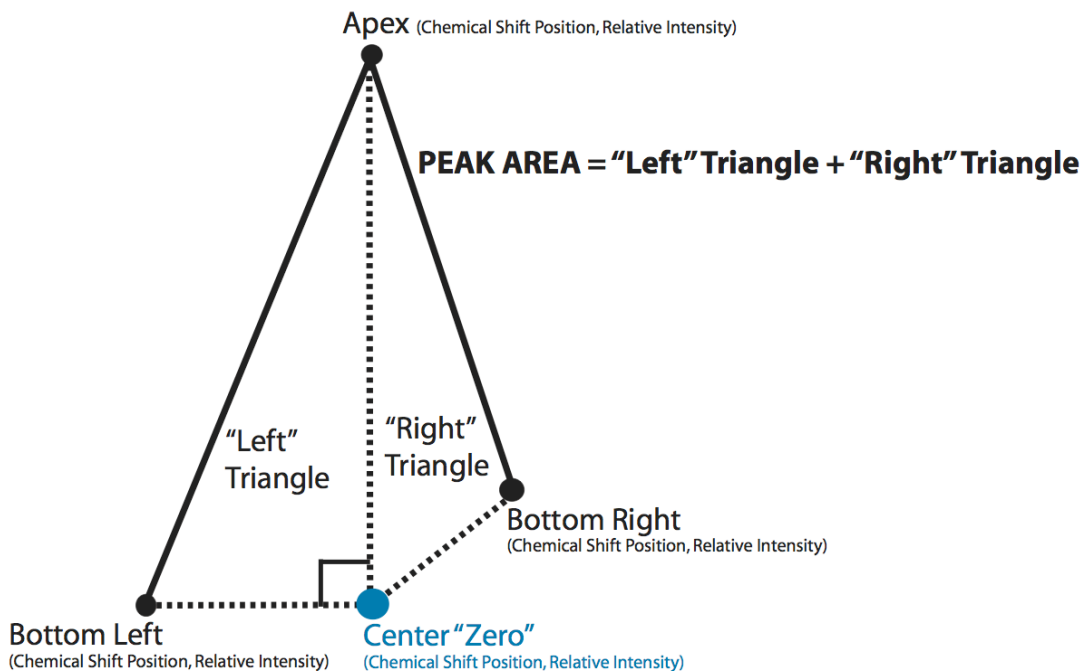


Update to PCANS software – Primary updates & additions:

1. New peak picking algorithms were added that are based upon a paper by Abdo et. al. that uses the statistical definition of an outliers to define which peaks are picked. See the following paper for details about the algorithm. Our algorithms are based upon this method except they allow for the user to define different cutpoints beside the standard of 3.0. The user also has the option to choose an optimal cutpoint based upon predefined statistics under the assumption that the optimal cutpoint is the one that chooses similar numbers of peaks from each spectrum. Our algorithms also differ from Abdo et. al.'s because they allow you to define peaks by Area alone (Abdo), Height alone, or BOTH Area and Height.
 - Abdo, Z., Schuette, U., Bent, B. J., Williams, C. J., Forney, L. J., and Joyce, P. (2006). Statis- tical methods for characterizing diversity of microbial communities by analysis of terminal restriction fragment length polymorphisms of 16s rrna genes. *Environmental Microbiology*, 8(5):929 – 938.
2. All peak picking algorithms (new & old) now output a file that contains the Area of each peak. This area calculation is based calculating the area of each side of the peak as a triangle and summing those areas. The area of each triangle is calculated using Heron's formula where a triangle's area is calculated based upon the length of the 3 sides. Note that the position of the three points that compose each triangle are based upon the 'Chemical Shift Position' and the Relative Intensity (height). The Apex, Bottom Left and Bottom Right are all determined when the peaks are picked. The Center "Zero" point has the same chemical shift position as the Apex but its Relative Intensity (height) is defined as the lower point between the bottom Left and the Bottom Right. The illustration below depicts how a peak's area is defined and calculated.



3. Bug fix that allows user to input the original data matrix with Chemical Shifts sorted in descending or ascending order. Originally PCANS required the Chemical Shifts be sorted in ascending order (0.0005, 0.0055, 0.0060, ...) and will throw an error if the input matrix Chemical Shifts are sorted in descending order (9.5000, 9.4995, 9.4990, ...).

4. Ability to just pick peaks (NEW & OLD Algorithms) instead of having to pick peaks and align those results.
5. Runs on Python 2.5 or 2.7

Required Software to Run PCANS GUI:

1. Python 2.5/Python 2.7 (version 2.5.4 is available for download here → <http://www.python.org/download/releases/2.5.4/> version 2.7.3 is available for download here → <http://www.python.org/download/releases/2.7.3>)
2. Python Modules: sys, time, cPickle, Tkinter, string, os, **numpy**, **scipy**, csv, random, math. All the modules are included in the standard 2.5 Python package except **Tkinter(?)**, **numpy & scipy**. **numPy** can be found here → <http://sourceforge.net/projects/numpy/files/NumPy/> and **SciPy** can be found here → <http://sourceforge.net/projects/scipy/files/> python may be missing Tkinter/tk package directions for installing it can be found here → http://tkinter.unpythonic.net/wiki/How_to_install_Tkinter

* Enthought Python Distribution (EPD Py25 v4.3.0 and EPD v7.3) installed for free for academic use includes Python 2.5.4/Python 2.7.3 and all needed modules

* I have also created an Amazon 'ami' that has ubuntu 12.04 loaded along with Python 2.7 and the appropriate python packages. I will add detailed directions as to how to use the AMI to run the software given that you have an amazon aws account. Note that for testing purposes I used a 'micro' instance or only a short period of time; thus, running PCANS in aws was free to me. I am not sure what will constitute Amazon's free tier in the future, but for less than \$2/\$4 you can use a small or medium instance for a 24 hour period. To save the hassle of having to load the software on a computer and setup the appropriate software environment this might be an easy option (see pricing here <http://aws.amazon.com/ec2/pricing/>). You will need the ability to ssh -X -i in so that the GUI windows will appear. On a windows OS I would suggest using an Xterm emulator like MobaXterm (<http://mobaxterm.mobatek.net/>) you can get a freeware version that will allow you to ssh into your amazon account along with sftp to upload your spectra files. To do so in MobaXterm you would need to set the "Extra Options" on the SSH settings to be -X -i /drives/c/<location of *.pem file> and when prompted for login use 'ubuntu' the IP address comes from the name the instance is given by amazon. For example if your instance is name ec2-108-22-50-201.compute-1.amazonaws.com, then your IP address is 108.22.50.201. For using terminal window on OS X/Unix/Linux ssh -X -i <location of *.pem file> <ubuntu@ec2-your-machine.compute-1.amazonaws.com>

* How to Login & use the Amazon AMI –

1. Create an amazon aws account here: <http://aws.amazon.com/> (you will have to give them a valid credit card number – they will only charge you if you use beyond the free tier see here for costs <http://aws.amazon.com/ec2/pricing/>)
2. Log into your aws account and from the AWS console - select INSTANCES and click on "Launch instance" use the Classic Wizard and click continue, click on the "My AMIs" tab and search "All Images" for the words "PCANS"
3. Select the PCANS AMI to launch your instance. Pick size (micro, small, medium) to suite your needs – micro options was free by amazon – BUT you need to verify prices based on instance size <http://aws.amazon.com/ec2/pricing/> Pick your time zone. The instance comes

with 10 GB root drive but if you require more size you need to adjust it now. Have Amazon create a key pair for you and note where the *.pem file is stored on your computer. Using the default security setting should work fine. Launch your instance.

* How to Login into my EC2/Instance that I just created –

From a terminal window (Unix/Linux/Mac) type the following ssh command filling in location of your saved private key (*.pem file) and the correct name of your EC2/Instance. Note the instance name is found by clicking the box next to the your running instance in the My Instances section of the AWS console and your *.pem file was saved to your computer when you had amazon create a key pair for you when you launched your EC2/Instance.

1.

SSH command with -X -i options:

ssh -X -i <Location of my *.pem key file> ubuntu@ec2-myEC2-Instance.compute-1.amazonaws.com>

SSH command with -X -i options – FOR an example instance:

ssh -X -i /Users/pcan_user/pcans_key.pem ubuntu@ec2-108-22-50-201.compute-1.amazonaws.com

For Windows users download and install a Xterminal window emulator like MobaXterm (<http://mobaxterm.mobatek.net/>). Unzip the file you get from MobaXterm and click on the MobaXterm.exe file. This opens a XTerminal session after you allow the program to run. Click on the tab to create a “New session”. Click on the “SSH” tab. For the “Remote hostname” enter your EC2/Instance’s IP address. If my was named ec2-108-22-50-201.compute-1.amazonaws.com, then my IP address is 108.22.50.201. Note the instance name is found by clicking the box next to the your running instance in the My Instances section of the AWS console and your *.pem file was saved to your computer when you had amazon create a key pair for you when you launched your EC2/Instance. In the “Extra options” box you need to put the location of your *.pem file

“Extra Options” with -X -i options:

-X -i </drives/...Location of my *.pem>

“Extra Options with -X -i options – FOR an example instance:

-X -i /drives/c/Users/pcan_user/pcans_key.pem

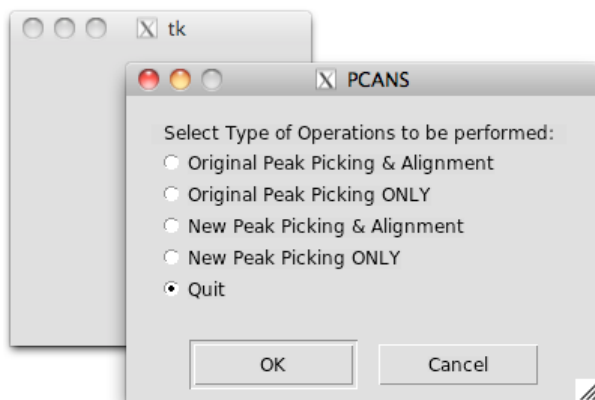
After clicking “OK” the new session will launch – be sure to Login as ubuntu

2. Once you are in you are logged in as user ‘ubuntu’ – you can sftp in as ubuntu and place your *.cvs file that contains the spectra you want to align in /home/PCANSv2/data . If you are using MobaXterm logged in as ubuntu you will be able to use the tools provided on the left-hand panel to move around directories and upload files directly into /home/PCANSv2/data directory.

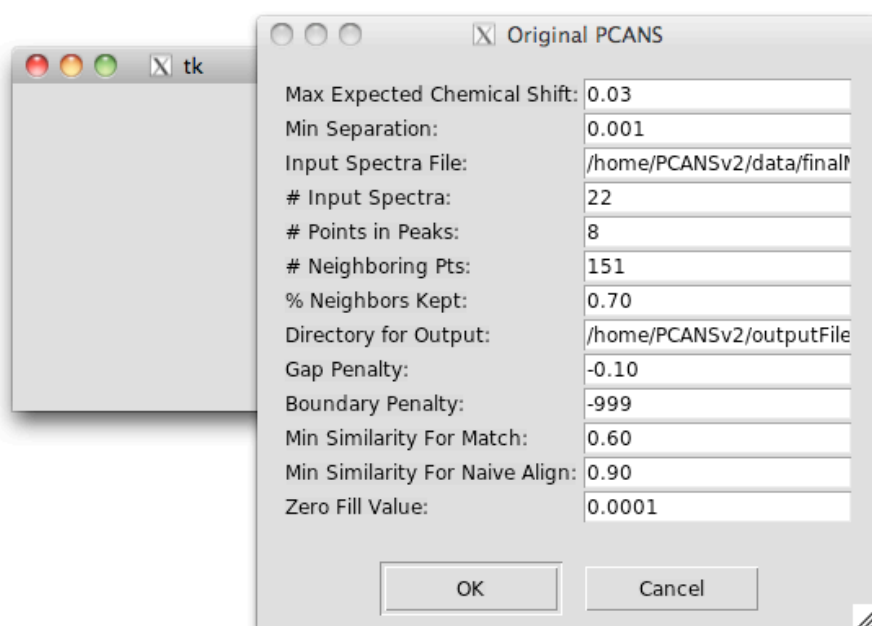
How to Run PCANS:

1. Download & unzip PCANSv2.zip (Windows users ignore the __MACOSX folder that gets created when you unzip PCANSv2.zip).
2. Load required software (Python 2.5/2.7 and then Numpy)
3. Open Terminal window (Command window for Windows users) and move to the .../PCANSv2/src/ (windows OS: ...\\PCANSv2\\src\\) directory. You are in the correct directory if it contains python program PCANS_AlignmentGUI.py

4. NOTE: The following directions were written for Python 2.5 – same should apply to Python 2.7 except replace “Python25” with “Python27”. In the terminal window Type “python PCANS_AlignmentGUI.py” to run alignment program – The GUI window below will appear. **This will not work for Windows OS.** → For a windows OS: type “C:\Python25\python PCANS_AlignmentGUI.py” where the “C:\Python25\” directory contains the python.exe file. If you have done the standard load of Python 2.5, then python.exe should be located in the “C:\Python25\” directory. If you have loaded python.exe somewhere else on your Windows OS you need to replace “C:\Python25\” with that location.
5. First window to appear will be the following – select the type of PCANS you want to run (Original or New) and if you want to run Peak Picking ONLY or Peak Picking & the Alignment



6. **IF you pick Original Peak Picking & Alignment – the following appears:** Default values are contained within the GUI window. The default values will run the alignment on .../PCANS/data/finalMouse_data.csv file (windows OS: .../PCANS\data\ytfinalMouse_data.csv) and output the results to .../PCANS/outputFiles/ directory (windows OS: .../PCANS\outputFiles\). To run an alignment on a different CSV file, one would need to replace “.../PCANS/data/finalMouse_data.csv”(#3) with the location and file name of the different CSV file. In addition, to save the alignment results in a different location one would need to replace “.../PCANS/outputFiles/”(#8) with that new location. The other alignment options can also be changed as described below. Once one has the desired values in the GUI click the ‘OK’ button to run alignment program. Clicking ‘Cancel’ closes the program without running. NOTE THE GUI DOES NO ERROR CHECKING ON INPUTS, MISTAKES WILL LEAD TO A RUNTIME FAILURE. BE AWARE OF WHAT IT ENTERED INTO GUI TEXT BOXES. CLICKING “OK” WITHOUT CHANGING THE DEFAULTS WILL ALWAYS RUN THE PROGRAM.



Detailed Description of GUI Inputs - ORIGINAL PCANS with Alignment:

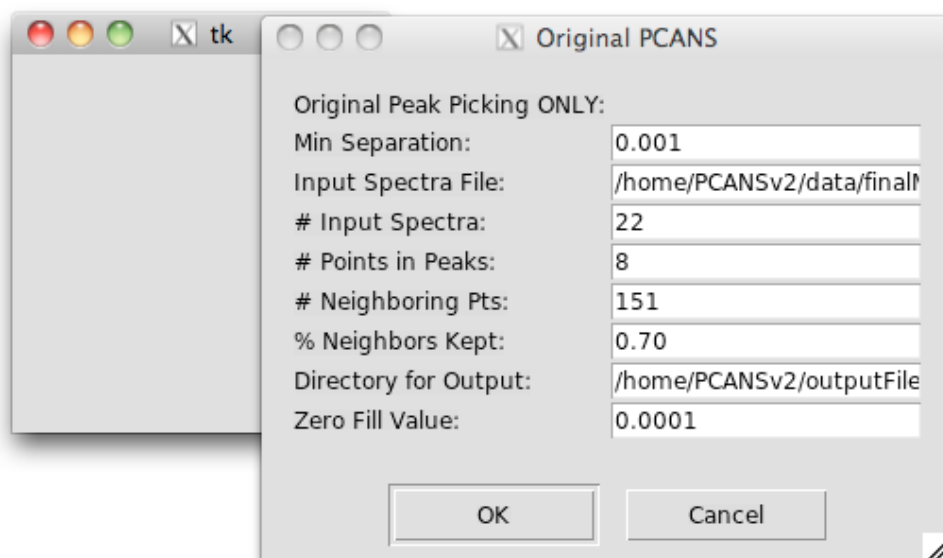
1. Maximum expected chemical shift. Should be a real number, typically 0.03 or 0.04 ppm depending on the data.
2. Minimum separation expected between peaks within each spectra. Generally 0.001 ppm but is data dependent, could be as small as 0.0001 ppm.
3. Input Spectra File location and file name. Defaults to .../PCANS/data/finalMouse_data.csv OR .../PCANS\data\finalMouse_data.csv, where ... represents where PCANS.zip was unzipped. To change to a different input file format should be /MyPath/MySpectraData.csv or C:\MyPath\MySpectraData.csv dependent upon one's OS. Expects the data to be a csv file, where first column contains chemical shift positions and the rest of the columns contain the spectra. The format should be the same as what can be found in .../PCANS/data/finalMouse_data.csv use it as a guide to creating your input spectra files
4. Number of input spectra in the input spectra file (integer), .../PCANS/data/finalMouse_data.csv has 23 columns first is chemical shift rest are spectra so its input is 22
5. Number of points required to form a peak within a spectra using the peak picking algorithm (integer). Data dependent, but 8 worked well with finalMouse_data.csv.
6. Number of points that define the region immediately surrounding a possible peak (neighbors of a possible peak) within the peak picking algorithm. Select an odd integer of 101 or greater points. 151 worked well with finalMouse_data.csv, which contained a total of 11,901 points from 0.5 to 9.5 ppm
7. Based on relative intensity of the points immediately surrounding a possible peak (neighbors). Peaks that are kept must have relative intensity (height) greater than X proportion of points surrounding it (its neighbors), where X is the proportion entered in this text box. Typically a real number of 0.60 or greater, 0.70 worked well with finalMouse_data.csv.
8. Output directory (including the last backward or forward slash) where output files will be stored. Defaults to .../PCANS/outputFiles/ OR .../PCANS\outputFiles\, where ... represents where PCANS.zip was unzipped. To change to a different output directory, the format should be /MyPath/OutputFolder/ or C:\MyPath\OutputFolder\ dependent upon one's OS.

9. Gap penalty should be a small real number that is less than the minimum expected similarity for a match. Can range from a small negative number to a value that is less than the minimum expected similarity for a match value (#11). -0.10 worked well for finalMouse_data.csv.
10. Boundary Penalty should be a value that is smaller than the gap penalty. If the user would prefer that the algorithm compute this penalty, then set the value to the default of -999. For finalMouse_data.csv this value was set to -999.
11. Minimum expected similarity between two peaks to allow for a match. Should be some proportion that ranges between 0.0 and 1.0. Typical value would be 0.60 or higher, but is data dependent. For finalMouse_data.csv 0.60 was used.
12. Minimum similarity required for naïve alignment, if two peaks are at least this similar then they will be aligned using the naïve alignment scheme. Typically this value would be 0.90 or higher, 0.90 was used for finalMouse_data.csv.
13. Zero fill value is the value that will be used to fill in the output files when there is no peak for a given chemical shift value. Typically we have used 0.0001 instead of zero for the OPLS analysis.

Expected Output Files in Output Directory - ORIGINAL PCANS with Alignment:

1. **origSpectra.csv** – comma separated file that contains input peak profiles (spectra) prior to alignment after peaks have been picked. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input spectra prior to alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
2. **origSpectra.py** – Prior to alignment after peaks have been picked, the input peak profiles (spectra) saved as a spectra object.
3. **pickPeakSpectra.py** – Before alignment of the picked peaks, the picked peak profiles (spectra) saved as a spectra object.
4. **pickPeakSpectraAr.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the peak area of the picked peaks for the input peak profiles (spectra) before alignment. The Area was calculated using Heron's formula as indicated in the "UPDATES" section of this document. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
5. **pickPeakSpectraHt.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input peak profiles (spectra) before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
6. **pickPeakSpectraWd.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the width at half-height (width) of the picked peaks for the input spectra before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input peak profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input profiles (spectra).
7. **alignSpectra.py** – After alignment of the picked peaks, the aligned peak profiles (spectra) saved as a spectra object.

8. **alignedSpectraAr.csv** -- comma separated file that contains the input peak profiles (spectra) after PCANS alignment. The first column contains chemicals shift position, next columns contain the peak area of the picked peaks for the input peak profiles (spectra) after alignment. The Area was calculated using Heron's formula as indicated in the "UPDATES" section of this document. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
 9. **alignedSpectraHt.csv** -- comma separated file that contains the input peak profiles (spectra) after PCANS alignment. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input peak profiles (spectra) after alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
 10. **alignedSpectraWd.csv** -- comma separated file that contains the input peak profiles (spectra) after PCANS alignment. The first column contains chemicals shift position, next columns contain the width at half-height (width) of the picked peaks for the input spectra after alignment. Zero fill values appear where there are no peaks for a given chemical shift and input peak profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input profiles (spectra).
 11. **finalConsensusSpectrum.csv** -- comma separated file that contains the final consensus spectrum post alignment. The first column contains chemicals shift position of the peaks in the final consensus spectrum. The values in this column represent the median chemical shift position of all the peaks that were aligned to that chemical shift position using PCANS. The next column contains the average relative intensity (height) of the peaks that were aligned to that chemical shift position in the consensus spectrum, the next column contains the average width at half-height (width) of all the peaks that were aligned to that chemical shift position in the consensus spectrum, and the final column contains the number of peaks that were aligned to that chemical shift position in the consensus spectrum.
 12. **finalConsensusSpectrum.py** – Final consensus spectrum post-alignment saved as a spectra object.
7. **IF you pick Original Peak Picking ONLY – the following appears:** Default values are contained within the GUI window. The default values will run the alignment on .../PCANS/data/finalMouse_data.csv file (windows OS: ...\\PCANS\\data\\ytfinalMouse_data.csv) and output the results to .../PCANS/outputFiles/ directory (windows OS: ...\\PCANS\\outputFiles\\). To run an alignment on a different CSV file, one would need to replace ".../PCANS/data/finalMouse_data.csv" (#3) with the location and file name of the different CSV file. In addition, to save the alignment results in a different location one would need to replace ".../PCANS/outputFiles/" (#8) with that new location. The other alignment options can also be changed as described below. Once one has the desired values in the GUI click the 'OK' button to run alignment program. Clicking 'Cancel' closes the program without running. NOTE THE GUI DOES NO ERROR CHECKING ON INPUTS, MISTAKES WILL LEAD TO A RUNTIME FAILURE. BE AWARE OF WHAT IT ENTERED INTO GUI TEXT BOXES. CLICKING "OK" WITHOUT CHANGING THE DEFAULTS WILL ALWAYS RUN THE PROGRAM.

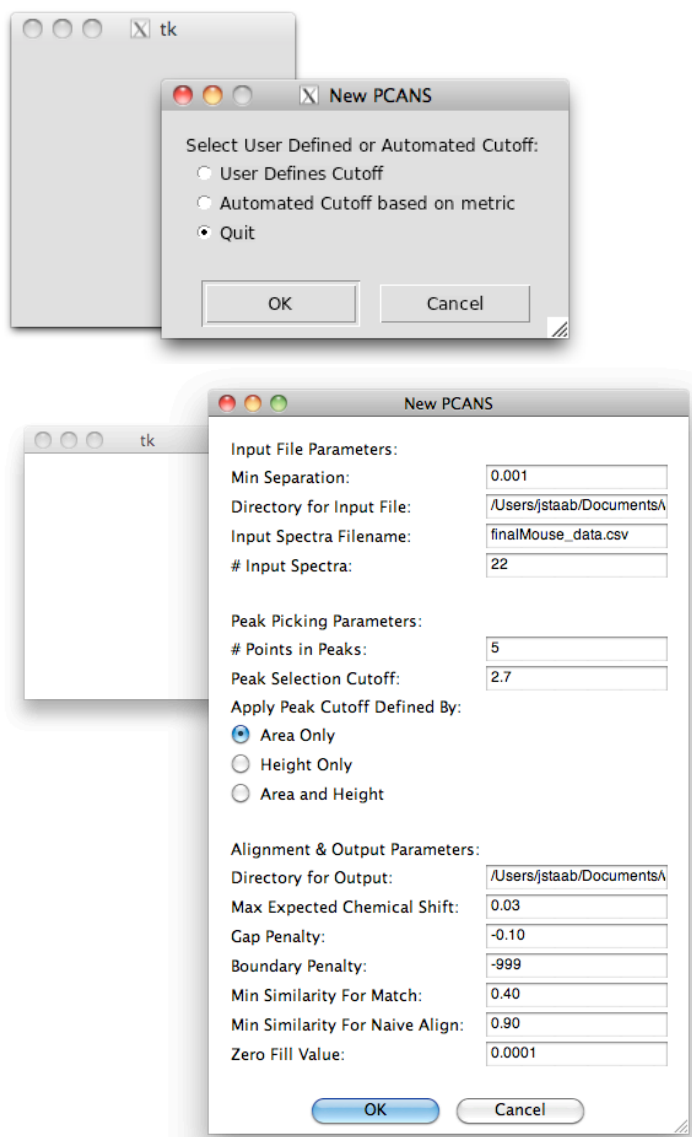


Detailed Description of GUI Inputs - ORIGINAL PCANS with Alignment:

1. Minimum separation expected between peaks within each spectra. Generally 0.001 ppm but is data dependent, could be as small as 0.0001 ppm.
2. Input Spectra File location and file name. Defaults to .../PCANS/data/finalMouse_data.csv OR .../PCANS\data\finalMouse_data.csv, where ... represents where PCANS.zip was unzipped. To change to a different input file format should be /MyPath/MySpectraData.csv or C:\MyPath\MySpectraData.csv dependent upon one's OS. Expects the data to be a csv file, where first column contains chemical shift positions and the rest of the columns contain the spectra. The format should be the same as what can be found in .../PCANS/data/finalMouse_data.csv use it as a guide to creating your input spectra files
3. Number of input spectra in the input spectra file (integer), .../PCANS/data/finalMouse_data.csv has 23 columns first is chemical shift rest are spectra so its input is 22
4. Number of points required to form a peak within a spectra using the peak picking algorithm (integer). Data dependent, but 8 worked well with finalMouse_data.csv.
5. Number of points that define the region immediately surrounding a possible peak (neighbors of a possible peak) within the peak picking algorithm. Select an odd integer of 101 or greater points. 151 worked well with finalMouse_data.csv, which contained a total of 11,901 points from 0.5 to 9.5 ppm
6. Based on relative intensity of the points immediately surrounding a possible peak (neighbors). Peaks that are kept must have relative intensity (height) greater than X proportion of points surrounding it (its neighbors), where X is the proportion entered in this text box. Typically a real number of 0.60 or greater, 0.70 worked well with finalMouse_data.csv.
7. Output directory (including the last backward or forward slash) where output files will be stored. Defaults to .../PCANS/outputFiles/ OR .../PCANS\outputFiles\, where ... represents where PCANS.zip was unzipped. To change to a different output directory, the format should be /MyPath/OutputFolder/ or C:\MyPath\OutputFolder\ dependent upon one's OS.
8. Zero fill value is the value that will be used to fill in the output files when there is no peak for a given chemical shift value. Typically we have used 0.0001 instead of zero for the OPLS analysis.

Expected Output Files in Output Directory - ORIGINAL PCANS with Alignment:

1. **origSpectra.csv** – comma separated file that contains input peak profiles (spectra) prior to alignment after peaks have been picked. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input spectra prior to alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
 2. **origSpectra.py** – Prior to alignment after peaks have been picked, the input peak profiles (spectra) saved as a spectra object.
 3. **pickPeakSpectra.py** – Before alignment of the picked peaks, the picked peak profiles (spectra) saved as a spectra object.
 4. **pickPeakSpectraAr.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the peak area of the picked peaks for the input peak profiles (spectra) before alignment. The Area was calculated using Heron's formula as indicated in the "UPDATES" section of this document. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
 5. **pickPeakSpectraHt.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input peak profiles (spectra) before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
 6. **pickPeakSpectraWd.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the width at half-height (width) of the picked peaks for the input spectra before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input peak profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input profiles (spectra).
8. **IF you pick NEW Peak Picking & Alignment – USER DEFINED CUTOFF – the following appears:** You must first pick if you want "User Defines Cutoff" OR "Automated Cutoff based on a Metric". This assumes you are using "User Defined Cutoff". Default values are contained within the GUI window. The default values will run the alignment on .../PCANS/data/finalMouse_data.csv file (windows OS: ...\\PCANS\\data\\ytfinalMouse_data.csv) and output the results to .../PCANS/outputFiles/ directory (windows OS: ...\\PCANS\\outputFiles\\). To run an alignment on a different CSV file, one would need to replace ".../PCANS/data/finalMouse_data.csv" (#3) with the location and file name of the different CSV file. In addition, to save the alignment results in a different location one would need to replace ".../PCANS/outputFiles/" (#8) with that new location. The other alignment options can also be changed as described below. Once one has the desired values in the GUI click the 'OK' button to run alignment program. Clicking 'Cancel' closes the program without running. NOTE THE GUI DOES NO ERROR CHECKING ON INPUTS, MISTAKES WILL LEAD TO A RUNTIME FAILURE. BE AWARE OF WHAT IT ENTERED INTO GUI TEXT BOXES. CLICKING "OK" WITHOUT CHANGING THE DEFAULTS WILL ALWAYS RUN THE PROGRAM

**Detailed Description of GUI Inputs - NEW PCANS with Alignment – USER DEFINED CUTOFF:**

1. Minimum separation expected between peaks within each spectra. Generally 0.001 ppm but is data dependent, could be as small as 0.0001 ppm.
2. Input Spectra File location. Defaults to .../PCANS/data/ OR ...\\PCANS\\data\\, where ... represents where PCANS.zip was unzipped. To change to a different input file format should be /MyPath/ or C:\\MyPath\\ dependent upon one's OS.
3. Input Spectra file name. Defaults to finalMouse_data.csv. To change to a different input file format should be MySpectraData.csv dependent upon one's OS. Expects the data to be a csv file, where first column contains chemical shift positions and the rest of the columns contain the spectra. The format should be the same as what can be found in .../PCANS/data/finalMouse_data.csv use it as a guide to creating your input spectra files
4. Number of input spectra in the input spectra file (integer), .../PCANS/data/finalMouse_data.csv has 23 columns first is chemical shift rest are spectra so its input is 22
5. Number of points required to form a peak within a spectra using the peak picking algorithm (integer). Data dependent, but 5 worked well with finalMouse_data.csv.

6. Peak Selection Cutoff – This is the cutpoint used to define what is a peak and what is noise in the new peak selection algorithm. For 5 points to define a peak the value 2.7 worked well with our finalMouse_data.csv. This value could range between 2.5 – 3.5 with 3.0 being what was used with the original paper about this new method.
7. Apply cutoff define by Area Only (as in original paper), Height (Relative Intensity) Only, or Both Area & Height. This is how you are defining a peak with regards to determining if a selected potential peak is “noise” or a real peak object.
8. Output directory (including the last backward or forward slash) where output files will be stored. Defaults to ../PCANS/outputFiles/ OR ../PCANS\outputFiles\, where ... represents where PCANS.zip was unzipped. To change to a different output directory, the format should be /MyPath/OutputFolder/ or C:\MyPath\OutputFolder\ dependent upon one's OS.
9. Maximum expected chemical shift. Should be a real number, typically 0.03 or 0.04 ppm depending on the data.
10. Gap penalty should be a small real number that is less than the minimum expected similarity for a match. Can range from a small negative number to a value that is less than the minimum expected similarity for a match value (#11). -0.10 worked well for finalMouse_data.csv.
11. Boundary Penalty should be a value that is smaller than the gap penalty. If the user would prefer that the algorithm compute this penalty, then set the value to the default of -999. For finalMouse_data.csv this value was set to -999.
12. Minimum expected similarity between two peaks to allow for a match. Should be some proportion that ranges between 0.0 and 1.0. Typical value would be 0.60 or higher, but is data dependent. For finalMouse_data.csv 0.40 was used.
13. Minimum similarity required for naïve alignment, if two peaks are at least this similar then they will be aligned using the naïve alignment scheme. Typically this value would be 0.90 or higher, 0.90 was used for finalMouse_data.csv.
14. Zero fill value is the value that will be used to fill in the output files when there is no peak for a given chemical shift value. Typically we have used 0.0001 instead of zero for the OPLS analysis.

Expected Output Files in Output Directory - NEW PCANS with Alignment – USER DEFINED CUTOFF:

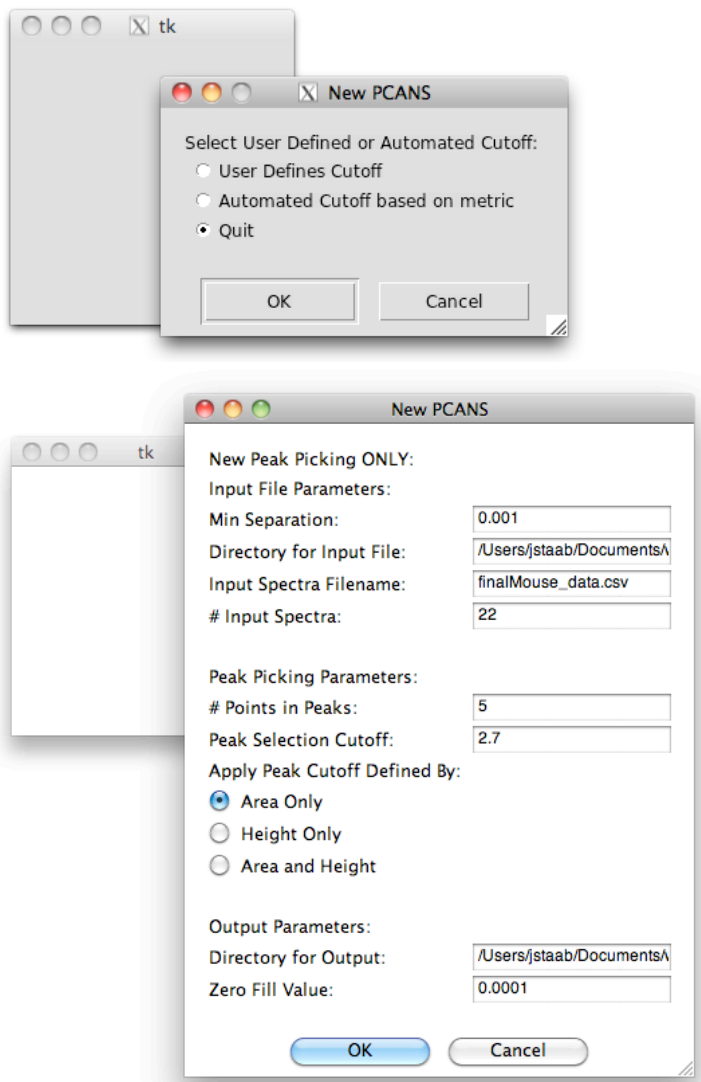
NOTE most files are labeled with PxxAyy where xx indicated the number of points that determine a peak and yy indicates the cutoff used.

1. **origPickPeaks.pkl** – Spectra object containing all the peaks picked prior to using the new peak picking algorithm to determine which peaks are real and which are noise based upon the selected cutoff.
2. **pickPeakSpectraPxxAyy.py** – Before alignment of the picked peaks, the picked peak profiles (spectra) saved as a spectra object.
3. **pickPeakSpectraArPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the peak area of the picked peaks for the input peak profiles (spectra) before alignment. The Area was calculated using Heron's formula as indicated in the “UPDATES” section of this document. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
4. **pickPeakSpectraHtPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input peak profiles (spectra) before alignment. Zero fill values appear where there are no peaks for a given

chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).

5. **pickPeakSpectraWdPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the width at half-height (width) of the picked peaks for the input spectra before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input peak profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input profiles (spectra).
 6. **alignSpectraPxxAyy.py** – After alignment of the picked peaks, the aligned peak profiles (spectra) saved as a spectra object.
 7. **alignedSpectraArPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) after PCANS alignment. The first column contains chemicals shift position, next columns contain the peak area of the picked peaks for the input peak profiles (spectra) after alignment. The Area was calculated using Heron's formula as indicated in the "UPDATES" section of this document. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
 8. **alignedSpectraHtPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) after PCANS alignment. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input peak profiles (spectra) after alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
 9. **alignedSpectraWdPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) after PCANS alignment. The first column contains chemicals shift position, next columns contain the width at half-height (width) of the picked peaks for the input spectra after alignment. Zero fill values appear where there are no peaks for a given chemical shift and input peak profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input profiles (spectra).
 10. **finalConsensusSpectrumPxxAyy.csv** -- comma separated file that contains the final consensus spectrum post alignment. The first column contains chemicals shift position of the peaks in the final consensus spectrum. The values in this column represent the median chemical shift position of all the peaks that were aligned to that chemical shift position using PCANS. The next column contains the average relative intensity (height) of the peaks that were aligned to that chemical shift position in the consensus spectrum, the next column contains the average width at half-height (width) of all the peaks that were aligned to that chemical shift position in the consensus spectrum, and the final column contains the number of peaks that were aligned to that chemical shift position in the consensus spectrum.
 11. **finalConsensusSpectrumPxxAyy.py** – Final consensus spectrum post-alignment saved as a spectra object.
9. **IF you pick NEW Peak Picking ONLY– USER DEFINED CUTOFF – the following appears:** You must first pick if you want "User Defines Cutoff" OR "Automated Cutoff based on a Metric". This assumes you are using "User Defined Cutoff". Default values are contained within the GUI window. The default values will run the alignment on .../PCANS/data/finalMouse_data.csv file (windows OS: ...\\PCANS\\data\\ytfinalMouse_data.csv) and output the results to .../PCANS/outputFiles/ directory (windows OS: ...\\PCANS\\outputFiles\\). To run an alignment on a different CSV file, one would need

to replace ".../PCANS/data/finalMouse_data.csv" (#3) with the location and file name of the different CSV file. In addition, to save the alignment results in a different location one would need to replace ".../PCANS/outputFiles/" (#8) with that new location. The other alignment options can also be changed as described below. Once one has the desired values in the GUI click the 'OK' button to run alignment program. Clicking 'Cancel' closes the program without running. NOTE THE GUI DOES NO ERROR CHECKING ON INPUTS, MISTAKES WILL LEAD TO A RUNTIME FAILURE. BE AWARE OF WHAT IT ENTERED INTO GUI TEXT BOXES. CLICKING "OK" WITHOUT CHANGING THE DEFAULTS WILL ALWAYS RUN THE PROGRAM



Detailed Description of GUI Inputs - NEW PCANS PEAK PICKING ONLY – USER DEFINED CUTOFF:

1. Minimum separation expected between peaks within each spectra. Generally 0.001 ppm but is data dependent, could be as small as 0.0001 ppm.
2. Input Spectra File location. Defaults to .../PCANS/data/ OR ...\\PCANS\\data\\, where ... represents where PCANS.zip was unzipped. To change to a different input file format should be /MyPath/ or C:\\MyPath\\ dependent upon one's OS.
3. Input Spectra file name. Defaults to finalMouse_data.csv. To change to a different input file format should be MySpectraData.csv dependent upon one's OS. Expects the data to be a csv file, where first column contains chemical shift positions and the rest of the columns contain the

spectra. The format should be the same as what can be found in

.../PCANS/data/finalMouse_data.csv use it as a guide to creating your input spectra files

4. Number of input spectra in the input spectra file (integer), .../PCANS/data/finalMouse_data.csv has 23 columns first is chemical shift rest are spectra so its input is 22
5. Number of points required to form a peak within a spectra using the peak picking algorithm (integer). Data dependent, but 5 worked well with finalMouse_data.csv.
6. Peak Selection Cutoff – This is the cutpoint used to define what is a peak and what is noise in the new peak selection algorithm. For 5 points to define a peak the value 2.7 worked well with our finalMouse_data.csv. This value could range between 2.5 – 3.5 with 3.0 being what was used with the original paper about this new method.
7. Apply cutoff define by Area Only (as in original paper), Height (Relative Intensity) Only, or Both Area & Height. This is how you are defining a peak with regards to determining if a selected potential peak is “noise” or a real peak object.
8. Output directory (including the last backward or forward slash) where output files will be stored. Defaults to .../PCANS/outputFiles/ OR ...\\PCANS\\outputFiles\\, where ... represents where PCANS.zip was unzipped. To change to a different output directory, the format should be /MyPath/OutputFolder/ or C:\\MyPath\\OutputFolder\\ dependent upon one’s OS.
9. Zero fill value is the value that will be used to fill in the output files when there is no peak for a given chemical shift value. Typically we have used 0.0001 instead of zero for the OPLS analysis.

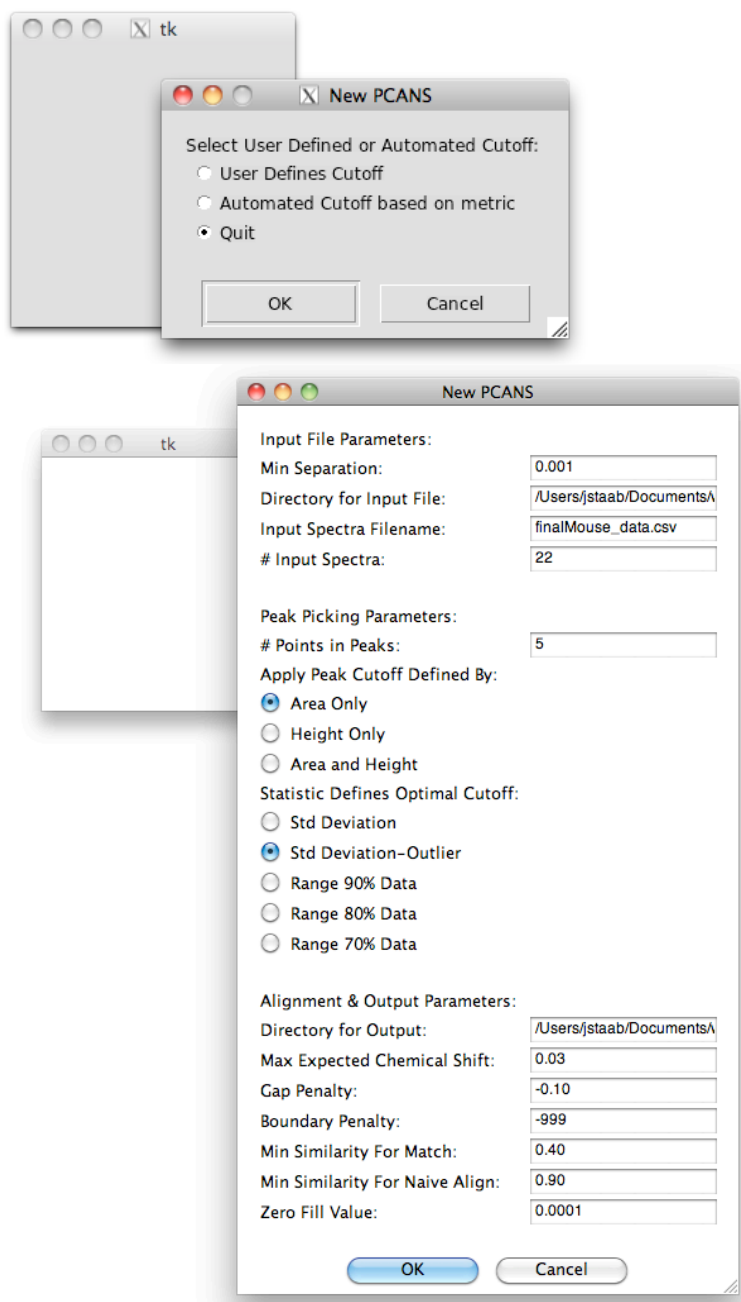
Expected Output Files in Output Directory - NEW PCANS with Alignment – USER DEFINED CUTOFF:

NOTE most files are labeled with PxxAyy where xx indicated the number of points that determine a peak and yy indicates the cutoff used.

1. **origPickPeaks.pkl** – Spectra object containing all the peaks picked prior to using the new peak picking algorithm to determine which peaks are real and which are noise based upon the selected cutoff.
2. **pickPeakSpectraPxxAyy.py** – Before alignment of the picked peaks, the picked peak profiles (spectra) saved as a spectra object.
3. **pickPeakSpectraArPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the peak area of the picked peaks for the input peak profiles (spectra) before alignment. The Area was calculated using Heron’s formula as indicated in the “UPDATES” section of this document. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
4. **pickPeakSpectraHtPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input peak profiles (spectra) before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
5. **pickPeakSpectraWdPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the width at half-height (width) of the picked peaks for the input spectra before alignment. Zero fill values appear where there are no peaks for a given chemical shift

and input peak profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input profiles (spectra).

10. IF you pick NEW Peak Picking & Alignment – AUTOMATED CUTOFF – the following appears: You must first pick if you want “User Defines Cutoff” OR “Automated Cutoff based on a Metric”. This assumes you are using “Automated Cutoff based on a Metric”. Default values are contained within the GUI window. The default values will run the alignment on .../PCANS/data/finalMouse_data.csv file (windows OS: ...\\PCANS\\data\\finalMouse_data.csv) and output the results to .../PCANS/outputFiles/ directory (windows OS: ...\\PCANS\\outputFiles\\). To run an alignment on a different CSV file, one would need to replace “.../PCANS/data/finalMouse_data.csv”(#3) with the location and file name of the different CSV file. In addition, to save the alignment results in a different location one would need to replace “.../PCANS/outputFiles/”(#8) with that new location. The other alignment options can also be changed as described below. Once one has the desired values in the GUI click the ‘OK’ button to run alignment program. Clicking ‘Cancel’ closes the program without running. NOTE THE GUI DOES NO ERROR CHECKING ON INPUTS, MISTAKES WILL LEAD TO A RUNTIME FAILURE. BE AWARE OF WHAT IT ENTERED INTO GUI TEXT BOXES. CLICKING “OK” WITHOUT CHANGING THE DEFAULTS WILL ALWAYS RUN THE PROGRAM

**Detailed Description of GUI Inputs - NEW PCANS with Alignment – AUTOMATED CUTOFF:**

1. Minimum separation expected between peaks within each spectra. Generally 0.001 ppm but is data dependent, could be as small as 0.0001 ppm.
2. Input Spectra File location. Defaults to .../PCANS/data/ OR ...\\PCANS\\data\\, where ... represents where PCANS.zip was unzipped. To change to a different input file format should be /MyPath/ or C:\\MyPath\\ dependent upon one's OS.
3. Input Spectra file name. Defaults to finalMouse_data.csv. To change to a different input file format should be MySpectraData.csv dependent upon one's OS. Expects the data to be a csv file, where first column contains chemical shift positions and the rest of the columns contain the

spectra. The format should be the same as what can be found in

.../PCANS/data/finalMouse_data.csv use it as a guide to creating your input spectra files

4. Number of input spectra in the input spectra file (integer), .../PCANS/data/finalMouse_data.csv has 23 columns first is chemical shift rest are spectra so its input is 22
5. Number of points required to form a peak within a spectra using the peak picking algorithm (integer). Data dependent, but 5 worked well with finalMouse_data.csv.
6. Apply cutoff define by Area Only (as in original paper), Height (Relative Intensity) Only, or Both Area & Height. This is how you are defining a peak with regards to determining if a selected potential peak is "noise" or a real peak object.
7. Statistic/Metric to use to find optimal cutoff under the assumption that the user expects about the same number of spectral peaks in each of the input spectrum. The first option is standard deviation of the number of peaks in each of the input spectra, the next option is standard deviation with the removal of up to 5% of the input spectra from the standard deviation calculation IF THE NUMBER OF PEAKS THEY HAVE IS LABELS THEM AS AN OUTLIER (3+ standard deviations from the mean). The next three options use 90%, 80% or 70% range of the peaks from the input spectra as the metrics. Note that using these metrics the cutoff that is deemed as optimal is the one that gives the smallest metric.
8. Output directory (including the last backward or forward slash) where output files will be stored. Defaults to .../PCANS/outputFiles/ OR ...\\PCANS\\outputFiles\\, where ... represents where PCANS.zip was unzipped. To change to a different output directory, the format should be /MyPath/OutputFolder/ or C:\\MyPath\\OutputFolder\\ dependent upon one's OS.
9. Maximum expected chemical shift. Should be a real number, typically 0.03 or 0.04 ppm depending on the data.
10. Gap penalty should be a small real number that is less than the minimum expected similarity for a match. Can range from a small negative number to a value that is less than the minimum expected similarity for a match value (#11). -0.10 worked well for finalMouse_data.csv.
11. Boundary Penalty should be a value that is smaller than the gap penalty. If the user would prefer that the algorithm compute this penalty, then set the value to the default of -999. For finalMouse_data.csv this value was set to -999.
12. Minimum expected similarity between two peaks to allow for a match. Should be some proportion that ranges between 0.0 and 1.0. Typical value would be 0.60 or higher, but is data dependent. For finalMouse_data.csv 0.40 was used.
13. Minimum similarity required for naïve alignment, if two peaks are at least this similar then they will be aligned using the naïve alignment scheme. Typically this value would be 0.90 or higher, 0.90 was used for finalMouse_data.csv.
14. Zero fill value is the value that will be used to fill in the output files when there is no peak for a given chemical shift value. Typically we have used 0.0001 instead of zero for the OPLS analysis.

Expected Output Files in Output Directory - NEW PCANS with Alignment – AUTOMATED CUTOFF:

NOTE most files are labeled with PxxAyy where xx indicated the number of points that determine a peak and yy indicates the cutoff used.

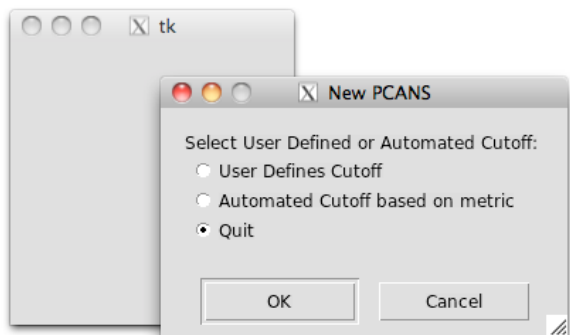
12. **origPickPeaks.pkl** – Spectra object containing all the peaks picked prior to using the new peak picking algorithm to determine which peaks are real and which are noise based upon the selected cutoff.
13. **pickPeakSpectraPxxAyy.py** – Before alignment of the picked peaks, the picked peak profiles (spectra) saved as a spectra object.

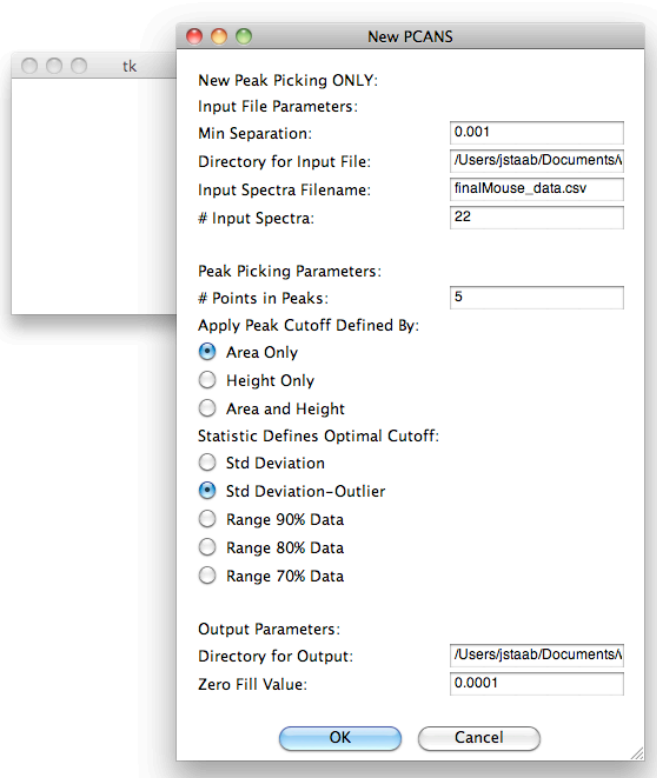
14. **pickPeakSpectraArPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the peak area of the picked peaks for the input peak profiles (spectra) before alignment. The Area was calculated using Heron's formula as indicated in the "UPDATES" section of this document. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
15. **pickPeakSpectraHtPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input peak profiles (spectra) before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
16. **pickPeakSpectraWdPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the width at half-height (width) of the picked peaks for the input spectra before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input peak profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input profiles (spectra).
17. **alignSpectraPxxAyy.py** -- After alignment of the picked peaks, the aligned peak profiles (spectra) saved as a spectra object.
18. **alignedSpectraArPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) after PCANS alignment. The first column contains chemicals shift position, next columns contain the peak area of the picked peaks for the input peak profiles (spectra) after alignment. The Area was calculated using Heron's formula as indicated in the "UPDATES" section of this document. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
19. **alignedSpectraHtPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) after PCANS alignment. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input peak profiles (spectra) after alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
20. **alignedSpectraWdPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) after PCANS alignment. The first column contains chemicals shift position, next columns contain the width at half-height (width) of the picked peaks for the input spectra after alignment. Zero fill values appear where there are no peaks for a given chemical shift and input peak profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input profiles (spectra).
21. **finalConsensusSpectrumPxxAyy.csv** -- comma separated file that contains the final consensus spectrum post alignment. The first column contains chemicals shift position of the peaks in the final consensus spectrum. The values in this column represent the median chemical shift position of all the peaks that were aligned to that chemical shift position using PCANS. The next column contains the average relative intensity (height) of the peaks that were aligned to that chemical shift position in the consensus spectrum, the next column contains the average width at half-height (width) of all the peaks that were aligned to that chemical shift position in the

consensus spectrum, and the final column contains the number of peaks that were aligned to that chemical shift position in the consensus spectrum.

22. finalConsensusSpectrumPxxAyy.py – Final consensus spectrum post-alignment saved as a spectra object.

11. IF you pick NEW Peak Picking ONLY – AUTOMATED CUTOFF – the following appears: You must first pick if you want “User Defines Cutoff” OR “Automated Cutoff based on a Metric”. This assumes you are using “Automated Cutoff based on a Metric”. Default values are contained within the GUI window. The default values will run the alignment on .../PCANS/data/finalMouse_data.csv file (windows OS: ...\\PCANS\\data\\ytfinalMouse_data.csv) and output the results to .../PCANS/outputFiles/ directory (windows OS: ...\\PCANS\\outputFiles\\). To run an alignment on a different CSV file, one would need to replace “.../PCANS/data/finalMouse_data.csv”(#3) with the location and file name of the different CSV file. In addition, to save the alignment results in a different location one would need to replace “.../PCANS/outputFiles/”(#8) with that new location. The other alignment options can also be changed as described below. Once one has the desired values in the GUI click the ‘OK’ button to run alignment program. Clicking ‘Cancel’ closes the program without running. NOTE THE GUI DOES NO ERROR CHECKING ON INPUTS, MISTAKES WILL LEAD TO A RUNTIME FAILURE. BE AWARE OF WHAT IT ENTERED INTO GUI TEXT BOXES. CLICKING “OK” WITHOUT CHANGING THE DEFAULTS WILL ALWAYS RUN THE PROGRAM



**Detailed Description of GUI Inputs - NEW PCANS peak picking ONLY – AUTOMATED CUTOFF:**

1. Minimum separation expected between peaks within each spectra. Generally 0.001 ppm but is data dependent, could be as small as 0.0001 ppm.
2. Input Spectra File location. Defaults to .../PCANS/data/ OR ...\\PCANS\\data\\, where ... represents where PCANS.zip was unzipped. To change to a different input file format should be /MyPath/ or C:\\MyPath\\ dependent upon one's OS.
3. Input Spectra file name. Defaults to finalMouse_data.csv. To change to a different input file format should be MySpectraData.csv dependent upon one's OS. Expects the data to be a csv file, where first column contains chemical shift positions and the rest of the columns contain the spectra. The format should be the same as what can be found in .../PCANS/data/finalMouse_data.csv use it as a guide to creating your input spectra files
4. Number of input spectra in the input spectra file (integer), .../PCANS/data/finalMouse_data.csv has 23 columns first is chemical shift rest are spectra so its input is 22
5. Number of points required to form a peak within a spectra using the peak picking algorithm (integer). Data dependent, but 5 worked well with finalMouse_data.csv.
6. Apply cutoff define by Area Only (as in original paper), Height (Relative Intensity) Only, or Both Area & Height. This is how you are defining a peak with regards to determining if a selected potential peak is "noise" or a real peak object.
7. Statistic/Metric to use to find optimal cutoff under the assumption that the user expects about the same number of spectral peaks in each of the input spectrum. The first option is standard deviation of the number of peaks in each of the input spectra, the next option is standard deviation with the removal of up to 5% of the input spectra from the standard deviation calculation IF THE NUMBER OF PEAKS THEY HAVE IS LABELS THEM AS AN OUTLIER (3+ standard deviations from the mean). The next three options use 90%, 80% or 70% range of the peaks

from the input spectra as the metrics. Note that using these metrics the cutoff that is deemed as optimal is the one that gives the smallest metric.

8. Output directory (including the last backward or forward slash) where output files will be stored. Defaults to .../PCANS/outputFiles/ OR ...\\PCANS\\outputFiles\\, where ... represents where PCANS.zip was unzipped. To change to a different output directory, the format should be /MyPath/OutputFolder/ or C:\\MyPath\\OutputFolder\\ dependent upon one's OS.
9. Zero fill value is the value that will be used to fill in the output files when there is no peak for a given chemical shift value. Typically we have used 0.0001 instead of zero for the OPLS analysis.

Expected Output Files in Output Directory - NEW PCANS Peak Picking ONLY – AUTOMATED CUTOFF:
NOTE most files are labeled with PxxAyy where xx indicated the number of points that determine a peak and yy indicates the cutoff used.

1. **origPickPeaks.pkl** – Spectra object containing all the peaks picked prior to using the new peak picking algorithm to determine which peaks are real and which are noise based upon the selected cutoff.
2. **pickPeakSpectraPxxAyy.py** – Before alignment of the picked peaks, the picked peak profiles (spectra) saved as a spectra object.
3. **pickPeakSpectraArPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the peak area of the picked peaks for the input peak profiles (spectra) before alignment. The Area was calculated using Heron's formula as indicated in the "UPDATES" section of this document. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
4. **pickPeakSpectraHtPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the relative intensity (height) of the apex of the picked peaks for the input peak profiles (spectra) before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input peak profiles (spectra).
5. **pickPeakSpectraWdPxxAyy.csv** -- comma separated file that contains the input peak profiles (spectra) before PCANS alignment. The first column contains chemicals shift position, next columns contain the width at half-height (width) of the picked peaks for the input spectra before alignment. Zero fill values appear where there are no peaks for a given chemical shift and input peak profiles (spectra). The final column contains a count of how many peaks exist for a given chemical shift position across all input profiles (spectra).

Uses UNC's Copyright and Permission Notice:

COPYRIGHT AND PERMISSION NOTICE

UNC Software: PCANS

Copyright (C) 2009 The University of North Carolina at Chapel Hill

All rights reserved

The University of North Carolina at Chapel Hill ("UNC") and the developers ("Developers") of PCANS ("Software") give recipient ("Recipient") and Recipient's Institution ("Institution") permission to use and copy the software in source and binary forms, with or without modification for **non-commercial purposes only** provided that the following conditions are met:

- 1) All copies of Software in binary form and/or source code, related documentation and/or other materials provided with the Software must reproduce and retain the above copyright notice, this list of conditions and the following disclaimer.
- 2) Recipient and Institution shall not distribute Software to any third parties.
- 3) The Software is provided "As Is." The Developers can not guarantee the provision of technical support or consultation for the Software. The Developers may provide a location on a UNC Web Site for Recipients to post comments, questions, and suggestions at some time in the future. Recipient may provide the Developers with feedback on the use of the Software in their research at that time. The Developers and UNC are permitted to use any information Recipient provides in making changes to the Software.
- 4) Recipient acknowledges that the Developers, UNC and its licensees may develop modifications to Software that may be substantially similar to Recipient's modifications of Software, and that the Developers, UNC and its licensees shall not be constrained in any way by Recipient in UNC's or its licensees' use or management of such modifications. Recipient acknowledges the right of the Developers and UNC to prepare and publish modifications to Software that may be substantially similar or functionally equivalent to your modifications and improvements, and if Recipient or Institution obtains patent protection for any modification or improvement to Software, Recipient and Institution agree not to allege or enjoin infringement of their patent by the Developers, UNC or any of UNC's licensees obtaining modifications or improvements to Software from the UNC or the Developers.
- 5) Recipient and Developer will acknowledge in their respective publications the contributions made to each other's research involving or based on the Software. The current citations for Software are:

<<LIST current citations here.>>>
- 6) Any party desiring a license to use the Software for commercial purposes shall contact The Office of Technology Development at UNC at 919-966-3929.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS, CONTRIBUTORS, AND THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER, CONTRIBUTORS OR THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT

LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.