Summer Undergraduate Research Program (SURP) Symposium

2020 Summer Undergraduate Research Program (SURP) Symposium

Jul 31st, 1:00 PM - 3:30 PM

Analyzing and Classifying Single Molecule Microscopy Data

Troy Buzynski
University of Northern Iowa

Ali Tabei
University of Northern Iowa

Recommended Citation

Buzynski, Troy and Tabei, Ali, "Analyzing and Classifying Single Molecule Microscopy Data" (2020). *Summer Undergraduate Research Program (SURP) Symposium*. 11.
https://scholarworks.uni.edu/surp/2020/all/11

Analyzing and Classifying Single Molecule Microscopy Data

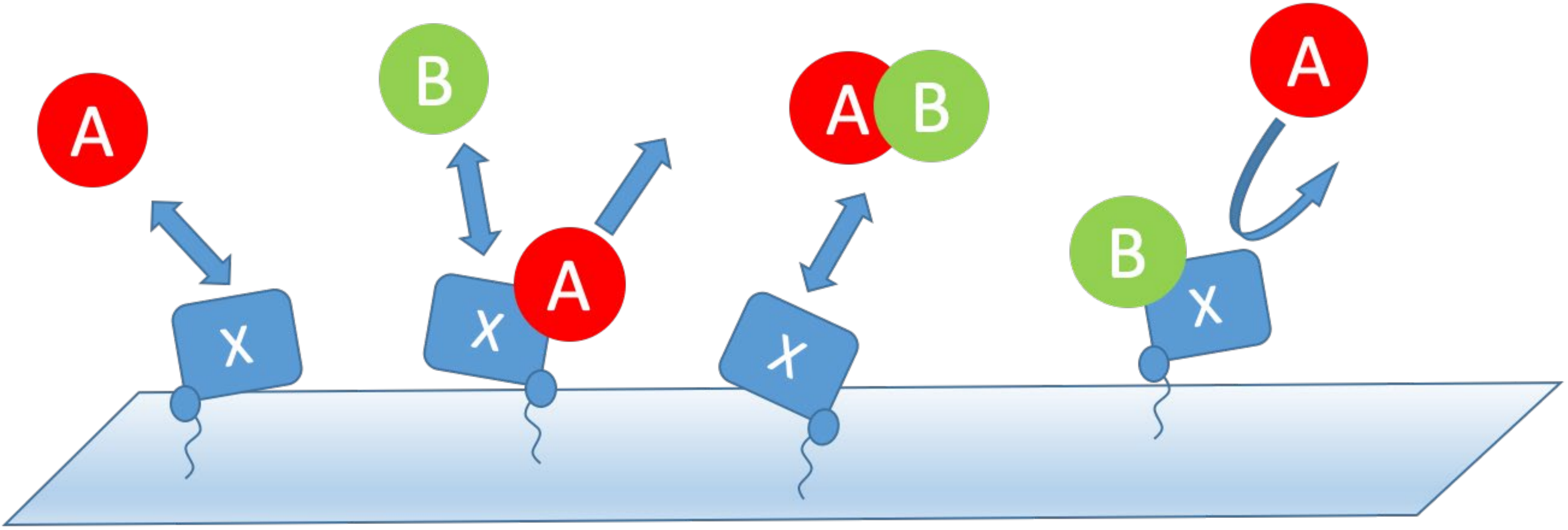
Troy Buzynski University of Northern Iowa

Purpose

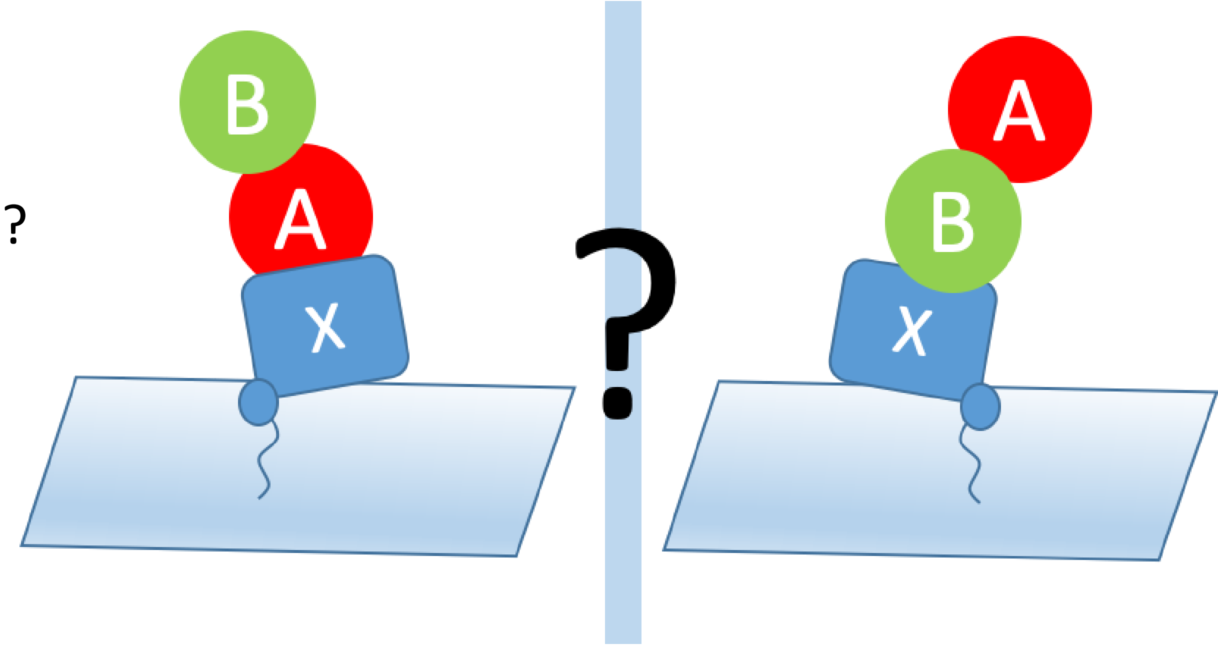
- **Discover the equations that describe the individual interactions between proteins and DNA. Determine the probabilities of each type of interaction.**

- DNA damage is a common occurrence. Organisms have a natural process to repair this damage, but when this process does not function properly it can cause disease, namely cancer. We are attempting to discover more about this process in hope that it leads to better techniques for dealing with diseases such as cancer.

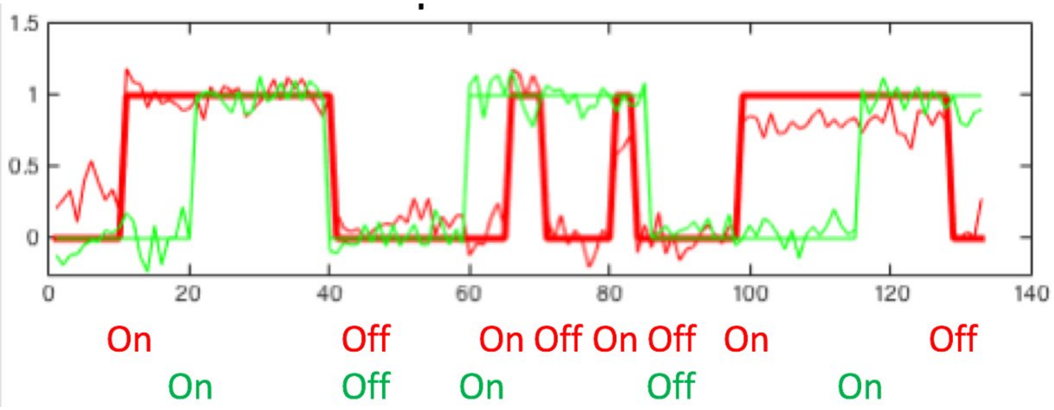
How do these proteins interact?



Does the binding of “B” cause “A” to unbind?
Do “A” and “B” bind together and then bind to the DNA?
If “B” is bound, does the inhibit the binding of “A”?



What kind of interactions are most common?

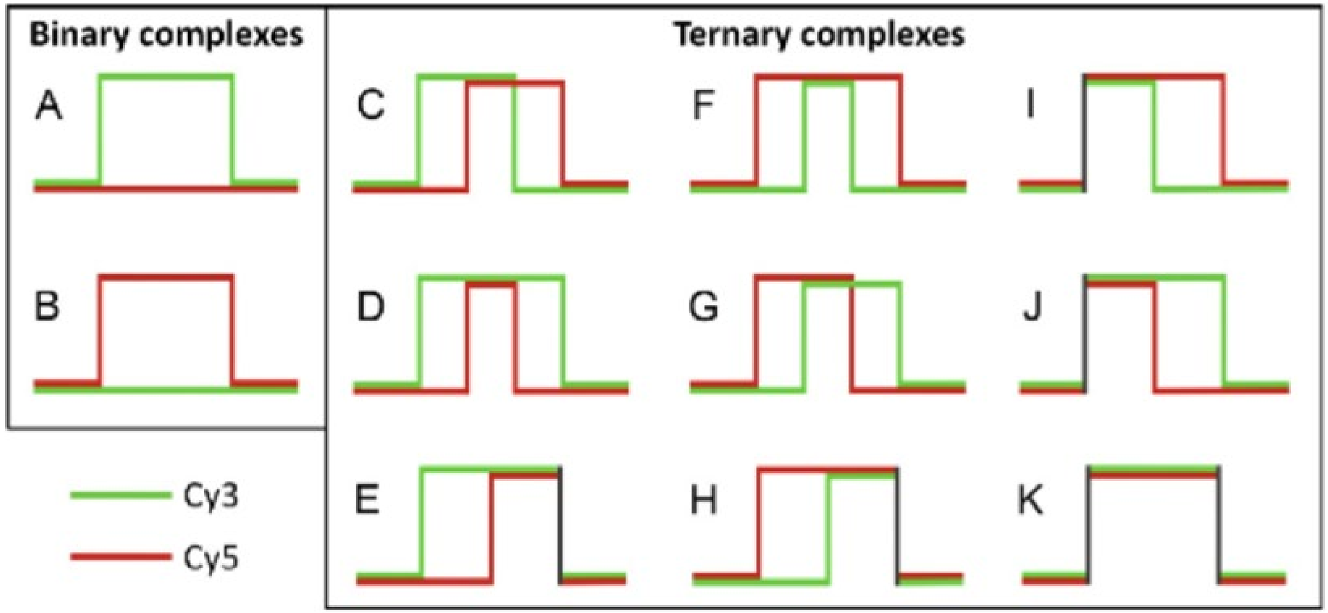


Signal and Noise

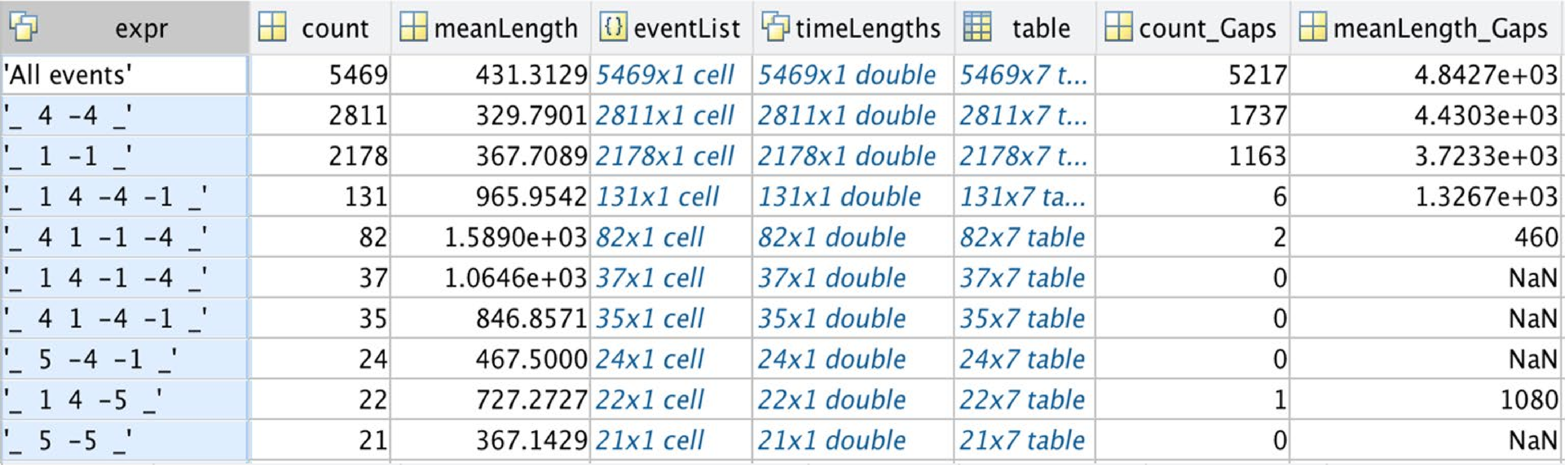
We process the DATA using KERA (Kinetic Event Resolving Algorithm) to give us an Idealized signal.

Search for Patterns

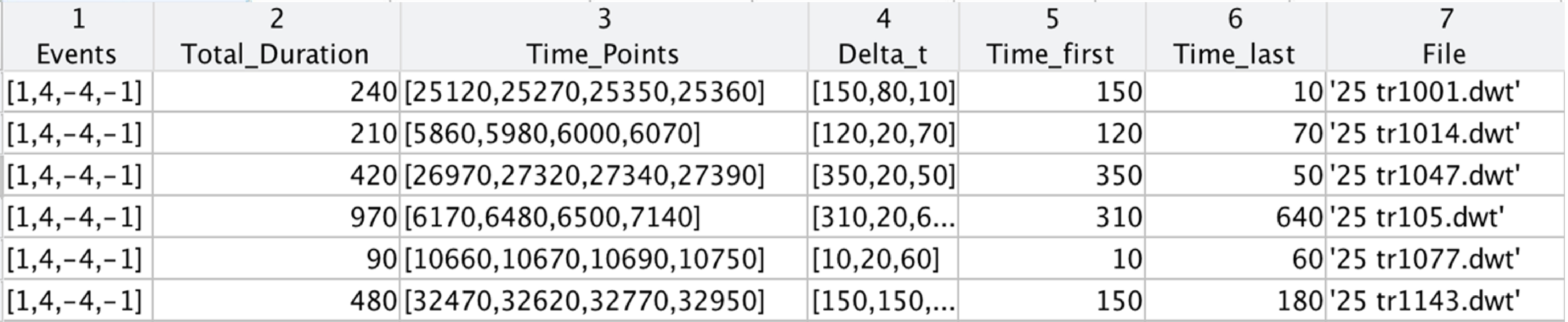
We record the occurrence of certain patterns or events.



The DATA



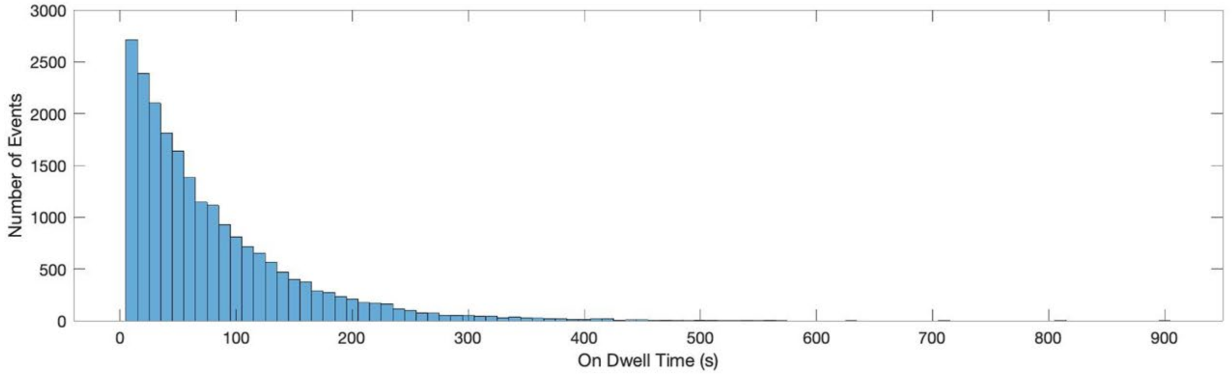
expr	count	meanLength	eventList	timeLengths	table	count_Gaps	meanLength_Gaps
'All events'	5469	431.3129	*5469x1 cell*	*5469x1 double*	*5469x7 t...*	5217	4.8427e+03
'_ 4 -4 _'	2811	329.7901	*2811x1 cell*	*2811x1 double*	*2811x7 t...*	1737	4.4303e+03
'_ 1 -1 _'	2178	367.7089	*2178x1 cell*	*2178x1 double*	*2178x7 t...*	1163	3.7233e+03
'_ 1 4 -4 -1 _'	131	965.9542	*131x1 cell*	*131x1 double*	*131x7 ta...*	6	1.3267e+03
'_ 4 1 -1 -4 _'	82	1.5890e+03	*82x1 cell*	*82x1 double*	*82x7 table*	2	460
'_ 1 4 -1 -4 _'	37	1.0646e+03	*37x1 cell*	*37x1 double*	*37x7 table*	0	NaN
'_ 4 1 -4 -1 _'	35	846.8571	*35x1 cell*	*35x1 double*	*35x7 table*	0	NaN
'_ 5 -4 -1 _'	24	467.5000	*24x1 cell*	*24x1 double*	*24x7 table*	0	NaN
'_ 1 4 -5 _'	22	727.2727	*22x1 cell*	*22x1 double*	*22x7 table*	1	1080
'_ 5 -5 _'	21	367.1429	*21x1 cell*	*21x1 double*	*21x7 table*	0	NaN

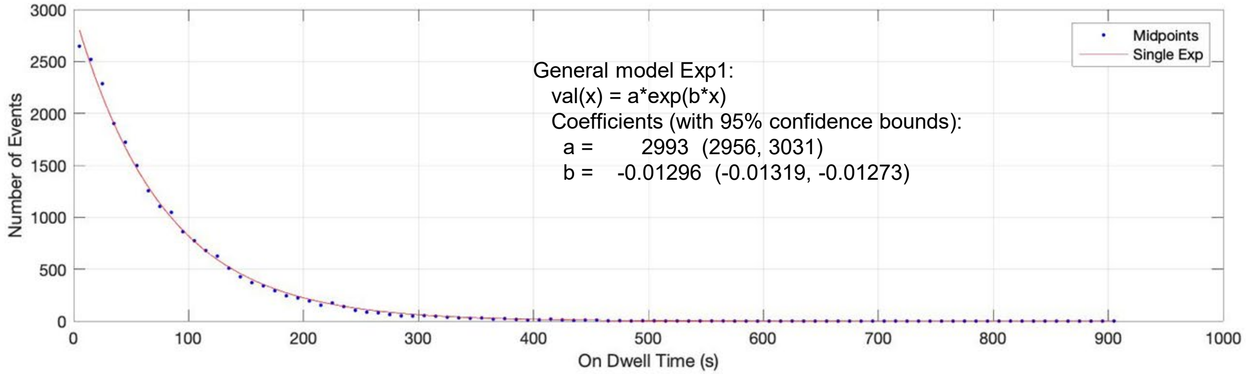


1 Events	2 Total_Duration	3 Time_Points	4 Delta_t	5 Time_first	6 Time_last	7 File
[1,4,-4,-1]	240	[25120,25270,25350,25360]	[150,80,10]	150	10	'25 tr1001.dwt'
[1,4,-4,-1]	210	[5860,5980,6000,6070]	[120,20,70]	120	70	'25 tr1014.dwt'
[1,4,-4,-1]	420	[26970,27320,27340,27390]	[350,20,50]	350	50	'25 tr1047.dwt'
[1,4,-4,-1]	970	[6170,6480,6500,7140]	[310,20,6...	310	640	'25 tr105.dwt'
[1,4,-4,-1]	90	[10660,10670,10690,10750]	[10,20,60]	10	60	'25 tr1077.dwt'
[1,4,-4,-1]	480	[32470,32620,32770,32950]	[150,150,...	150	180	'25 tr1143.dwt'

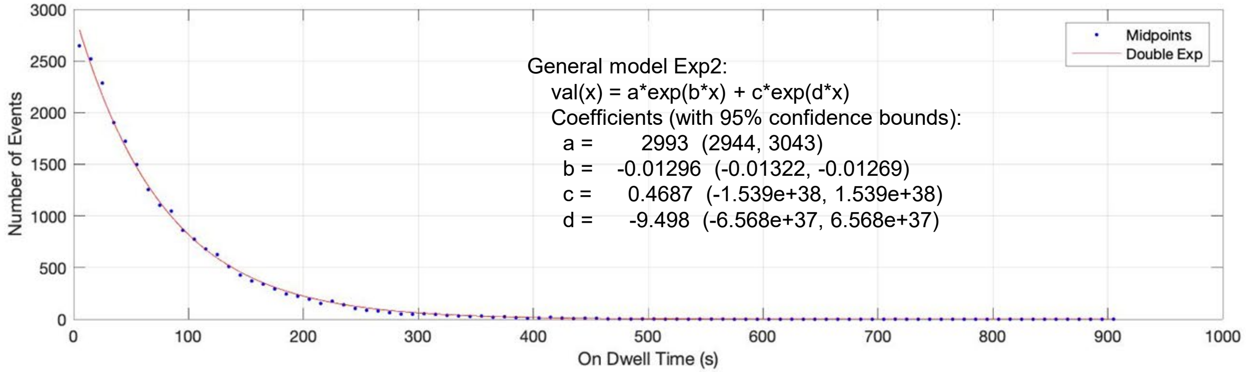
Methods & Data

- What model fits the DATA better?
    - Single or Double Exponential
- What are the constants that result in the best fit?
    - Rate constants determine probability of the event





```
General model Exp1:
   val(x) = a*exp(b*x)
   Coefficients (with 95% confidence bounds):
     a =        2993  (2956, 3031)
     b =    -0.01296  (-0.01319, -0.01273)
```



```
General model Exp2:
   val(x) = a*exp(b*x) + c*exp(d*x)
   Coefficients (with 95% confidence bounds):
     a =        2993  (2944, 3043)
     b =    -0.01296  (-0.01322, -0.01269)
     c =      0.4687  (-1.539e+38, 1.539e+38)
     d =      -9.498  (-6.568e+37, 6.568e+37)
```

Any Questions?