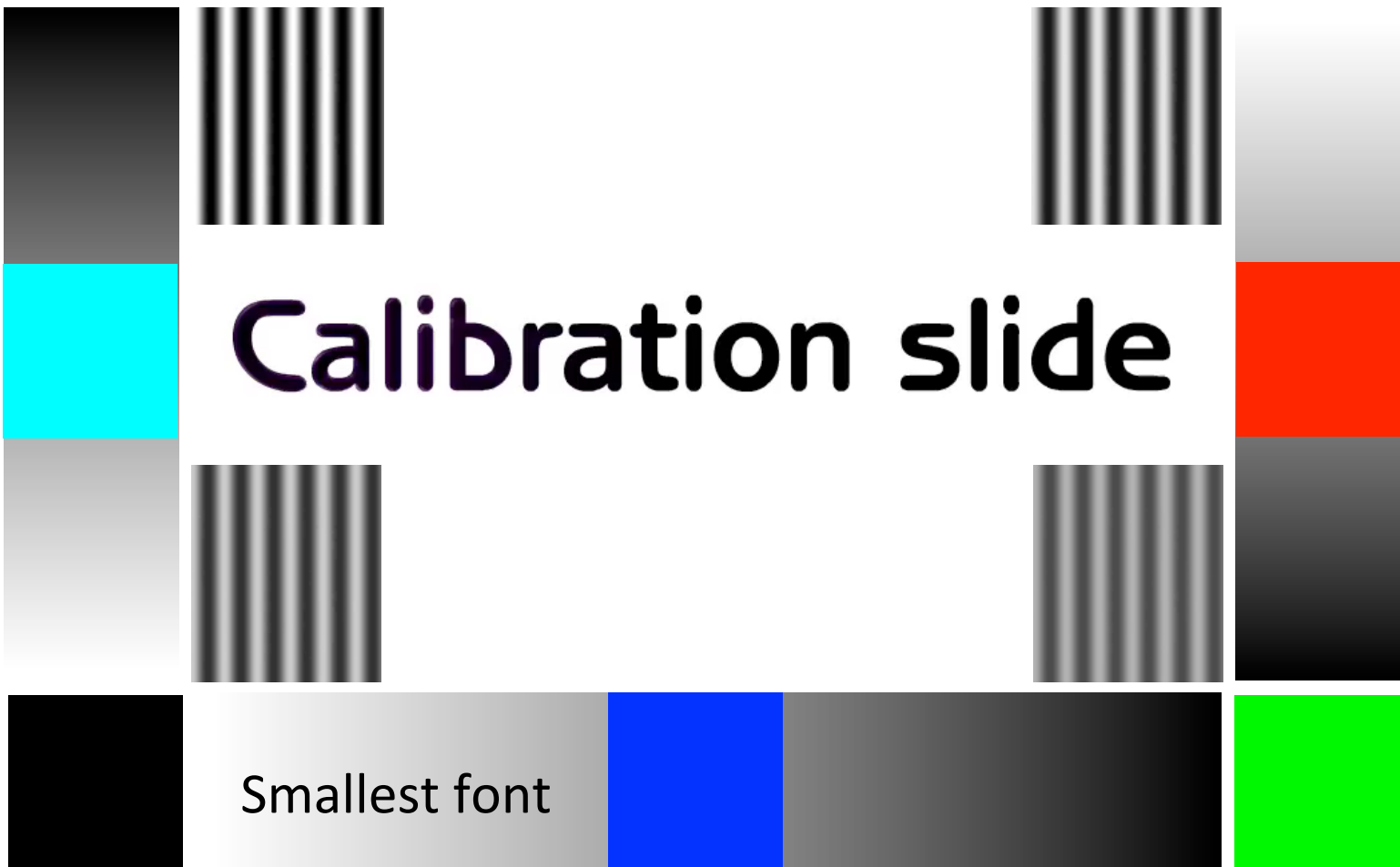


Smallest font



Smallest font

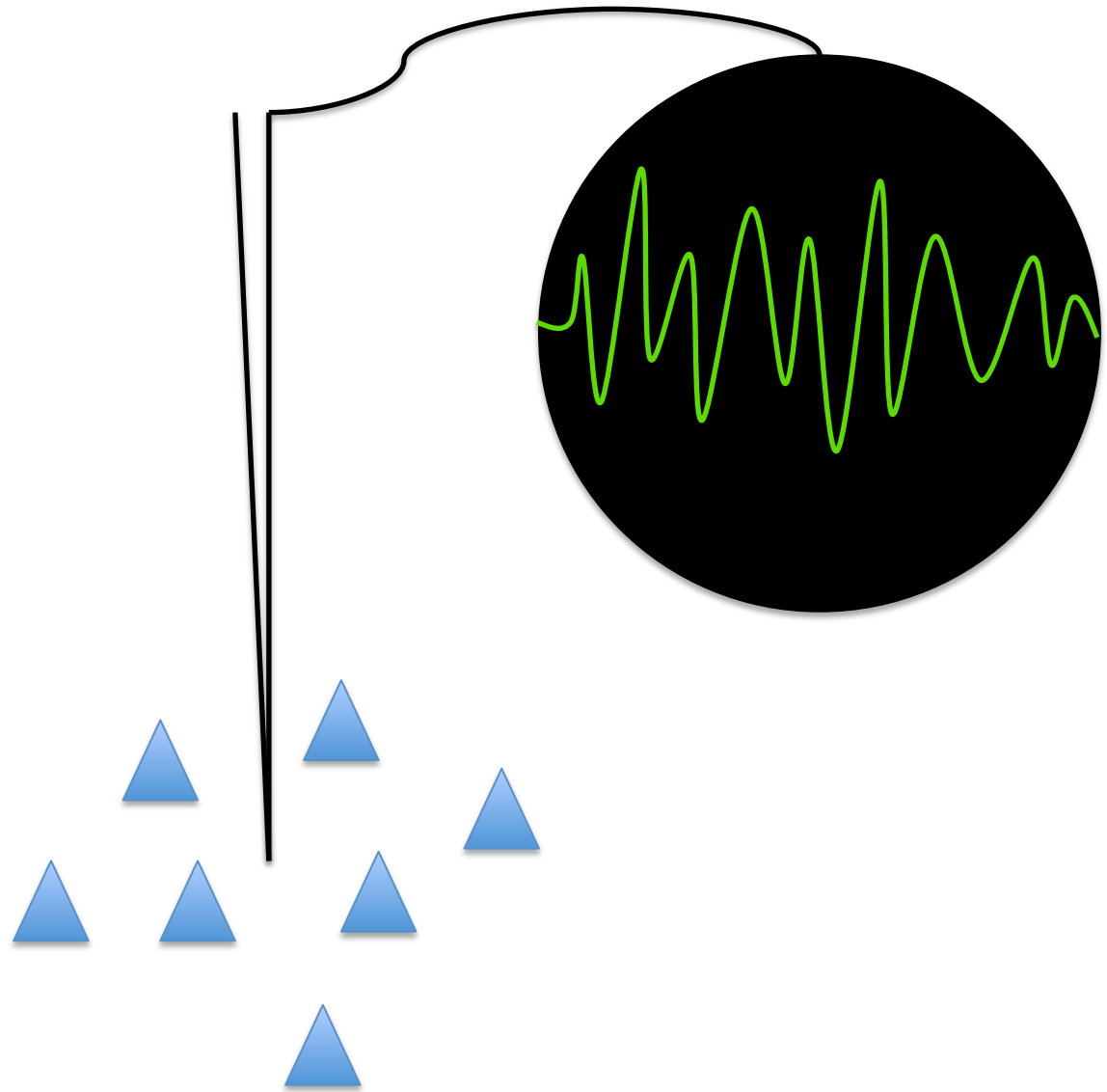
# Spike sorting

## - A serious problem -

Index in position 1 exceeds array bounds (must not exceed 1000)

Pascal Wallisch  
New York University

# The problem



# Why it is tricky

- This is an instantiation of a larger class of problems:  
An inverse problem.
- Solutions are – in principle – underdetermined.

# On inverse problems...

- $3 + 4 = 7$
- $4 + 3 = 7$
- $1 + 6 = 7$
- $6 + 1 = 7$
- $7 = ?$
- Working backwards from the results is inherently ambiguous.
- The same output could have been produced by an infinite number of inputs/sources
- The only way to deal with that is to nail down one of the inputs as a constraint.
- For instance, if I tell you that the result is 7 and one of the inputs is 2, you know that the other input is...



# What people do

0. Don't do it
1. Online spike sorting (manual)
2. Offline spike sorting (manual)
3. Offline spike sorting (increasingly fully automated)

# 0. Don't do it

- Declare that good spike sorting is extremely time consuming, not critical to the question of the study and cannot be done well.
- Solution: Just don't do it at all.
- How?
- Declare that all signals from a single electrode are a “multi-unit cluster” and move on.

# Assessment

## Advantages

- The most time-efficient solution.

## Disadvantages

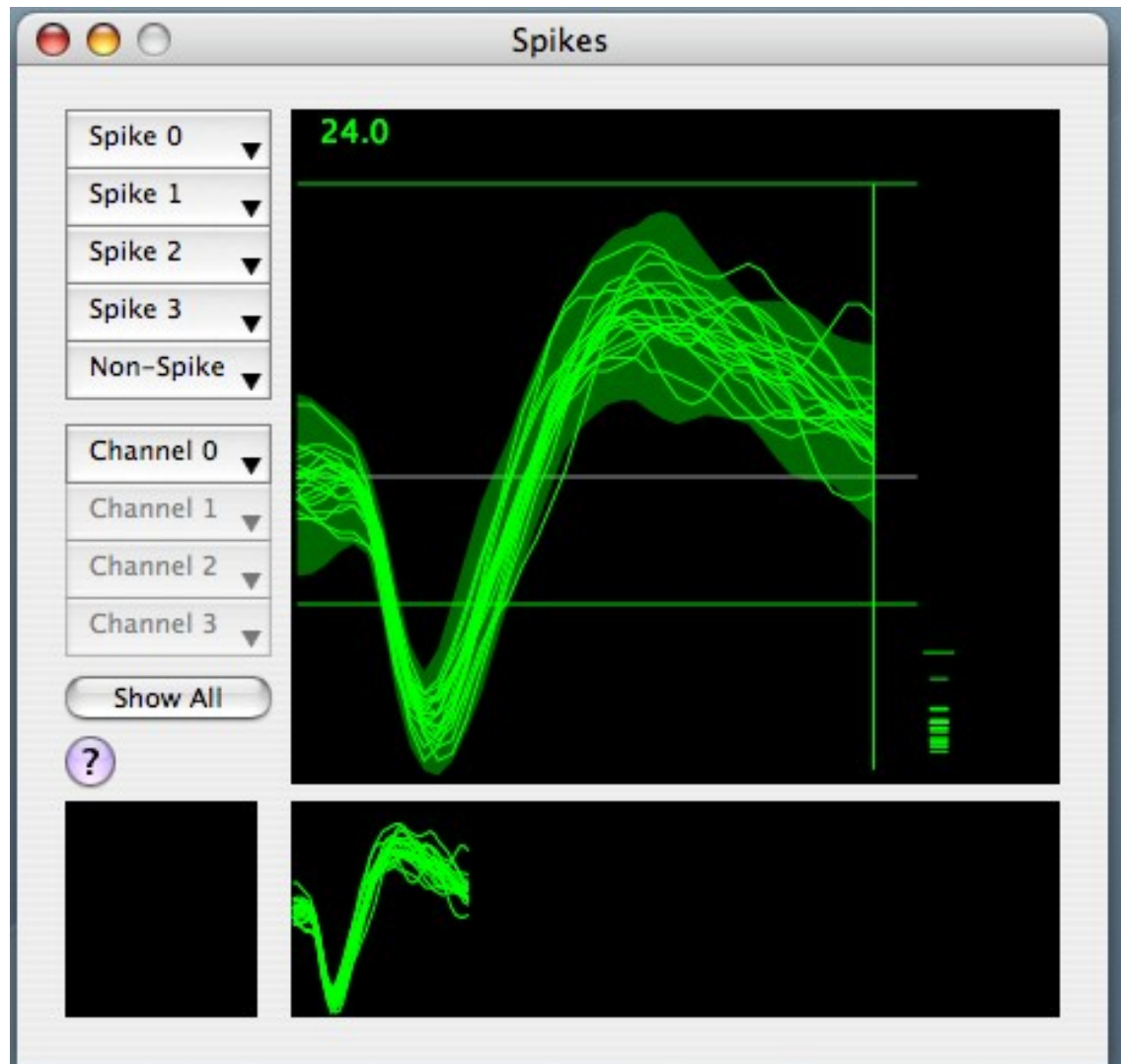
- Sacrificing the key advantages of doing spike recordings in the first place (in contrast to LFP).
- If the question does rely on resolving signals from single neurons, this gives misleading results (e.g. absolute firing rates, relative tuning width, multi-peak tuning curves, ...)



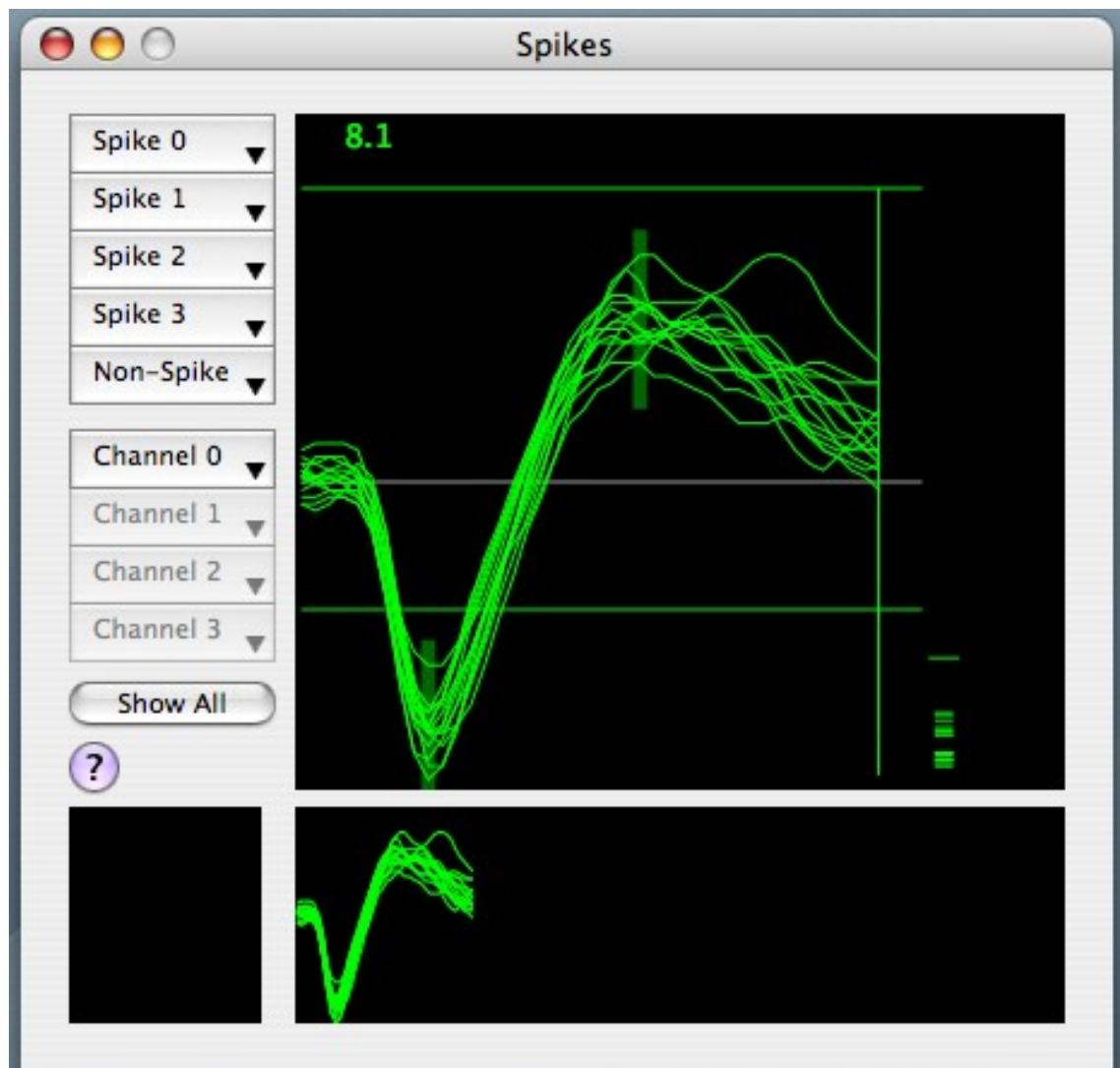
# 1. Online spike sorting

- Implicit heuristic: Because spike sorting is an unsolved problem, it is best avoided.

# Thresholding



# Window discrimination



Live demo (EXPO)

# Assessment

## Advantages

- Straightforward
- Yields extremely clean data if done properly.

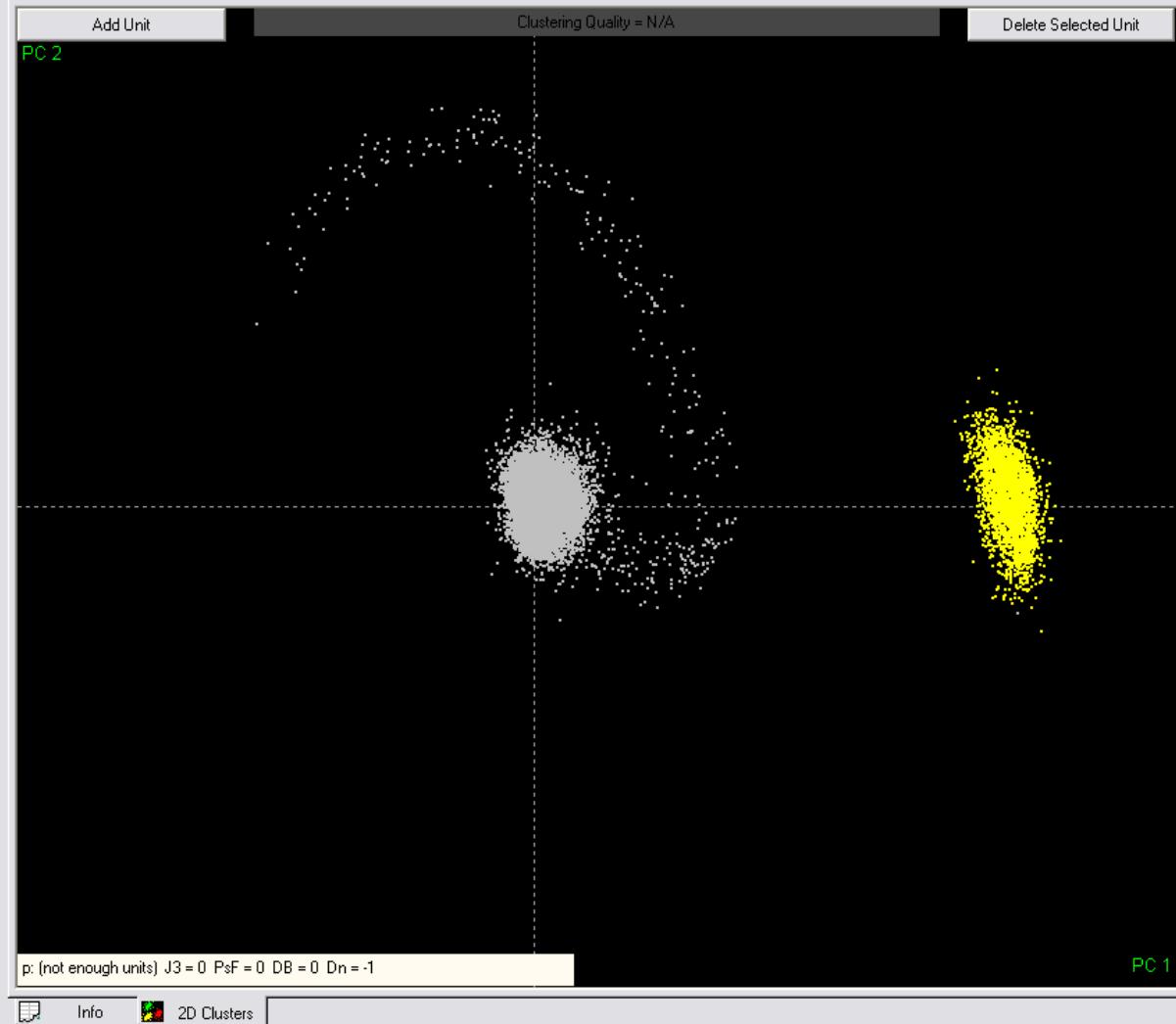
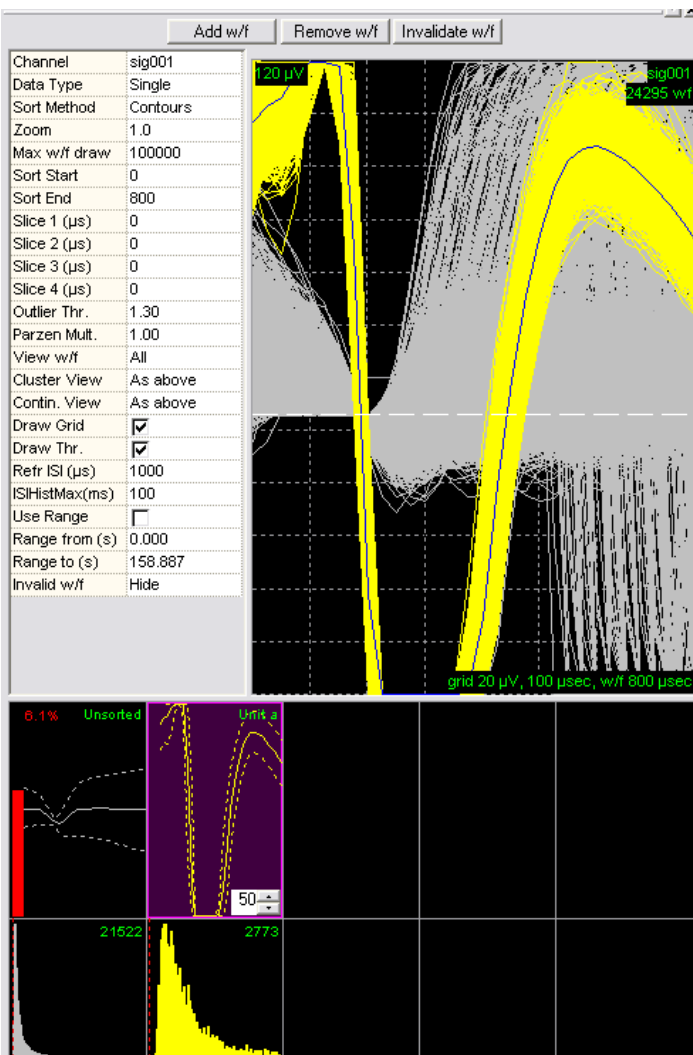
## Disadvantages

- Requires constant vigilance to “maintain isolation” (new neurons “pop up”, others die down, most change shape over time).
- Losing most of the potential signal if criteria are stringent enough.
- No clear metrics for what is “good enough” other than “this looks good enough”
- Not practical for arrays

## 2. Offline spike sorting

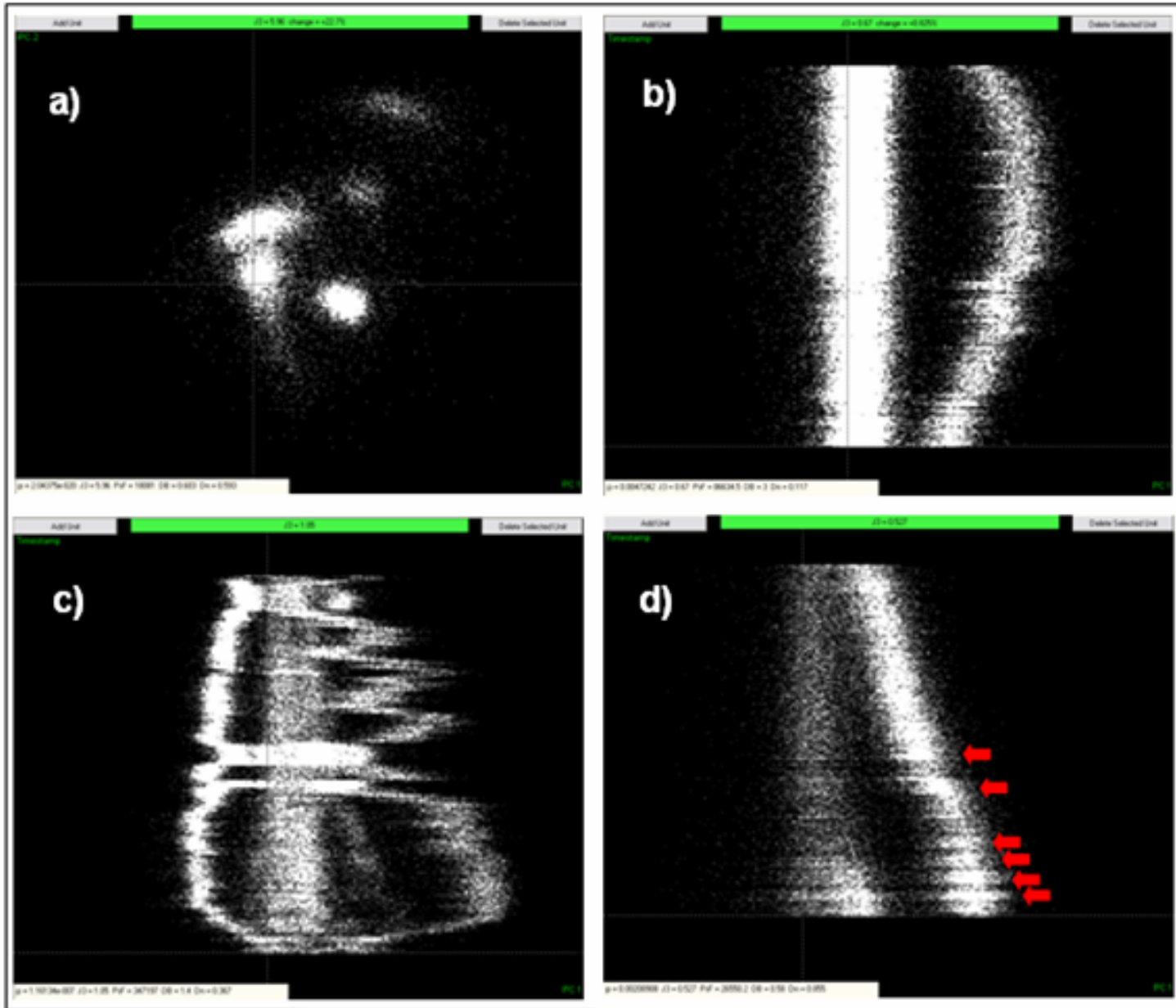
- Implicit assumptions: Spikes from the same neuron have roughly the same shape (allowing for measurement noise).
- So if one does a PCA of the shape, signals from different neurons will resolve into different, distinguishable clusters that can be resolved visually (“cluster-cutting”)

# “Cluster-cutting”



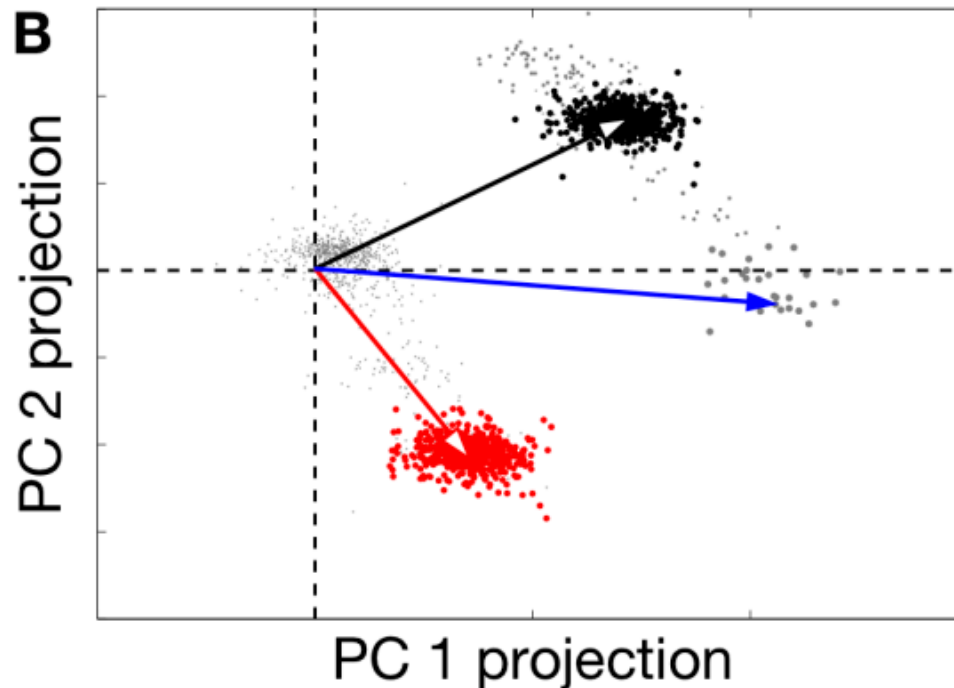
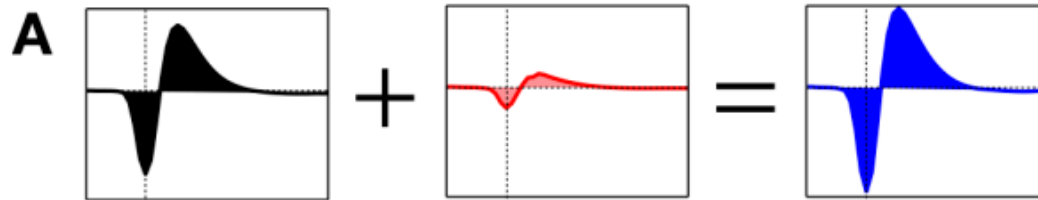


# Documented non-stationarities



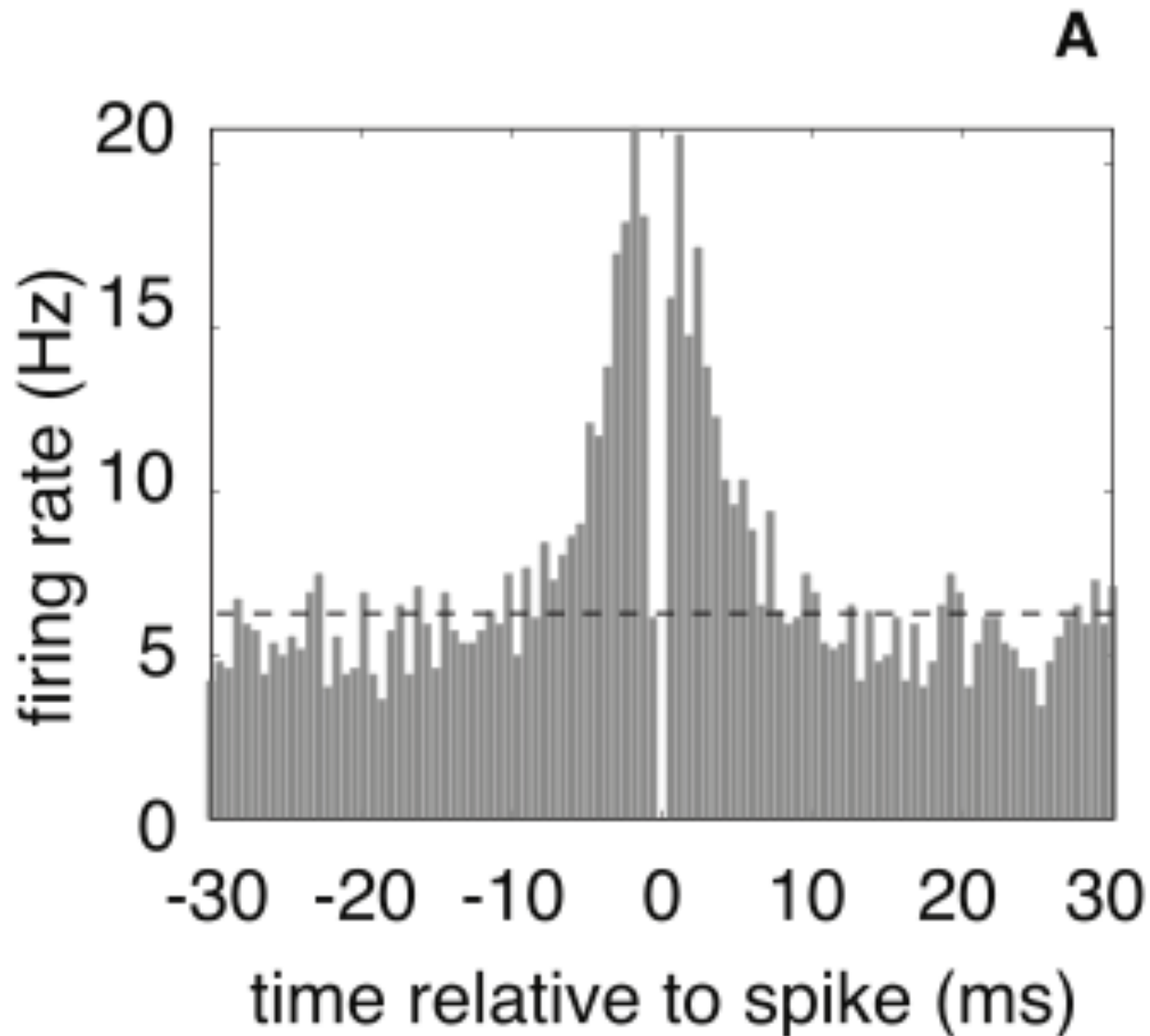
# Synchronous spike superimpose linearly

synchronous spiking



Pillow et al. (2013)

# Why is this a concern?



# Matlab offline sorter

# Assessment

## Advantages

- Keeps most of the signal.
- Allows to do due diligence offline.
- Allows to sort huge arrays.

## Disadvantages

- Extremely time-consuming and tedious if done manually.
- No clear or binding metrics as to what constitutes a good sort across labs (or even within).
- Can't recover a good sort from a bad recording.
- We know that the key assumption (stationarity of the waveform) is not met in many recordings, particularly long ones.

# 3. Algorithmic approaches

- The future? (The end of the beginning)

		<u>HerdingSpikes2</u>	<u>IronClust</u>	<u>JRClust</u>	<u>KiloSort</u>	<u>KiloSort2</u>	<u>Klusta</u>	<u>MountainSort4</u>	<u>SpykingCircus</u>	<u>Tridesclous</u>	<u>Waveclus</u>
+	PAIRED_BOYDEN		0.63	0.59	0.44	0.76	0.34*	0.47	0.54	0.32*	
+	PAIRED_CRCNS_HC1		0.67	0.7*	0.57*	0.75*	0.78	0.83	0.78	0.37	
+	PAIRED_KAMPFF	0.84*	0.87	0.94*	0.8	0.96	0.29*	0.82	0.73	0.63*	
+	PAIRED_MEA64C_YGER	0.85*	0.87	0.86	0.82	0.83*	0.66*	0.81	0.87	0.8*	
+	PAIRED_MONOTRODE		0.26				0.54	0.6	0.49	0.36	0.52
+	SYNTH_BIONET		0.86	0.75	0.82	0.85		0.74	0.66*	0.33	
+	SYNTH_MAGLAND		0.95	0.88	0.83	0.9*	0.85	0.94	0.7	0.75	
+	SYNTH_MEAREC_NEURONEXUS		0.97	0.93	0.91	0.97*	0.87	0.89	0.9	0.85*	
+	SYNTH_MEAREC_TETRODE		0.96	0.65	0.91	0.69*	0.84	0.93	0.87	0.88	
+	SYNTH_MONOTRODE		0.28				0.04*	0.84	0.35	0.25	0.77
+	SYNTH_VISAPY	0.86	0.92		0.82*	0.97	0	0.85	0.95	0.64	
+	HYBRID_JANELIA	0.66	0.92	0.73	0.86	0.9	0.77*	0.7	0.79	0.47	
+	MANUAL_FRANKLAB		0.4	0.31	0.42	0.52	0.57	0.64	0.42	0.03	