# Get spikes

* MCD files won’t open in OfflineSorter unless MC\_Rack is installed also
* Detect spikes
  + 1 sample = 1 tick = 100µs
  + Waveform length: 5 ms = 50 samples
  + Prethreshold period: 1 ms = 10 samples
  + Min dead time b/w waveforms: 0 ms = 0 samples
    - 0 because separate cells could be firing around the same time and we don’t miss those for later sorting analysis
  + Positive threshold of 5 σ’s
    - No negative threshold value because older Offline Sorter doesn’t have that option
* Offline Sorter will consider any threshold cross a spike, then extract “waveform length” samples from around that spike, with “prethreshold period” samples before the spike; if two spikes are within “dead time” of each other, then the second one is removed
* Remove artifacts
  + Width of artifact: 1ms = 10 ticks
  + Occurs on 70% of channels ( 42 / 60)
* Make Excel spreadsheet cataloguing each file’s recording length, approx timestamps of any artifacts, and (optionally) any channels that were really busy
* Export as .PLX files
  + Only timestamp data (no continuous)

# Data Validation on Channels

* Open .PLX files for each retina’s latest recording (so they had time to settle on MEA)
  + If last dark and last light were consecutive, look at both and compare
  + Don’t use any recordings shorter than 600 sec
* Align waveforms
  + Use global minimum or global maximum
  + Adjust timestamps?
  + Try without, try with, see which gives fewest false positives
    - Test on light AND dark files
* Sort waveforms into units
  + Use a T-dist E-M sort on 3D feature space (Chalupa used Shoham EM cluster algorithm)
  + Features are principal components 1-3 (in keeping with Chalupa)
* See which channels have 1 cell
  + 1 unit has the vast majority of spikes or there are pretty few spikes over all
  + All other channels will not be considered during inter-genotype T-tests (doesn’t affect *intra*-genotype tests though)
* See which channels don’t have an mRGC (so we’re not just seeing light-evoked spiking)
  + Really high spiking the whole time = mRGC, so scrap channel
  + Really low spiking the whole time, even without obvious waves = no mRGC, so keep channel
  + Really high spiking, even with obvious waves = mRGC, so scrap it
  + High spiking right when the light is turned on is expected either way
  + Obvious bursts with steady, low back ground firing = good judgment!
* Save this information into Excel spreadsheet from above…DO NOT save these changes to the .PLX files
  + I.e., only PLX files with every waveform unsorted should be kept

# Prune Spikes and WABs

* At the end of this, there will be folders for each experimental population (genotype / light condition)
  + Each population folder will have:
    - A folder for each retina, containing text files of the recordings (generated below)
    - A wave-analysis summary workbook for each retina
    - And nothing else!!!
  + The text files and Excel workbooks are generated as follows
* For each recording
  + Open its .PLX files in NeuroExplorer
    - If no data shows up when PLX is opened, go to View > Data Import Options > Plexon Files and select all data types that you want imported
  + Do burst analysis
    - Use all data, not just 30-600 sec or anything like that (Excel will do this later)
    - Don’t use short recordings (names like Dark30001.plx)
    - Chalupa’s parameters:
      * Max ISI to start burst: 0.1 s
      * Min ISI to end burst: 1 s
      * Min interval between bursts: 5 s
      * Min duration of burst: 0.5 s
      * Min number of spikes in burst: 4
    - My parameters (shorter bursts, closer together, with longer ISIs, basically just yielding *more* bursts):
      * Max ISI to start burst: 0.5 s
      * Min ISI to end burst: 0.5 s
      * Min interval between bursts: 1 s
      * Min duration of burst: 0.1 s
      * Min number of spikes in burst: 4
    - Surprise algorithm parameters:
      * Min Surprise = 0 (smaller number seems to include more spikes per burst)
  + Export *interval* and *neuron* (spike) data to TXT files and import these to Excel workbook
    - These TXT files will be sorted into the retinal folders
    - Don’t import any files whose end timestamps are before time\_end (600s)
  + Remove bad data:
    - Spikes with timestamps before time\_start (30s) or after time\_end (600s)
    - Bursts with start timestamps before time\_start or end timestamps after time\_end
    - Bursts where *peak* in-burst spike freq is not different from *mean* in-burst spike freq by at least 1000% (i.e., (peak-mean)/mean >= 10 )
    - Non-wave-associated bursts
      * For each channel
        + For each burst interval on that channel

If each nth (=10th)of a burst is within the required time range of some nth of a burst on at least k (=3) neighboring channels, then it’s a WAB

Time range is half a bin duration in plus half the neighbor’s bin duration

* + Get burst property values of every channel, similar to NeuroExplorer’s burst analysis summary results
    - Mean spike frequency = (num spikes) / (recording time = 600 – 30)
    - Mean inter-spike interval = (recording time
    - Mean burst frequency = (num bursts) / (recording time = 600 – 30)
    - Mean inter-burst interval = AVG(burst2Start – burst1End)
    - Mean burst duration = AVG(burstStart – burstEnd)
    - Percent of Spikes in Bursts = SUM(spikes in bursts) / spikes
    - Mean in-burst spike frequency
      * Count all spikes on a channel that are within a burst and divide by that burst’s duration
      * Average these values over all bursts on a channel
    - Mean peak in-burst spike frequency
      * Get the shortest interspike interval between spikes in a burst
      * That burst’s peak spike freq is then (1 / shortest ISI)
      * Average these values over all bursts on a channel
* Create a new sheet for Averages
  + Channel IDs in row headers, property names in column headers
  + Place averages of the above properties over all recordings into corresponding cells of the Averages sheet
  + Keep all recording sheets visible and catalogued on a Contents sheet
  + Remove bad channels
    - Channels with fewer than 3 (?) WABs on average, channels with average burst durations longer than 10s (?)

# Do statistics

* Minitab ANOVA configuration
  + Do Stat > ANOVA > One-Way
    - Options > check “Assume equal variances”, use “Confidence level” of 95, with Two-sided “Type of confidence interval”
    - Comparisons > check “Tukey”, use “Error rate for comparisons” of 5, uncheck “Interval plot for differences of means” and check “Grouping information”
    - Graphs > uncheck “Interval plot”
    - Results > uncheck “Method” for all but the first ANOVA tests being run, then uncheck “Factor information,” “Model summary,” and “Means” for all ANOVA tests
  + Copy the text output from Minitab to the ANOVA Results worksheet for saving
* Prepare a single workbook for stats in the project folder
  + It has a Results sheet, a Main sheet with the macro button, a sheet for each of the two experimental populations being compared
  + Copy data from the Averages sheet of the summary workbook of every retina in an experimental population, and paste into the corresponding sheet of the stats workbook
  + If this is an inter-genotype comparison, delete channels that recorded more than one neuron (found in steps above)
  + Results from a given comparison will be outputted to the Main sheet, which can then be copied to the Results sheet to summarize all comparisons ☺
* Determine the number of channels to resample per retina
  + 70% of number of channels on retina with the fewest remaining channels
  + Same for both experimental populations?
* Do the following 1000 times
  + Resample the calculated number of channels from every retina
  + For each burst property
    - Use the combined values from the resampled channels to conduct a T-test between the two populations on the given property
    - Use Bonferroni correction
* P-value for this comparison equals the average (?) p-value over all 1000 bootstraps
* Compare the two populations
  + Run a multiple hypothesis test (Hotelling’s T-square, p=0.05)
    - Use 70% of the number of channels on retina with fewest valid channels
* Do post-hoc analysis to find which populations differ
  + Multivariate T-tests (=Hotelling’s T2) for all 4C2 = 6 pairwise comparisons to find which ones specifically differ and in what direction
    - Use family-wise error rate correction (Bonferroni? P=0.05/6=.008)
    - Each group consists of average channel-values pulled from retinal populations by second round of bootstrapping
  + For each significant comparison, do T-tests for all 5 variables to see which variables contributed to group differences
    - Use family-wise error rate correction (Bonferroni? P=0.05/5=0.01) for each significant comparison
    - One-tailed or two-tailed?
    - Paired, if same genotype
    - Each group consists of average channel-values pulled from retinal populations by third round of bootstrapping
  + Provide graphs of data (with error bars)