

# Simulating Spectral Kurtosis Mitigation Against Realistic RFI Signals

---

Evan Smith<sup>1,2</sup>

ECMWF RFI 2022 Conference

||

18 Feb 2022

R. Lynch<sup>3</sup>, D.J. Pisano<sup>1,2,4</sup>

<sup>1</sup>West Virginia University; <sup>2</sup>Center for Gravitational Waves and Cosmology (WVU); <sup>3</sup>Green Bank Observatory; <sup>4</sup>University of Cape Town

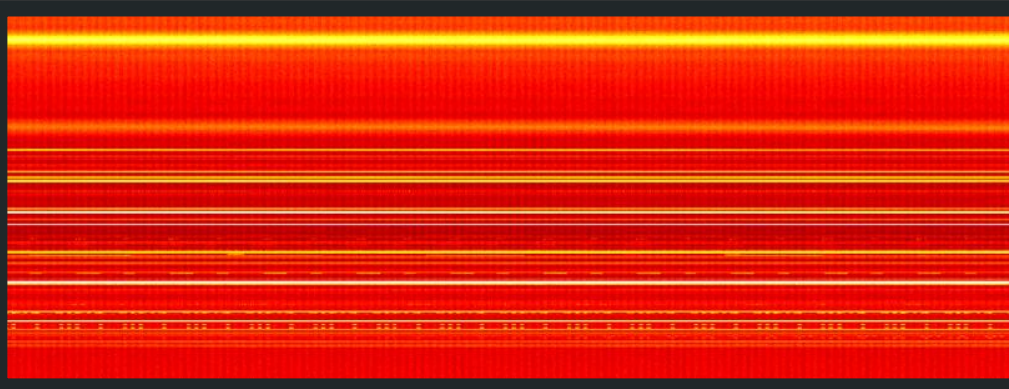
# Aims

---

- Develop Python toolkit of realistic RFI signal generators
- Apply Spectral Kurtosis mitigation
- Analyze when it does/doesn't work well and contextualize how to best use  $\widehat{SK}$



Green Bank Telescope baseband data



# Spectral Kurtosis

---

- 1) Evaluate the sums of the power values and squared power values:

$$S_1(k) = \sum_{i=1}^M (P_i(k)) \quad S_2(k) = \sum_{i=1}^M (P_i(k))^2$$

- 1) Calculate estimator

$$\widehat{SK} = \frac{Md + 1}{M - 1} \left( \frac{MS_2}{S_1^2} - 1 \right)$$

- $\widehat{SK}$  values deviating outside  $1 \pm$  calculated thresholds are considered RFI
- Thresholds are calculated by numerically solving PDF at  $3\sigma$  level

# Multi-scale Spectral Kurtosis

---

- Single-scale (SS) SK is weak to certain signal characteristics:
  - 50% duty cycle
  - High data rate (much faster than  $M$  spectra)
- $S_1$  and  $S_2$  values can be further binned into small rectangles
  - 2x1, 2x2, 4x2, 3x5, etc.
- Has the effect of 'blending' duty cycles of signals in nearby time-frequency pixels and push SK values of RFI-affected pixels away from unity
- Can apply MS-SK on top of conventional SS-SK

# Simulating RFI

---

- Python script written to simulate an RFI signal embedded in noise and perform SK mitigation
- RFI characteristics:
  - Modulation type (BPSK, QPSK, ASK, BFSK, etc.)
  - Amplitude is ramped up from 0 to 1 over the course of the test observation
  - Data rate (kilo-symbols per second)
  - Duty cycle and period (0-100%, over some amount of ms)
  - SK  $M$
  - MS-SK bin shape
- Comparison RFI mask
  - Convert Signal-only data to dB using noise as reference
  - Flag as RFI when signal is greater than -10 dB
  - Record true positive and false positive rates

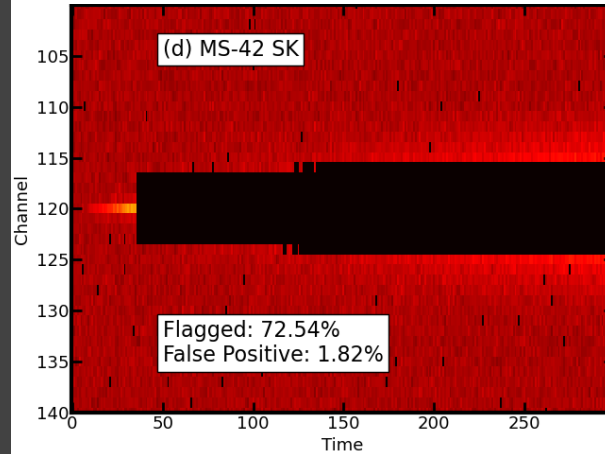
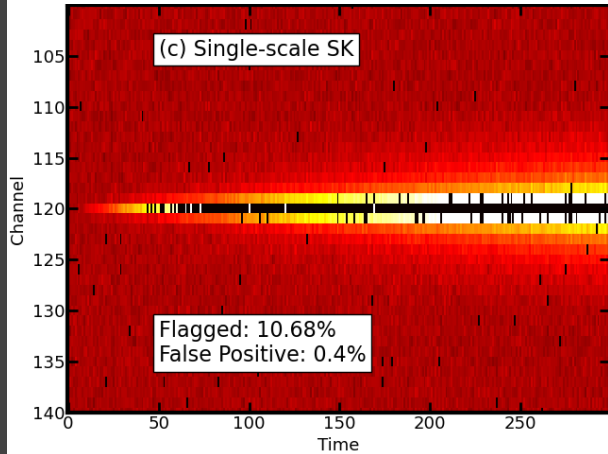
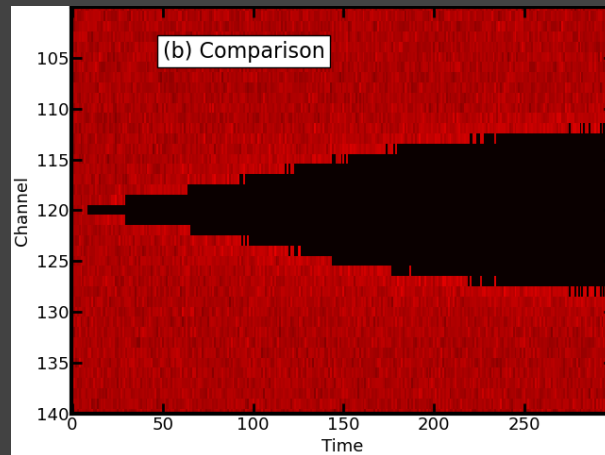
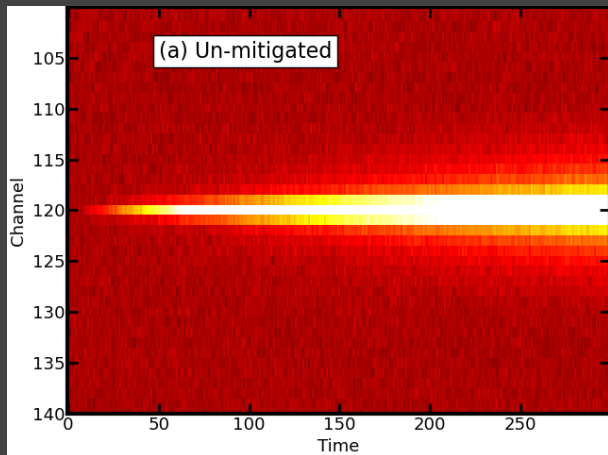
# Simulating RFI

```
178 def bpsk(nbits,symbol_rate,wincut,fc,Ebit,fs=800e6):
179     #binary phase shift keyed
180     #nbits :    number of bits to generate
181     #symbol_rate : how many kilobits per second (binary=1 bit per symbol)
182     #wincut :    cutoff frequency for scipy.signal.firwin
183     #fc :        baseband carrier frequency (Hz)
184     #Ebit :      Energy per bit (not used, need to implement signal strength in a way that makes sense)
185     #fs :        sampling rate (samples per second)
186     #=====
187
188     #number of samples per bit
189     tbit = int( fs / (symbol_rate*1e3) )
190
191     #index array
192     x = np.arange(nbits*tbit)
193
194     #generate random bits and extend into symbols
195     bit_seq = np.random.randint(0,2,size=(int(len(x)/tbit)+1,))
196     bit_seq = (2*bit_seq)-1
197     pulse = np.ones(tbit)
198     sym_seq = np.kron(bit_seq,pulse)[:len(x)]
199
200     #apply sinc FIR with a width of 20% bit length
201     fir_sz = int(0.2*tbit)
202     sinc = scipy.signal.firwin(fir_sz, cutoff=wincut/fir_sz, window="rectangular")
203     sym_seq = scipy.signal.convolve(sym_seq,sinc,mode='same',method='fft')
204
205     #apply carrier signal
206     ts = 1/fs
207     e_vec = np.exp(2.j*np.pi*fc*x*ts)
208
209     #binary phase shift means 180deg flip - a simple +/-1 multiply
210     sig = sym_seq * e_vec
211
212     #return time samples as well as bit data stream (for verification/sanity)
213     return sig,sym_seq,bit_seq
```

- Generate random bit sequence
- Apply FIR filter to symbol sequence to restrict sidelobe emission
- Use bits to inform carrier signal generation

Modeled after code shared by Dr. Chad Spooner at <https://cyclostationary.blog/>

# Simulating RFI

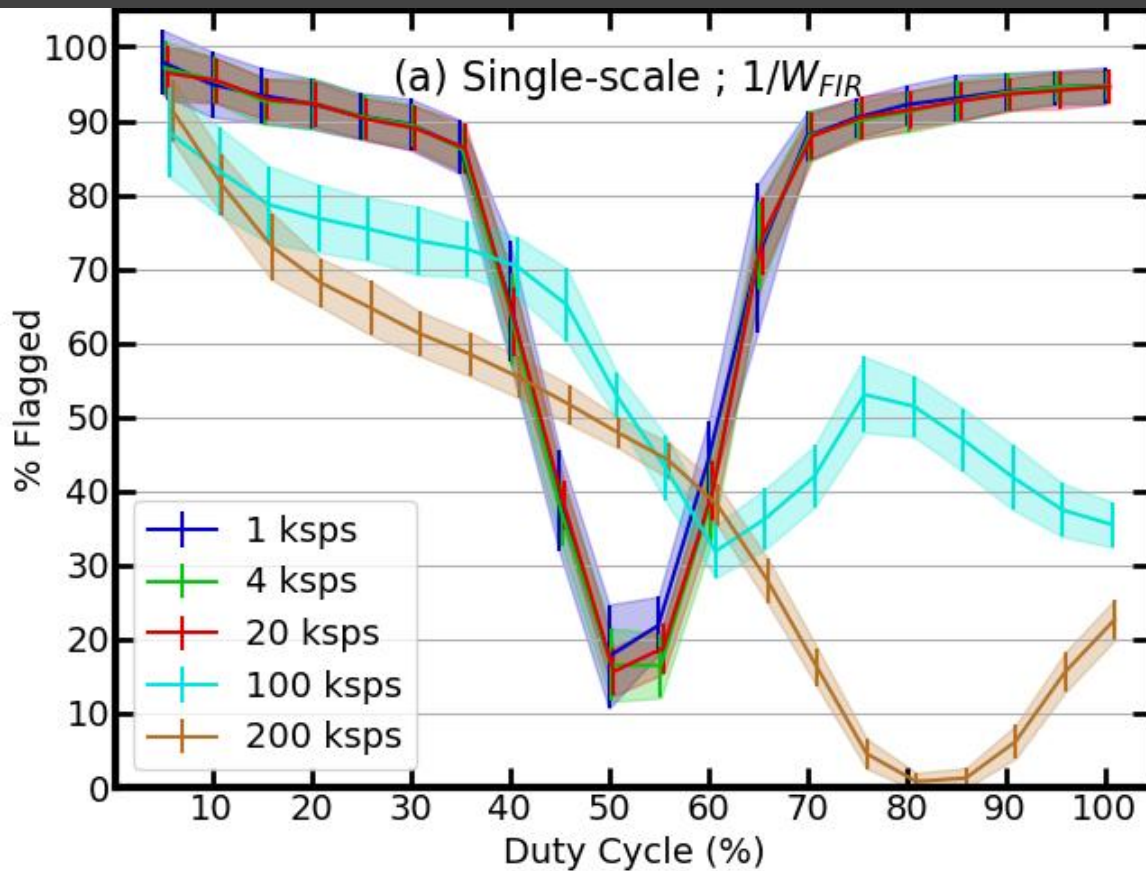


# Simulating RFI

- Pick two parameter ranges and simulate every combination 100x
- To match similar pulsar data taken with GBT:
  - 50 MHz sampling rate / 256 channels
  - 24-tap PFB channelization with hamming window smoothing
- Each simulation is  $60 \widehat{SK}$  bins long
- Record true/false positive rates and calculate mean/deviation

Data Rate (kps)	Duty Cycle (%)	\SKs $M$ (# spectra)	MS shape	Channel Center	FIR cutoff
1	5	128	★ Single-scale	★ 120	★ $1/W_{FIR}$
4	10	256	MS-12	120.25	$4/W_{FIR}$
★ 20	15	★ 512	MS-21	120.5	
100	...	1024	MS-22		
200		2048	MS-24		
	...	4096	MS-42		
	95				
	★ 100				

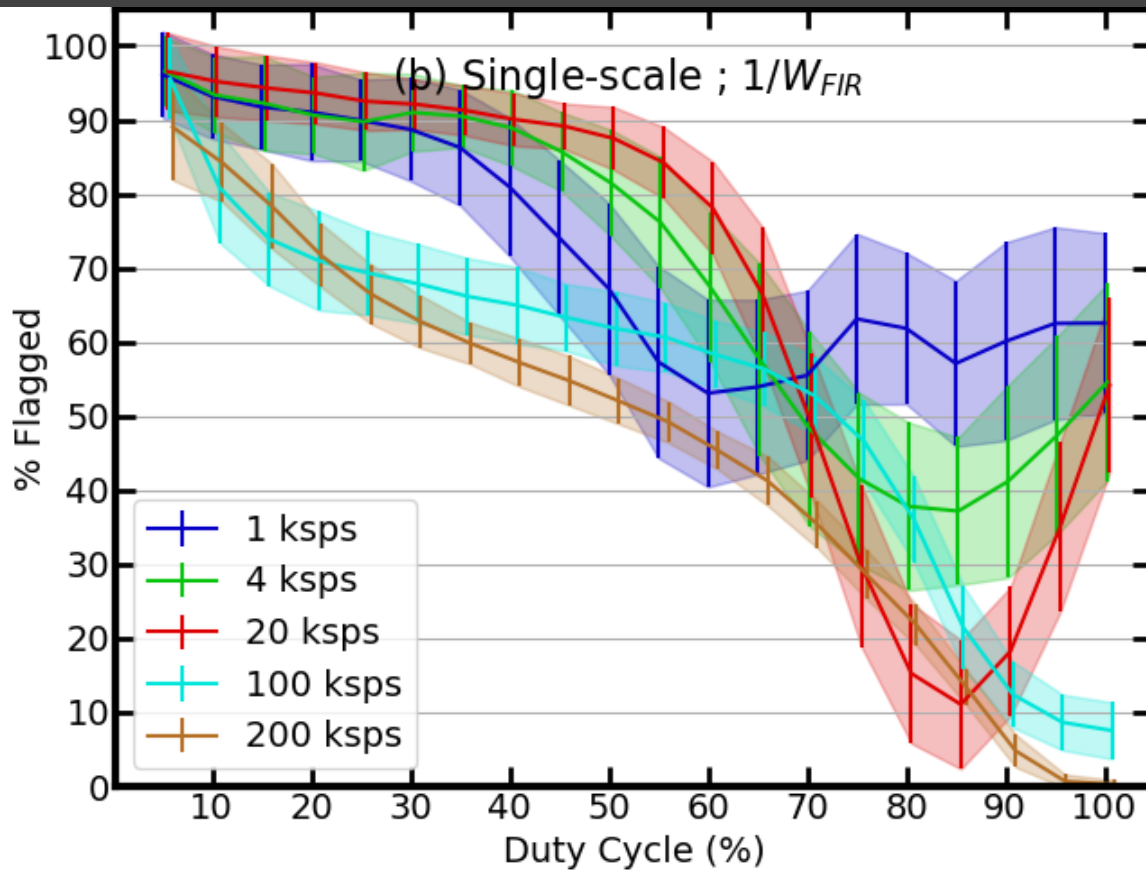
# Results



## Comparing duty cycle to data rate: BPSK

- Expected weakness to 50% duty cycle from low data rate signals
- Higher data rate signals are likely not represented accurately
  - Due to PFB/sampling rate
  - Sidelobes are not well detected

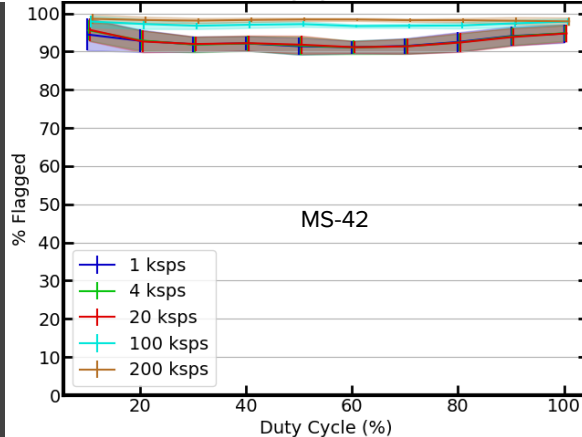
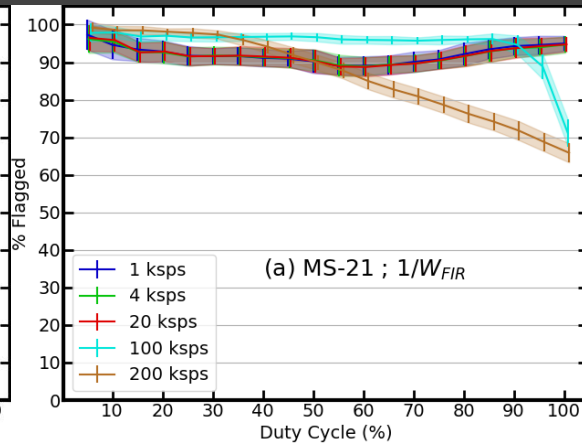
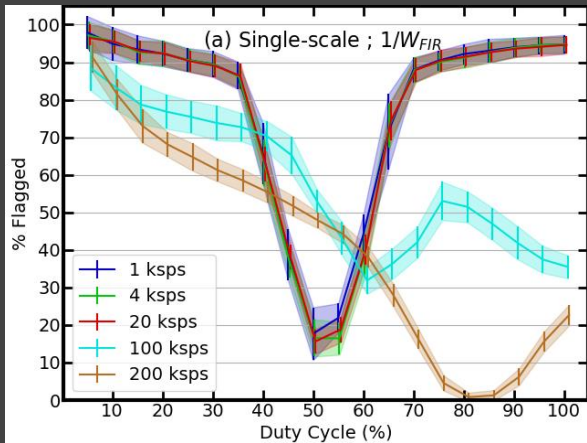
# Results



Comparing duty cycle to data rate: ASK

- Expected weakness to 50% duty cycle from low data rate signals
  - Effective duty cycle is different from intrinsic because of power level differences between symbols

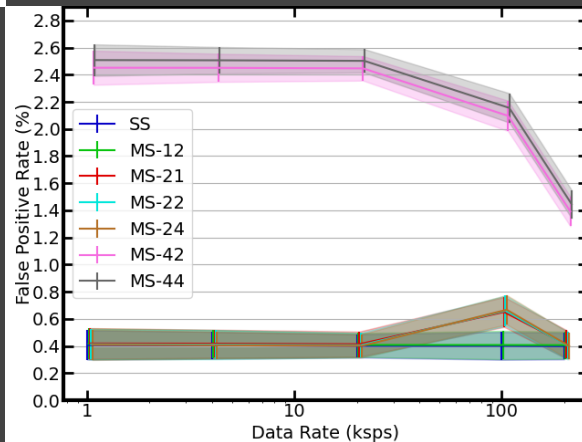
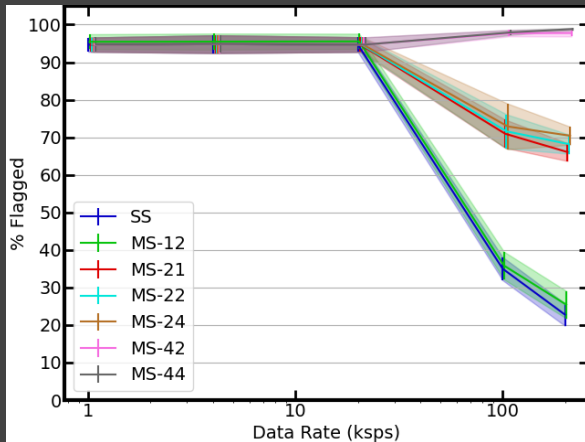
# Results



## Multi-scale SK: BPSK encoding

- Including at least one extra channel completely subverts the 50% duty cycle weakness
- Including more than one channel can flag >90% of the data, at the risk of overflagging
- Adding extra time bins does not affect flagging

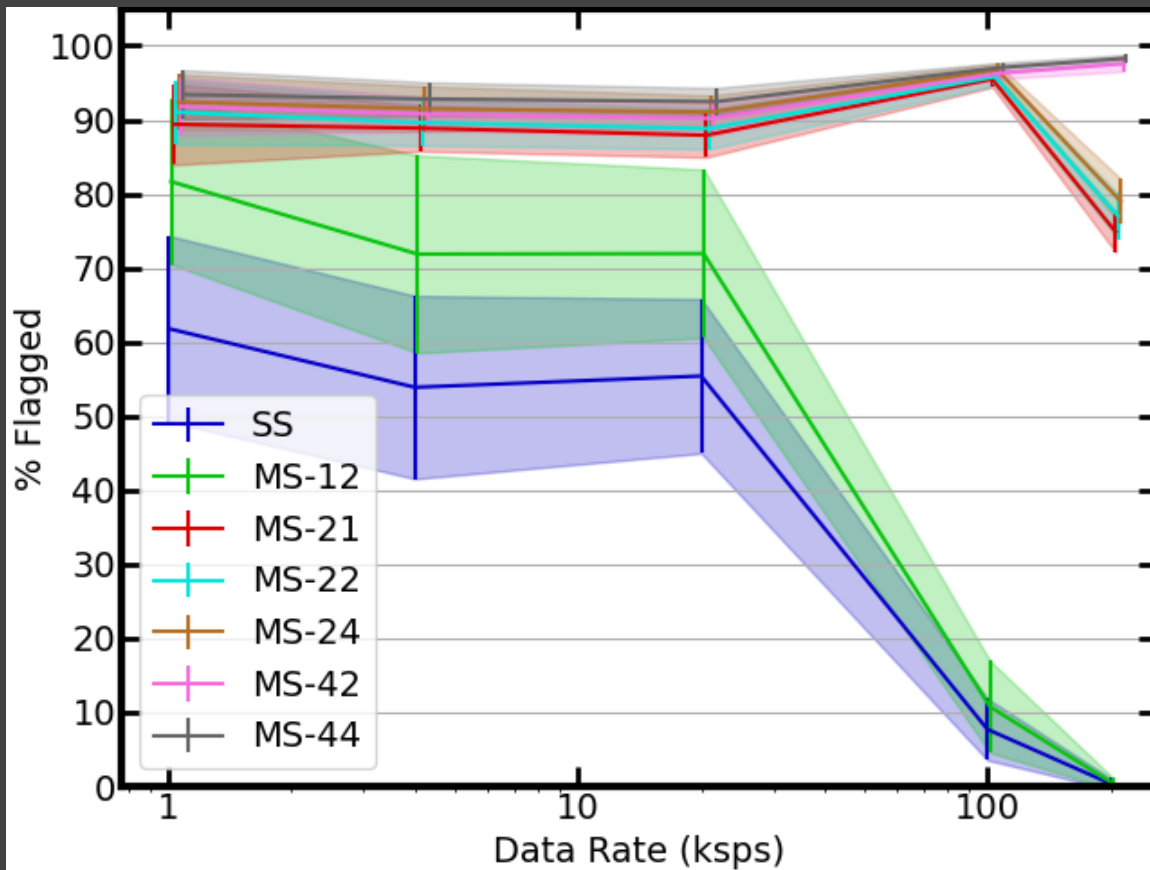
# Results



## Multi-scale SK: BPSK encoding

- Adding an extra time bin does not increase flagging effectiveness or FPR
- Adding a channel increases flagging, but overflags when the multi-scale channel width  $>$  signal width
- With  $M=512$ , theoretical lowest fpr is  $\sim 0.4\%$

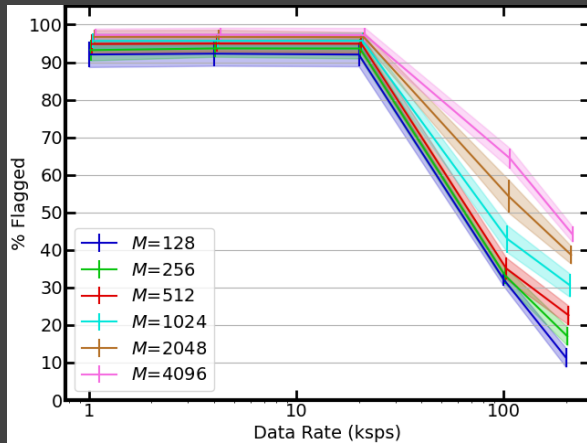
# Results



## Multi-scale SK: ASK/FSK

- In the case of ASK/FSK signals, including an extra time bin can raise the flagging
- Signal can change in power level from symbol to symbol
- Once an extra channel is added, the time bin dimension is less effective

# Results

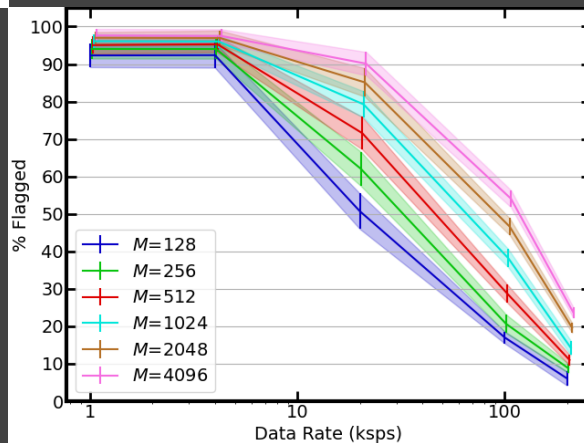


Cutoff:  $1 / W_{\text{FIR}}$

Comparing  $M$  to data rate

- Signals with no sidelobes are always flagged  $>90\%$
- As the data rate rises, higher  $M$  leads to higher flagging

Cutoff:  $4 / W_{\text{FIR}}$



# Summary

---

## Duty Cycle

---

- SK is weak to 50% *effective* duty cycles
- Can be subverted with multi-scale SK

## SK $M$

---

- Raising  $M$  leads to higher flagging
- Slightly increases false positive rate

## Modulation type

---

- BPSK and QPSK are flagged at similar rates
- ASK/FSK signals that change power level are detected differently from PSK

## Multi-scale SK

---

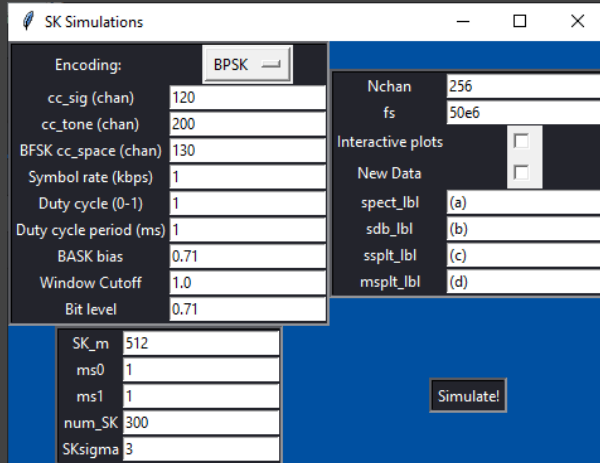
- Extra channels always increase flagging - can dramatically increase false positive rate if MS channel width is higher than signal width
- Extra time bin increases flagging only in the cases of amplitude-modulated signals

# Works in progress

## RFI Simulator GUI

[https://github.com/etsmit/RFI\\_Simulator](https://github.com/etsmit/RFI_Simulator)

1. Fill in the boxes
2. Generate RFI
3. Flag it and examine plots



The screenshot shows the SK Simulations GUI with the following parameters:

Encoding:	BPSK
cc_sig (chan)	120
cc_tone (chan)	200
BFSK cc_space (chan)	130
Symbol rate (kbps)	1
Duty cycle (0-1)	1
Duty cycle period (ms)	1
BASK bias	0.71
Window Cutoff	1.0
Bit level	0.71
Nchan	256
fs	50e6
Interactive plots	<input type="checkbox"/>
New Data	<input type="checkbox"/>
spect_lbl	(a)
sdb_lbl	(b)
ssplt_lbl	(c)
msplt_lbl	(d)
SK_m	512
ms0	1
ms1	1
num_SK	300
SKsigma	3

Simulate!

## More encoding techniques

- Current RFI generators are restricted to simpler modulation types
- There is a *expansive* list of digital signal encoding we would like to model in our simulations

### Digital modulation

ASK · APSK · CPM · FSK · MFSK · MSK · OOK · PPM ·  
PSK · QAM · SC-FDE · TCM · WDM

### Hierarchical modulation

QAM · WDM

### Spread spectrum

CSS · DSSS · FHSS · THSS

# Self-promotion

---



- Planning to finish Ph.D. work within the year and looking for jobs!

<https://etsmit.github.io/>

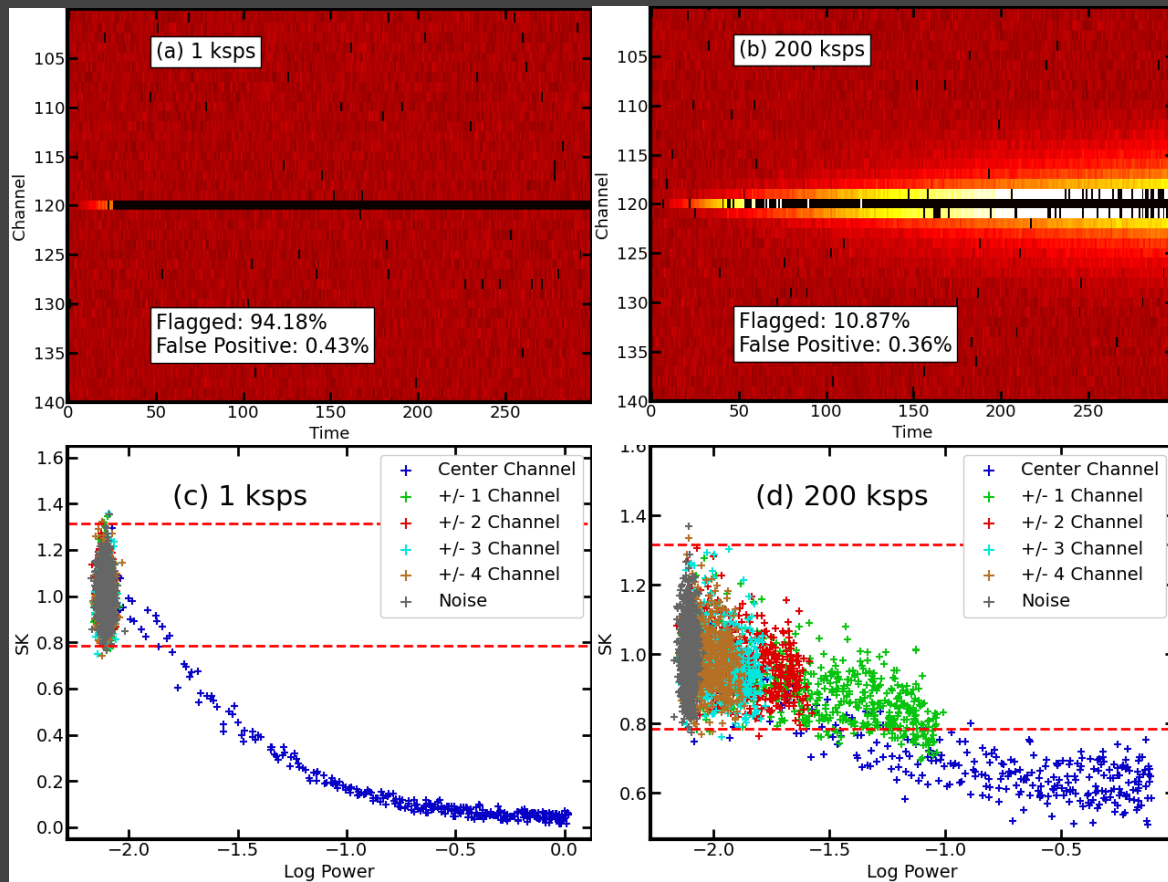
etsmit12@gmail.com



- Interested in:
  - Spatial nulling
  - Digital Signal Processing Development
  - Phased Array Feeds

(bonus slides)

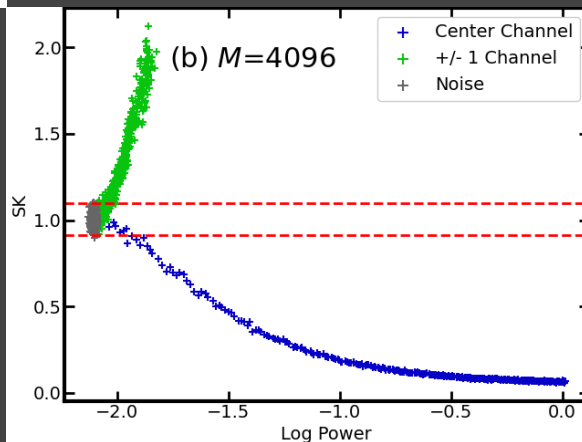
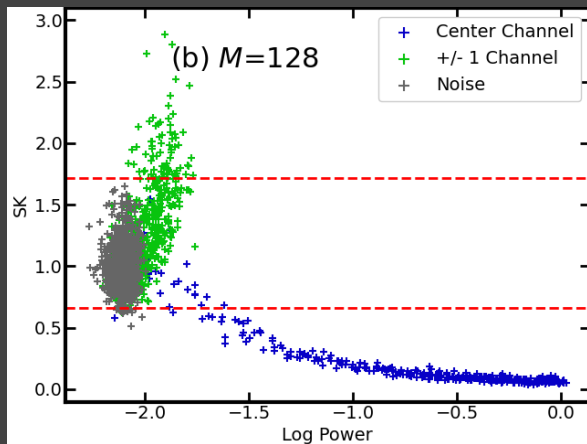
# Results



## Comparing $M$ to data rate

- As data rate rises, sidelobes grow in strength but do not affect SK enough to be flagged

# Results

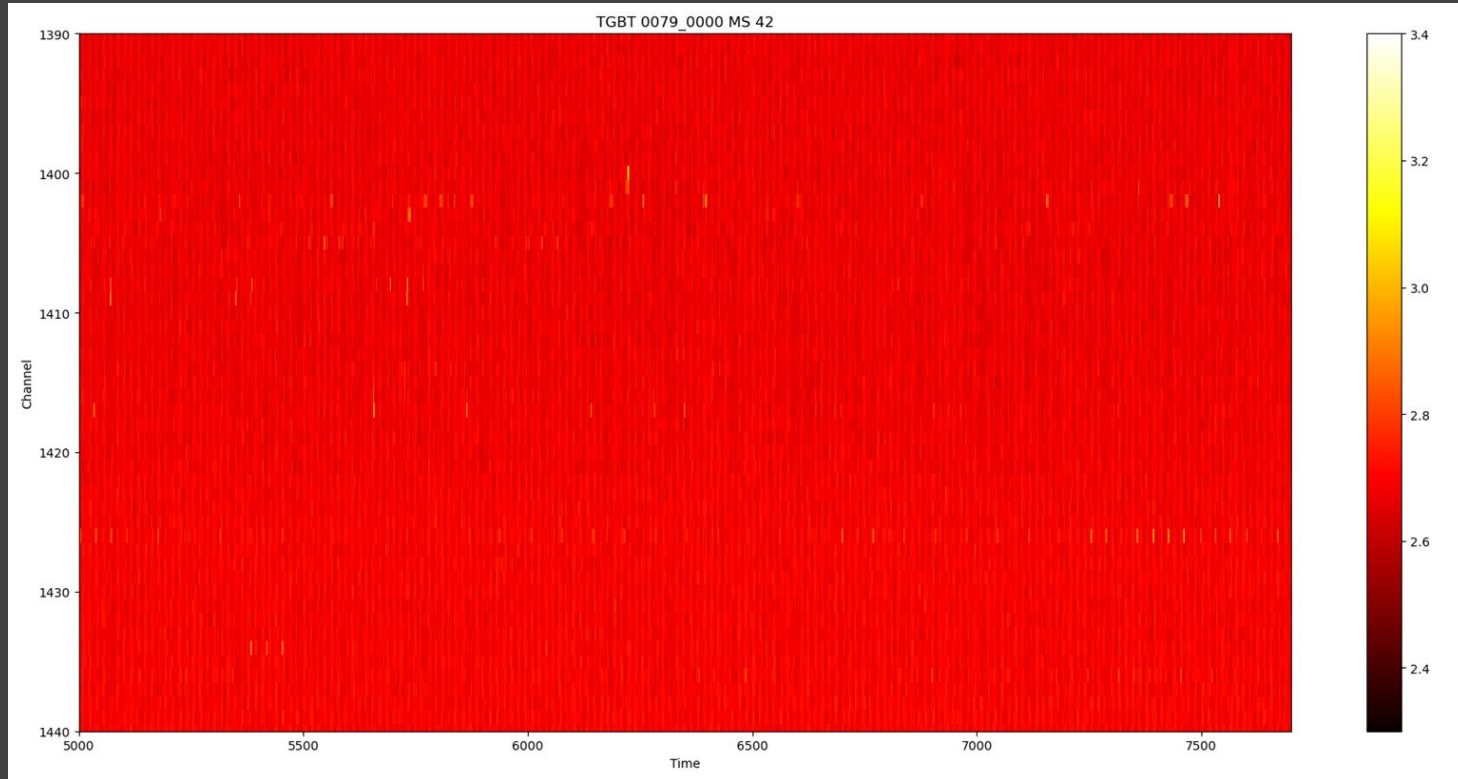


## Comparing $M$ to data rate

- As data rate rises, sidelobes grow in strength but do not affect SK enough to be flagged
- Raising  $M$  dramatically distorts RFI-affected SK values away from 1

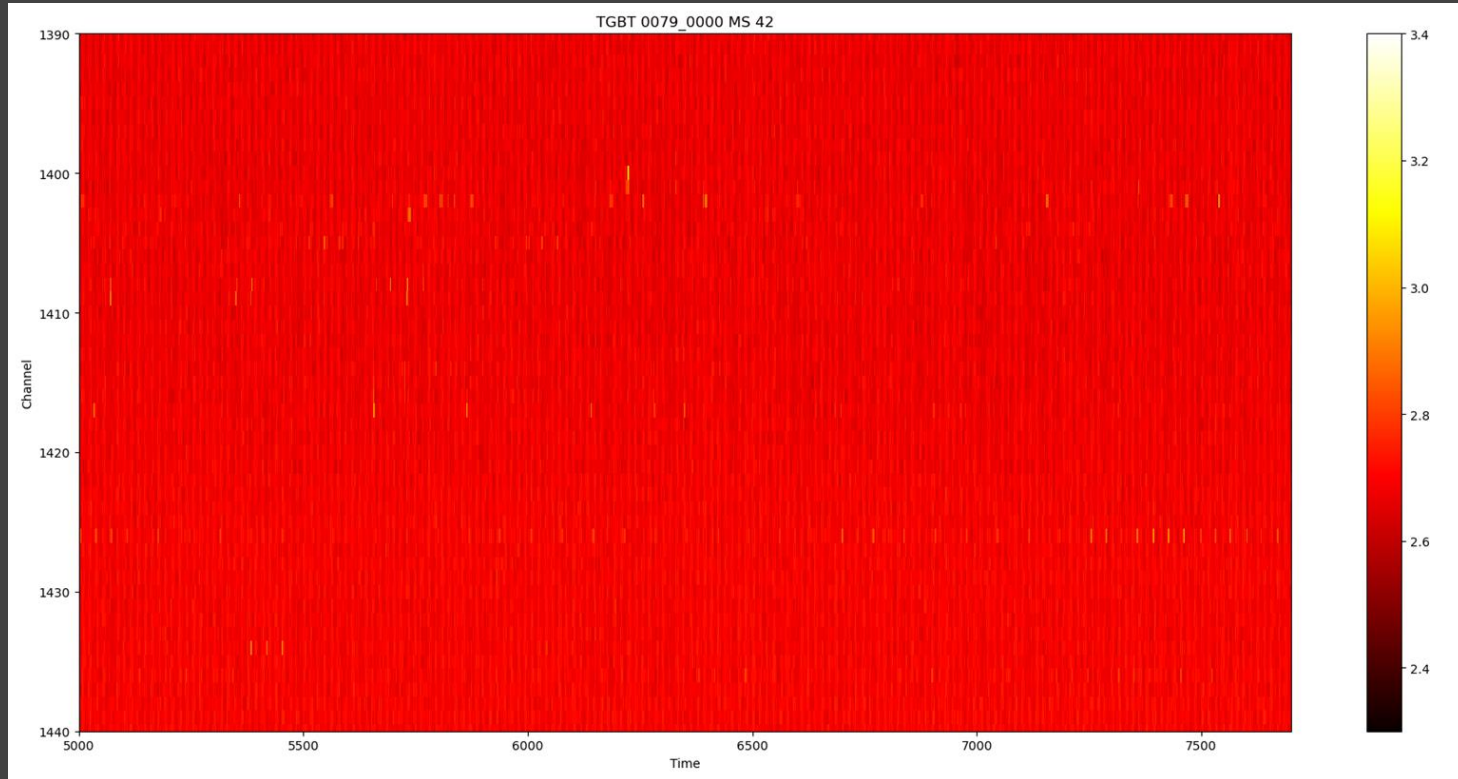
# Results - Iridium

---

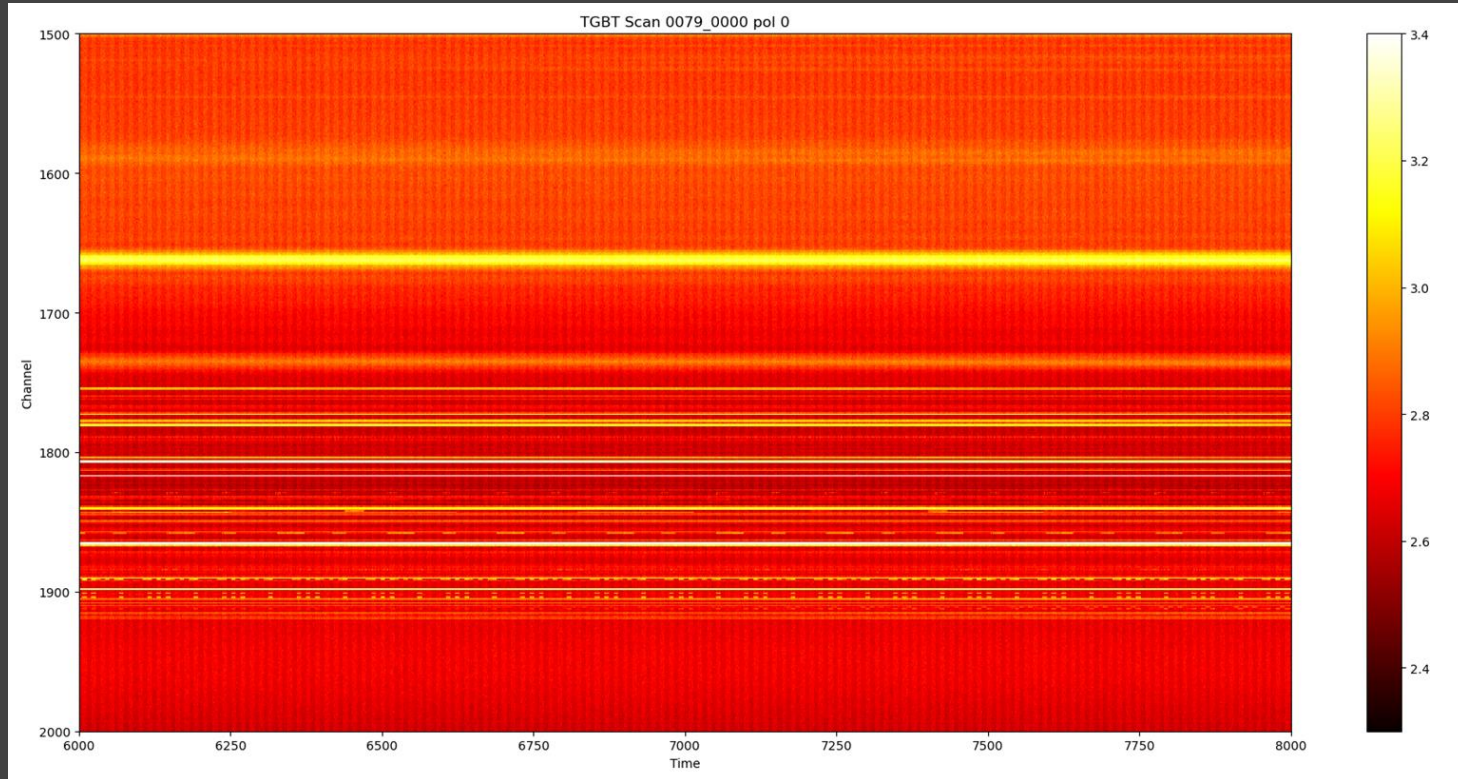


# Results - Iridium (MS-42)

---



# Results - Met. Sat.



# Results - Met. Sat. (MS-42)

