0.297534	7.83149e-012
1.2	3	6.64081	69852.5	0.681574	0.29223	5.69861e-012
1.4	3	6.79178	71440.5	0.703768	0.280848	1.87838e-012
1.6	3	6.79129	71435.3	0.704003	0.280708	1.88657e-012
1.8	2	2.48906	39273.7	0.708119	0.356318	2.50775e-007
2	2	2.49327	39340	0.697772	0.362784	2.31624e-007

	Scan Val	#Units	
Best by J3:	1.4	3	Re-sort Using This Value
Best by Pseudo-F:	1.4	3	Re-sort Using This Value
Best by D-B:	1	3	Re-sort Using This Value
Best by Dunn:	2	2	Re-sort Using This Value
Best by p:	1.4	3	Re-sort Using This Value

Channels / Chan Params / Sort Details / Global / Events / Stats / PCA / Templates / **Scan**

Info / Sort Summary / 2D Clusters

You can do several different types of scans, which you can select from the **Type** droplist. For each type, click **Start Scan** to cause the procedure indicated by the Type droplist to be performed at each step of the parameter over the range specified by the **Start**, **End** and **Step By** values. At each step, the Sorting Quality Statistics are calculated and displayed in a table, along with the number of units and percentage of waveforms sorted. By default, Offline Sorter sorts and calculates the Sorting Quality Statistics in 2D feature space, but you can use the 3D feature space instead by checking the **Use 3D** checkbox. When the scan completes, the “best” values of the parameter being scanned according to each of the Sorting Quality Statistics is displayed in a table. To re-sort the channel with any of these best values, click **Re-Sort Using This Value** on the line for the statistic. Also, double-click on any line in the scan table to re-run the procedure using that value of the scanned parameter.

**Note:** Other settings in Offline Sorter can also affect the results of the scan. For example, the **Include unsorted waveforms as a unit in the statistics calculations** in the **Tools | Options Stats** tab affects all the Sorting Quality Statistic values.

---

The available scan types are:

- Scan Parzen Radius Multiplier for Valley-Seeking : This scan type steps the Parzen Mult. parameter (as can be set via the Control Table) for the [Automatic Clustering using Valley-Seeking Method \(page 4-46\)](#) sorting algorithm over the range specified by the **Start**, **End**, and **Step By** values, re-running the sorting each time.

**Note:** The **Limit Number of Waveforms used in Valley Seeking Sorting** setting in the **Tools | Options Sort** tab can affect the clustering.

- Scan number of clusters for K-Means : This scan type runs the [Semi-Automatic Clustering using K-Means \(page 4-42\)](#) sorting algorithm, scanning the number of clusters over the range specified by the **Start**, **End** and **Step By** values. The initial cluster centers are automatically arranged along a diagonal line through the feature space origin for each step. This procedure is questionable for two reasons:
  - The results of the K-Means algorithm can be sensitive to the initial cluster centers, and the algorithm used to place the centers is not the optimal choice for the center locations.
  - The sorting quality statistics sometimes cannot be compared “apples-to-apples” across different numbers of clusters, which can cause the best values to be at the extremes of the scanned range.

**Note:** The current setting of the **Outlier Threshold** in the Control Table is used in the K-Means sorting at each step.

- Scan Outlier Threshold : This scan type runs the Remove Outliers tool (See [“Cleaning the Selection” on page 4-54.](#)), scanning the Outlier Threshold over the range specified by the **Start**, **End**, and **Step By** values. At each step, the sorting resets to the initial state, that is, the effects of the Remove Outliers tool are undone after each step. However, when you click **Re-Sort Using This Value** or double-click on a scan table line, remember to undo the action before you repeat it, or the Remove Outliers tool is applied to the results of the last operation, which is not the same initial condition that was in effect during the initial scan.
- Scan Waveform Limit for Valley-Seeking : This scan type runs the [Automatic Clustering using Valley-Seeking Method \(page 4-46\)](#) sorting algorithm while stepping the waveform limit over the range specified. The waveform limit, which you can set in the **Tools | Options Sort** dialog box, limits the number of points used by the Valley-Seeking sorting algorithm. A typical scan range for this type of scan might be 1000 to 10000 in steps of 1000. This is useful way for you to get a feeling as to how the number of points can affect the values of the sort quality metrics.
- Scan D.O.F. Multiplier for T-Dist E-M : This scan type runs the [Automatic Clustering using T-Distribution E-M Method \(page 4-49\)](#) sorting algorithm

over the range of D.O.F Multipliers you specify. The D.O.F. Multiplier affects the sorting as described in [Details of the Sorting Algorithms \(page 5-13\)](#).

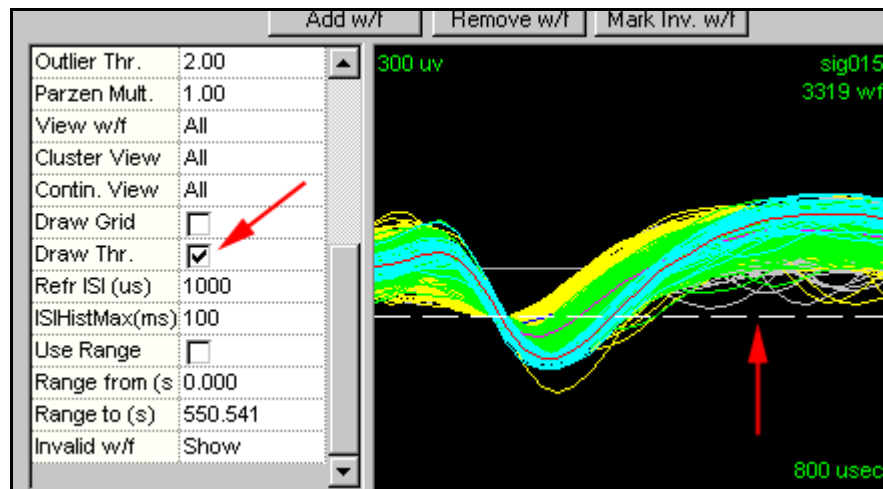
- **Scan Number of Clusters for Std E-M** : This scan type runs the [Semi-Automatic Clustering using Standard E-M \(page 4-45\)](#) sorting algorithm, which scans the number of clusters over the range specified. You can set the cluster center positions by using the same procedure described earlier for the “Scan number of clusters for K-Means”, and the same caveats apply.

## Re-thresholding Spike Data

In addition to the ability of Offline Sorter to threshold continuously-recorded data (e.g. DDT files), Offline Sorter can also simulate the effect of setting a different voltage threshold for the detection of spike waveforms. All spike waveforms that do not cross the new threshold are marked as invalid. Obviously, setting a threshold lower than what was in effect when the file was created does not put more waveforms into the file.

The first step in running the re-thresholding tool is to make the threshold marker line (a dashed white horizontal line) appear on the Waveform window. There are two ways to accomplish this:

- Check the **Draw Thr.** checkbox on the control panel. This permanently enables the display of the threshold marker line on the Waveform display
- Select the **Tools | Re-Threshold Waveforms** menu item. This temporarily shows the threshold marker line. The line disappears after the re-thresholding operation completes.



When the threshold marker line appears, you can click-and-drag the line up or down to establish a new threshold. As soon as you complete the drag, the waveforms are re-thresholded. You can undo the effects of re-thresholding the data either with the **Undo** command or with the **Waveforms | Mark All Waveforms Valid** menu selection.

---

## Assigning Unsorted Points to Closest Unit

After some operations (e.g. Valley-Seeking Sorting) it is desirable to assign unsorted waveforms membership in the nearest unit, subject to the Outlier Threshold. Select **Tools | Assign Unsorted To Closest Unit** to perform this operation.

The algorithm this tool uses is as follows:

- Calculate the centroid and standard deviations along each principal axis for each existing unit in the appropriate 2D or 3D feature space. If the 3D Cluster View was showing when the tool was selected, then 3D feature space is used, otherwise 2D feature space is used.
- Calculate the Mahalanobis distance from each unassigned waveform point to all the unit centroids, in order to find the closest unit. Note: The units of Mahalanobis distance is standard deviations.
- If the distance to the closest unit is less than the Outlier Threshold for that unit, assign the waveform to that unit.

**Note:** You can invoke this tool automatically following a Valley-Seeking automatic sort. This is useful if the Limit Number of Waveforms used in Valley-Seeking Algorithm option is used, as this can leave points unassigned that clearly belong to units. You can set these options in the **Sort** tab of the **Tools | Options** dialog.

**Note:** Temporarily setting **Outlier Thr.** to a large number in the Control Table before running this tool can force the algorithm to assign all unsorted points to a cluster, regardless of the distance of the point from the nearest cluster center.

## Batch File Processing

Offline Sorter provides the ability to execute a subset of its commands via a script. If you select a text file with the **File | Execute Batch Command File** menu item, it results in that file being read and interpreted as batch commands.

When the command file executes, it writes a log file with a .log extension on the end of the original command file name.

The format of the batch files (which are given the extension OFB by convention) is one command per line. `/'` will comment out a line, and lines are not case sensitive. A command consists of a main keyword, followed by sub-keywords or arguments.

The commands are:

**ForEachChannel.** Specifies an action to be done on each channel of the file(s).

Syntax: **ForEachChannel** *sub-command*

Sub-commands for the `ForEachChannel` command are:

- **ValleySeek, ValleySeek2d** – run valley seeking sorting on each channel, using 2D feature space
- **ValleySeek3d** – run valley seeking sorting on each channel, using 3D feature space
- **TDist, TDist2d** – run T-Distribution E-M sorting on each channel, using 2D feature space
- **TDist3d** – run T-Distribution E-M sorting on each channel, using 3D feature space
- **SortSummaryToPP** – after sorting, dumps the sort summary to PowerPoint for each channel
- **LogStats2d** – after sorting, dumps the 2D sort quality stats to the log file
- **LogStats3d** – after sorting, dumps the 3D sort quality stats to the log file
- **RemoveOutliers2d** – after sorting, run the Remove Outliers tool, using the currently set Outlier Threshold and 2D feature space
- **RemoveOutliers3d** – after sorting, run the Remove Outliers tool, using the currently set Outlier Threshold and 3D feature space
- **AssignUnsorted2d** – run the Assign Unsorted tool, using the currently set Outlier Threshold and 2D feature space
- **AssignUnsorted3d** – run the Assign Unsorted tool, using the currently set Outlier Threshold and 3D feature space
- **RemoveShortISI** – run the Remove Short ISI tool, using the currently set Refractory Period
- **Align** – run waveform alignment, using the currently set alignment parameters
- **KeepWfsBetweenEvents** – run an interval filter that keeps only the waveforms between two specified events, with an optional time offset from each event
- **KeepWfsInSpan** – run an interval filter that keeps only the waveforms within a time range around a specified event
- **KeepWfsBetweenTimes** – run an interval filter that keeps only the waveforms within a specified time range
- **InvalidateWaveformsBetweenTimes** – run an interval filter that invalidates the waveforms within a specified time range
- **Detect** – sets up waveform detection for each channel

e.g. `ForEachChannel ValleySeek`

---

**ForEachFile.** Specifies an action to be done for each file, after all channels are processed.

Syntax: **ForEachFile** *sub-command*

Sub-commands for the ForEachFile command are:

- **Save** – saves the file, with a numeric suffix added on to the original filename
- **SaveAsPlx** – saves the file as a new plx file (same as **File | Export to New .PLX** menu option), possibly with a numeric suffix
- **SaveToNex** – saves the timestamps etc. to a NEX file, possibly with a numeric suffix added on to the original filename

e.g **ForEachFile SaveToNex**

**Set.** Changes a parameter, immediately.

Syntax: **Set** *parameterName parameterValue*

The available parameters are:

- **ParzenRadiusMult** – sets the parzen radius multiplier for valley seeking sorting to the specified value
- **DOFMult** – sets the degrees of freedom multiplier for T-Distribution E-M sorting to the specified value
- **OutlierThreshold** – sets the outlier threshold to the specified value
- **RefractoryPeriod** – sets the refractory period to the specified value, for the **RemoveShortISI** channel command
- **WaveformLimit** – sets the limit for the number of waveforms used in the valley-seeking sorting. For more details, see [“Automatic Clustering using Valley-Seeking Method” on page 4-46](#)
- **FeatureX, FeatureY, FeatureZ** – set the feature for each dimension of feature space. See the following table for numerical values for the features
- **Slice1, Slice2, Slice3, Slice4** – set the slice position (in ticks) for the Slice n features
- **UseRangeStart, UseRangeEnd** – only consider the portion of the file with timestamps between the Start and End values
- **AlignType, AlignMaxShift** – parameters for waveform alignment. AlignMaxShift is the maximum amount a waveform will be shifted (in clock ticks). The AlignType determines the how the waveforms will be aligned; 0 = first local max, 1 = first local min, 2 = first local extremum, 3 = global max, 4 = global min

- **BetweenEvent1,BetweenEvent2** – event names to use for **KeepWfsBetweenEvents** interval filter
- **BetweenEvent1Offset,BetweenEvent2Offset** – time offsets in msec to use for **KeepWfsBetweenEvents** interval filter. Filter will invalidate waveforms outside of the time interval between (time of BetweenEvent1 + BetweenEvent1Offset) to (time of BetweenEvent2 + BetweenEvent2Offset).
- **SpanEvent** – event name to use for **KeepWfsInSpan** interval filter
- **SpanEventStartOffset,SpanEventEndOffset** – time offsets in msec to use for **SpanEvent** interval filter. Filter will invalidate waveforms outside of the time interval (time of SpanEvent + SpanEventStartOffset) to (time of SpanEvent + SpanEventEndOffset)
- **TimeStart,TimeEnd** – defines the time interval in msec to use for the **KeepWfsBetweenTimes** and **InvalidateWaveformsBetweenTimes** interval filters
- **DetectSigmas** – sets the detection threshold as a number of sigmas away from the mean of the Peak Height Histogram; the value can be negative for negative thresholds.

e.g **Set** *ParzenRadiusMult 2.0*

**File.** Queues a file for processing. The file is not actually processed until the Process command is encountered. Specify the full path to the files.

Syntax: **File** *filename*

e.g. **File** *c:\myfiles\test.plx* will queue the file test.plx for processing.

**Dir.** Queues all of the files that match a template for processing. The files are not actually processed until the Process command is encountered. The template should include the full path to the files.

Syntax: **Dir** *template*

e.g. **Dir** *c:\myfiles\\*.plx* will queue all of the .plx files in the c:\myfiles directory for processing.

**Process.** Processes all queued files. Resets the queued files list.

Syntax: **Process**

### Example Batch Command File

Here are the contents of example batch file, ofsbatch.ofb:

```
ForEachChannel ValleySeek  
ForEachChannel SortSummaryToPP  
ForEachChannel LogStats2d
```



---

```
//ForEachChannel LogStats3d
ForEachFile SaveToNex
Set FeatureX 0
Set FeatureY 1
Set ParzenRadiusMult 1.0
File C:\PlexonData\Plexon\plx\chan9threeunits3.plx
File C:\PlexonData\Plexon\plx\ev024.plx
Process
Set ParzenRadiusMult 2.0
File C:\PlexonData\Plexon\plx\chan9threeunits3.plx
File C:\PlexonData\Plexon\plx\ev024.plx
Process
```

The behavior of the example file is to:

- Set the parzen radius multiplier to 1.0
- Open chan9threeunits3.plx

For each channel:

- Do 2D valley seeking sort, using PC1 and PC2
- Send the sort summary to PowerPoint
- Log the 2D sort quality stats to the log file
- Save the timestamps to chan9threeunits3.nex
- Open ev024.plx

For each channel:

- Do 2D valley seeking sort , using PC1 and PC2
- Send the sort summary to PowerPoint
- Log the 2D sort quality stats to the log file
- Save the timestamps to ev024.nex
- Set the parzen radius multiplier to 2.0
- Do the same thing as described above to both files

## Notes

- Executing a command via a batch file is the same as executing it via menus/buttons in Offline Sorter. If a **Set** command changes a parameter, the new value of that parameter is in effect the next time Offline Sorter is started.

- Conversely, values not explicitly set by the batch commands obtain their values via the settings that were saved the last time that Offline Sorter was run. This can lead to non-reproducible results, so it is good policy to do explicit sets of everything that can affect the sorting results in the batch file.
- Currently, non-spike channels are skipped. This may change in the future.
- Numerical values to use in the Set FeatureXXX commands:
  - 0 PC 1
  - 1 PC 2
  - 2 PC 3
  - 3 Slice 1
  - 4 Slice 2
  - 5 Peak
  - 6 Valley
  - 7 Peak-Valley
  - 8 Energy
  - 9 Nonlinear Energy
  - 10 Timestamp
  - 11 PC 4
  - 12 PC 5
  - 13 PC 6
  - 14 PC 7
  - 15 PC 8
  - 16 Peak FWHM
  - 17 Valley FWHM
  - 18 Peak Tick
  - 19 Valley Tick

## Printing

Offline Sorter can print any view from the Multi-Display Window, and also the Waveform Window. Select **File | Print...** or **File | Print Preview...** to print or preview the currently active window in the Multi-Display Window.

The commands **File - Print All Sort Summaries** and **File - Send All Sort Summaries to PowerPoint** iterate through all channels and print or export the Sort Summary for each channel to a single PowerPoint file or a single print job.

---

If the computer has Adobe PDF Distiller installed, the **File - Print All Sort Summaries** offers a means to easily get the Sort Summaries for all channels in a file into a single PDF document. Simply select **Adobe PDF** as the printer in the **Print** dialog, and type the PDF file name.

The colors and the cluster view point size for printing can be different from the values used for display. Select the values to use for printing in the **Printing Options** tab in the **Tools | Options** dialog box. By default, the printing colors are the same as the display colors, except that the background color is set to white. Some of the display colors do not print with high contrast on some printers.

The 2D Cluster view has an option to dump the display to a Windows metafile, which you can then import into most graphics programs. To do so, from the right-click menu on the 2D Cluster view, select **Print to Windows Metafile**.

To capture the 3D Cluster view, you can use the built-in Windows screen capture method. With the 3D Cluster view showing, press ALT-PRINT SCREEN. Bring up a program such as Microsoft Paint, and select **Paste**. The image may have to be cropped to eliminate everything but the 3D Cluster view.

## Large File Tips

The following are some settings and usage tips that can significantly speed up Offline Sorter when working with large PLX files.

### Statistics Calculations

In general, Offline Sorter must recalculate sort quality statistics whenever any change is made to the sorting on a channel. This can take significant time, especially if unit pair-wise statistics are being calculated. The **Stats** tab of the **Tools | Options** panel contains checkboxes that enable or disable the calculation of the advanced sort-quality statistics, and also the unit pair-wise statistics. If the statistics are not of interest, you can uncheck these items to speed up sorting operations.

### Buffer Sizes

On the **File Import** tab of the **Options** page, there are two **Buffer Size** settings to be aware of. In both cases, these buffer sizes are specified in megabytes. These values should always be less than the amount of RAM in your computer, or severe performance degradation can occur.

The Channel Buffer Size setting under **Reading Buffer Size** is the size of the temporary cache Offline Sorter uses when reading waveforms from files. Increasing this number generally speeds up reading larger files, but might slow down the reading of smaller files.

The buffer size under **Continuous Data** is used for holding the data from continuous data channels. If you do not care to work with continuous data channels, or you are willing to wait for continuous data to be dynamically read in after a channel is initially loaded, set this to a small value. Note that internally, several buff-

ers of the size entered here are allocated, so entering larger values can quickly multiply the amount of memory used to hold continuous data.

### Ignoring Continuous Data Channels

Plexon PLX files often contain a large amount of continuous (slow channel) data that can be a significant fraction of the total file size. By default, Offline Sorter reads in this continuous data when the file is opened.

If you do not care to inspect or work with the continuous channel data, you can instruct Offline Sorter to bypass reading the continuous data altogether by checking the **Disable loading continuous data channels for PLX files** checkbox under the **Continuous Data** section on the **File Import** tab of the **Tools | Options** dialog box.



#### CAUTION

##### Continuous Data Channels Quietly Ignored

When **Disable loading continuous data channels for PLX files** setting is checked, continuous data channels are quietly ignored—Offline Sorter gives no indication that these continuous channels even exist in the data file, and they are not saved during a **Export to New .PLX** operation. Moreover, Offline Sorter retains this setting between startups. If you check this setting, Offline Sorter ignores analog channels without warning you until you specifically uncheck the setting.

When saving a NEX file, by default, Offline Sorter must read the continuous data channels in order to write them to the NEX file, which can be time-consuming. If you do not require that the continuous data channels be present in the NEX file, you can disable writing the continuous data channels by unchecking the **Export continuous data channels** checkbox in the **Save As NEX Options** dialog box.

### Initial PCA Calculation

By default, Offline Sorter performs a PCA calculation on each channel after it is loaded into memory. However, this calculation is wasted if the sorting is to be done in a feature space that does not involve principal component projections (e.g. using peak heights to sort tetrode data). You can effectively turn off the PCA calculation by entering a value of 0 in the edit box under **Maximum Number of Waveforms to Use in the PCA Calculation** in the **Feature Calc** tab of the **Options** window.

### Deferring Waveform Drawing

Sometimes, Offline Sorter's redrawing of the Waveform view and the Continuous Data views can be time-consuming. If the **Defer Waveform Drawing Until Clicked** checkbox in the **Display** tab of the **Options** dialog box is checked, the displays are not redrawn until you left-click in them. When you view animated sorting progress, the views are temporarily put into deferred drawing mode, otherwise it slows down the animation too much.

---

## Split PLX files (.plx0)

You can use an entirely different approach to speed up Offline Sorter operations on large PLX files. You can use a separate program called PlxUtil, which installs with Offline Sorter, to pre-process a large PLX file. PlxUtil pre-processes the PLX file into a new on-disk file format called PLX0. The PLX0 file format enables Offline Sorter to operate much more efficiently. The PLX0 file format consists of a PLX0 file that contains file header information, events, and continuous data channels, plus a number of .plxNNN files that contain the waveforms for each channel. The PLX0 file and all the accompanying .plxNNN channel files are generally stored together in their own directory.

To create a PLX0 file, start PlxUtil and select **Make separate files for each channel**. Select the original file, and the name of the output PLX0 file (probably in a new directory). The calculation of PCA information is optional; if it is not calculated and saved by PlxUtil, it is recalculated by Offline Sorter. Click **Create channel files now!** to write the PLX0 file and the associated .plxNNN files.

To use the split files, open the PLX0 file in Offline Sorter. Do not attempt to open the individual .plxNNN channel files. Sort the file as usual – all Offline Sorter operations are available, with the exception of the **File - Save As** file menu item. Interim sorting results (using **File – Save**) save much faster since only the .plxNNN file for the current channel needs to be written. You can write the sorting results to NEX files as usual, or you can save them to a PLX file by using the **File - Export to New .PLX** operation.



\* Plexon and the Plexon logo are trademarks of Plexon Inc. Microsoft, Windows, Windows NT, Windows 2000, Windows XP, Excel, and PowerPoint are trademarks of Microsoft Corporation. Sentinel, SentinelPro, and SentinelSuperPro are trademarks of Rainbow Technologies. Adobe and PDF are trademarks of Adobe Systems, Inc. MATLAB is a trademark of The Math Works, Inc. NI and NI-DAQ are trademarks of National Instruments. NeuroExplorer is a trademark of Nex Technologies.

Data Analysis Software  
**Offline Sorter**  
User's Guide

Plexon Inc  
6500 Greenville Avenue  
Suite 730  
Dallas, Texas 75206  
Telephone: (214) 369-4957  
Fax: (214) 369-1775

[www.plexoninc.com](http://www.plexoninc.com)

Copyright © 1983-2005 Plexon Inc., All Rights Reserved

**Plexon Inc. Proprietary**

The information contained herein is the property of Plexon Inc. and it is proprietary and restricted solely to assist Plexon Inc. customers. Neither this document nor the contents may be disclosed, copied, revealed or used in whole or in part for any other purpose without the prior written permission of Plexon Inc. This document must be returned upon request of Plexon Inc.

Information is subject to change without notice. Plexon Inc. reserves the right to make changes in equipment design or components as progress in engineering or manufacturing may warrant.

Plexon Inc. and the Plexon logo are trademarks of Plexon Inc.

Document number: 00000-C-5002-A

Document issue: 1.6

Document status: Standard

Software Version and Issue: 2.7

Date: January 2005

