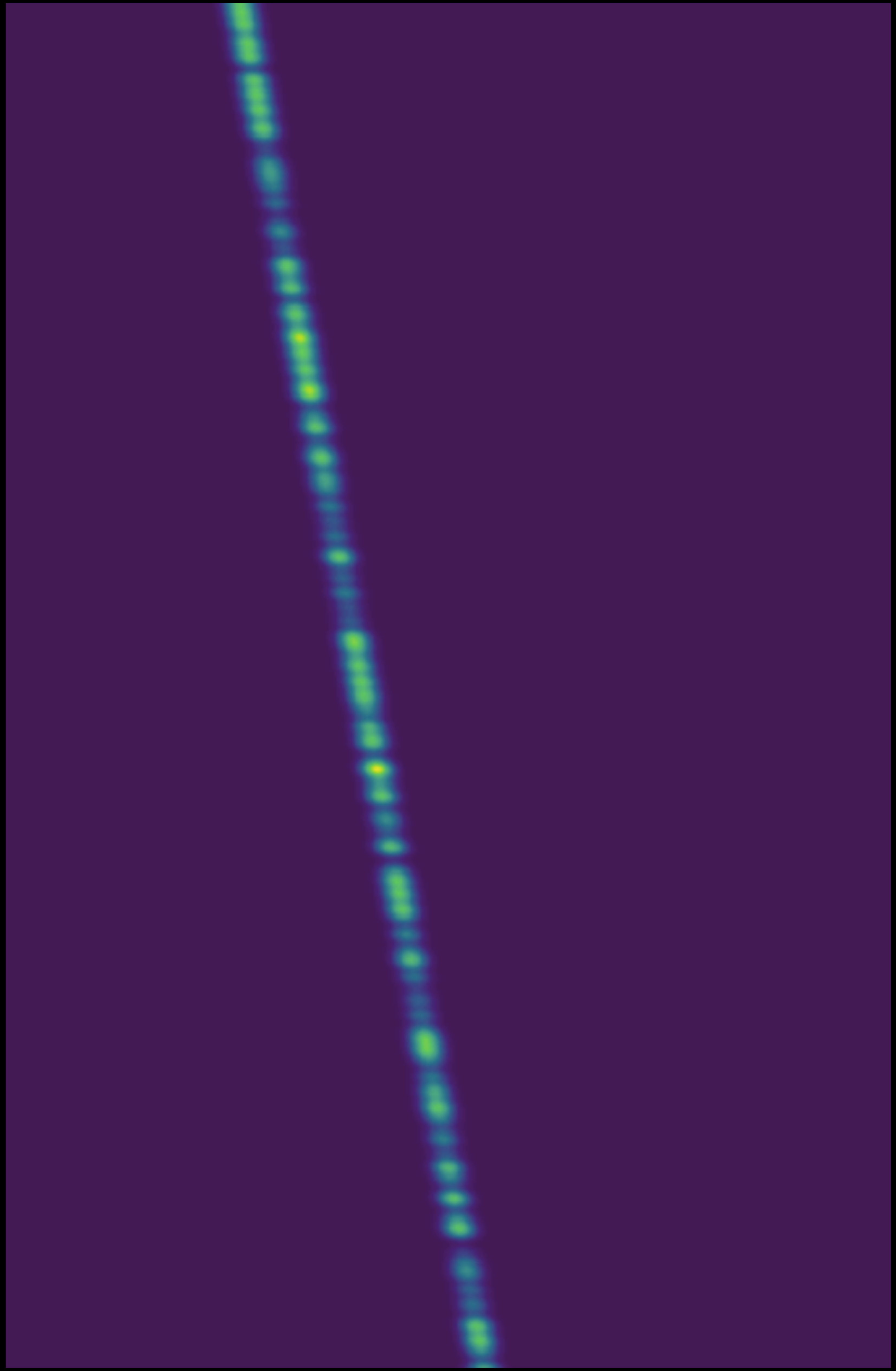


BREAKTHROUGH
LISTEN

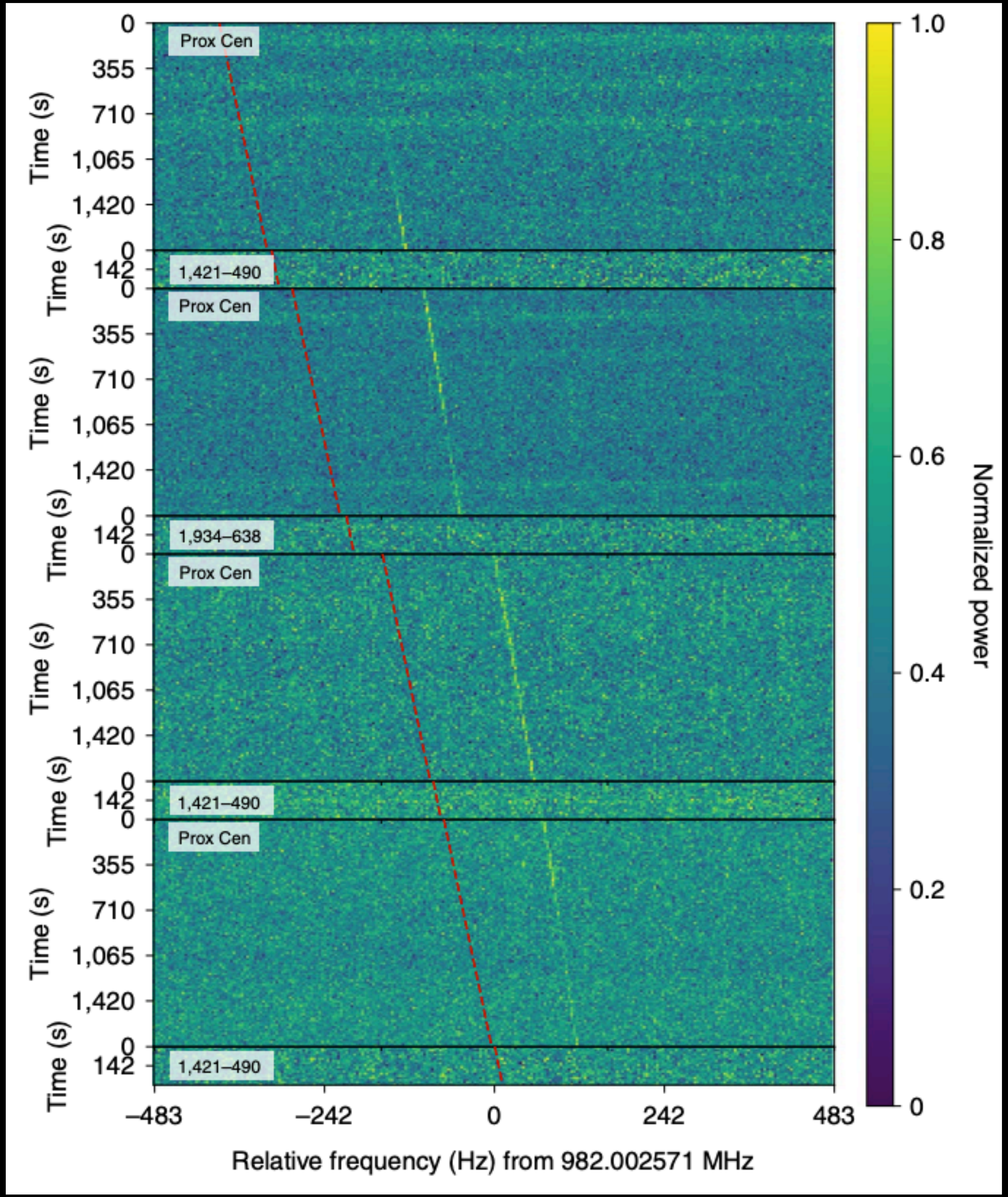
On Detecting Interstellar Scintillation in Narrowband Radio SETI

Bryan Brzycki
University of California Berkeley
AAS, January 12, 2023

BERKELEY SETI RESEARCH CENTER

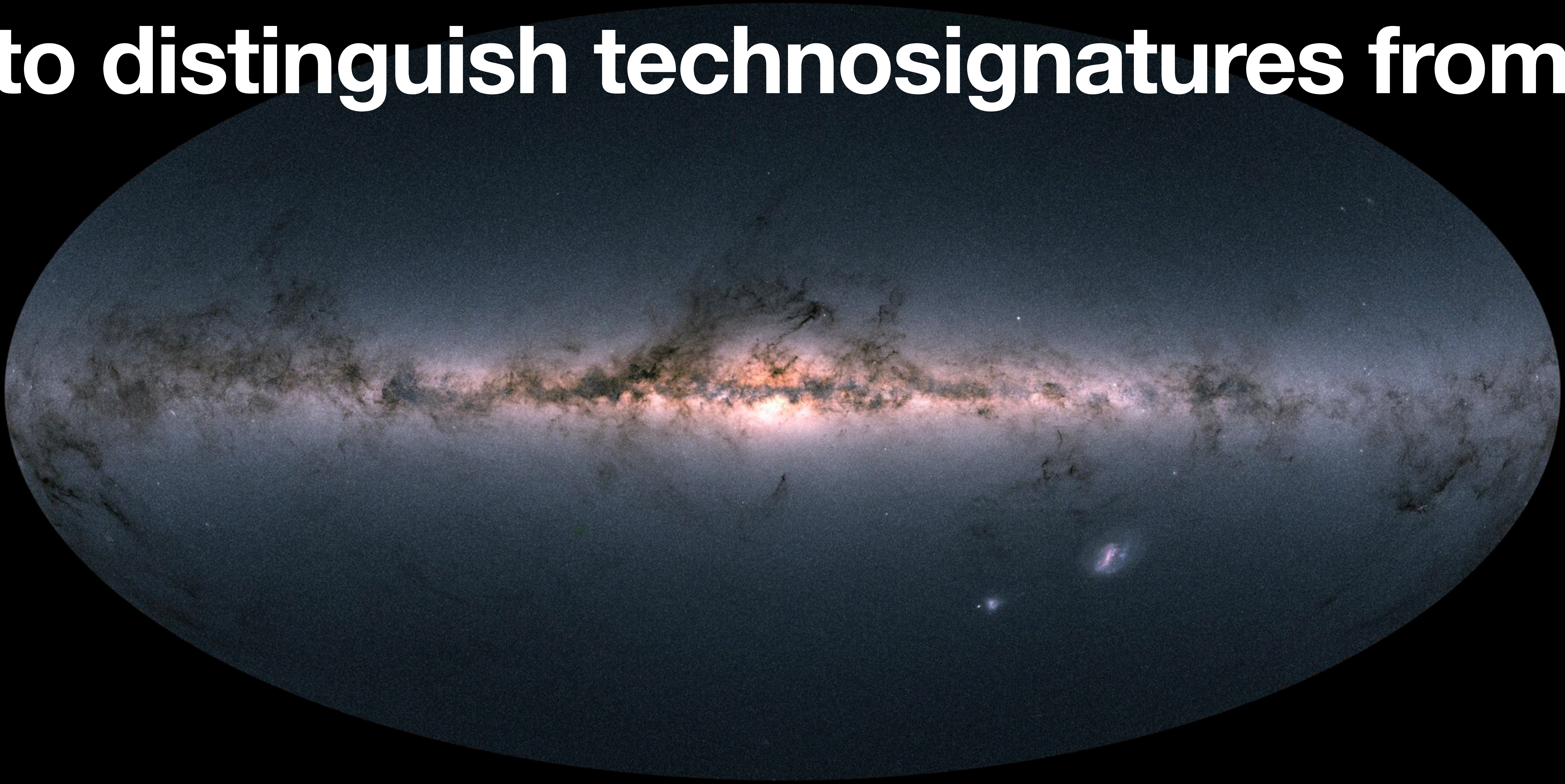
Standard filters used for radio technosignature candidates

- Narrowband (vs. astrophysical sources)
- Non-zero drift rate (vs. RFI)
- Sky localization (vs. RFI)



Smith et al. 2021

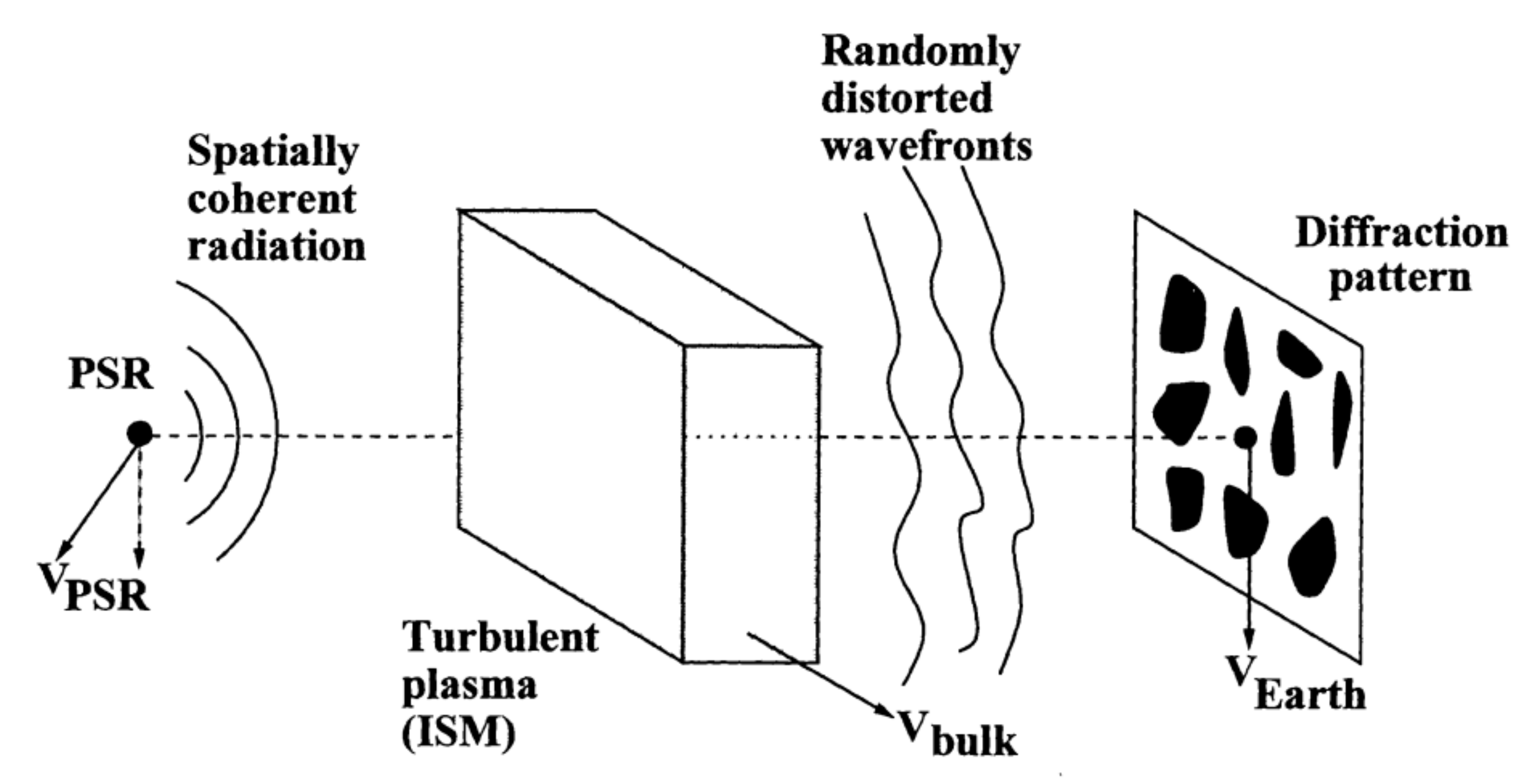
Can we use astrophysical phenomena as a filter to distinguish technosignatures from RFI?



ESA

Diffractive scintillation in the ISM

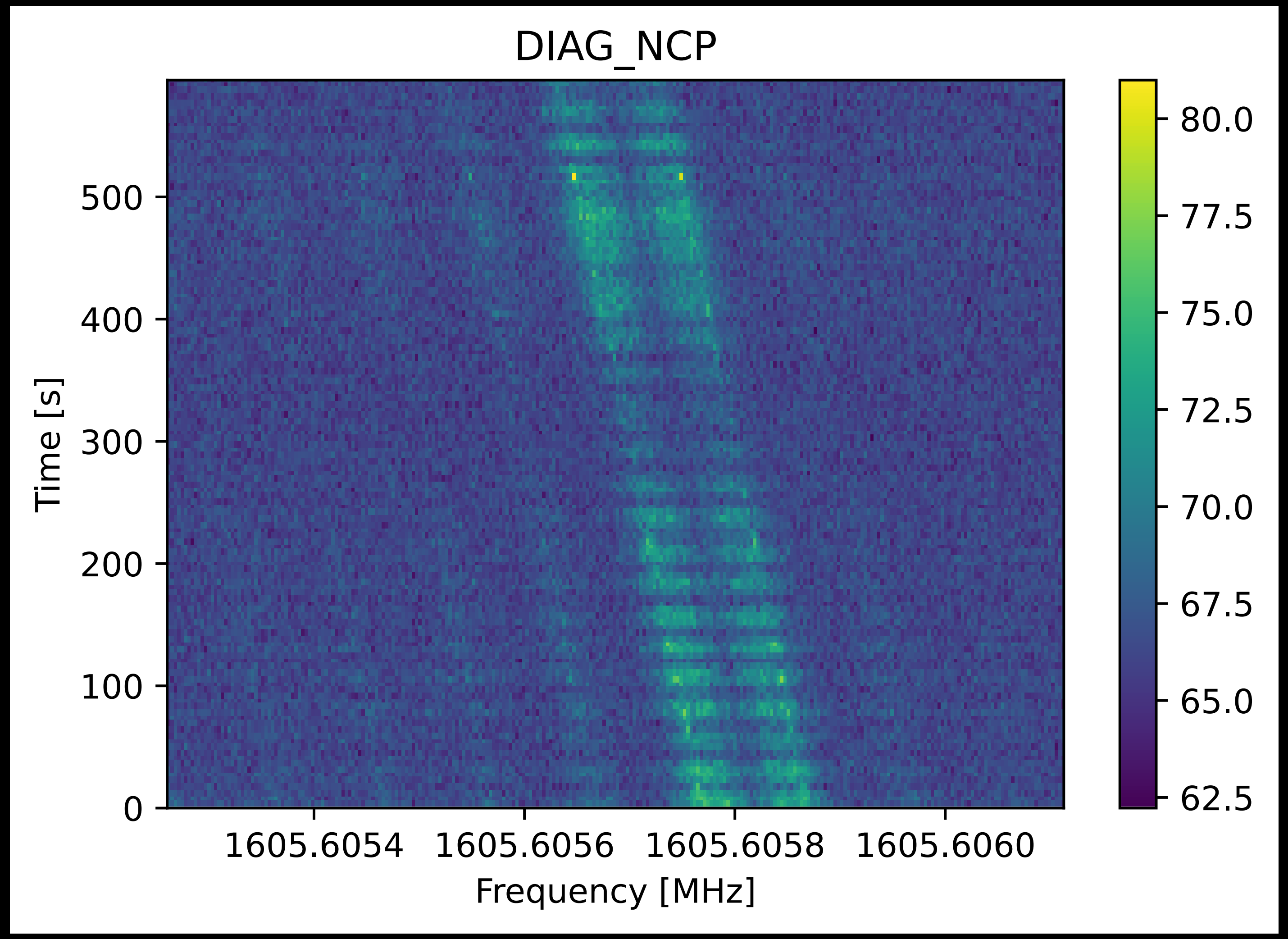
- Electron density fluctuations in ionized plasma give rise to phase fluctuations
- Interference pattern at observer plane with characteristic spatial and spectral scales
- Can lead to 100% intensity modulation on characteristic temporal scales Δt_d, especially towards the galactic center

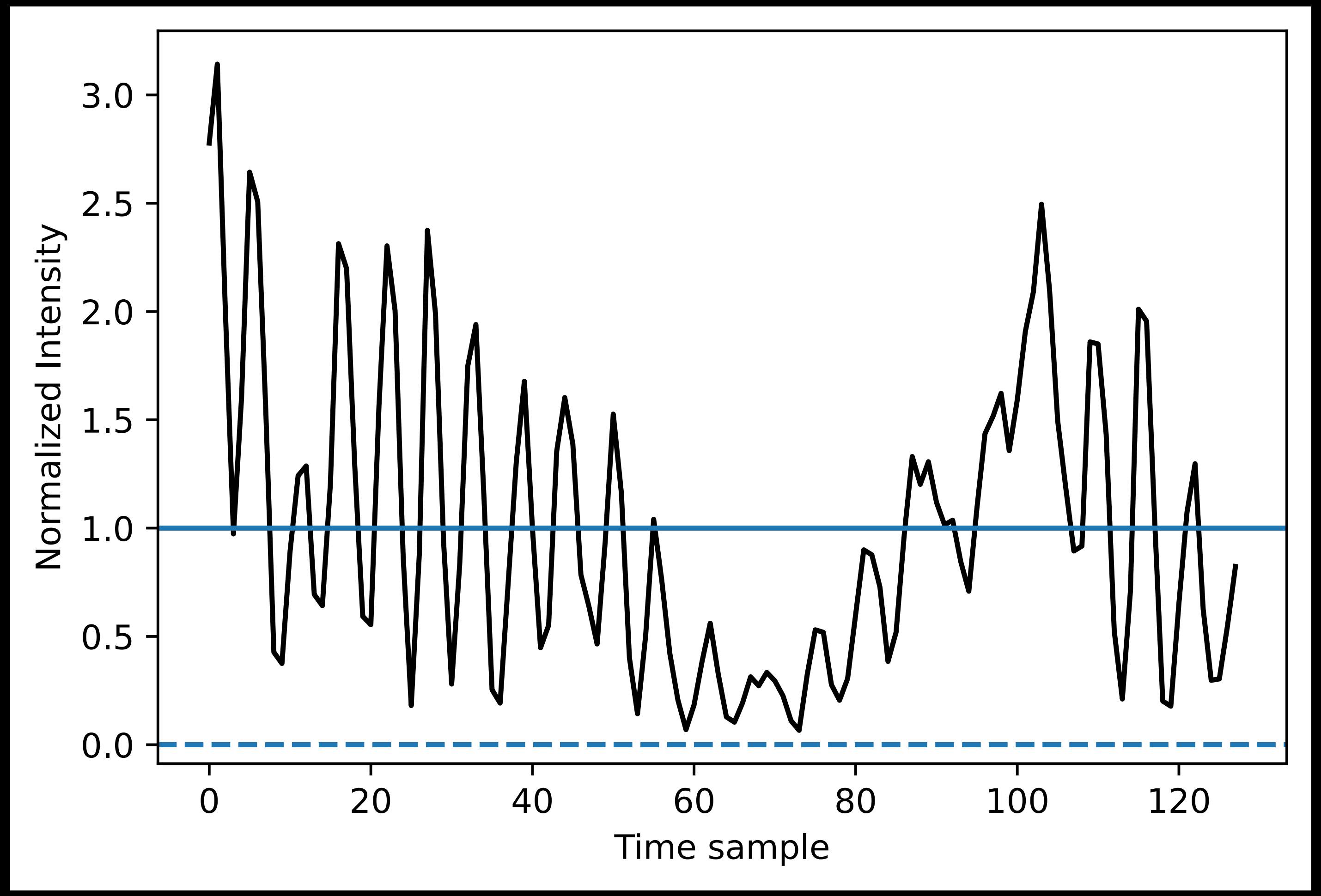


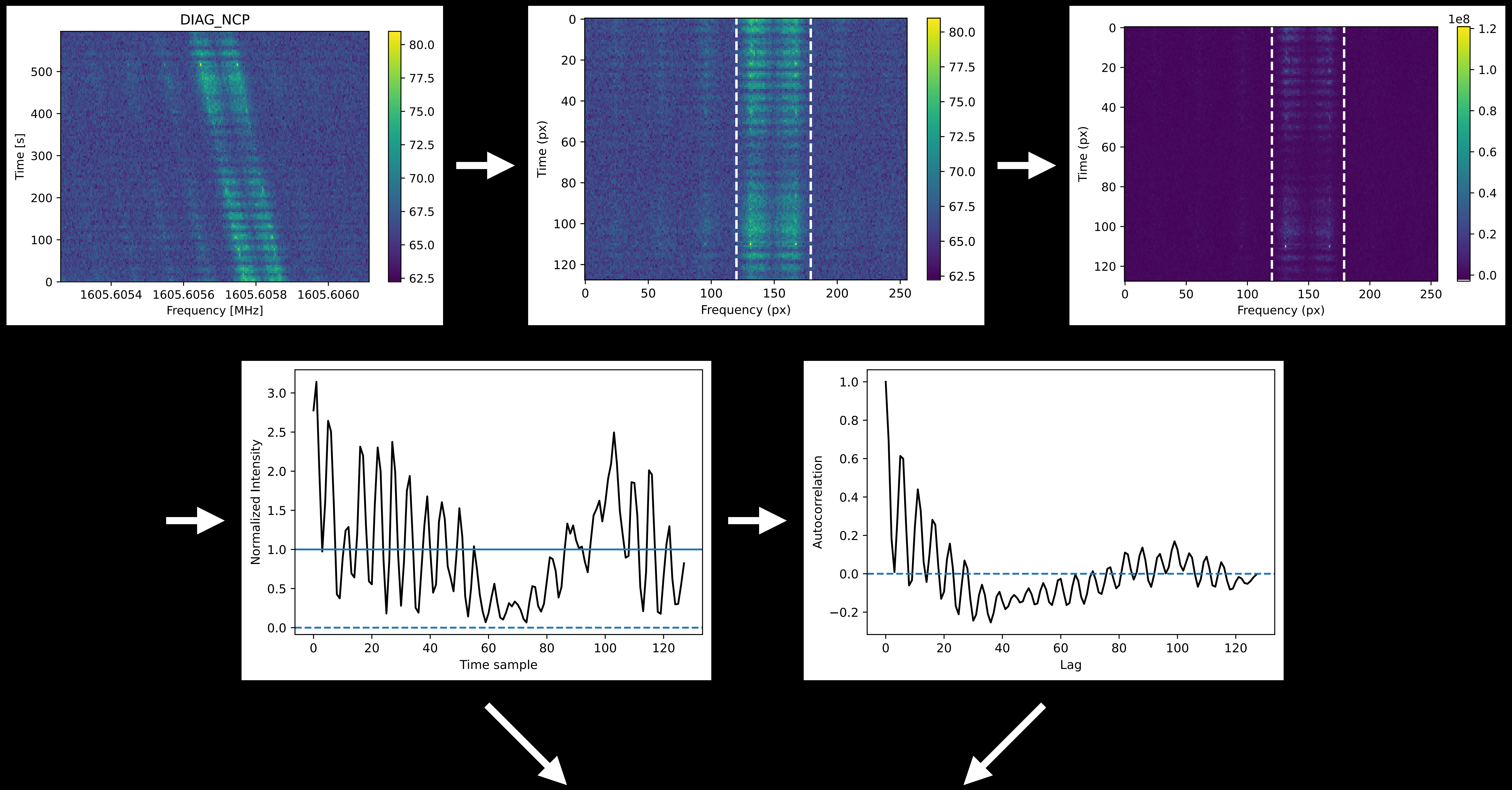
Cordes 2002

So how might we detect scintillation?

- Estimate intensity time series from detected signals for analysis
- Since scintillation is stochastic, identify measurable statistics for asymptotic behavior
    - E.g. standard deviation, Kolmogorov-Smirnoff statistic, fit to autocorrelation function
- Would existing RFI modulation confound real scintillation?
    - Develop process for creating synthetic scintillated data
    - Compare statistical distributions of detected signals with those of synthetic scintillated signals





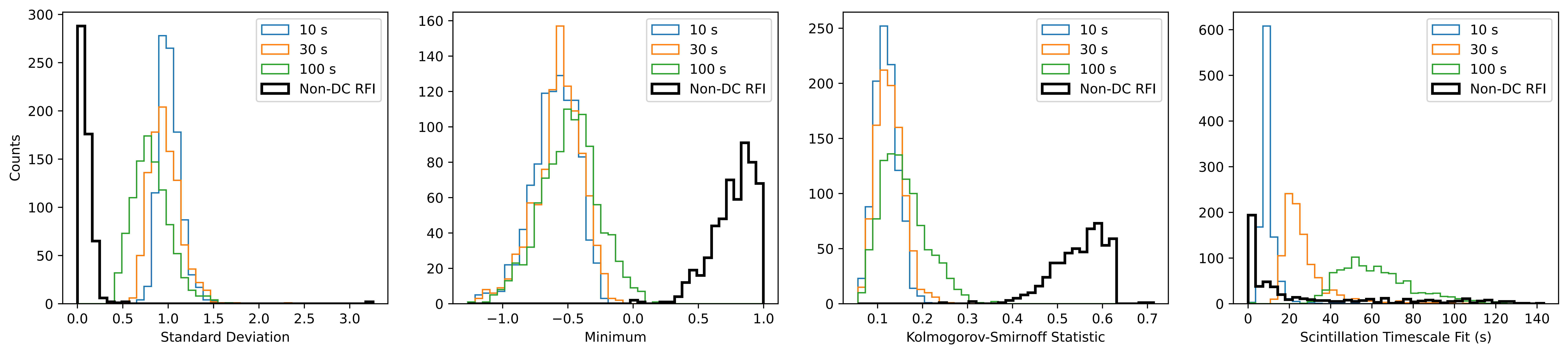


Diagnostic statistics

GBT RFI vs. synthetic scintillated signals

C band

S/N > 25



Standard Deviation

Minimum

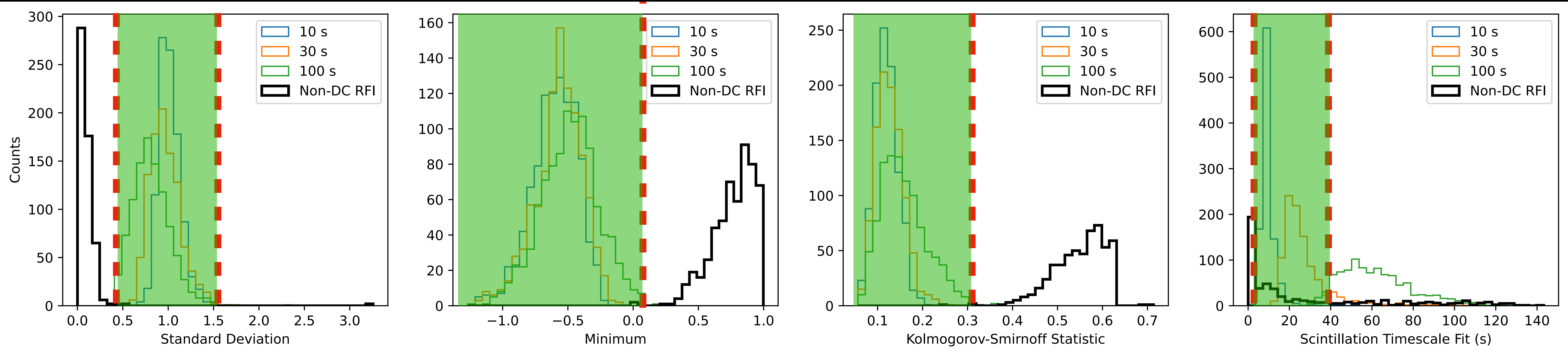
Kolmogorov-Smirnoff Statistic

Scintillation Timescale Fit

GBT RFI vs. synthetic scintillated signals

C band

S/N > 25



Standard Deviation

Minimum

Kolmogorov-Smirnoff Statistic

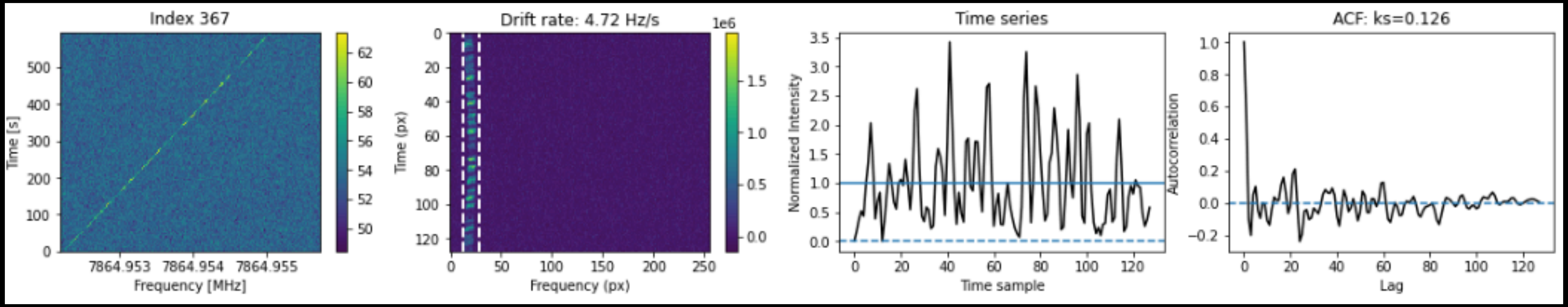
Scintillation Timescale Fit

Summary

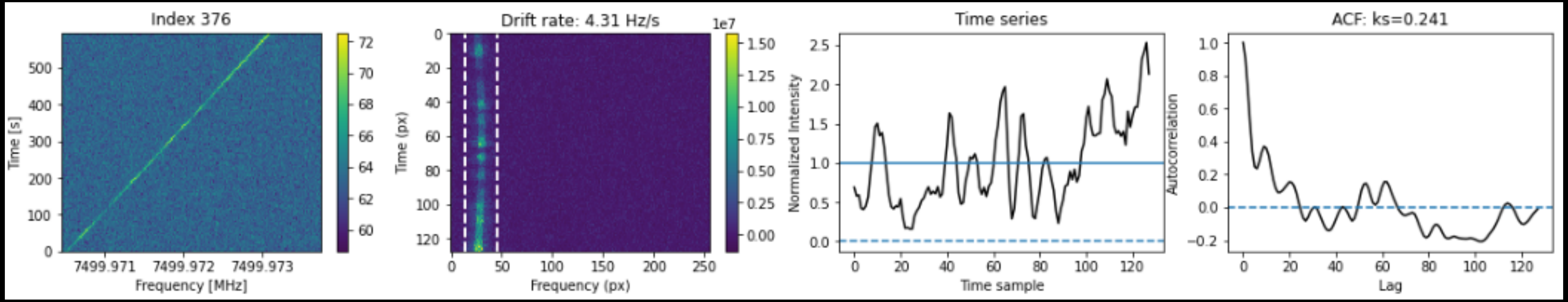
- Developed a methods framework for scintillation analysis, with accompanying codebase
- Based on the RFI environment, we get better thresholding at higher frequencies
- Better extraction & classification methods may lead to improvements
- Looking forward: dedicated survey to search for scintillated signals towards the Galactic Center

Extra Slides

What signals pass these thresholds?



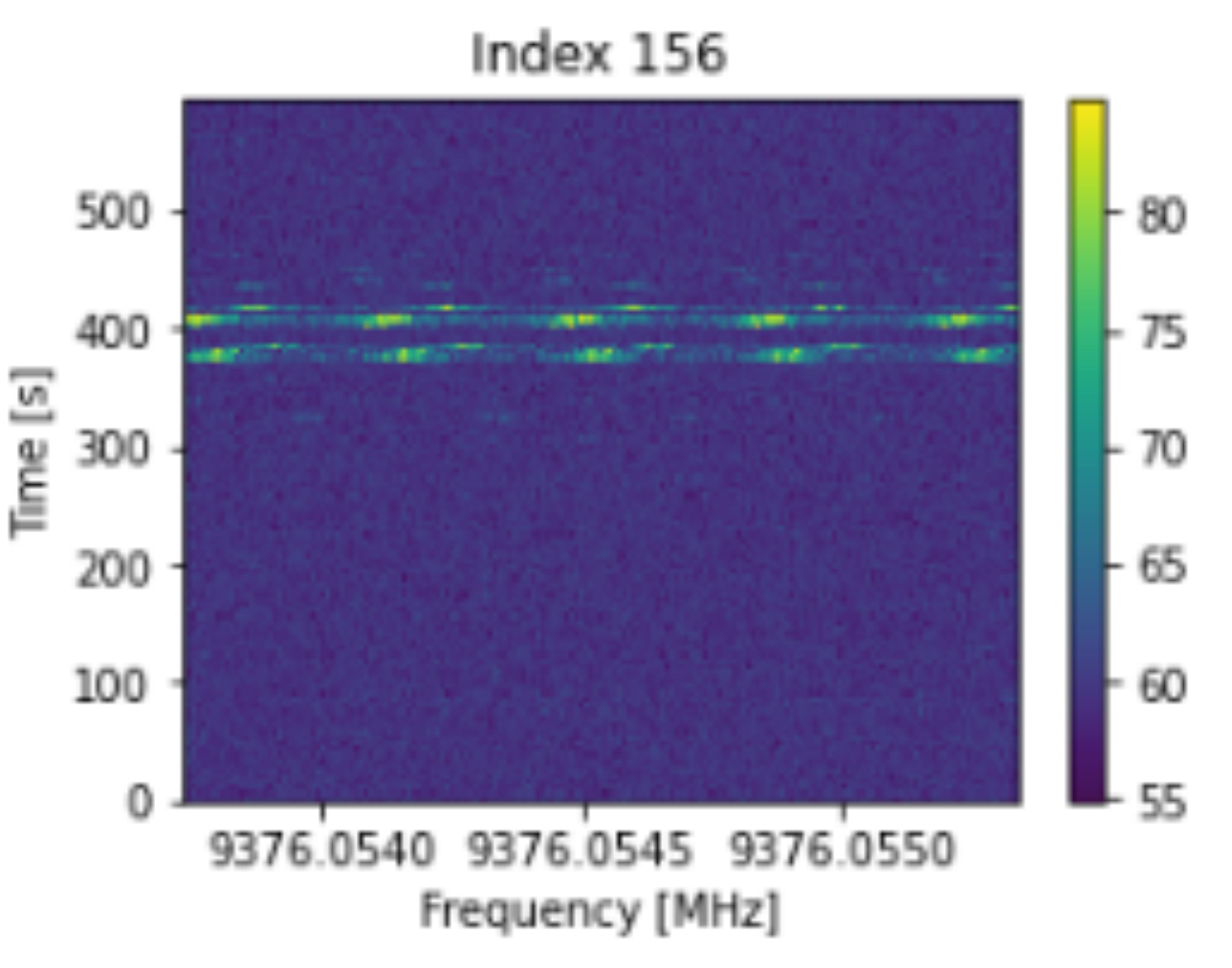
Timescale fit ~ 2 s

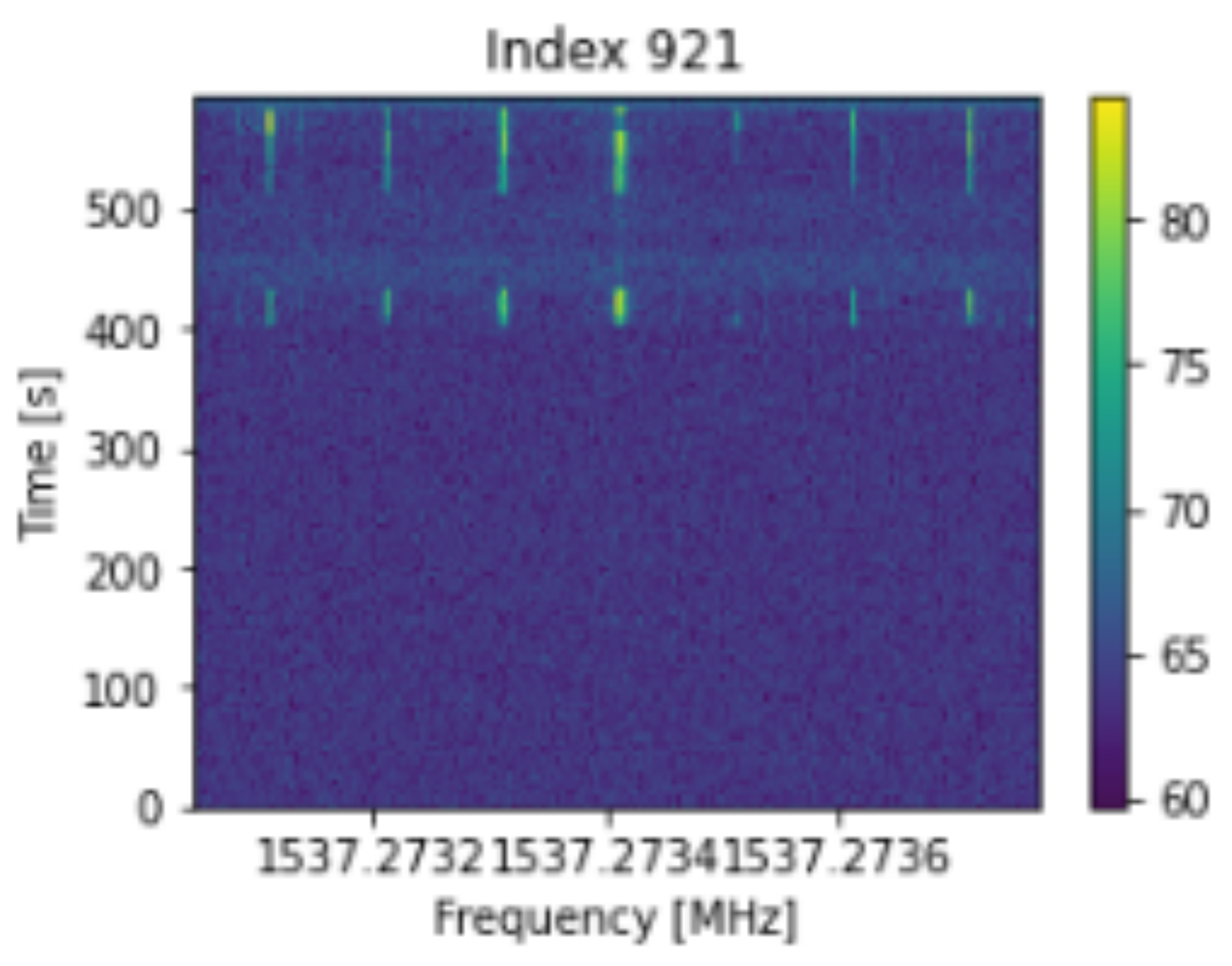


Timescale fit ~ 60 s

Limitations from RFI analysis

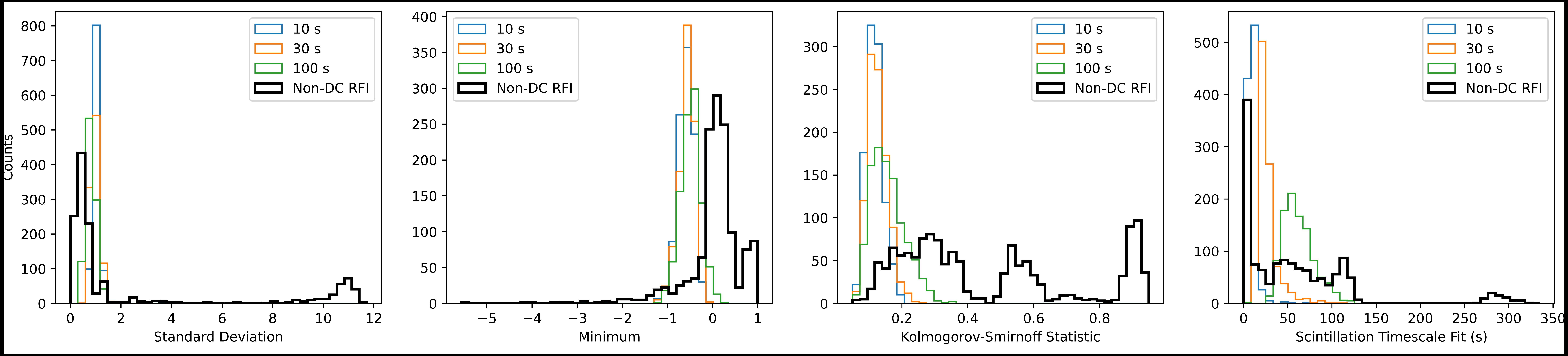
- L and S bands in particular are very noisy
- Non-narrowband signals detected just because they are above the SNR threshold
- Difficult to apply a one-size-fits-all bounding box method
- Perhaps ML can help!





L band

S/N = 25



Std. Dev.

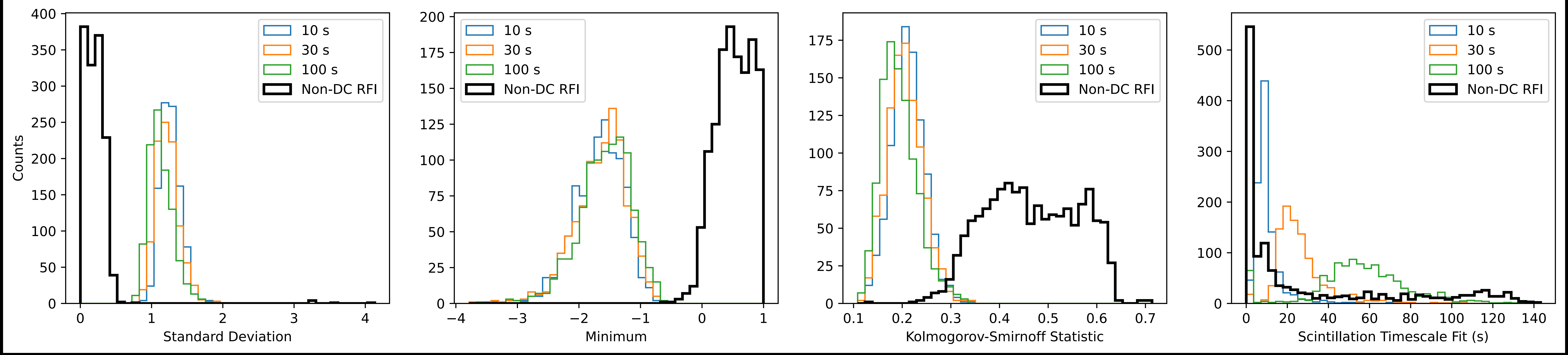
Minimum

KS Statistic

Timescale Fit

C band

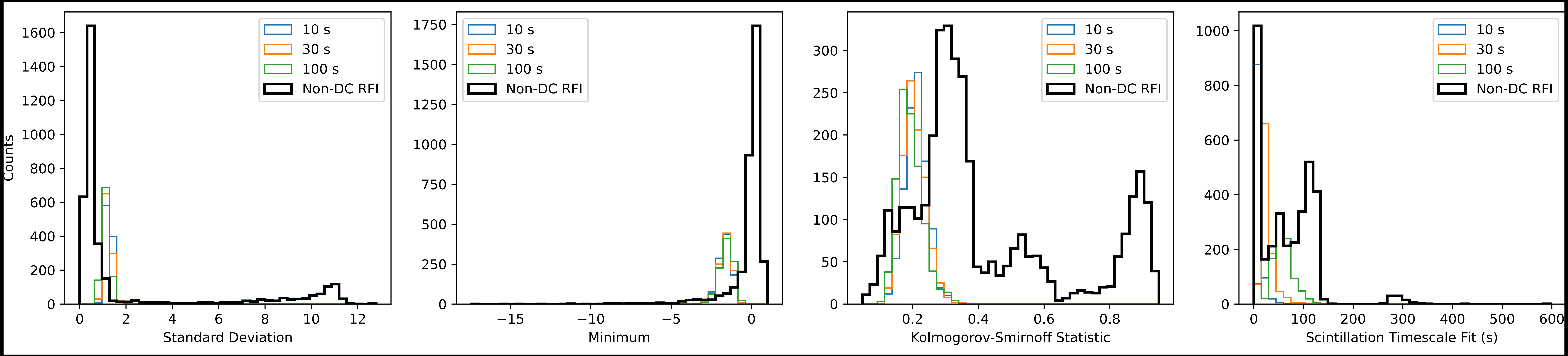
S/N = 10



Counts
Standard Deviation
Minimum
Kolmogorov-Smirnoff Statistic
Scintillation Timescale Fit (s)
10 s
30 s
100 s
Non-DC RFI

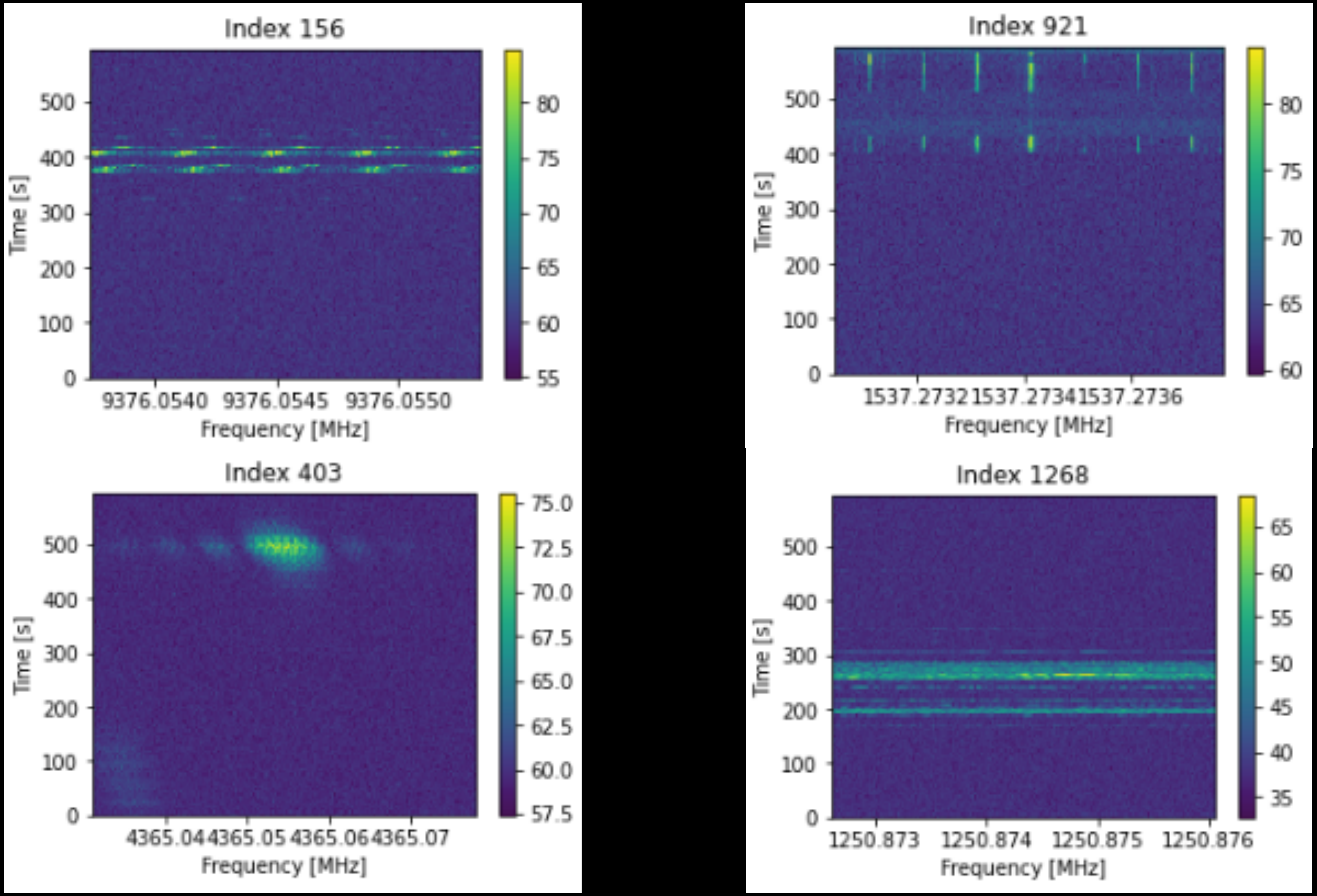
L band

S/N = 10



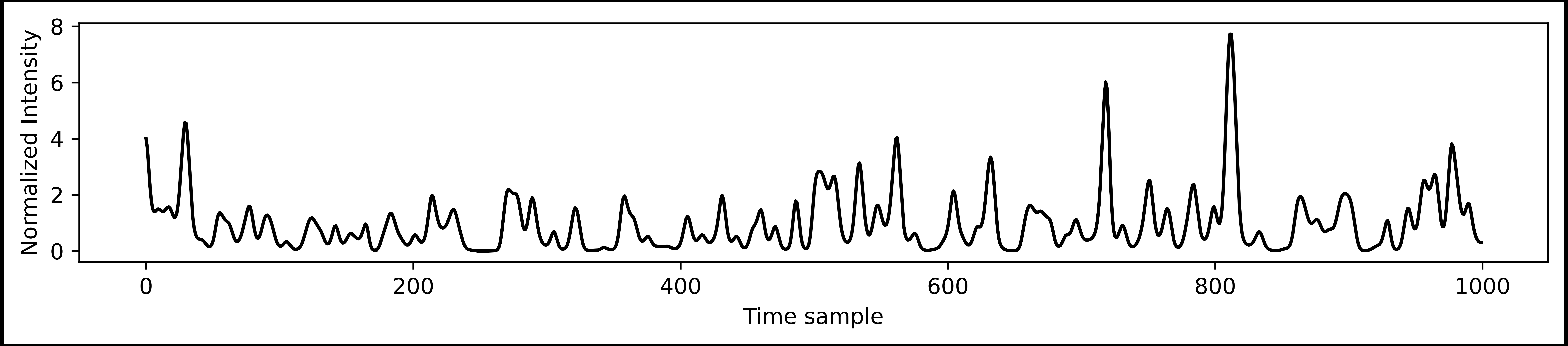
Counts
Standard Deviation
Minimum
Kolmogorov-Smirnoff Statistic
Scintillation Timescale Fit (s)
10 s
30 s
100 s
Non-DC RFI

High standard deviation (RMS) signals are pulsed - or broadband



Quick way to produce synthetic data with asymptotic statistics

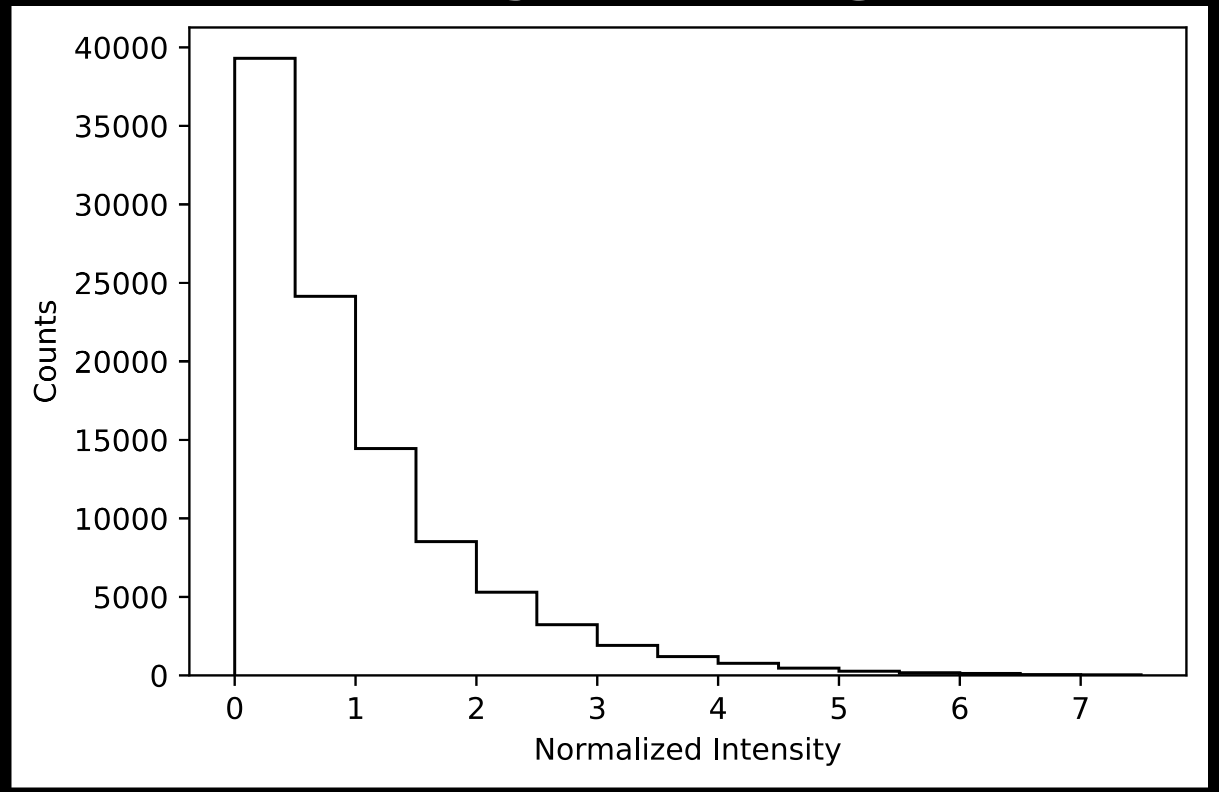
- (Cario & Nelson 1996) The ARTA random process matches:
  - Target intensity distribution
  - Target autocorrelation structure (with custom asymptotic precision)



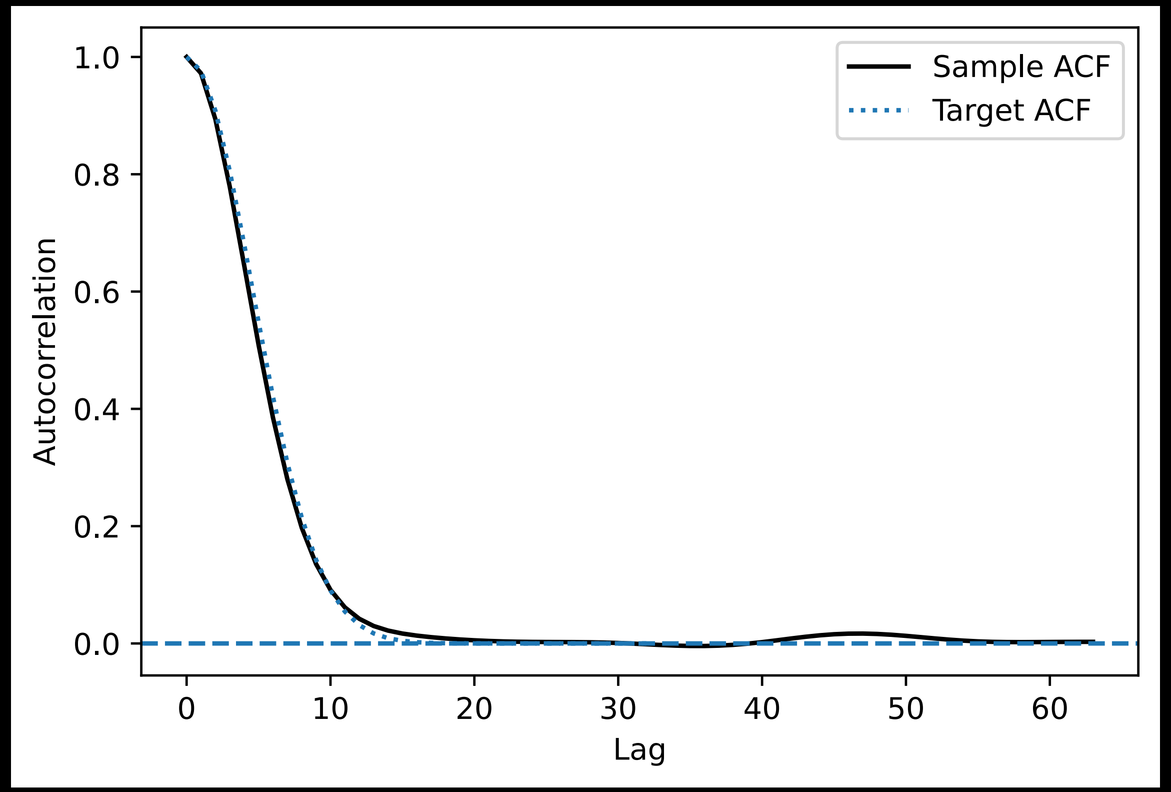
n = 1e5

$\Delta t_d = 30$ s

Intensity histogram



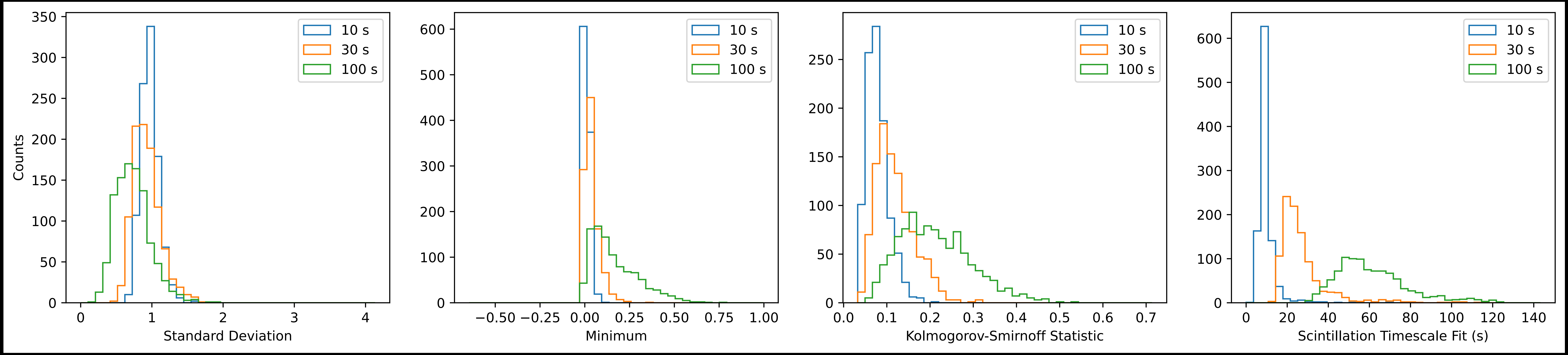
Autocorrelation



BERKELEY SETI RESEARCH CENTER

BREAKTHROUGH LISTEN

Statistics using low number of synthetic samples



Std. Dev. | Minimum | KS Statistic | Timescale Fit

10 min "observation", 4.65 s resolution

Estimating scattering strength

NE2001. I. A NEW MODEL FOR THE GALACTIC DISTRIBUTION OF FREE ELECTRONS AND ITS FLUCTUATIONS

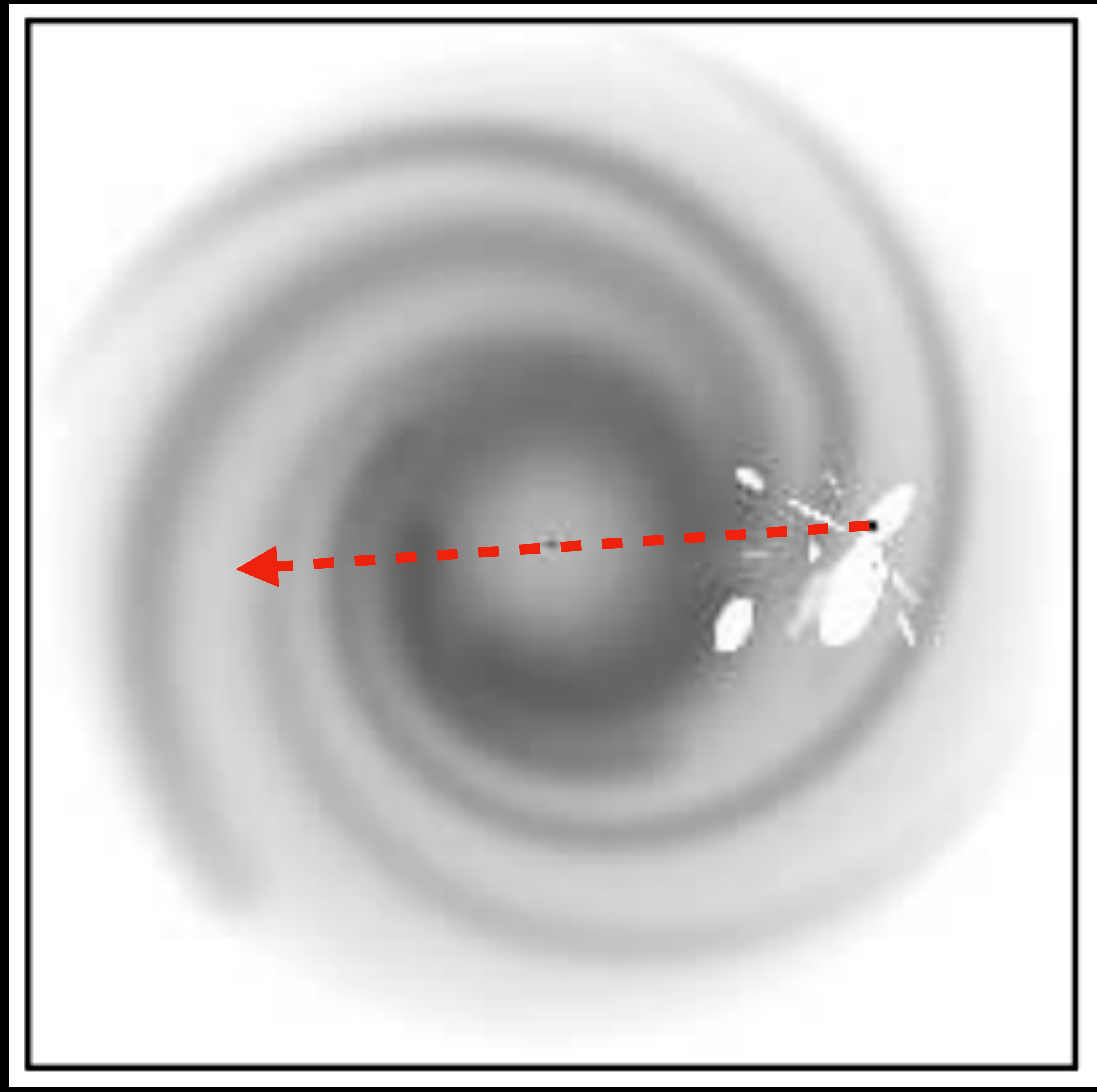
J. M. Cordes
Astronomy Department and NAIC, Cornell University, Ithaca, NY 14853
cordes@spacenet.tn.cornell.edu

T. Joseph W. Lazio
Naval Research Lab, Code 7213, Washington, D.C. 20375-5351
Joseph.Lazio@nrl.navy.mil

- NE2001 model estimates scattering parameters
- Assumes defaults of 1 GHz and 100 km/s — requires scaling!

$$\Delta t_d \propto \nu^{6/5} v_T^{-1}$$

- We use Monte Carlo sampling for unknown parameters

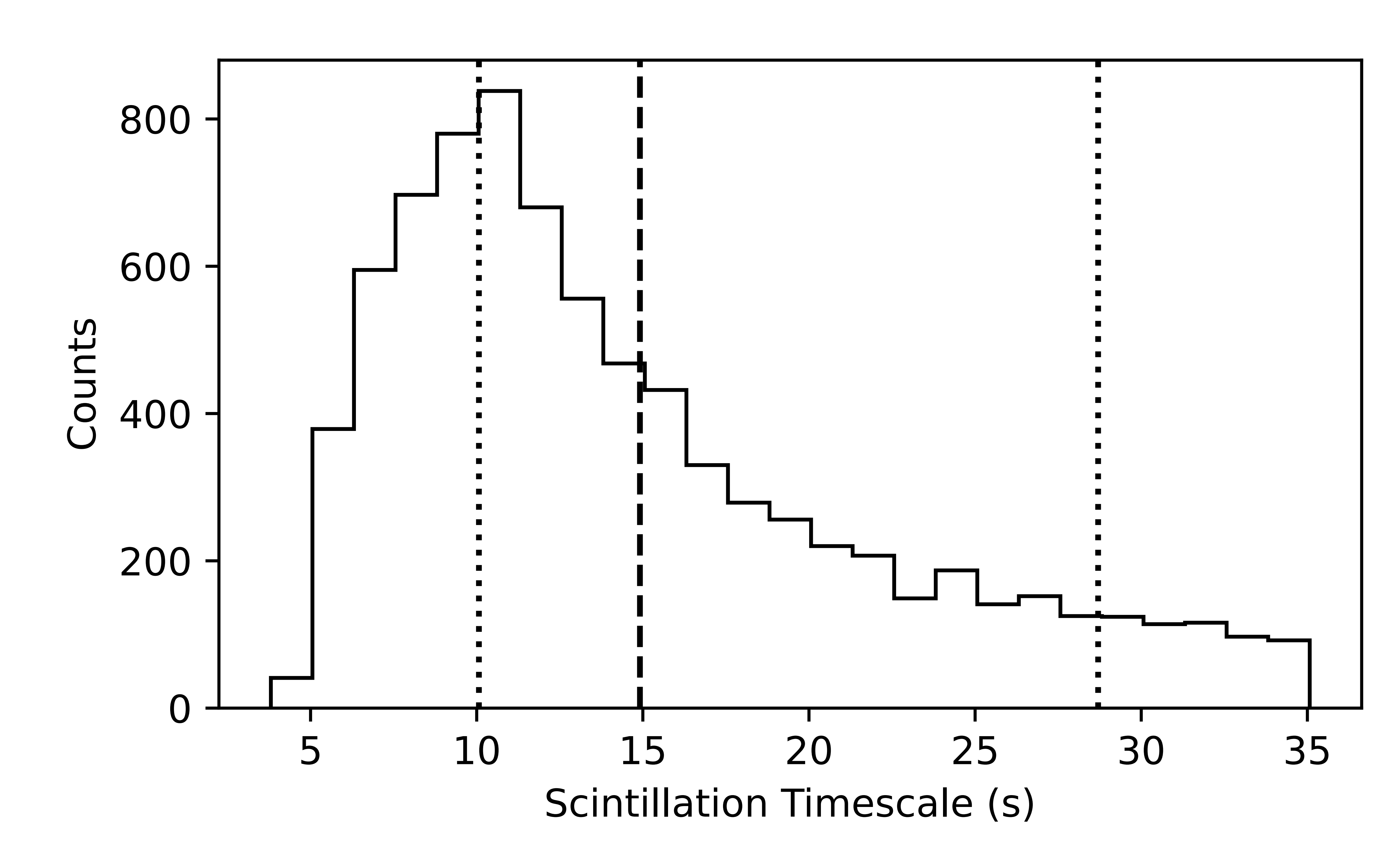


C-band

(l, b) = (1, 0)

Inter-quartile range

Median

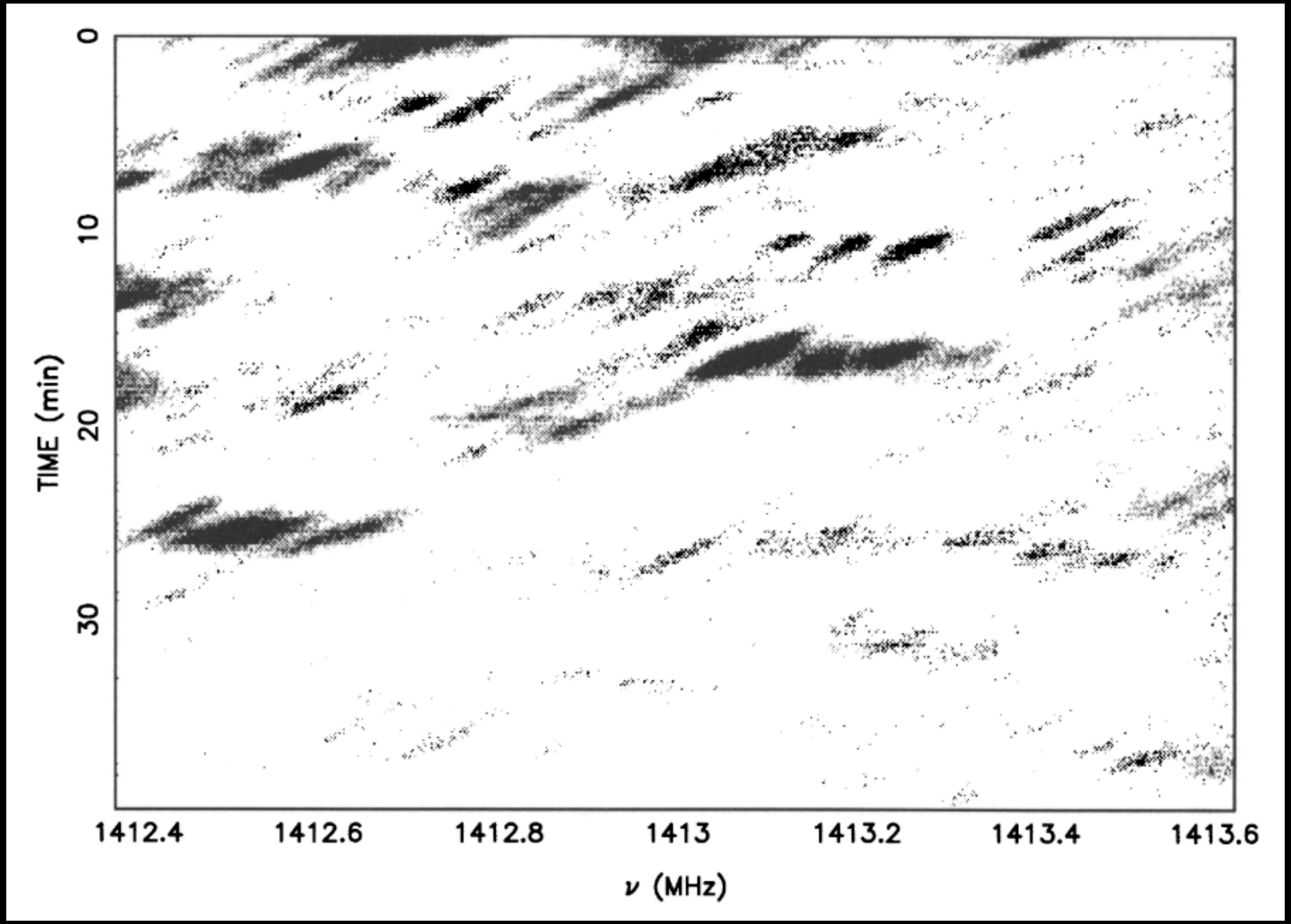


ISM Scattering & Scintillation

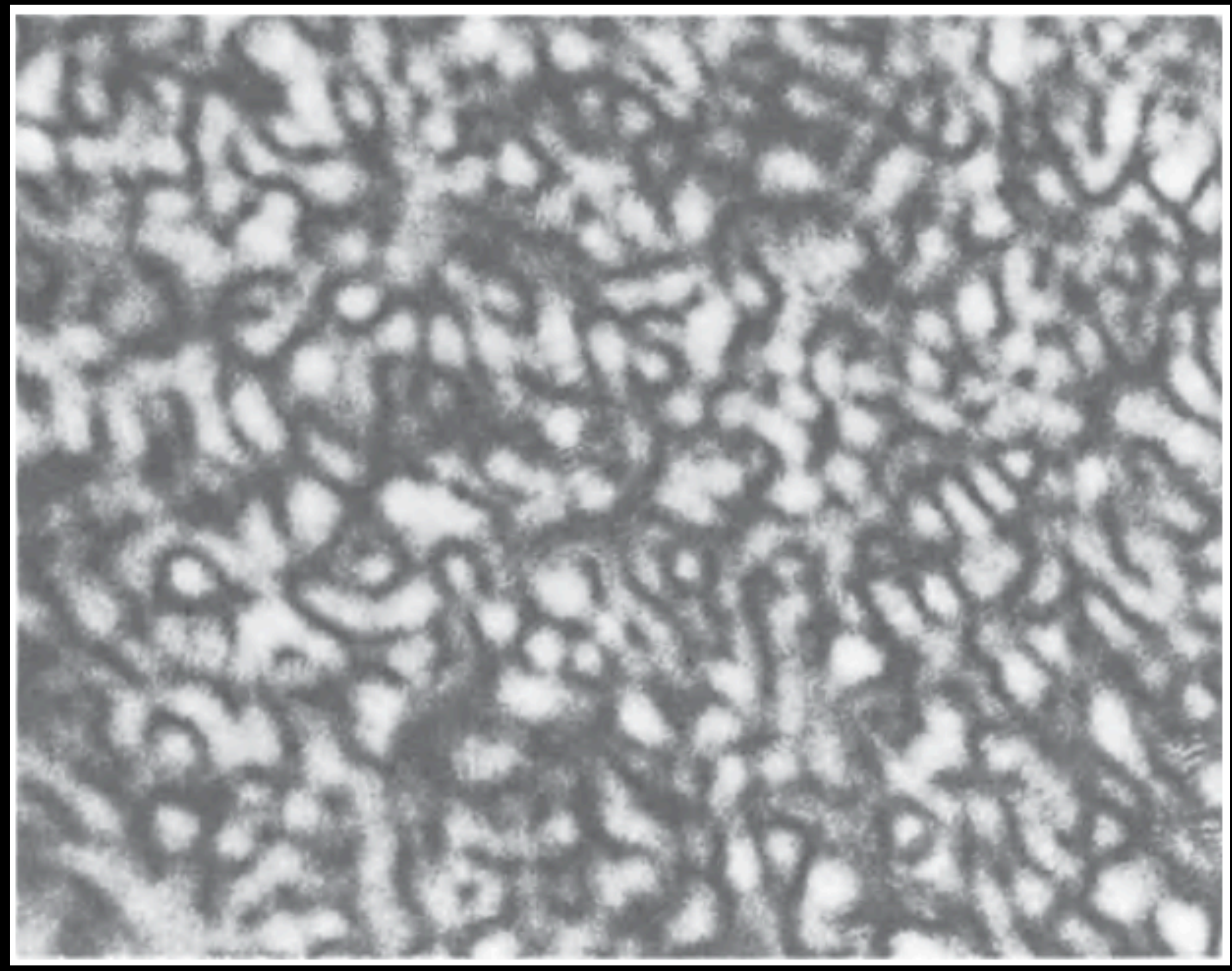
- Interaction between radio waves and free electrons in plasma

- Pulsar observations paved the way

- Parallels with laser speckle



Cordes & Lazio 1991



Goodman 1984

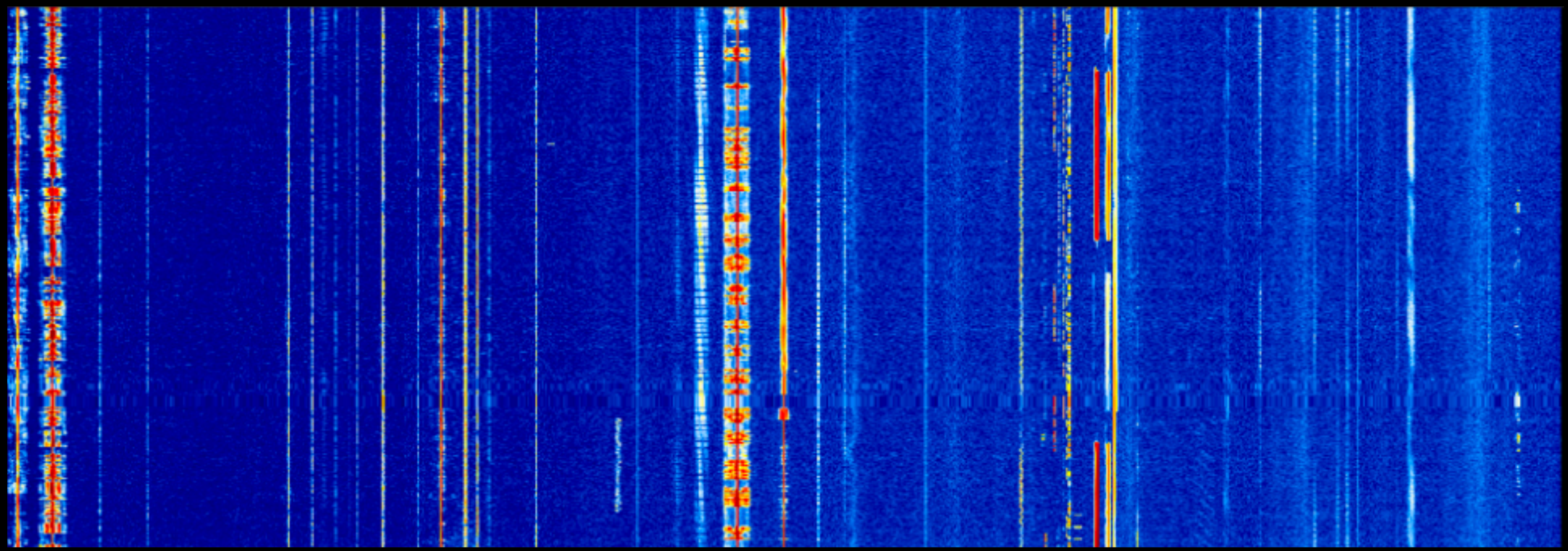
Scattering and SETI research

INTERSTELLAR SCATTERING EFFECTS ON THE DETECTION OF NARROW-BAND SIGNALS

JAMES M. CORDES AND T. JOSEPH LAZIO
National Astronomy and Ionosphere Center and Department of Astronomy, Cornell University, Ithaca, NY 14853

Received 1990 October 4; accepted 1991 January 15

- Many studies acknowledge scattering but attempt to avoid it
- Generally, SETI techniques aren't sensitive to detailed morphology
    - Noise, modulation, S/N
- Stochastic effects are hard to describe



Bigger picture: research goals

- Where and how should we look to target scintillated narrowband sources? Is this feasible and worth trying?

- Develop a overall methodology, coding, and analysis framework

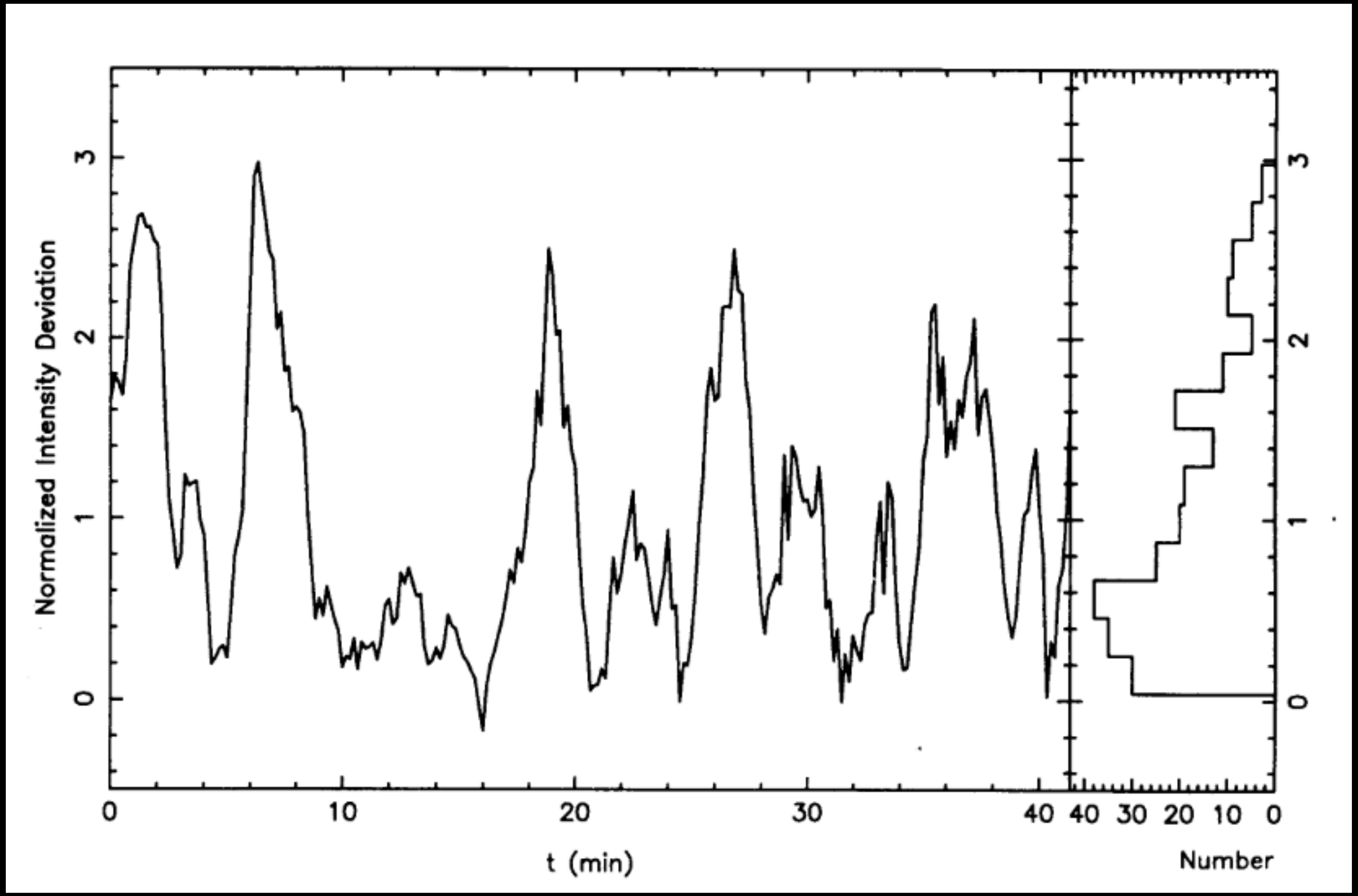
Can we detect scintillated narrowband technosignatures?

1. What scintillation timescales should we expect?
2. How can we probe asymptotic statistics?
3. Can we differentiate scintillated signals from existing RFI?

What would strongly scintillated signals look like?

Assuming 100% duty-cycle narrowband emission

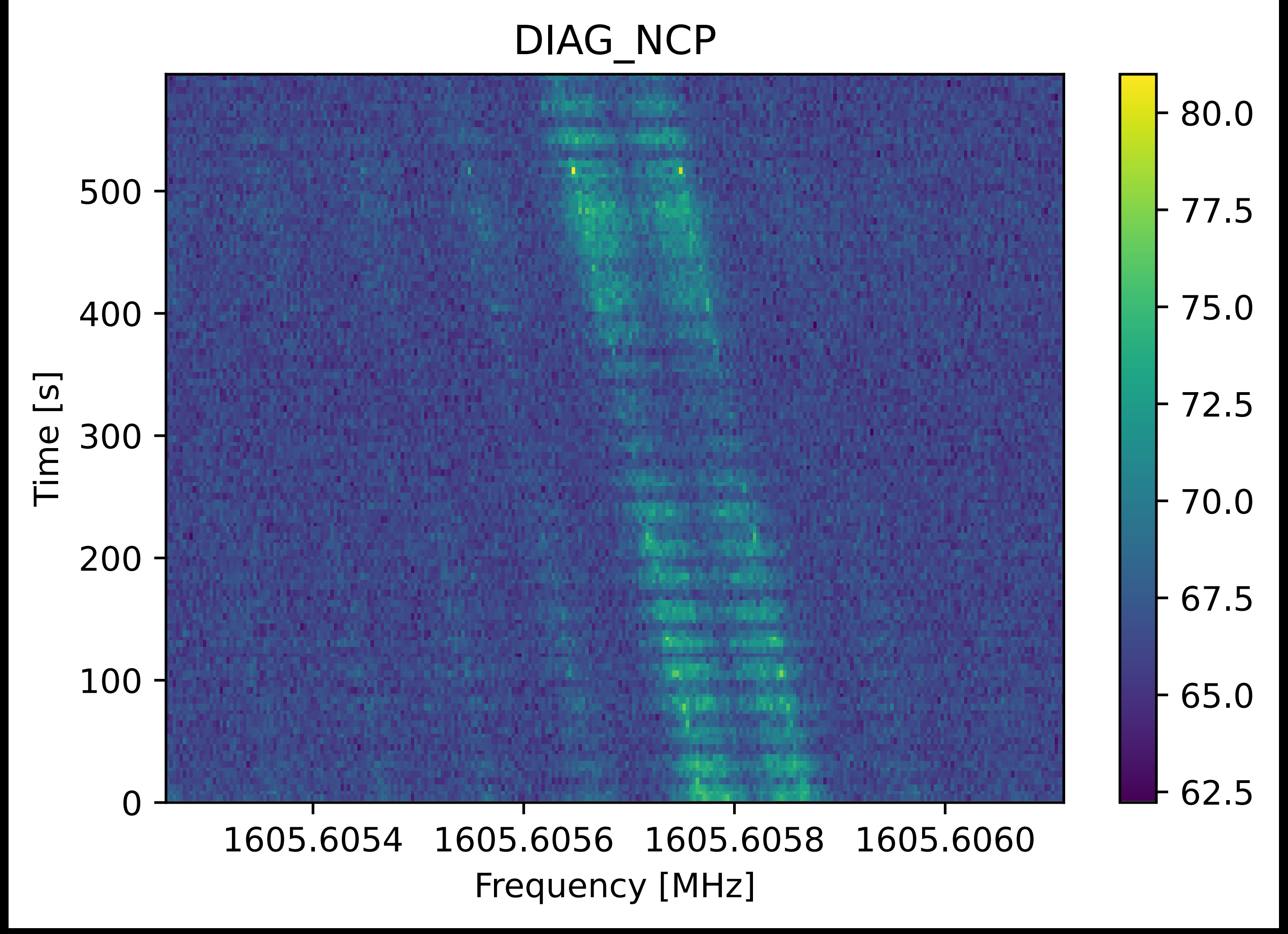
- Asymptotic behavior:
    - Exponential intensity distribution
    - Approximately Gaussian autocorrelation, with characteristic timescale

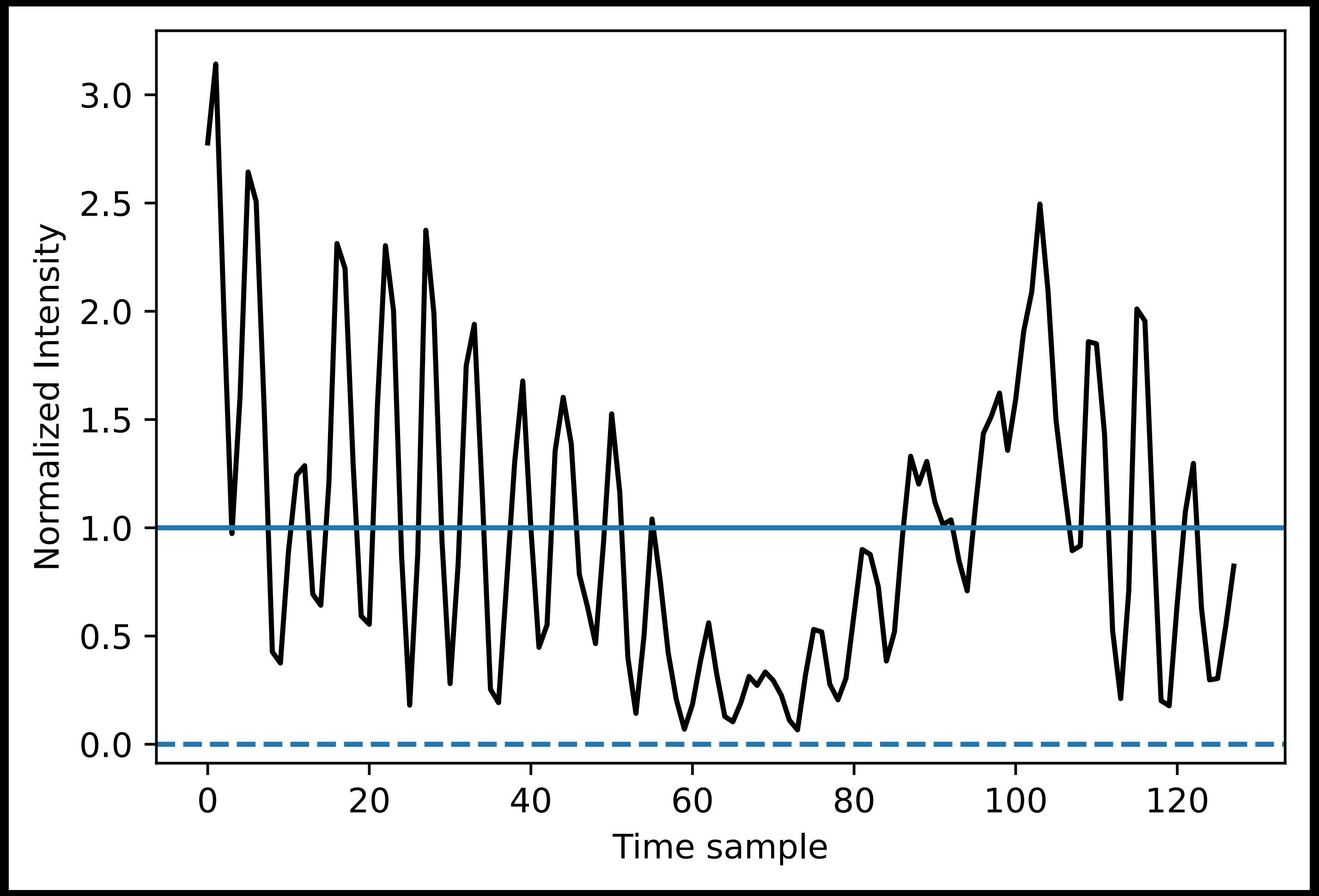


Cordes & Lazio 1991;
Cordes, Lazio, Sagan 1997

Given a signal… is it scintillated?

- Create bounding box around narrowband signal
- Estimate noise-subtracted intensity time series, normalized to mean 1
- Compute “diagnostic statistics” that pertain to asymptotic behavior
    - E.g. standard deviation, Kolmogorov-Smirnoff statistic, fit to autocorrelation function





What would strongly scintillated signals look like?

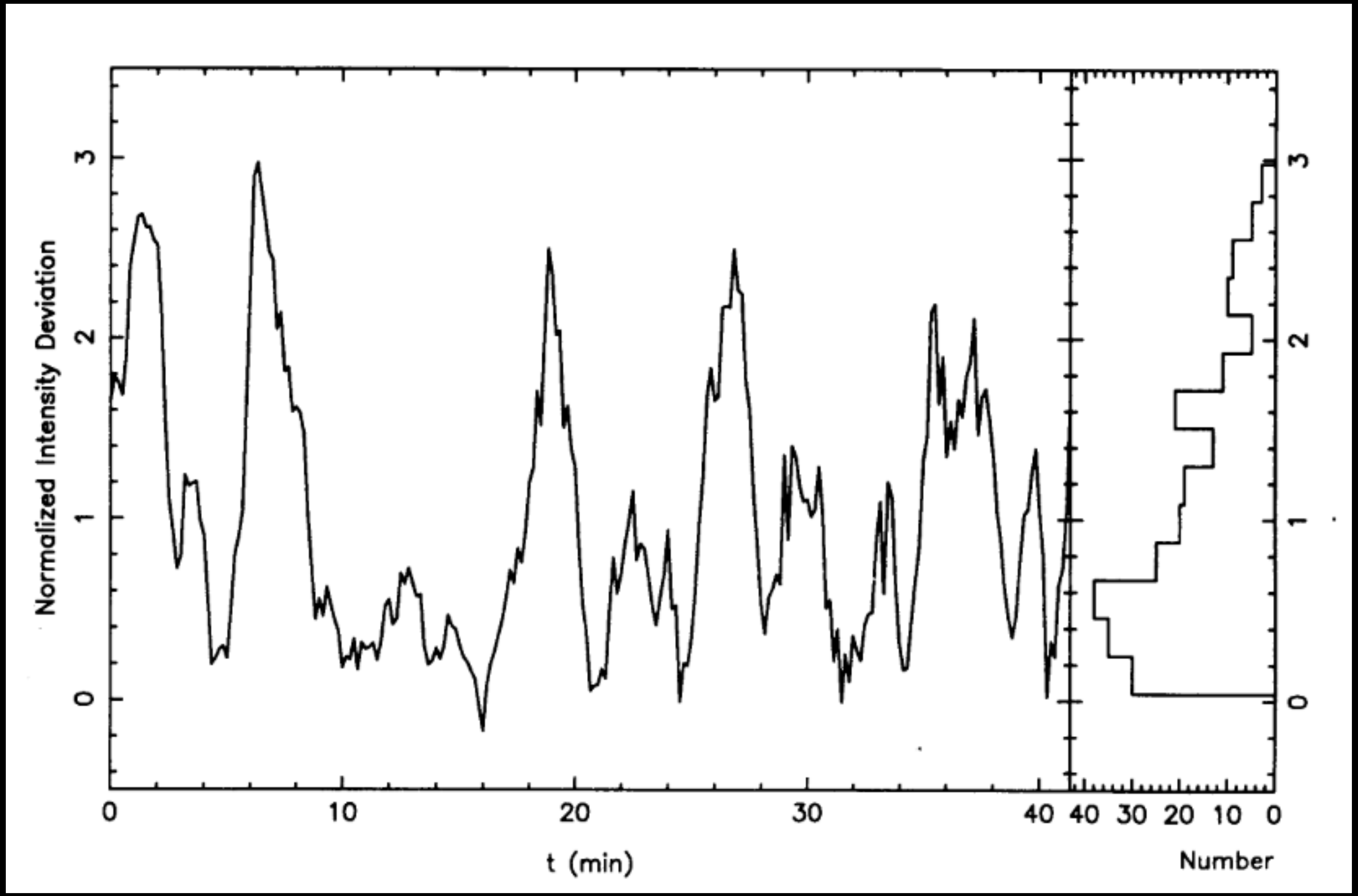
- Expected asymptotic behavior:
- Exponential intensity distribution

$$p(I) \propto e^{-I/\langle I \rangle}$$

- Near Gaussian autocorrelation, with characteristic timescale

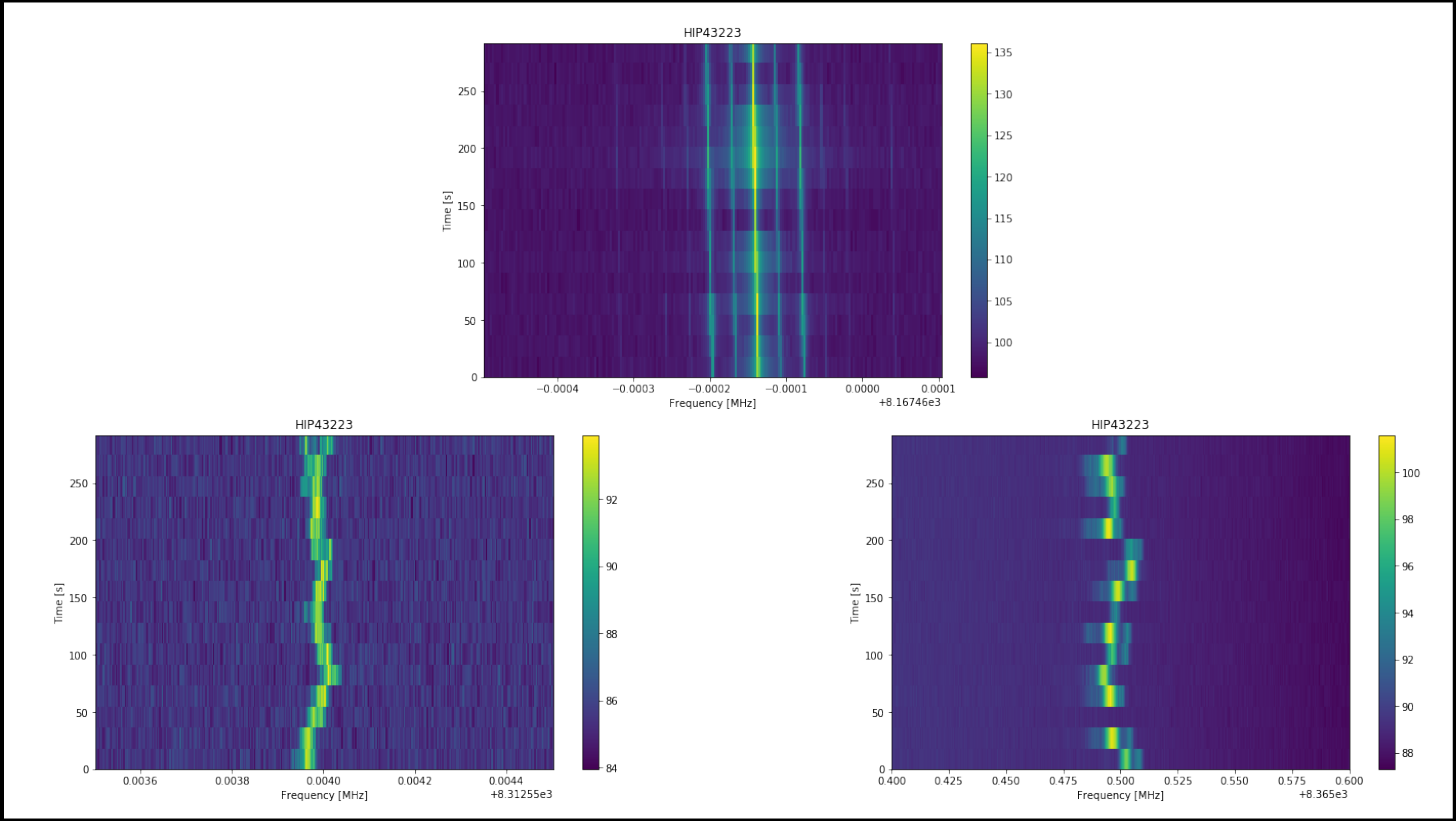
$$\rho(\tau) \sim e^{-(\tau/\Delta t_d)^2}$$

Assuming 100% duty-cycle narrowband emission



Cordes & Lazio 1991;
Cordes, Lazio, Sagan 1997

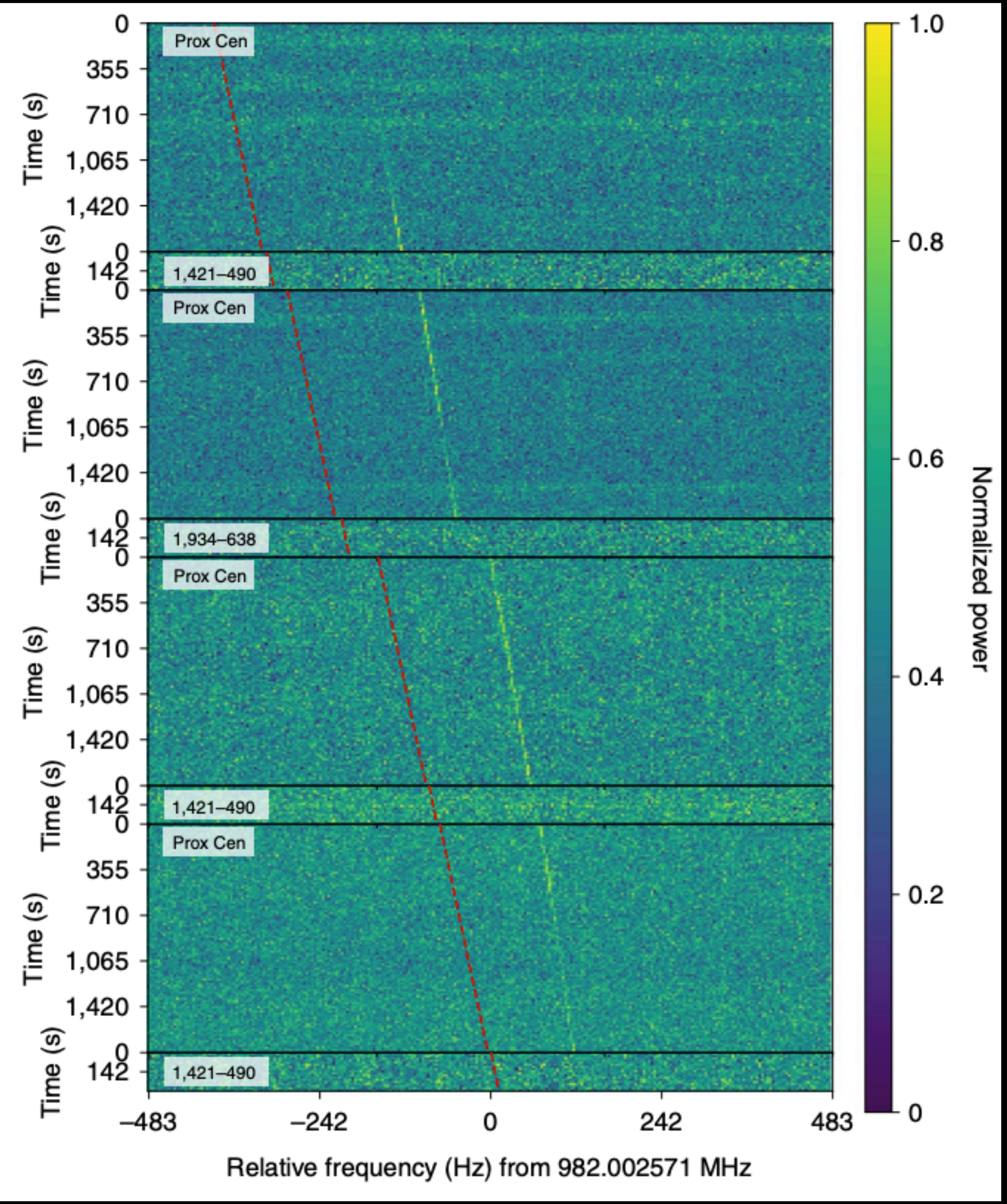
But what does the RFI environment look like?



Extra Slides x2

Plasma effects as a search filter for SETI

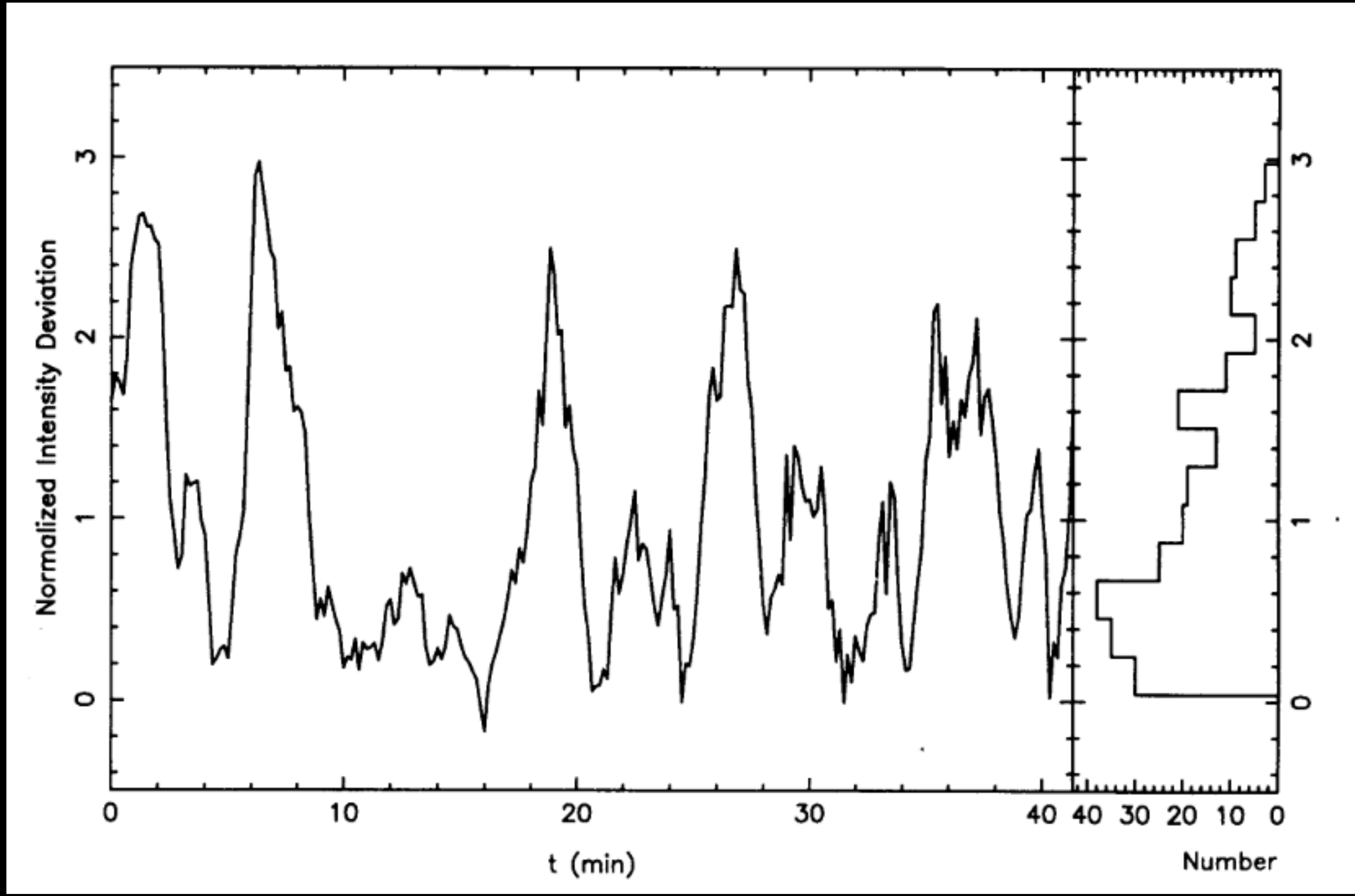
- Modern radio SETI involves detecting a vast number of signals and filtering likely candidates
- For a few reasons, most filters do not involve the effects on the signal itself
- We propose that in some cases, we can detect scintillation from the ISM in narrowband signals, which would heavily imply extrasolar origin



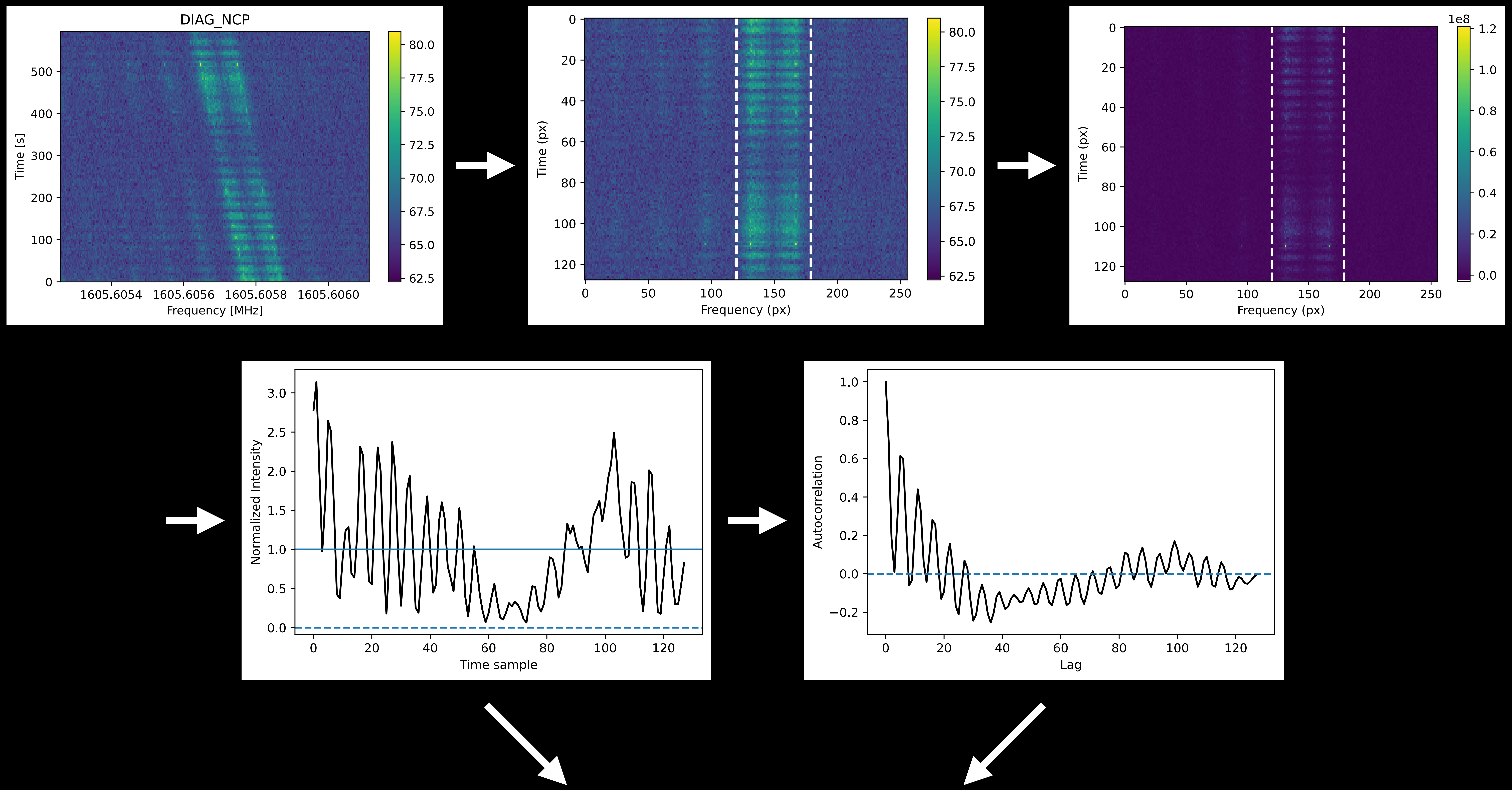
Smith et al. 2021

Methods

- Target 100% duty-cycle, narrowband transmitters
- Since scintillation is stochastic by definition, identify measurable statistics
- Estimate intensity time series from detected signals for analysis
- Use procedure on RFI in unlikely directions to probe the interference environment



Cordes & Lazio 1991;
Cordes, Lazio, Sagan 1997

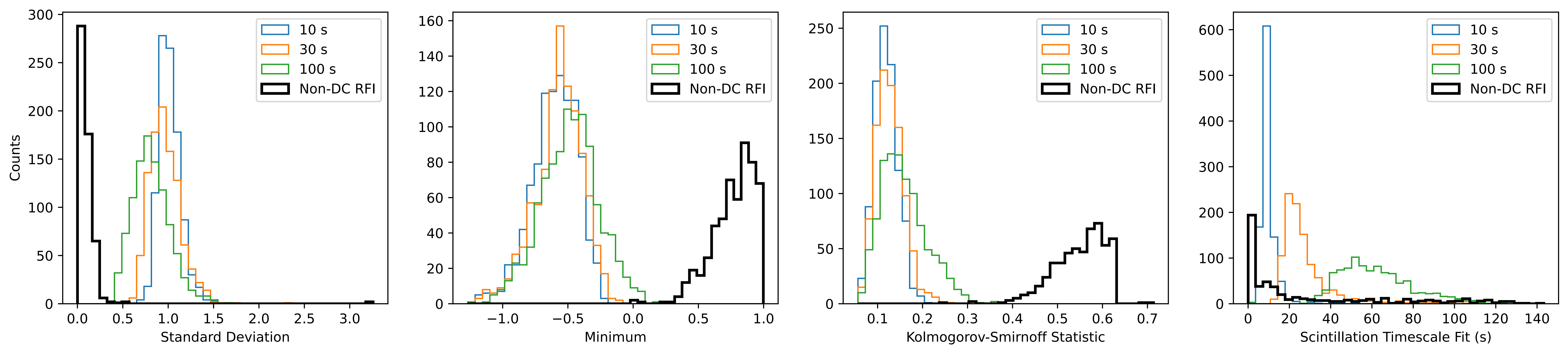


Diagnostic statistics

RFI Analysis

C band

S/N = 25



Standard Deviation

Minimum

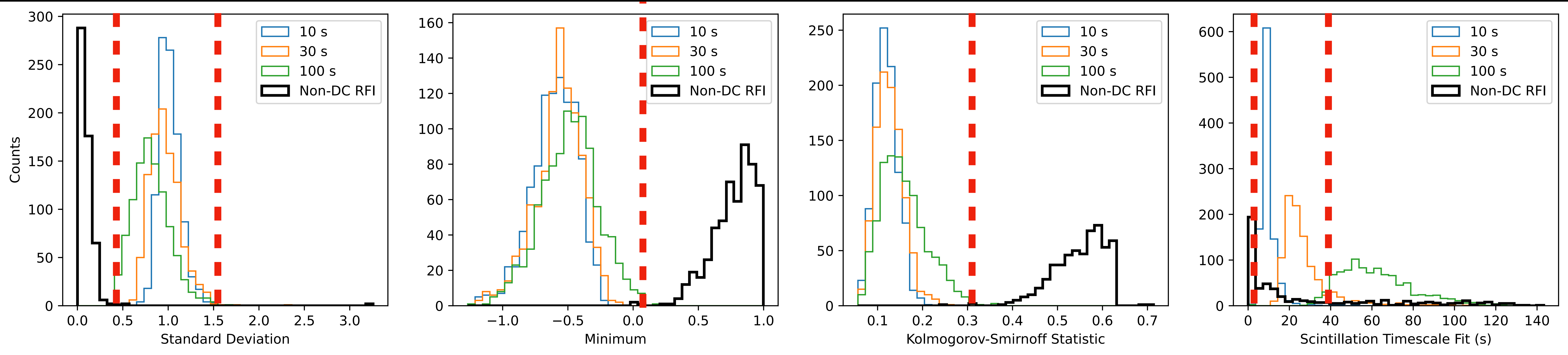
Kolmogorov-Smirnoff Statistic

Scintillation Timescale Fit

RFI Analysis

C band

S/N = 25



Standard Deviation

Minimum

Kolmogorov-Smirnoff Statistic

Scintillation Timescale Fit

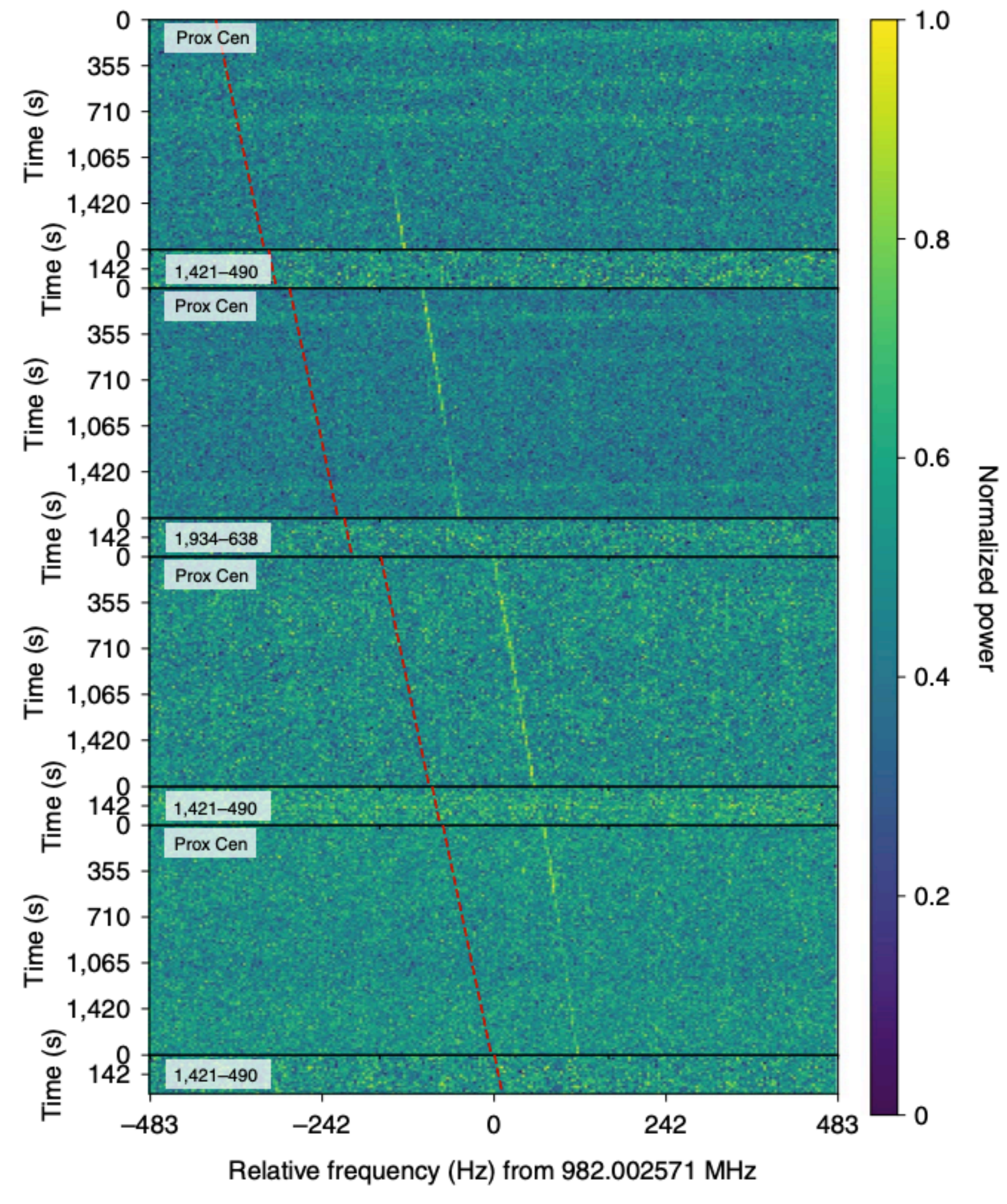
Summary

- Developed a framework for scintillation analysis, with accompanying code
- Because of RFI environment, higher frequencies are more amenable
- Looking forward: dedicated survey with custom resolution to search near the Galactic Center
- Better extraction / classification methods may lead to improvements

Extra Slides

Candidate identification and differentation

- Narrowband (vs. astrophysical sources)
- Non-zero drift rate (vs. RFI)
- Sky localization (vs. RFI)



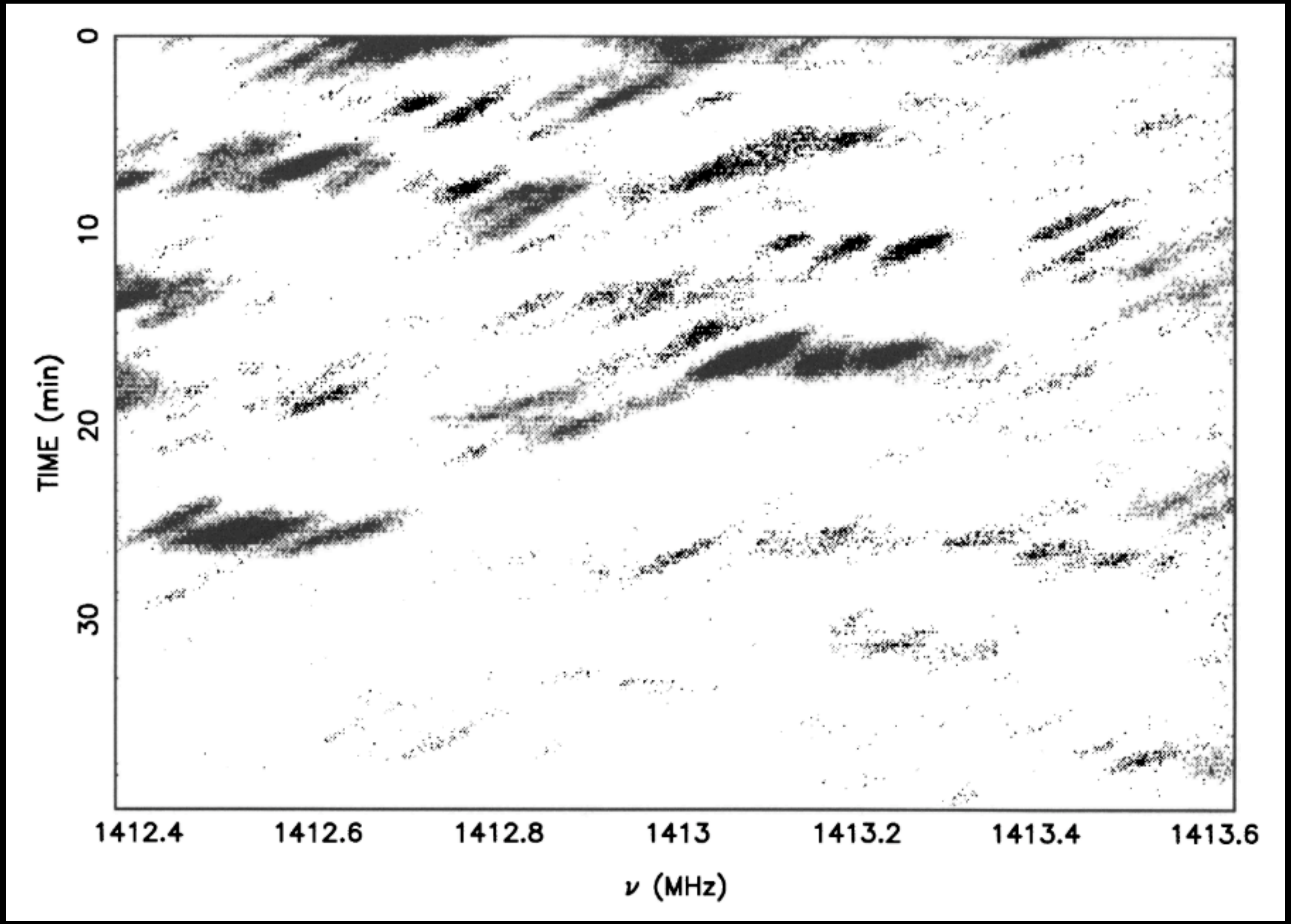
Smith et al. 2021

Can we use astrophysical phenomena to discriminate technosignatures from RFI?

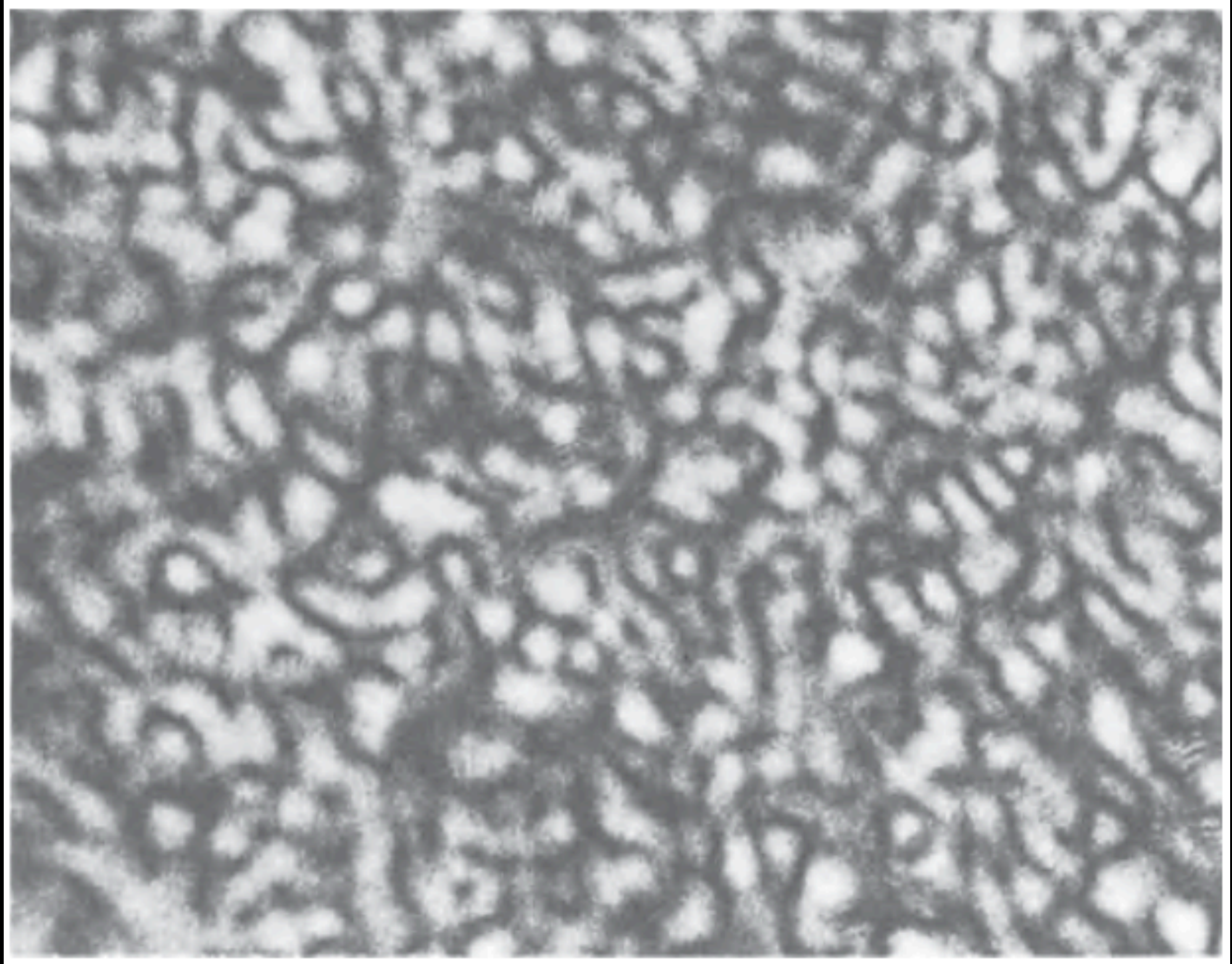
ESA

ISM Scattering & Scintillation

- Interaction between radio waves and free electrons in plasma
- Pulsar observations paved the way
- Parallels with laser speckle



Cordes & Lazio 1991



Goodman 1984

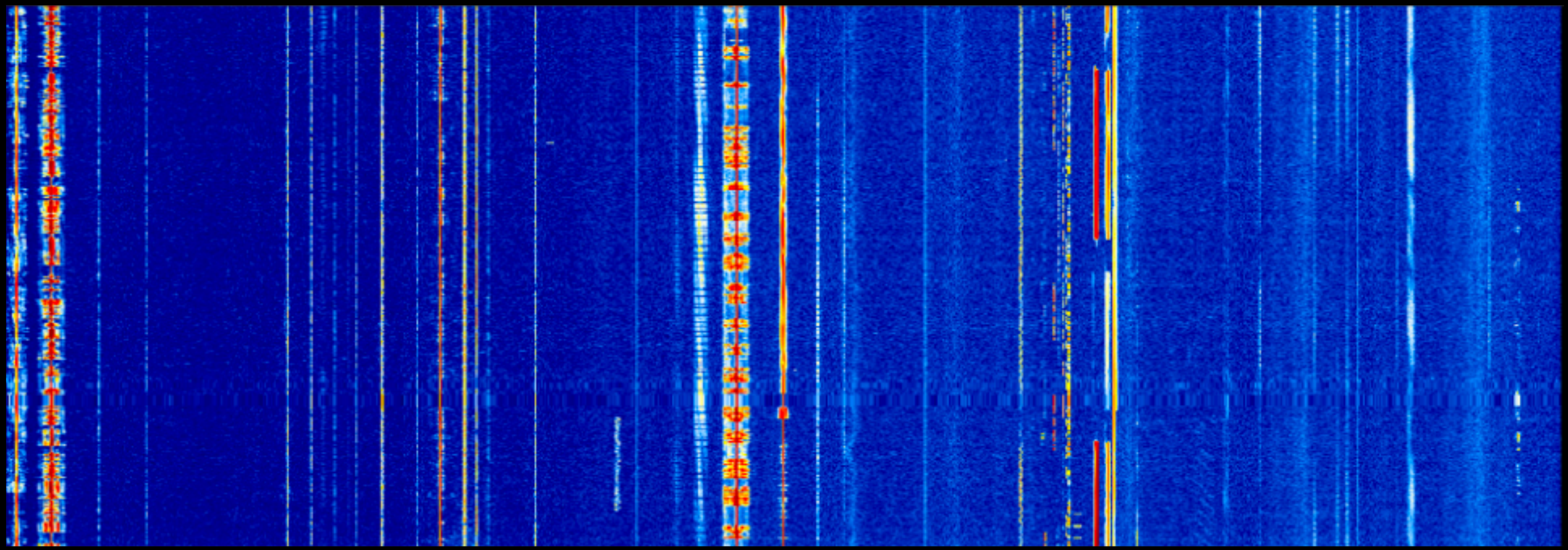
Scattering and SETI research

INTERSTELLAR SCATTERING EFFECTS ON THE DETECTION OF NARROW-BAND SIGNALS

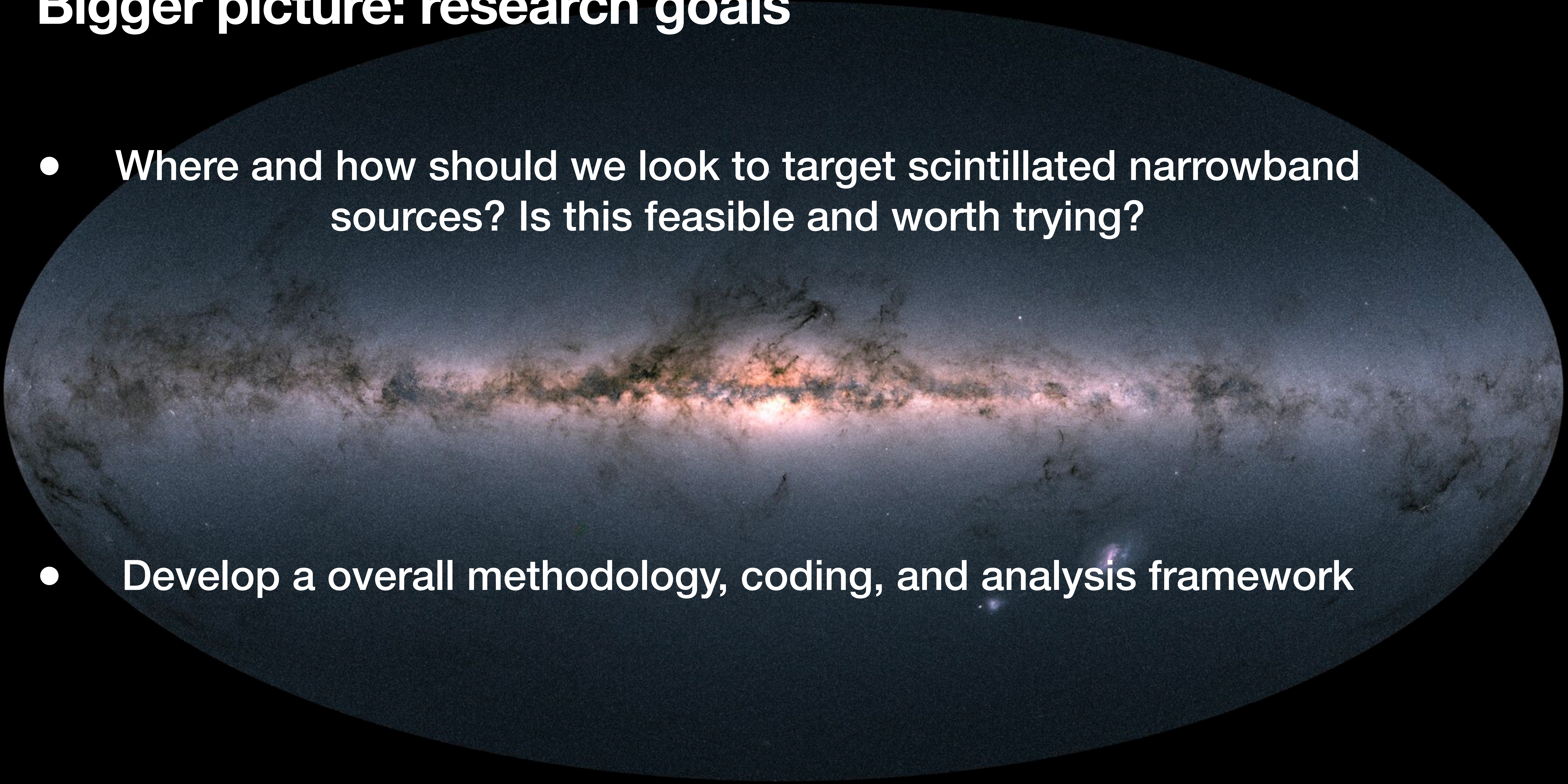
JAMES M. CORDES AND T. JOSEPH LAZIO
National Astronomy and Ionosphere Center and Department of Astronomy, Cornell University, Ithaca, NY 14853

Received 1990 October 4; accepted 1991 January 15

- Many studies acknowledge scattering but attempt to avoid it
- Generally, SETI techniques aren't sensitive to detailed morphology
  - Noise, modulation, S/N
- Stochastic effects are hard to describe



Bigger picture: research goals



- Where and how should we look to target scintillated narrowband sources? Is this feasible and worth trying?

- Develop a overall methodology, coding, and analysis framework

What would strongly scintillated signals look like?

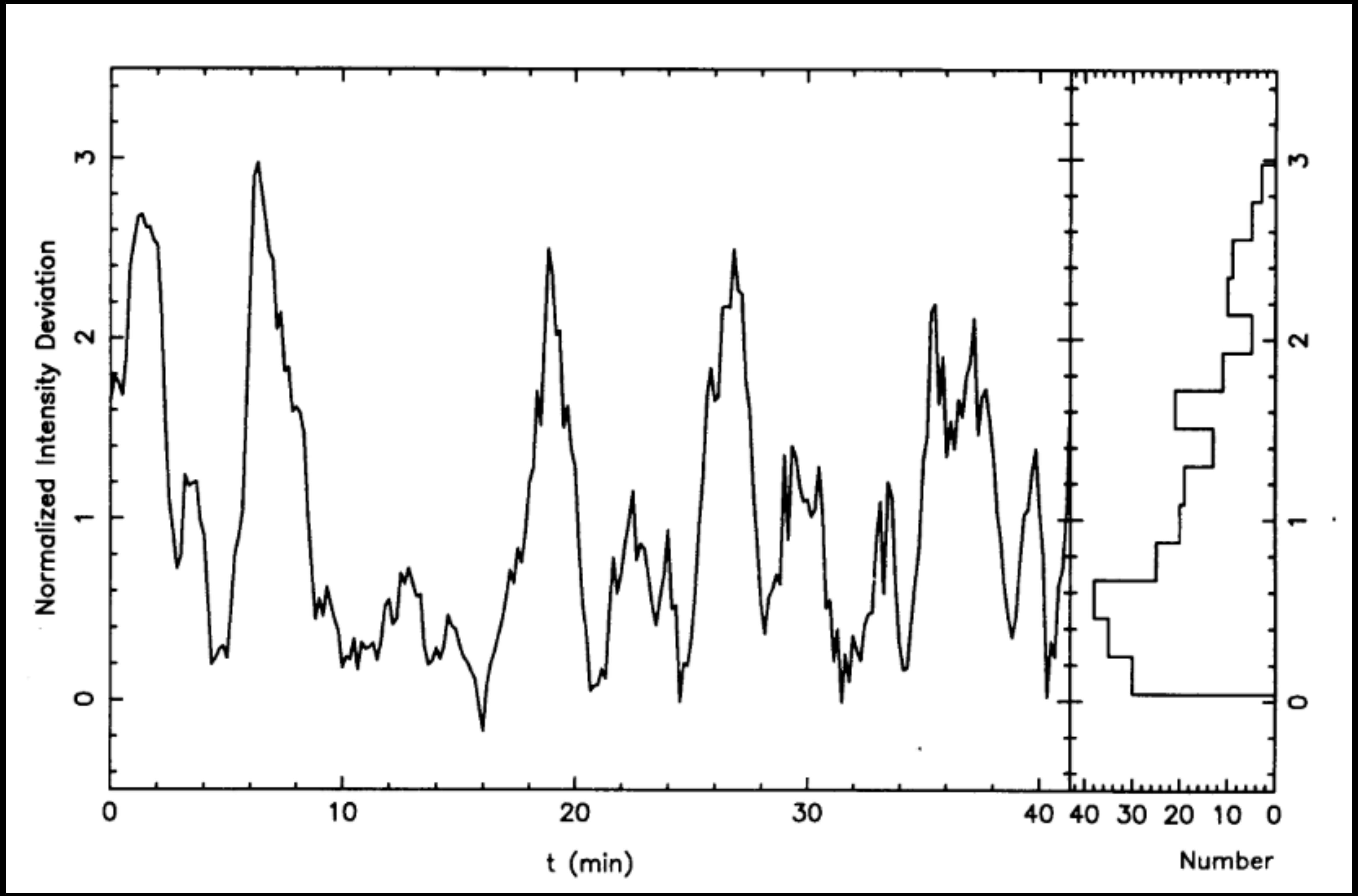
Assuming a 100% duty-cycle narrowband transmitter

- Exponential intensity distribution

$$p(I) \propto e^{-I/\langle I \rangle}$$

- Near Gaussian auto-correlation (ACF), with characteristic timescale

$$\rho(\tau) \sim e^{-(\tau/\Delta t_d)^2}$$



Cordes & Lazio 1991;
Cordes, Lazio, Sagan 1997

Can we detect scintillated narrowband technosignatures?

1. What scintillation timescales should we expect?

2. How can we probe asymptotic statistics?

3. Can we differentiate scintillated signals from existing RFI?

Estimating scattering strength

NE2001. I. A NEW MODEL FOR THE GALACTIC DISTRIBUTION OF FREE ELECTRONS AND ITS FLUCTUATIONS

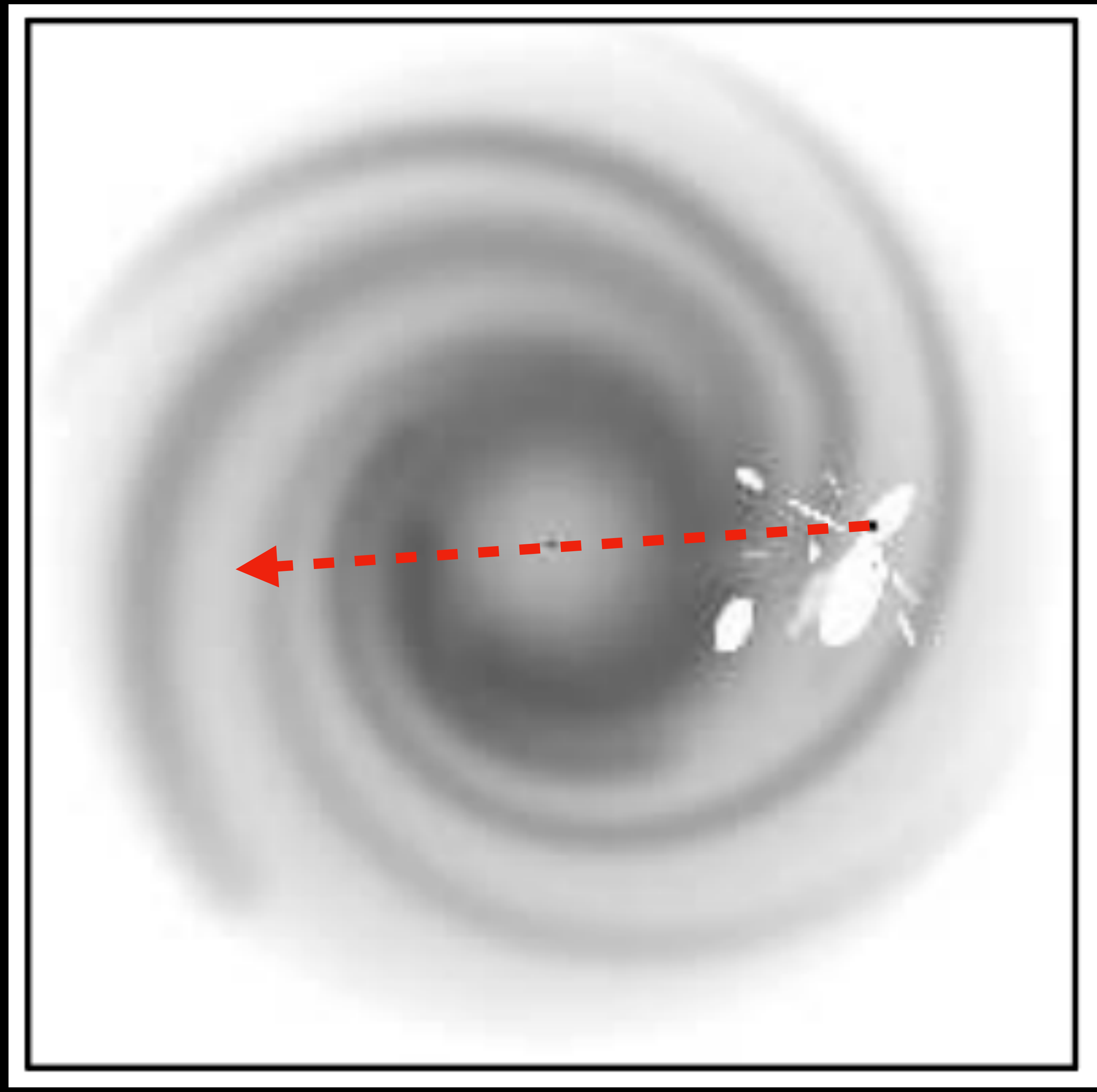
J. M. Cordes
Astronomy Department and NAIC, Cornell University, Ithaca, NY 14853
cordes@spacenet.tn.cornell.edu

T. Joseph W. Lazio
Naval Research Lab, Code 7213, Washington, D.C. 20375-5351
Joseph.Lazio@nrl.navy.mil

- NE2001 model estimates scattering parameters

- Assumes defaults of 1 GHz and 100 km/s — requires scaling!

$$\Delta t_d \propto \nu^{6/5} v_T^{-1}$$



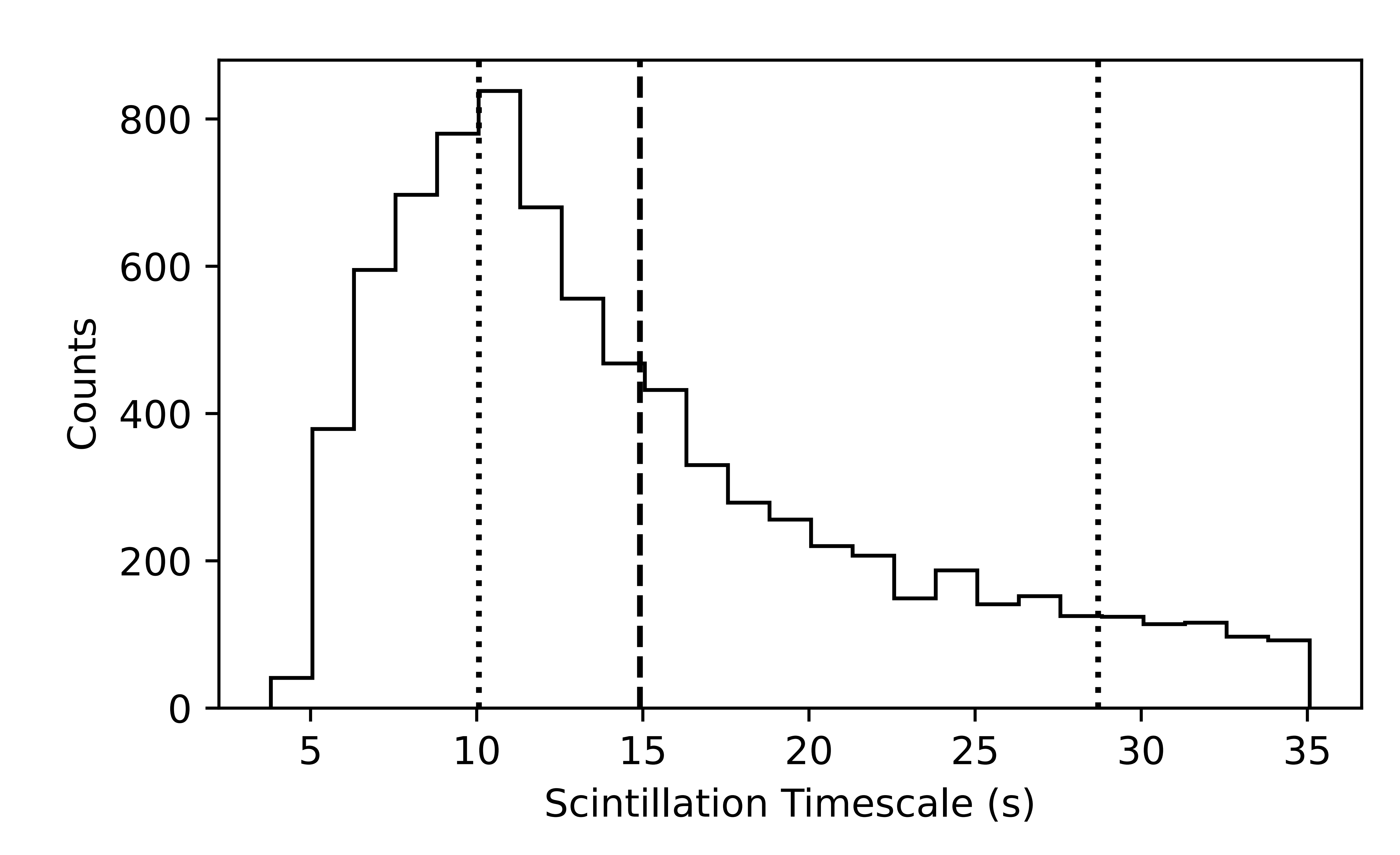
- We use Monte Carlo sampling for unknown parameters

C-band

(l, b) = (1, 0)

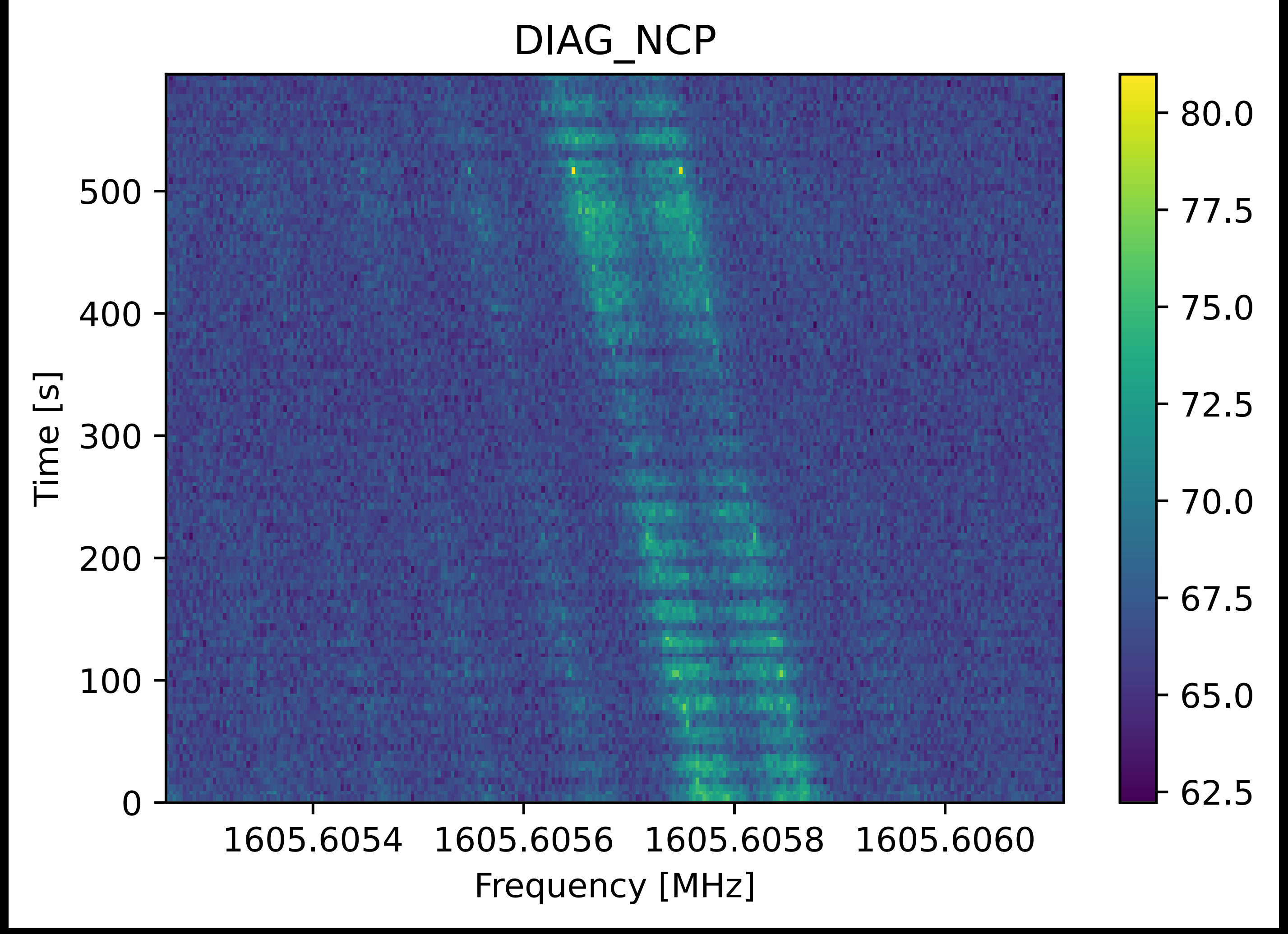
Inter-quartile range

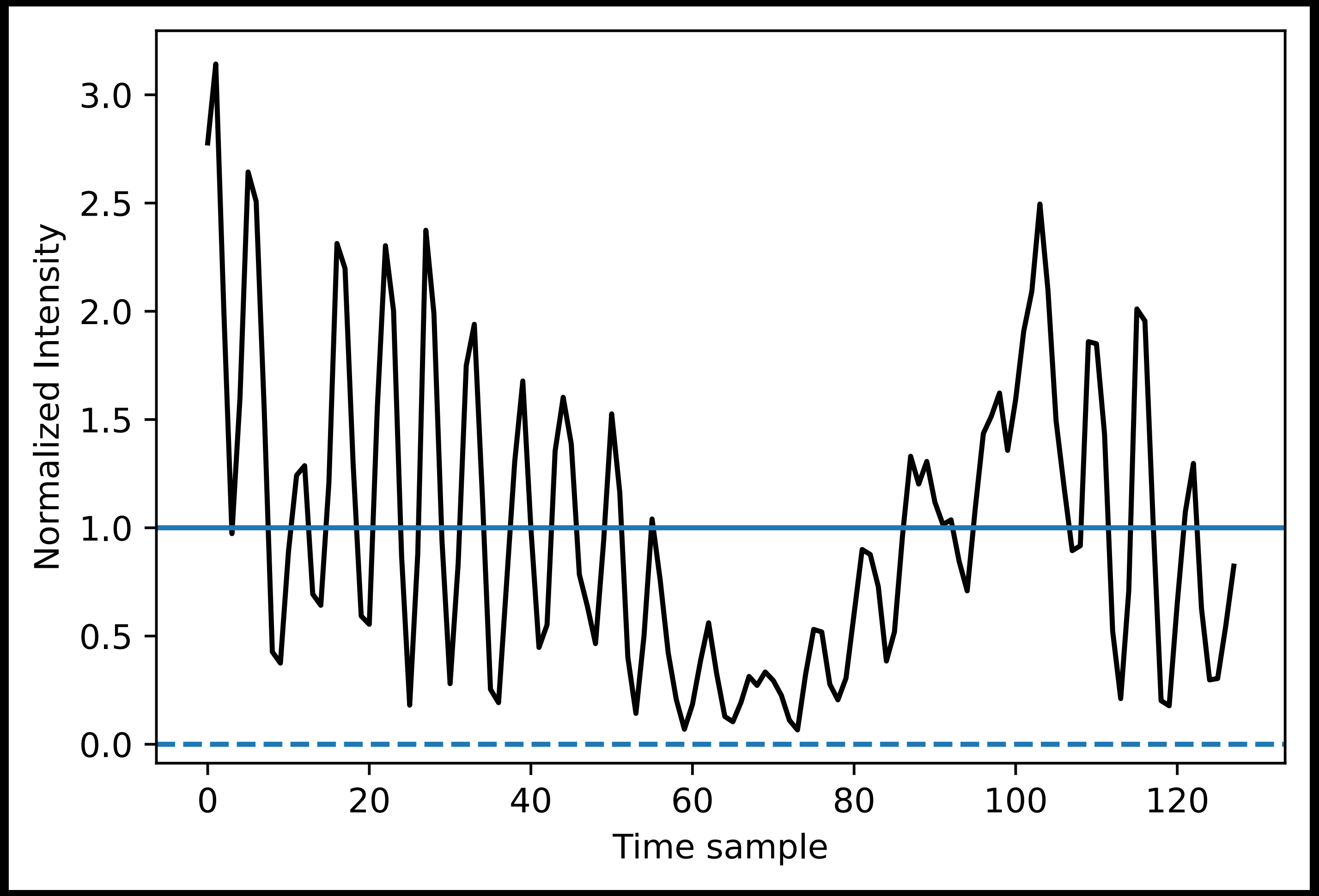
Median



Given a signal… is it scintillated?

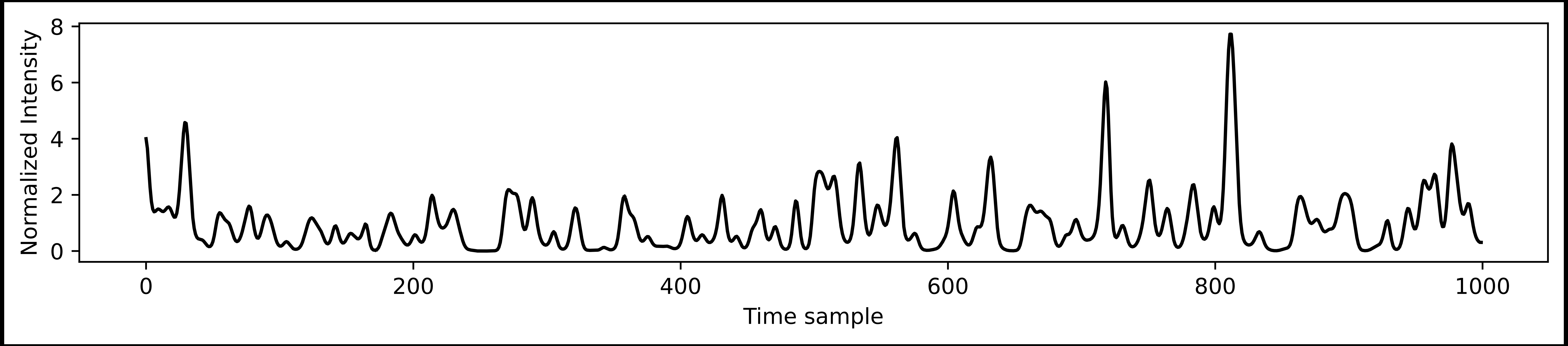
- Create bounding box around narrowband signal
- Estimate noise-subtracted intensity time series, normalized to mean 1
- Compute diagnostic statistics that pertain to asymptotic behavior
  - E.g. standard deviation, Kolmogorov-Smirnoff statistic, fit to autocorrelation function





Quick way to produce synthetic data with asymptotic statistics

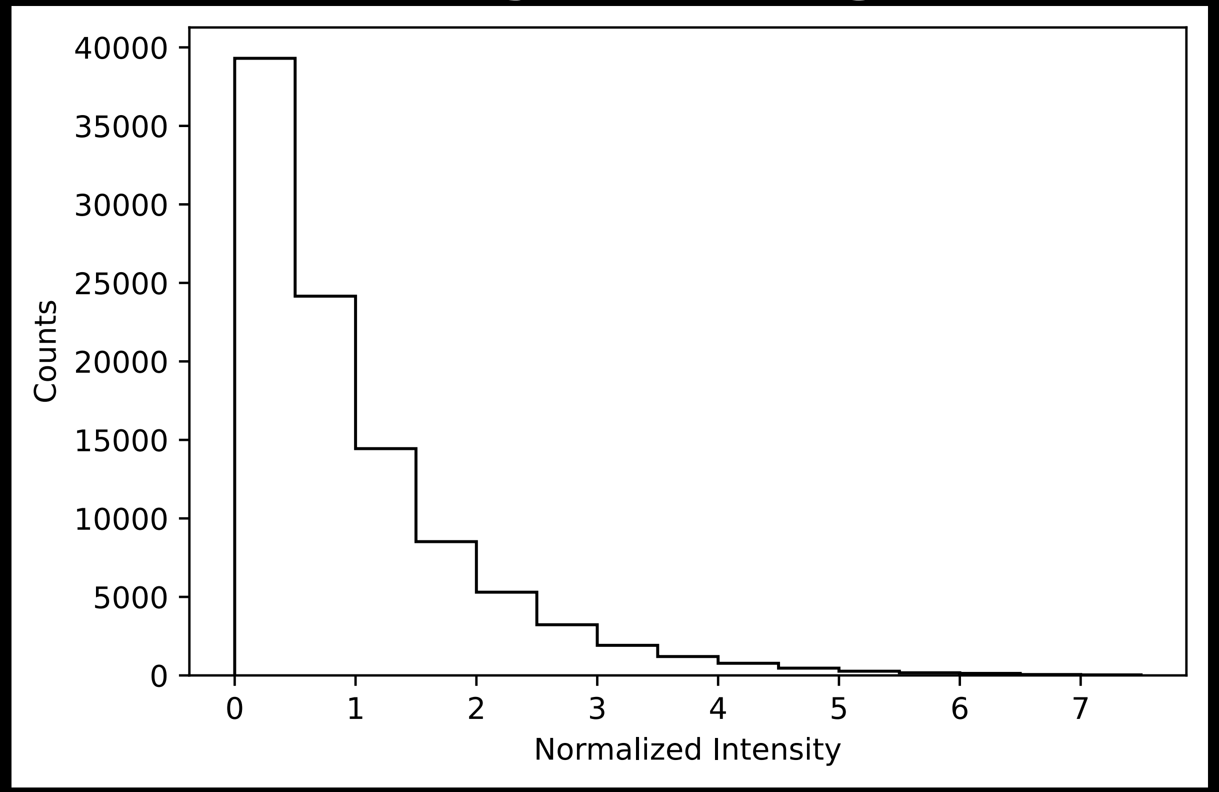
- (Cario & Nelson 1996) The ARTA random process matches:
    - Target intensity distribution
    - Target autocorrelation structure (with custom asymptotic precision)



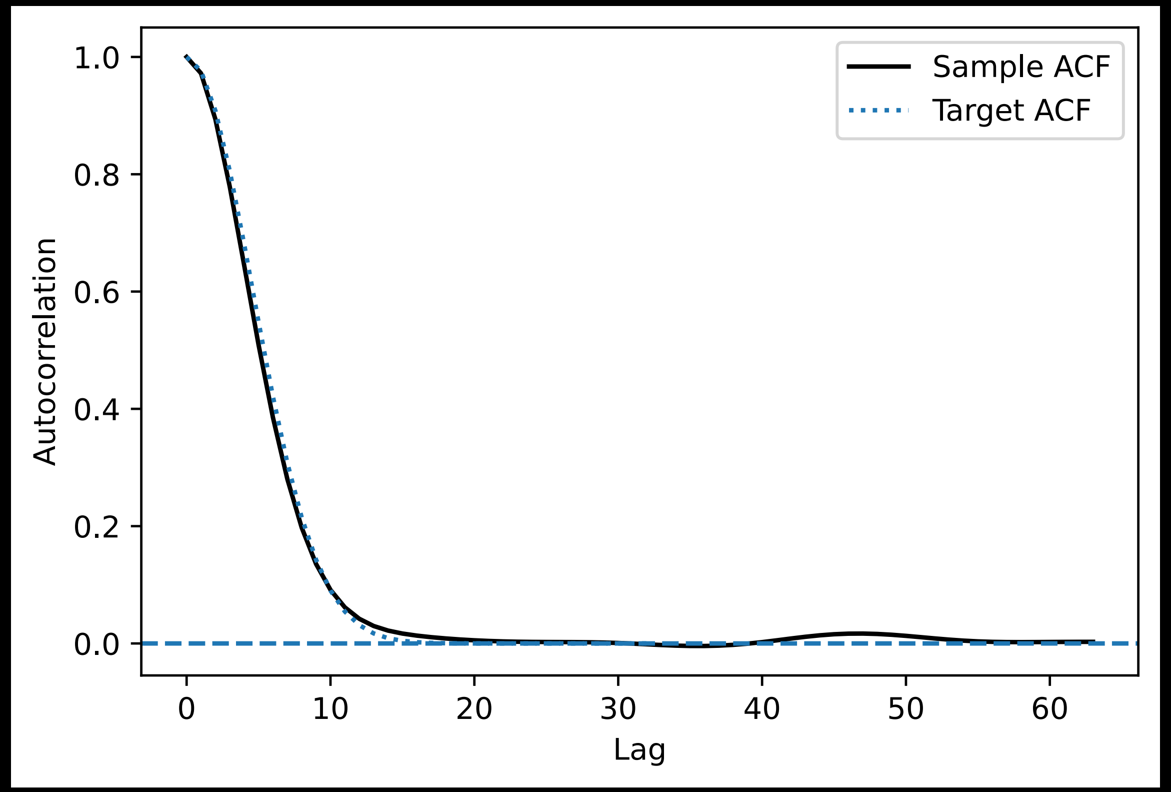
n = 1e5

$\Delta t_d = 30$ s

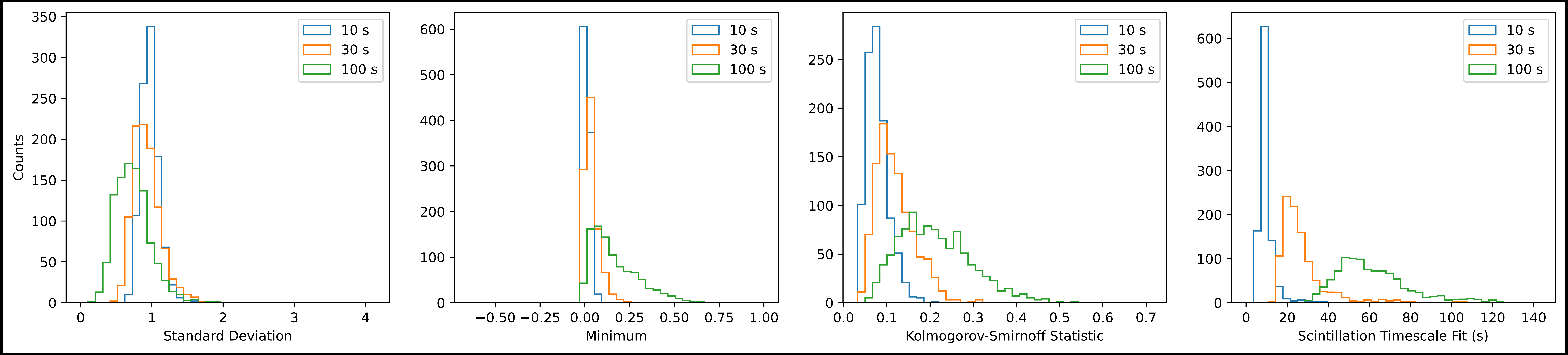
Intensity histogram



Autocorrelation



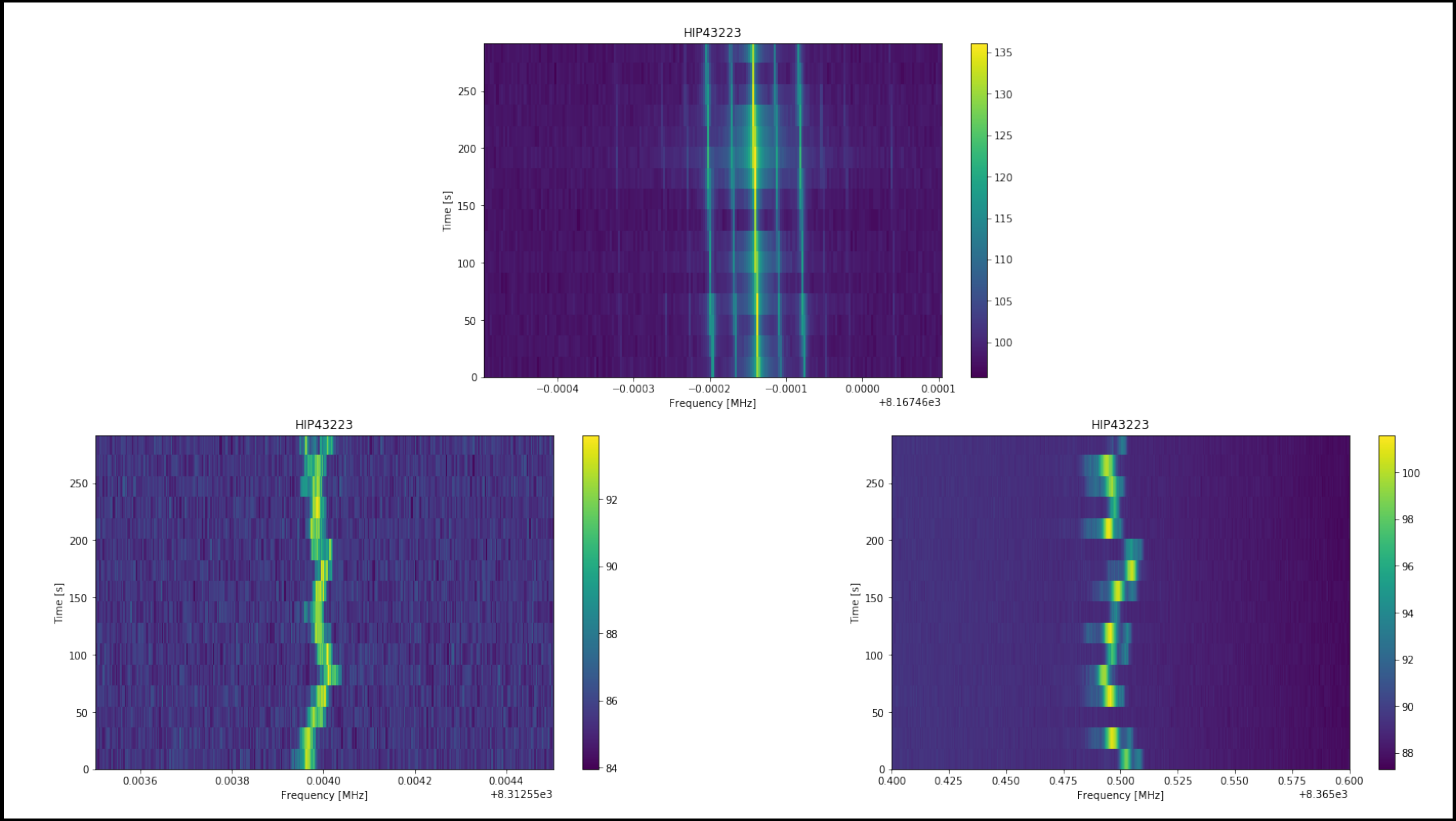
Statistics using low number of synthetic samples

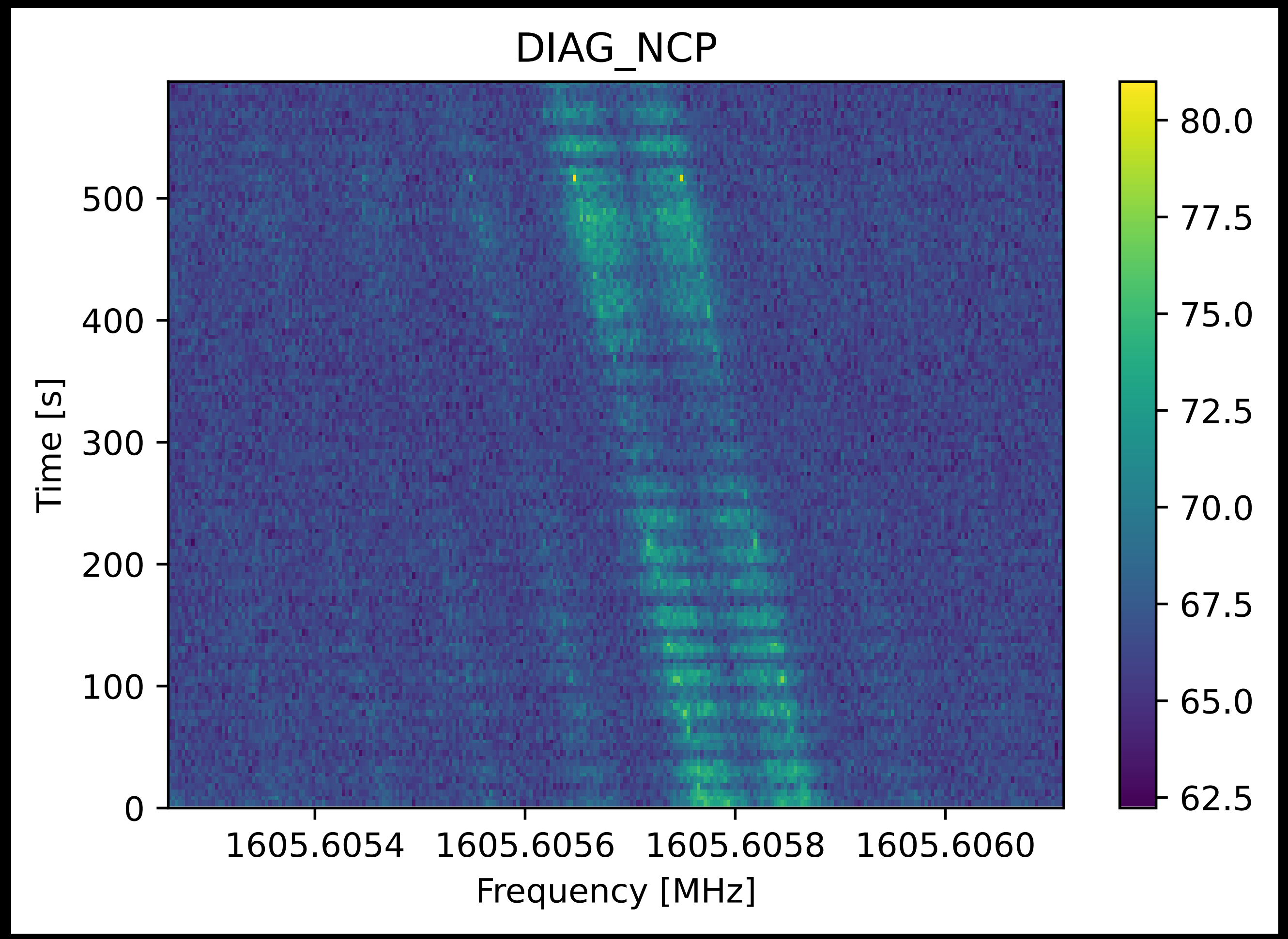


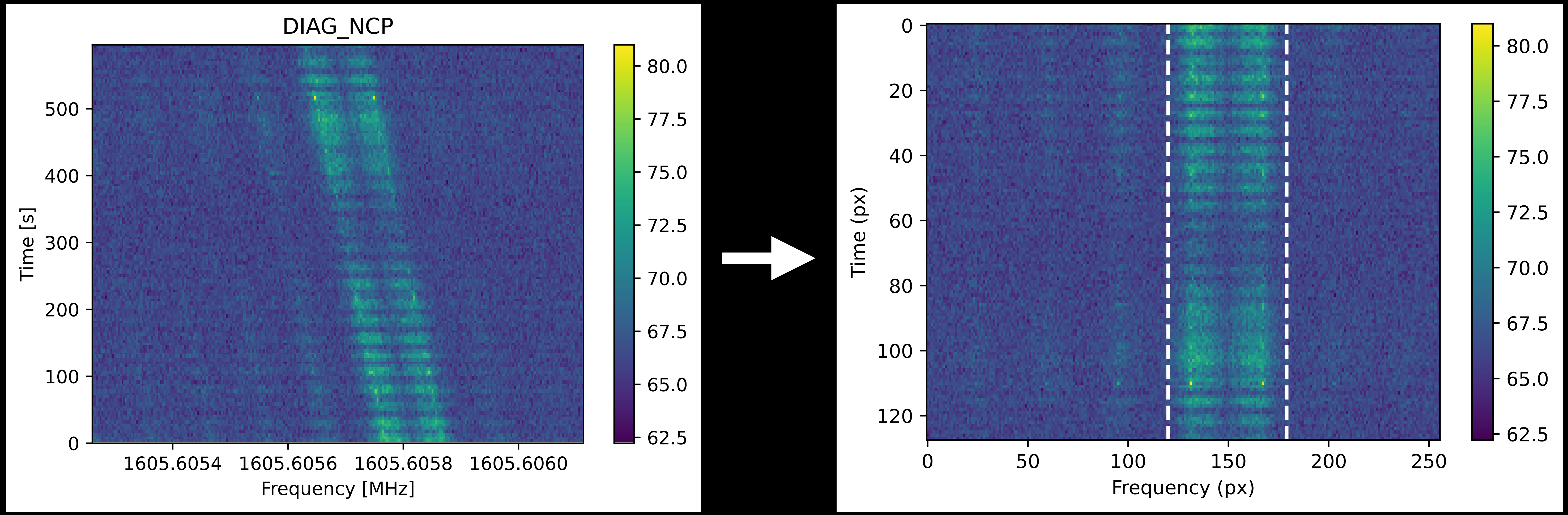
Std. Dev. | Minimum | KS Statistic | Timescale Fit

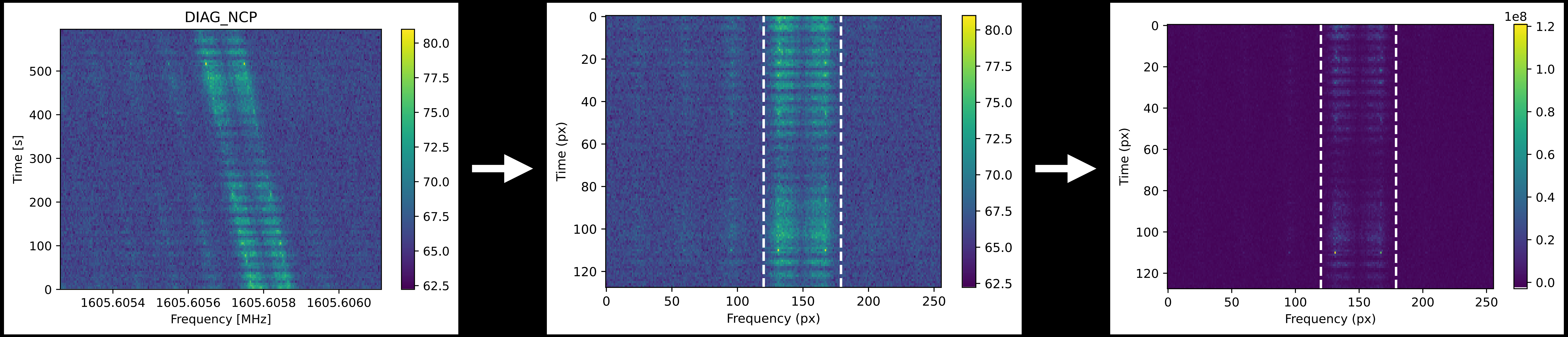
10 min “observation”, 4.65 s resolution

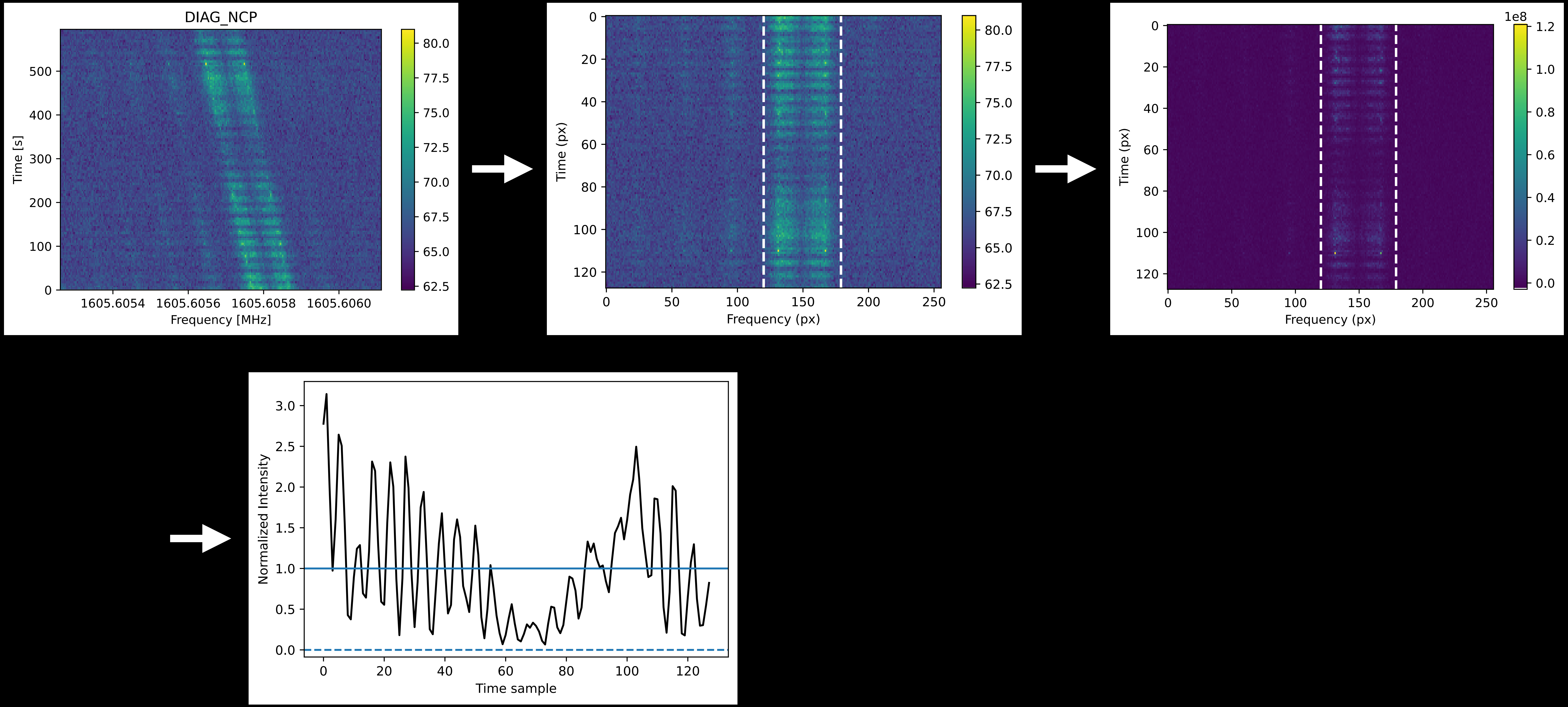
But what does the RFI environment look like?

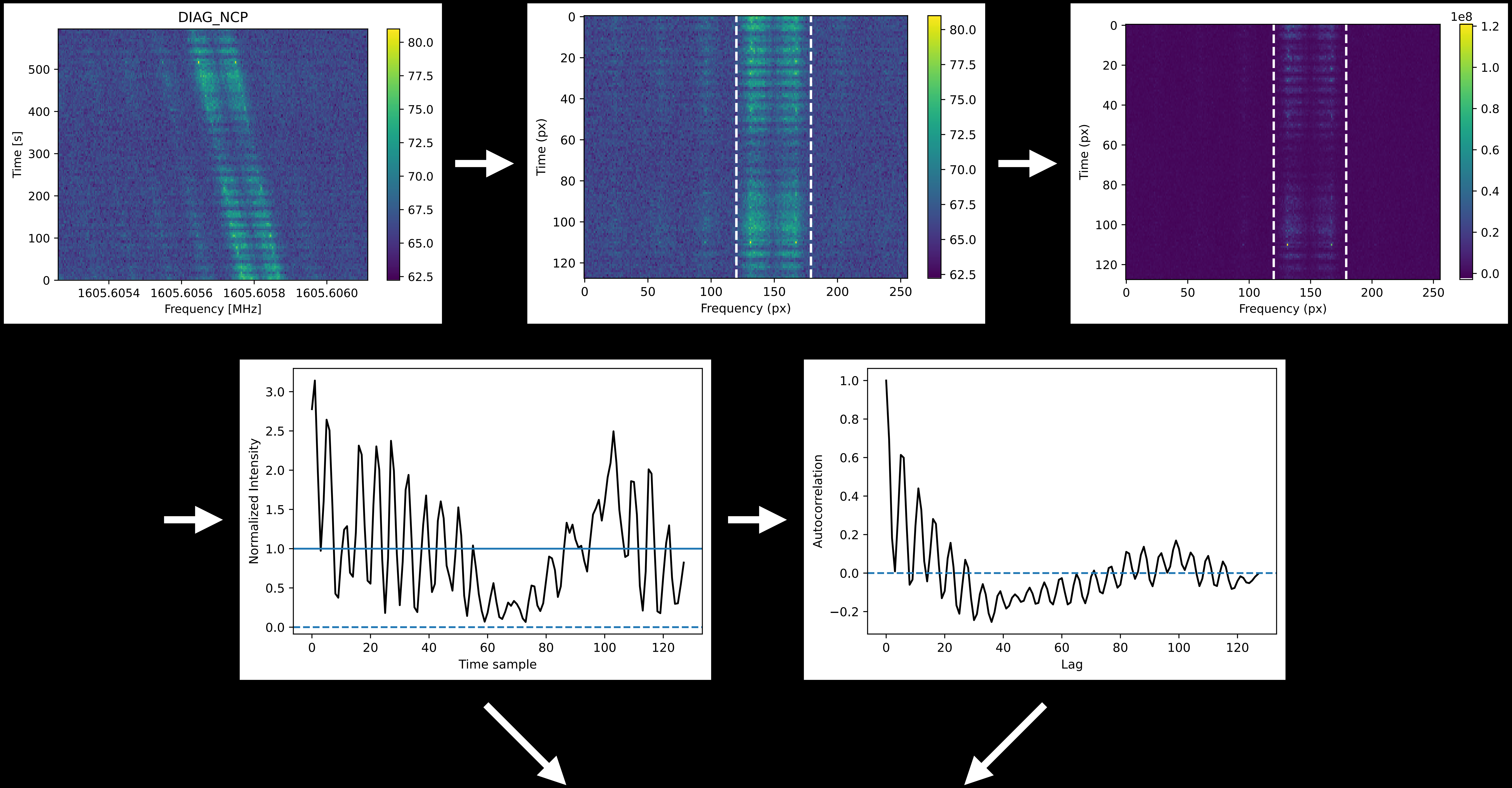



DIAG_NCP
Time [s]
500
400
300
200
100
0
1605.6054
1605.6056
1605.6058
1605.6060
Frequency [MHz]
80.0
77.5
75.0
72.5
70.0
67.5
65.0
62.5


DIAG_NCP
Time [s]
0
100
200
300
400
500
1605.6054
1605.6056
1605.6058
1605.6060
Frequency [MHz]
80.0
77.5
75.0
72.5
70.0
67.5
65.0
62.5
Time (px)
0
20
40
60
80
100
120
0
50
100
150
200
250
Frequency (px)
80.0
77.5
75.0
72.5
70.0
67.5
65.0
62.5


DIAG_NCP
Time [s]
Frequency [MHz]
1605.6054
1605.6056
1605.6058
1605.6060
Time (px)
Frequency (px)
1e8

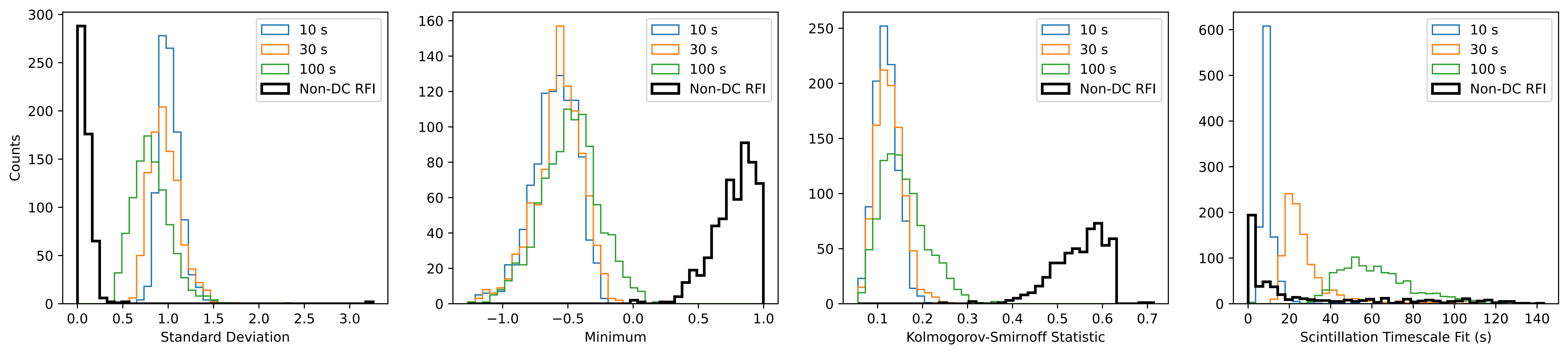

DIAG_NCP
Time [s]
Frequency [MHz]
1605.6054
1605.6056
1605.6058
1605.6060
Time (px)
Frequency (px)
1e8
Normalized Intensity
Time sample



Diagnostic statistics

C band

S/N = 25



Std. Dev.

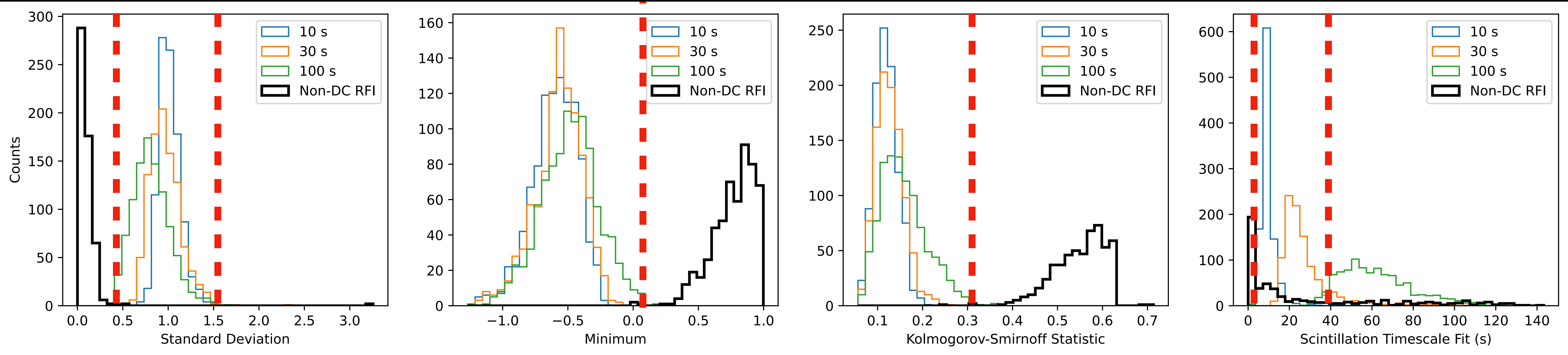
Minimum

KS Statistic

Timescale Fit

C band

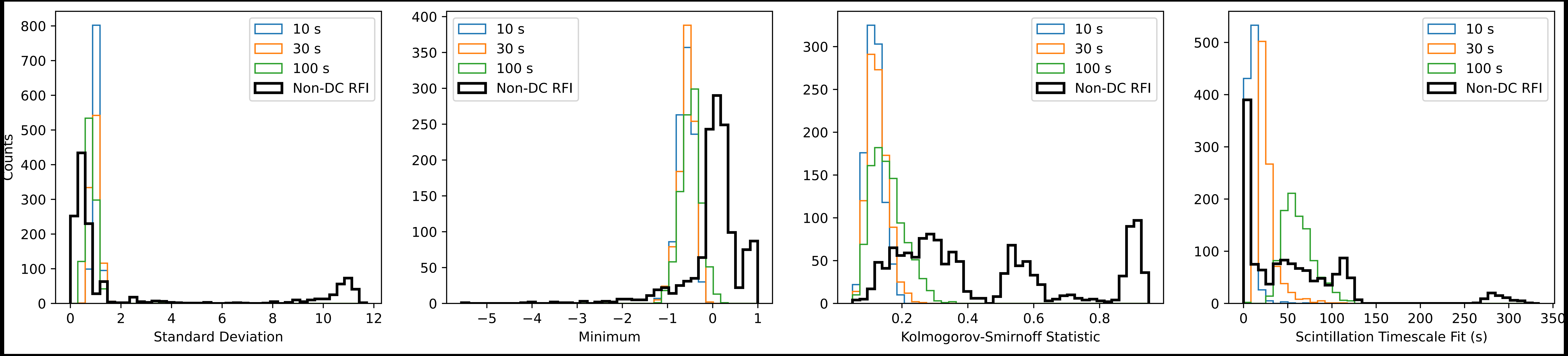
S/N = 25



Std. Dev. **Minimum** **KS Statistic** **Timescale Fit**

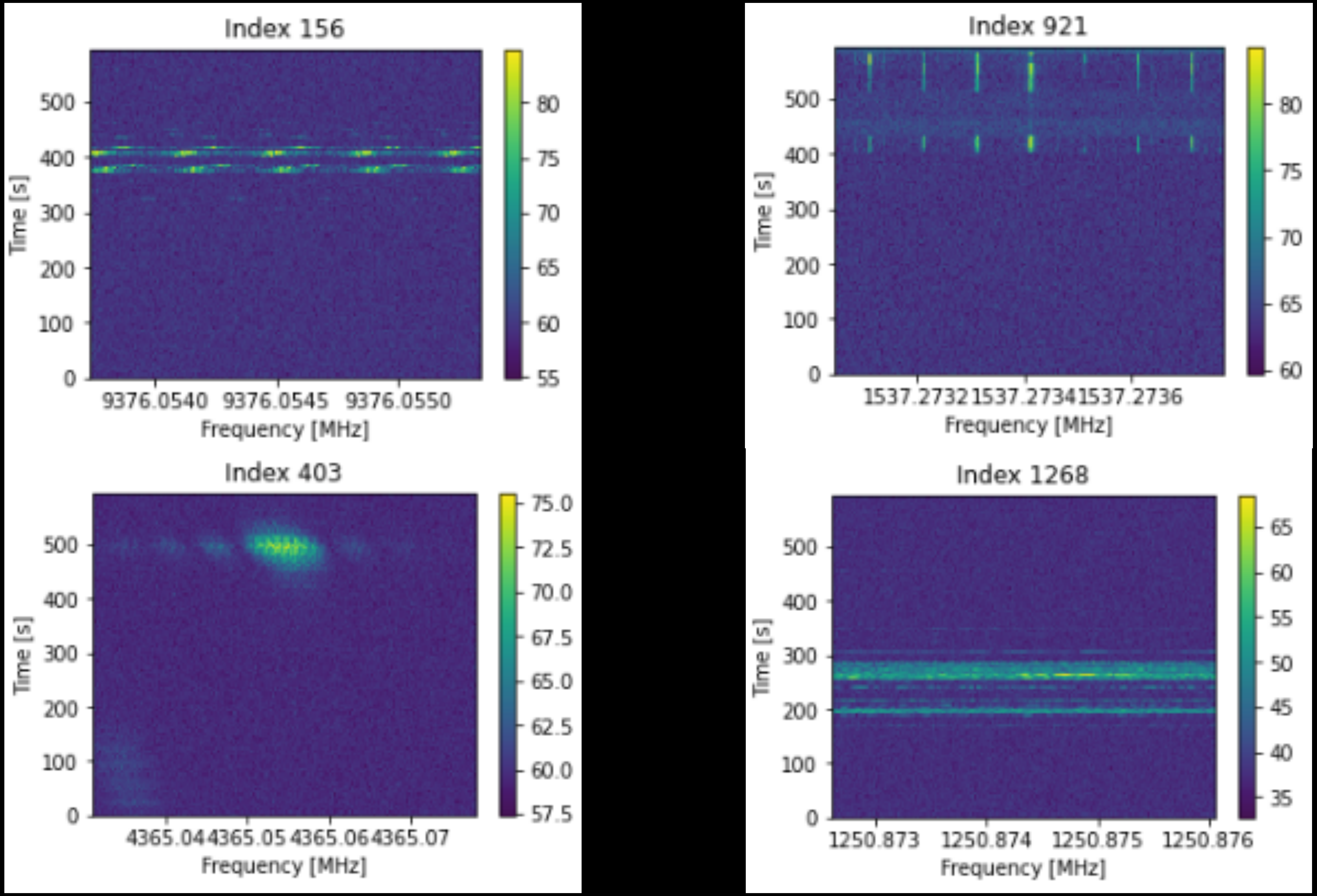
L band

S/N = 25

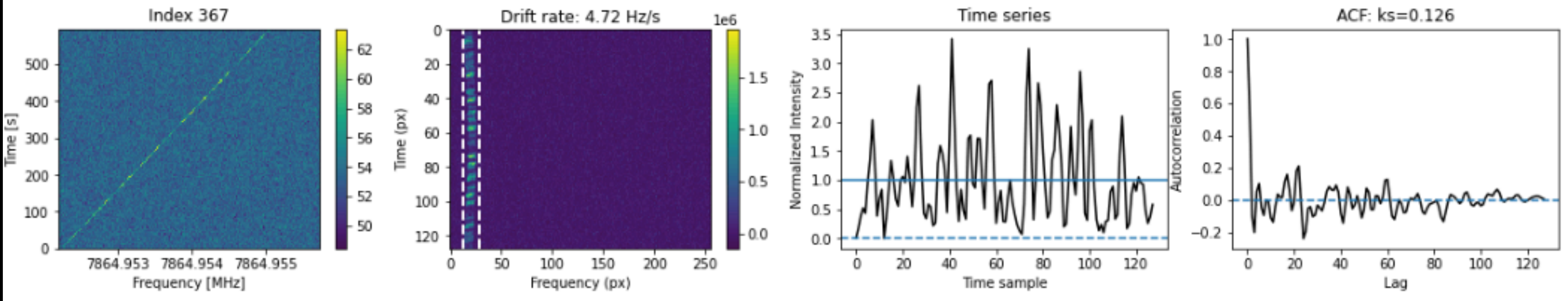


Std. Dev.   Minimum   KS Statistic   Timescale Fit

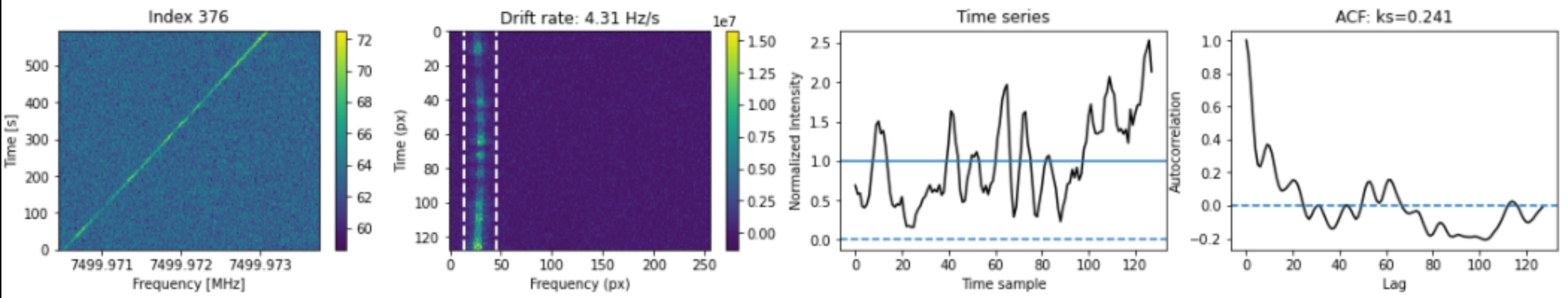
High standard deviation (RMS) signals are pulsed - or broadband



What signals pass the threshold?



Timescale fit ~ 2 s

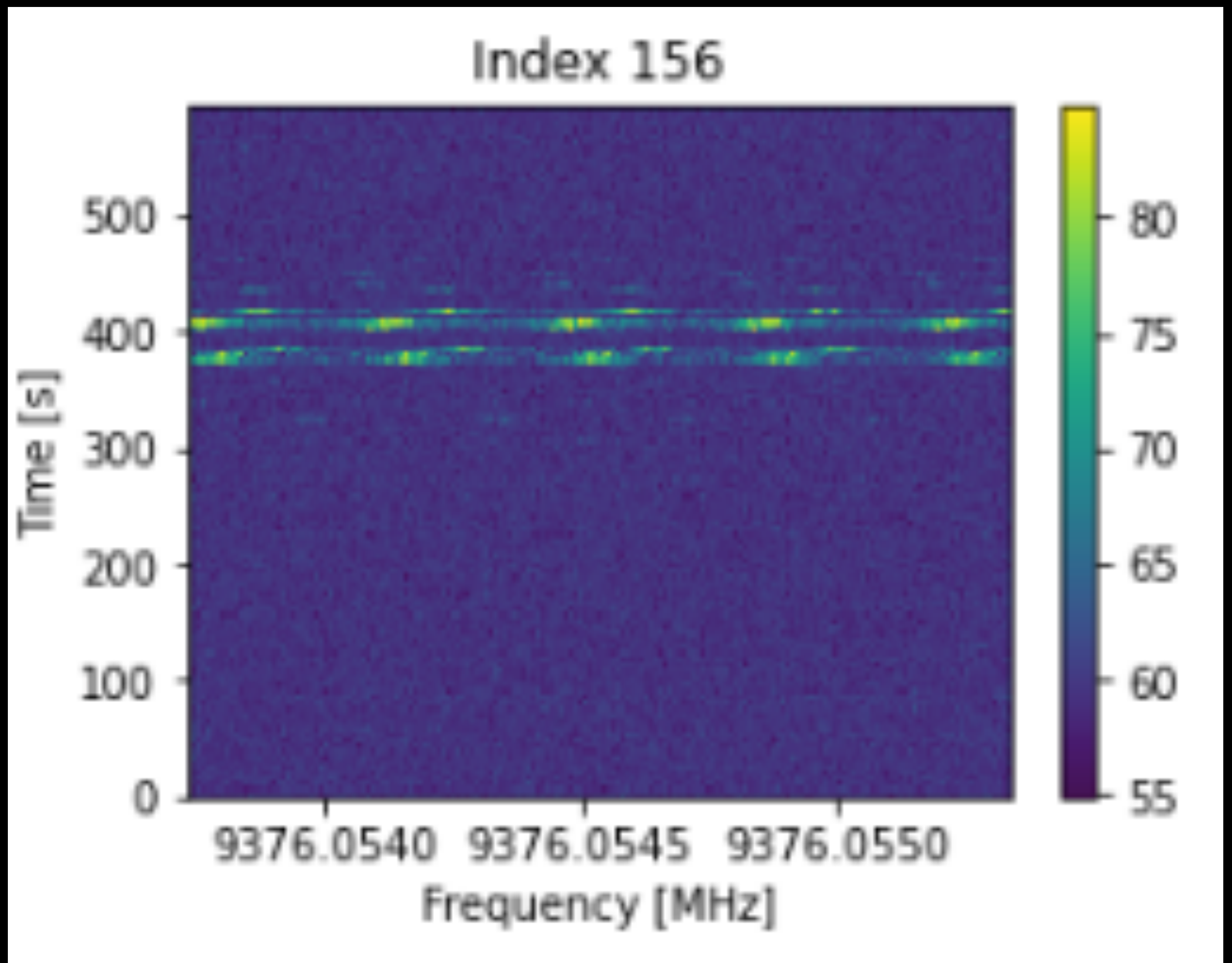


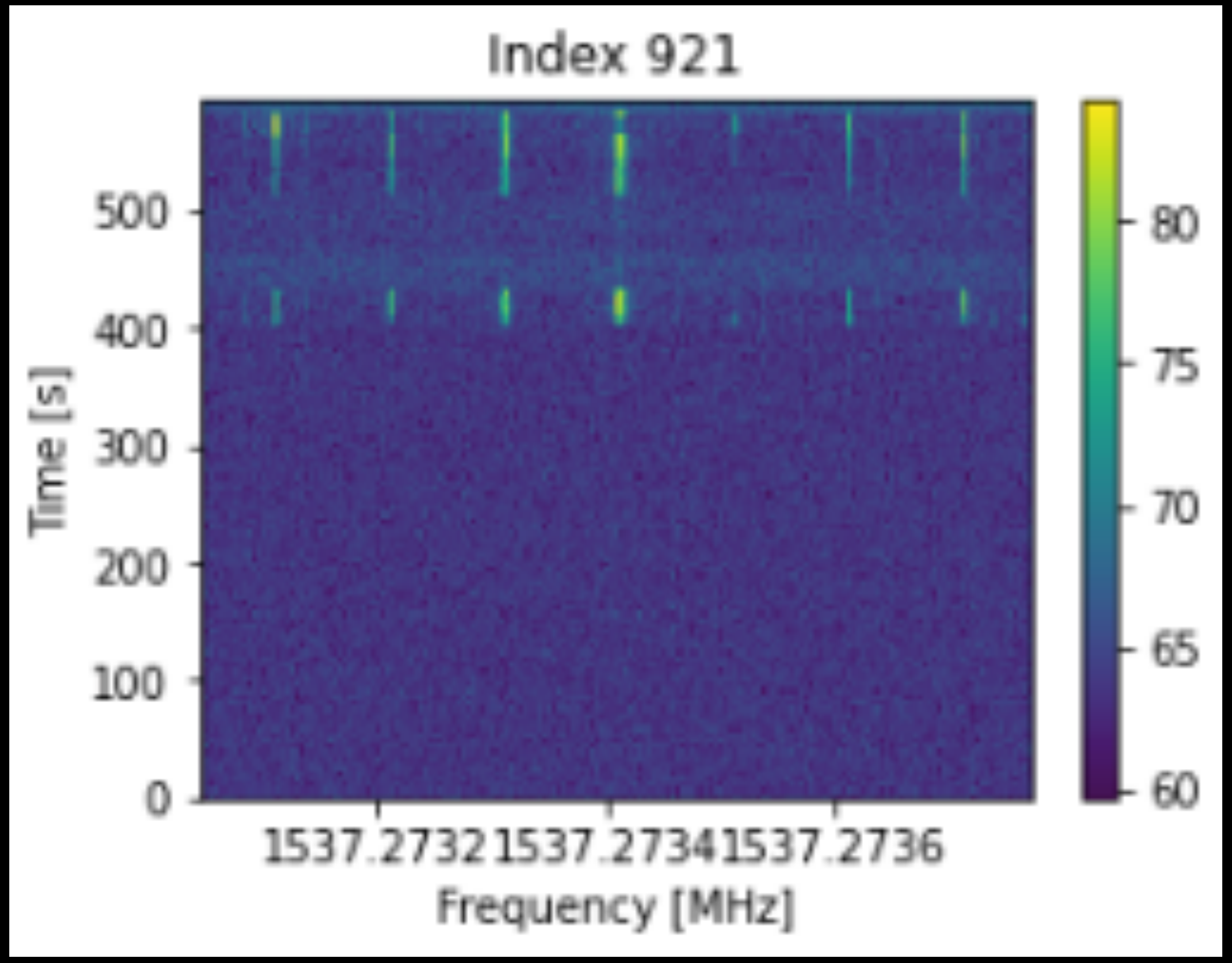
Timescale fit ~ 60 s

- At C-band, S/N > 25, 3 out of 1102

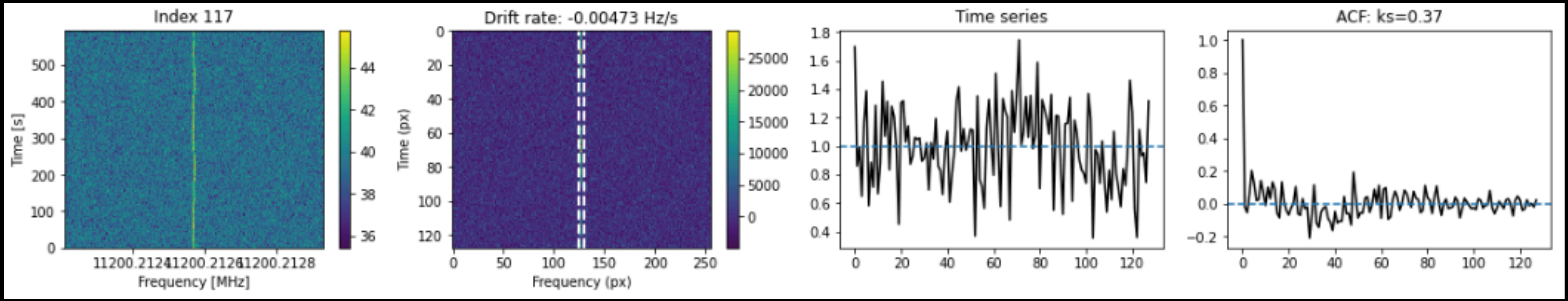
Limitations from RFI analysis?

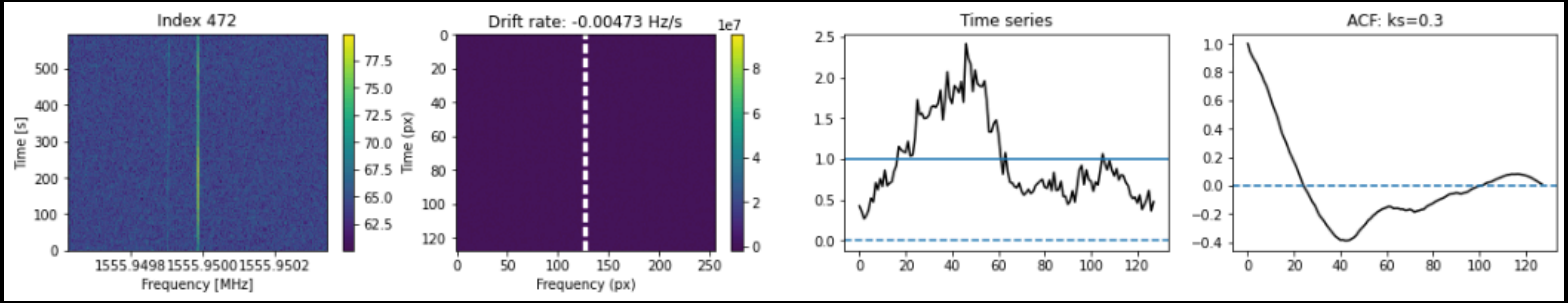
- L and S bands in particular are very noisy
- Non-narrowband signals detected just because they are above the SNR threshold
- Difficult to apply a one-size-fits-all bounding box method
- Perhaps ML can help!



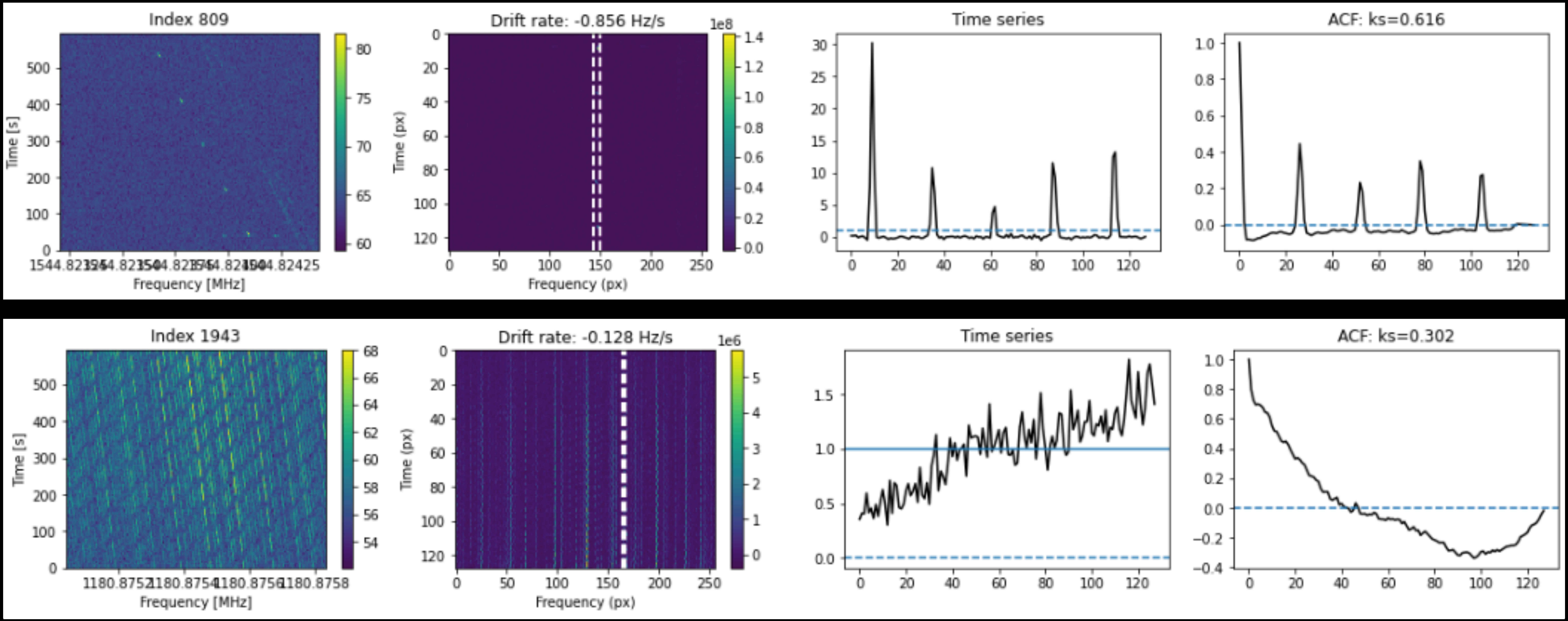


Some examples





Some more examples



Examples of diagnostic statistics

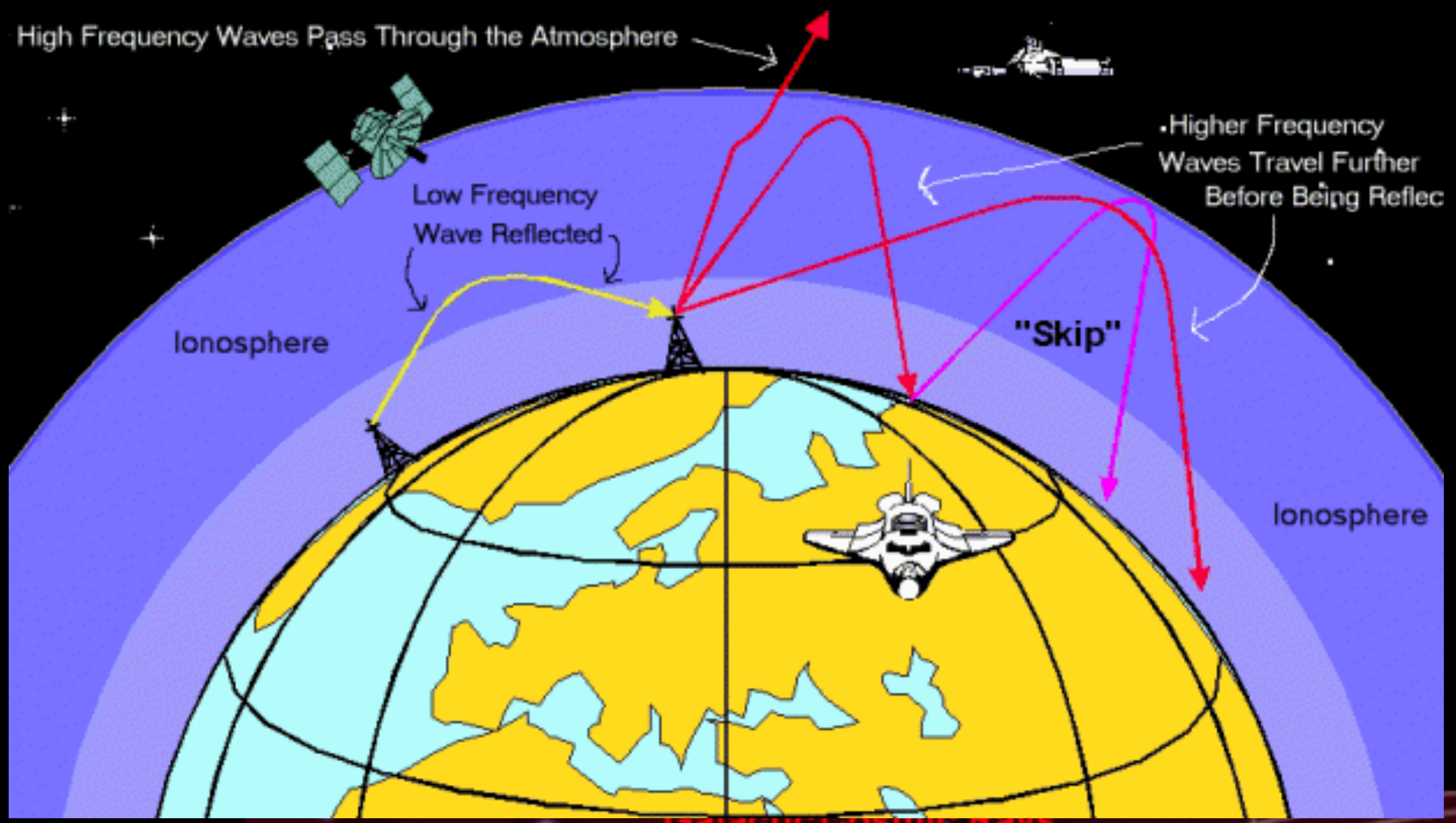
Statistic	Asymptotic Value (with no noise)	Target Distribution
Standard Deviation (RMS)	1	Intensity, exponential
Minimum	0	Intensity, exponential
Kolmogorov-Smirnoff statistic	0	Intensity, exponential
Autocorrelation lag	Variable	Autocorrelation, Gaussian
Least squares fit to autocorrelation	Variable	Autocorrelation, Gaussian

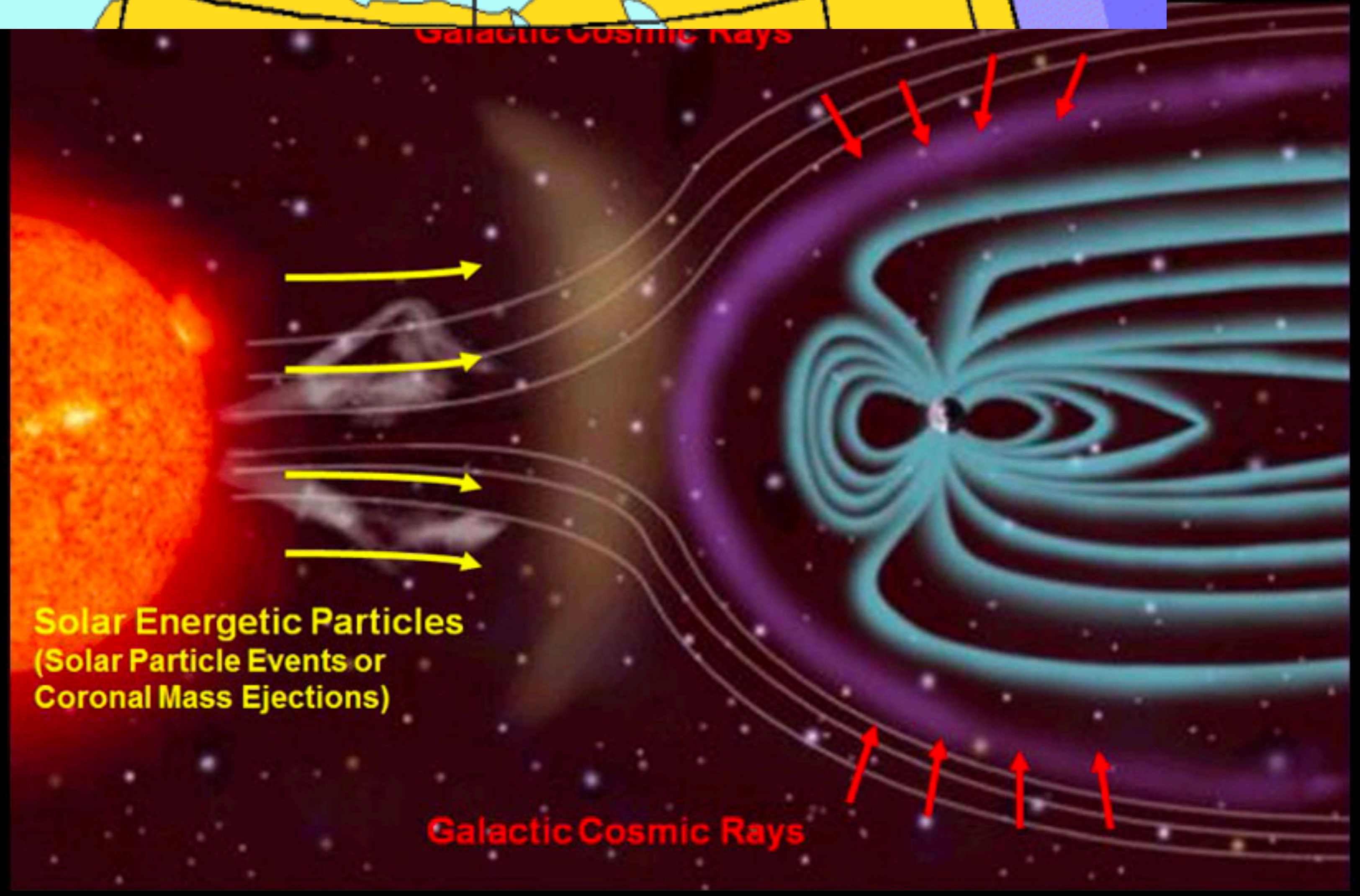
There are a number of constraints…

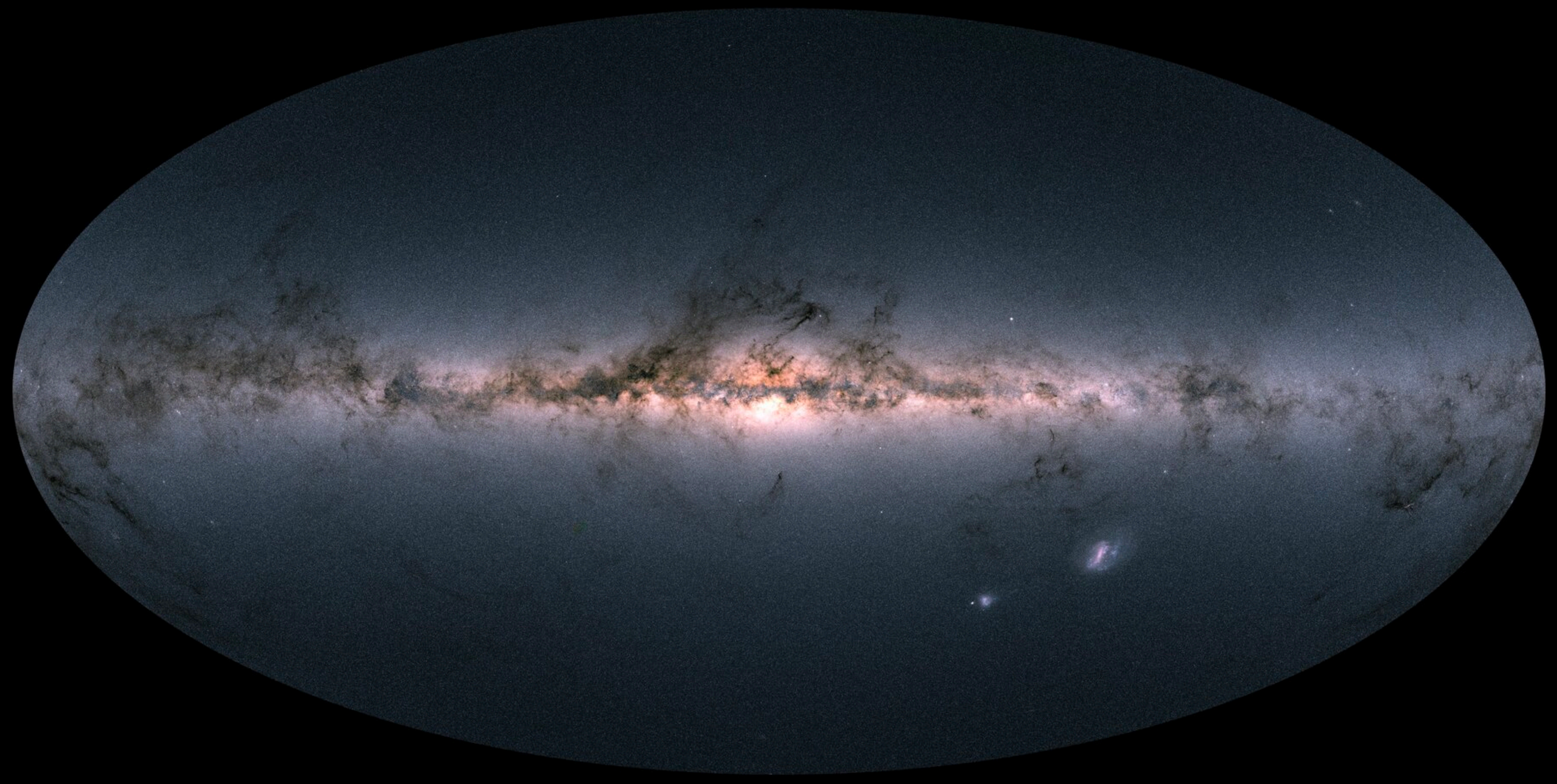
- Time resolution - Sufficiently resolve scintles
- Integration time - Collect enough scintles, gain stability
- Signal brightness - Compute accurate statistics embedded in noise
- RFI environment - Bad normalization, false narrowband detections, confounding modulation

Regions of ionized plasma

- Ionosphere
- Interplanetary Medium (IPM)
- Interstellar Medium (ISM)

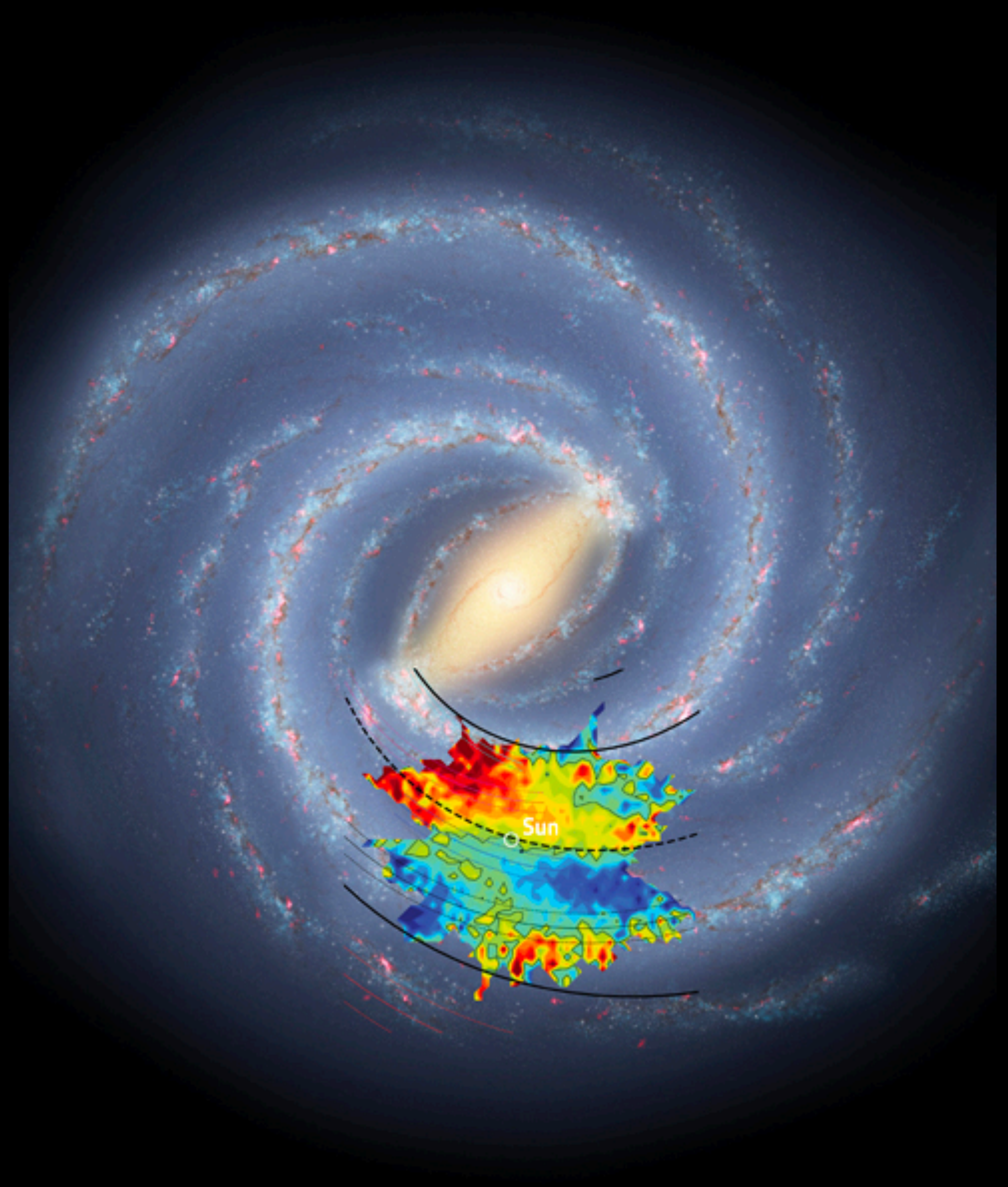






Next steps: a Galactic Center / Galactic Plane survey

- Target most promising sections of parameter space
- Survey of Galactic plane with interesting targets
- Gaia DR3?



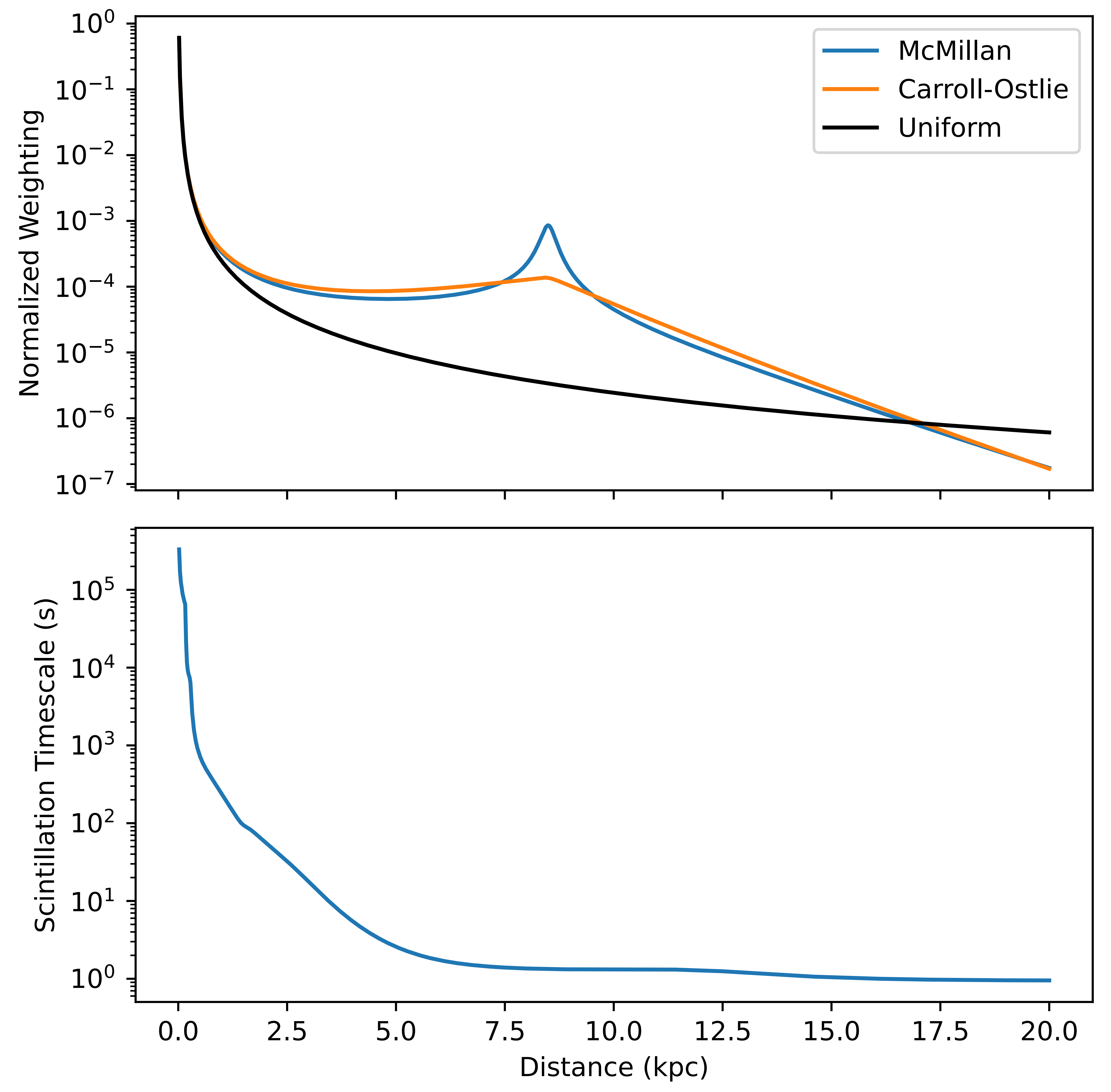
Gaia

Density-based sampling

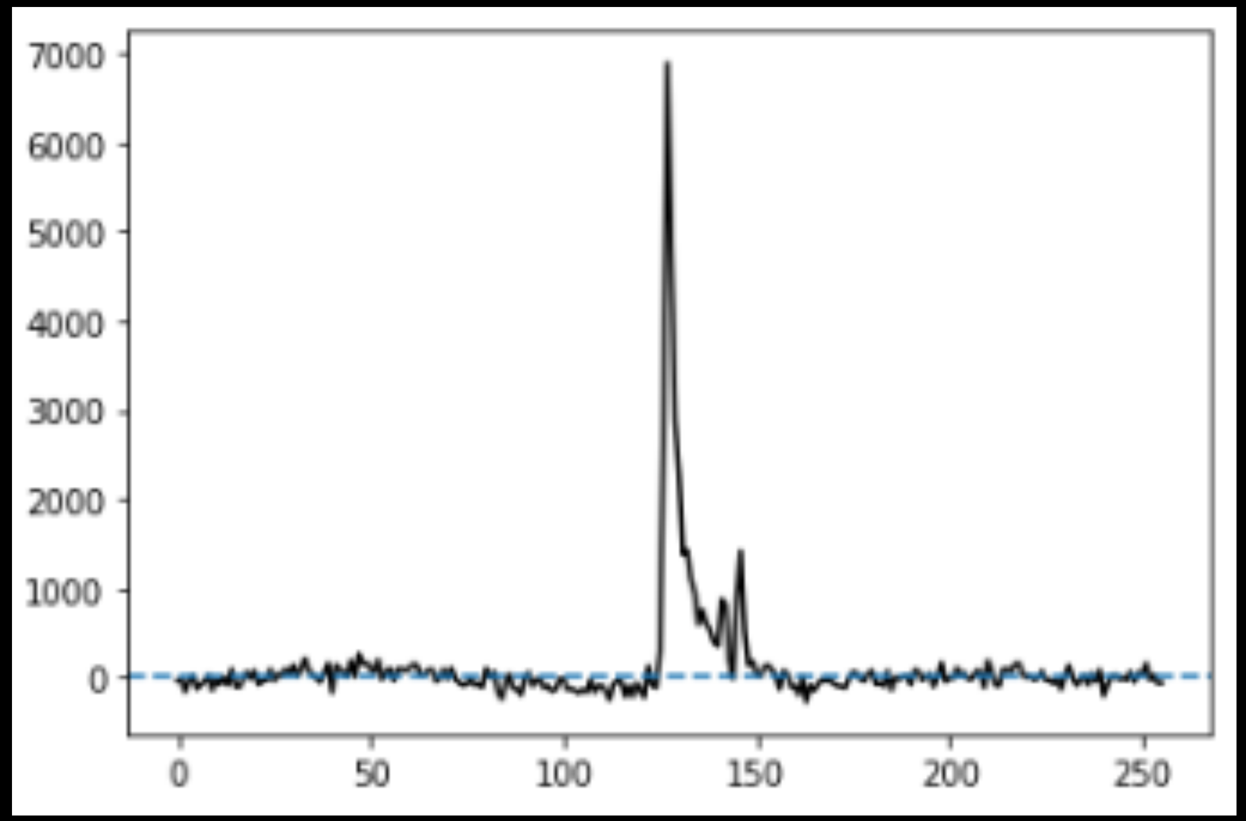
Modulating by the inverse square-law for detectability:

Depends on the assumptions made about transmission power and resources.

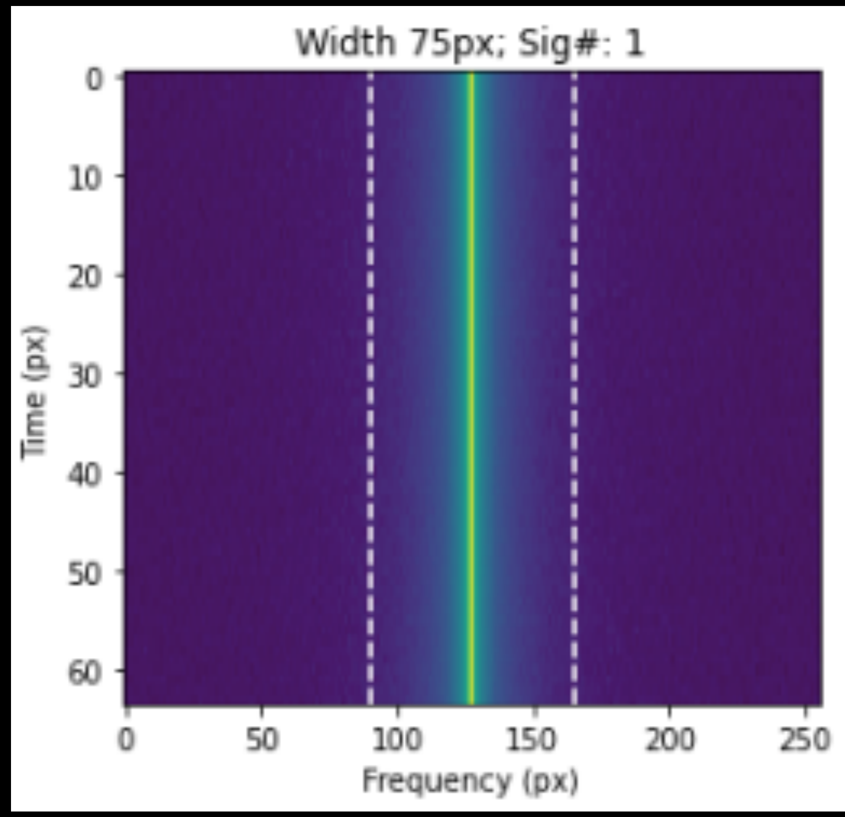
(l, b) = (1, 0)



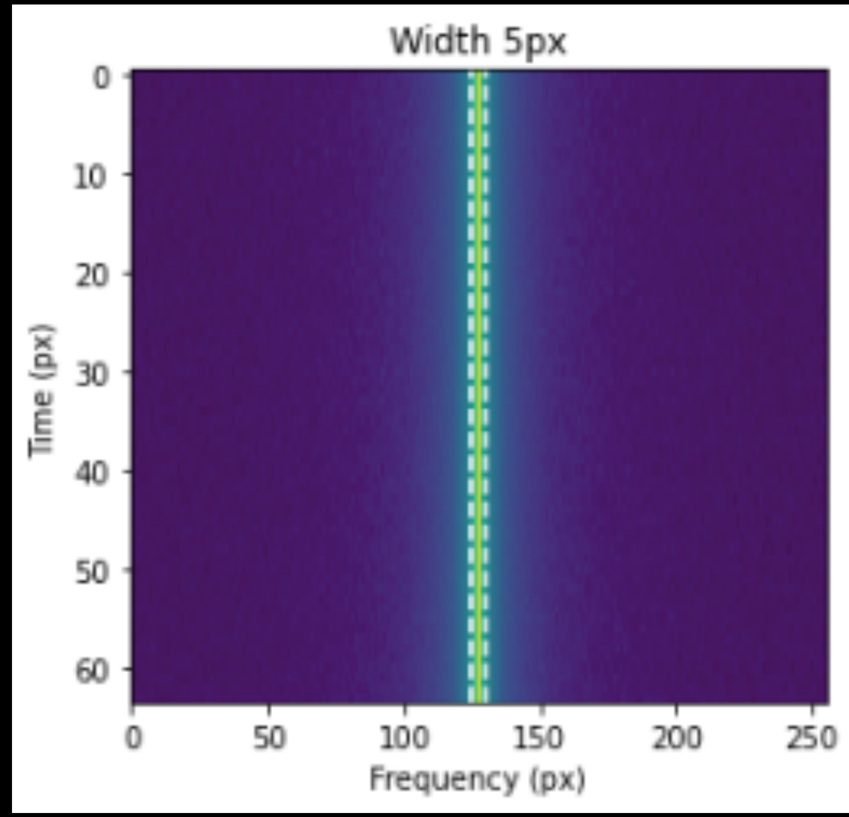
Selecting bounding boxes



- After experimentation with various methods, the final pipeline uses a combination of baseline fitting and peak detection to calculate the right size of frame to use
- The final bounds are created using a thresholding method, similar to PSRCHIVE
- Take the final bounded signal and integrate in the frequency direction to derive our raw time series — then we normalize to mean of 1 before calculating our scattering statistics



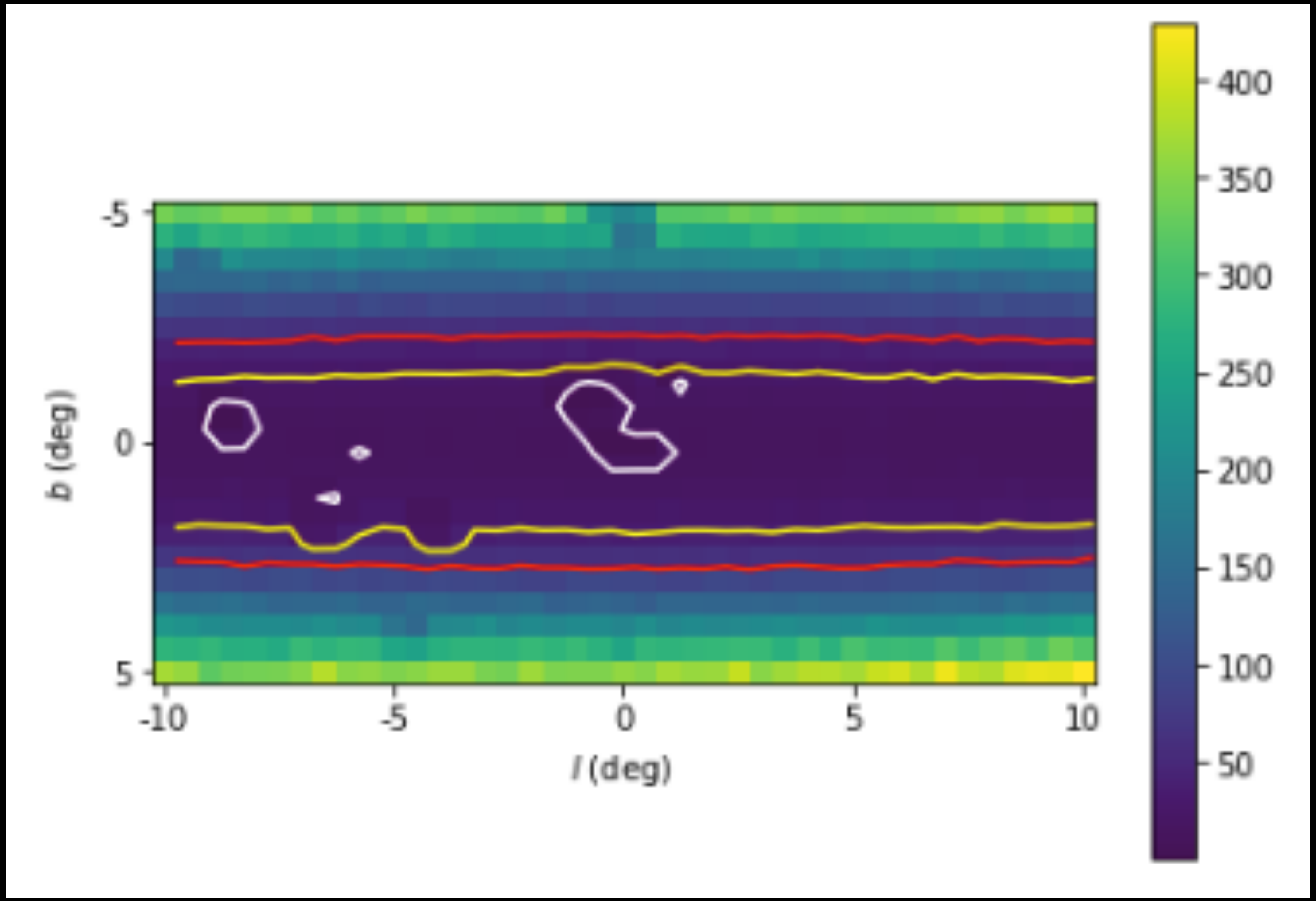
Polynomial fit



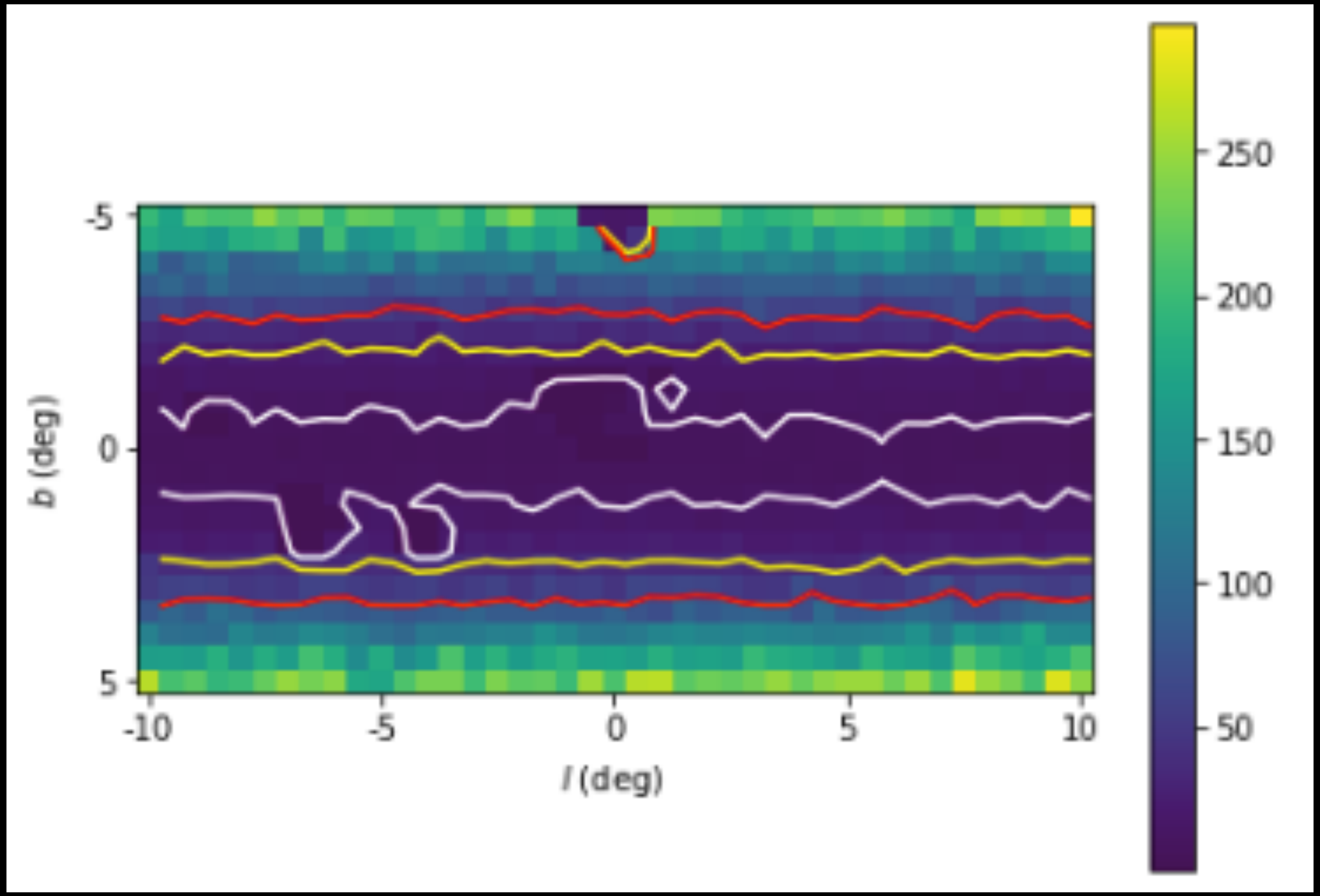
Threshold fit

Scintillation maps around the GC at C-band

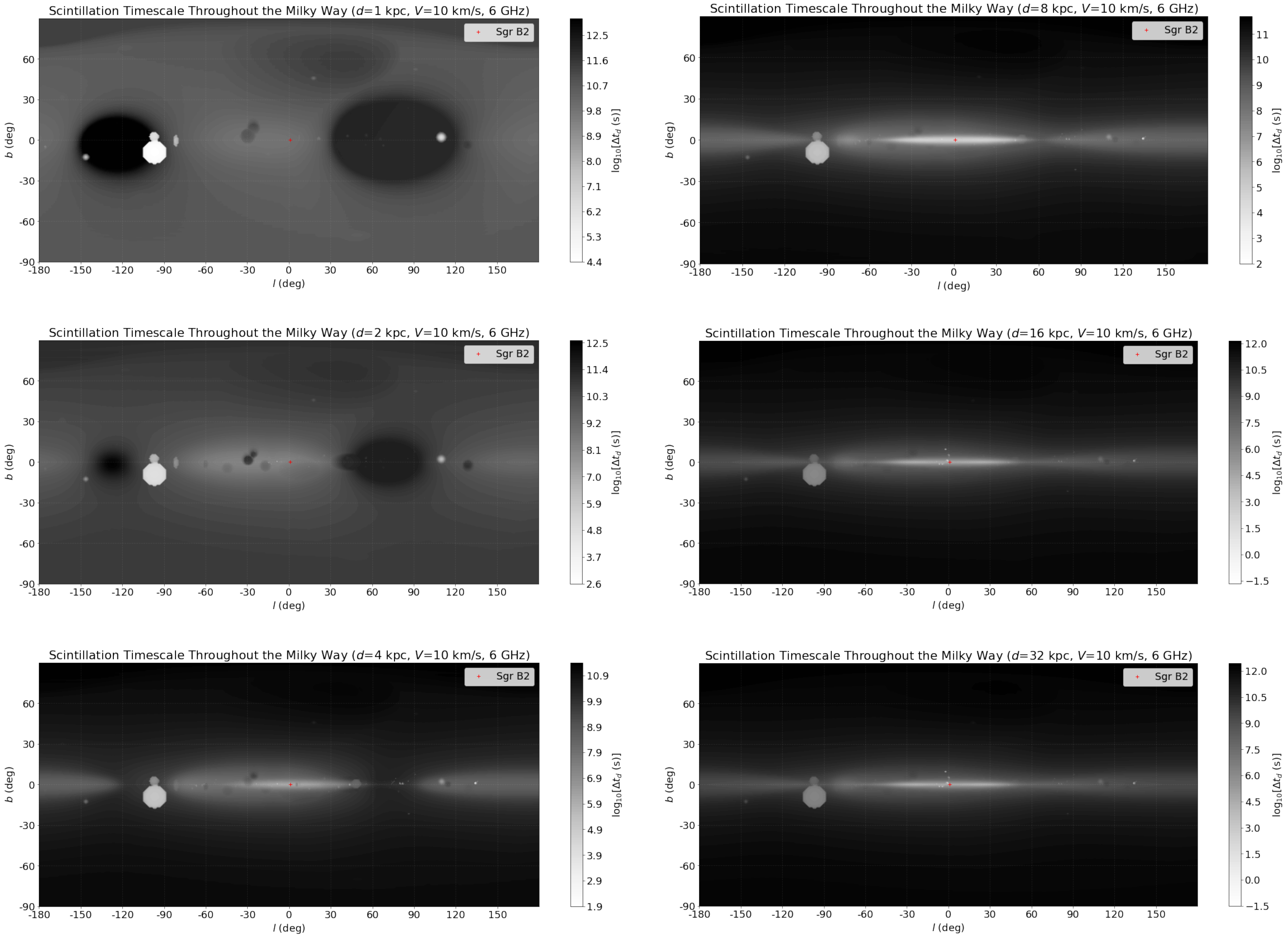
Median



Mode



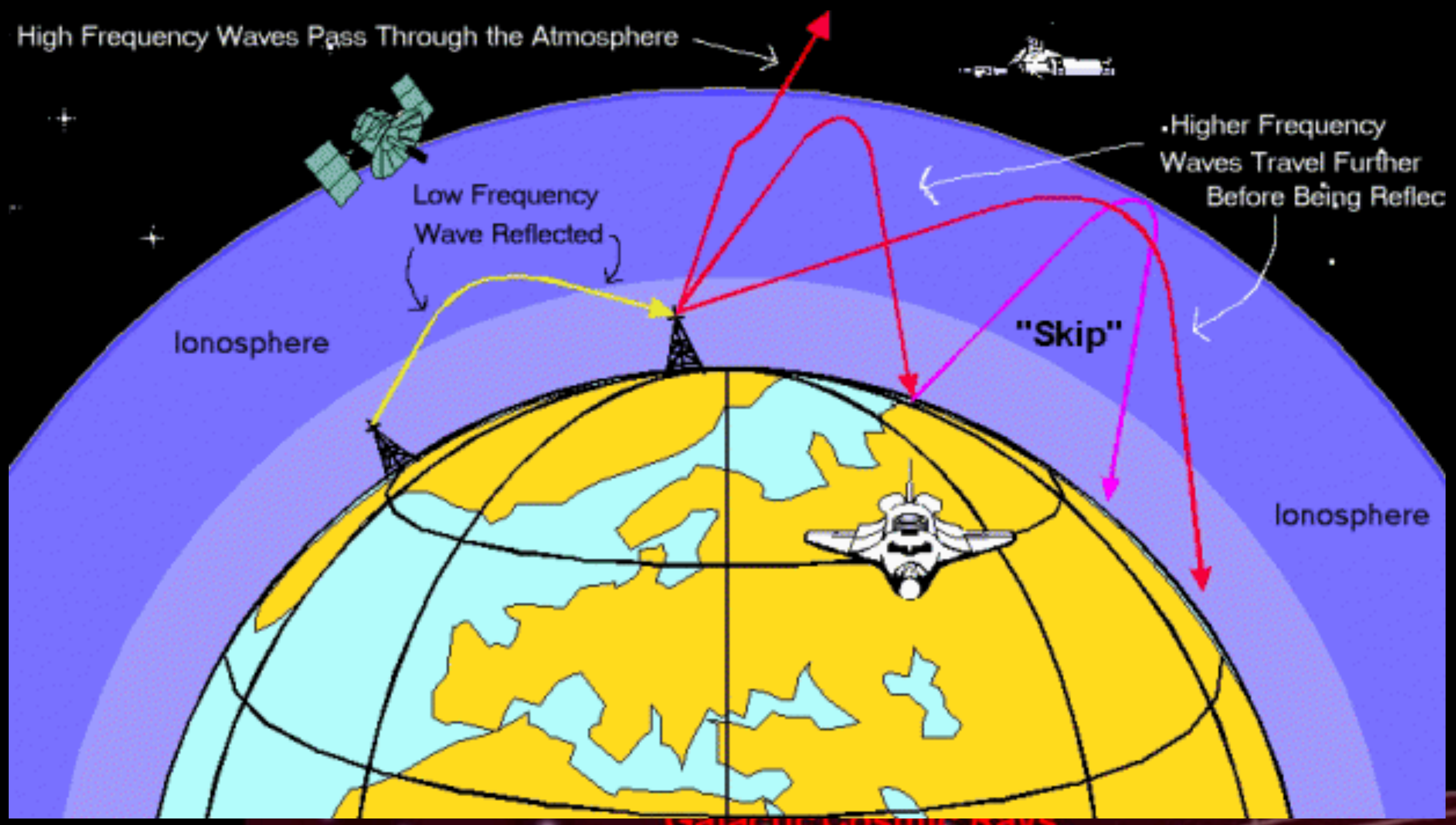
10 s, 30 s, 60 s contours

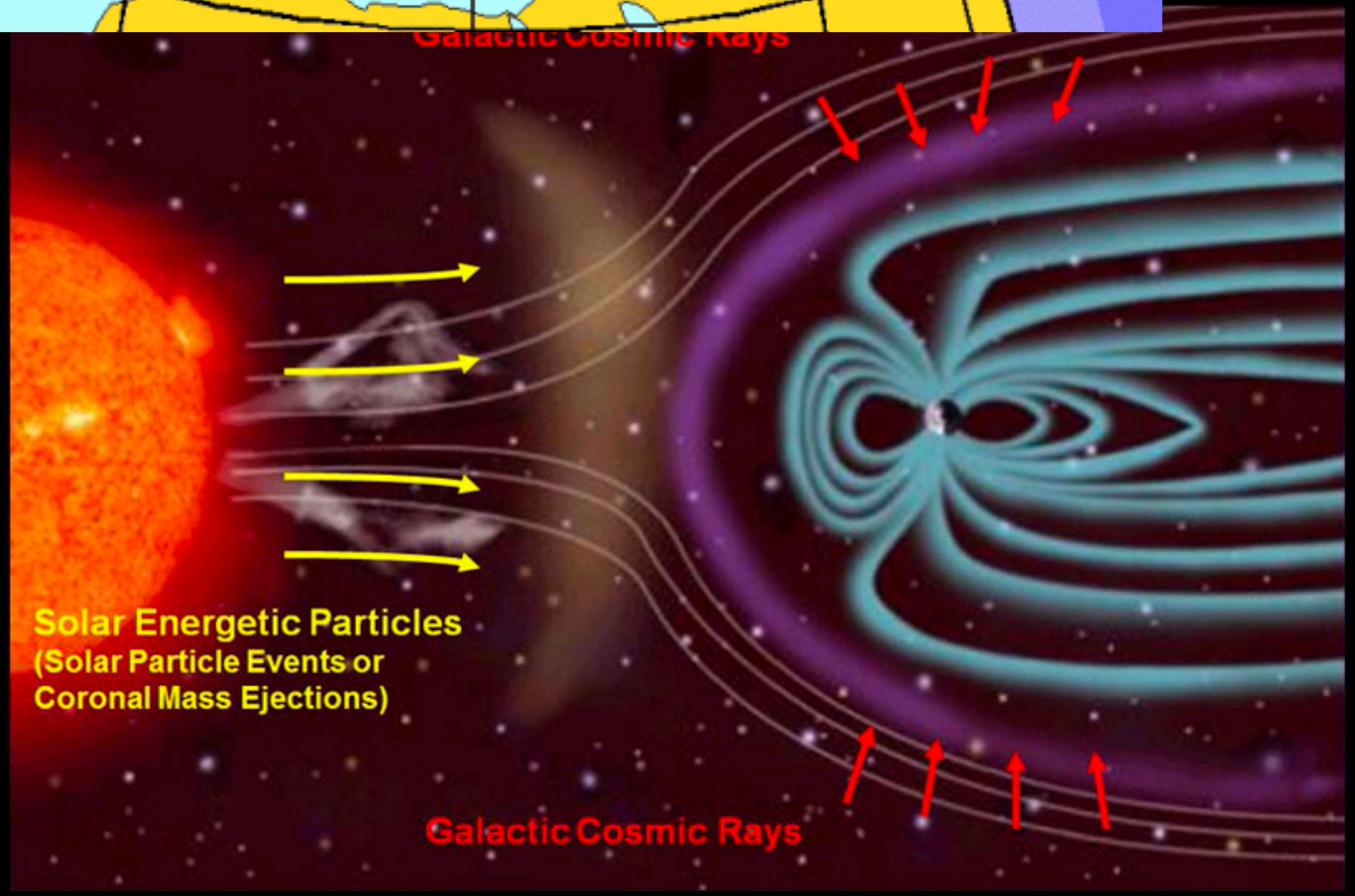


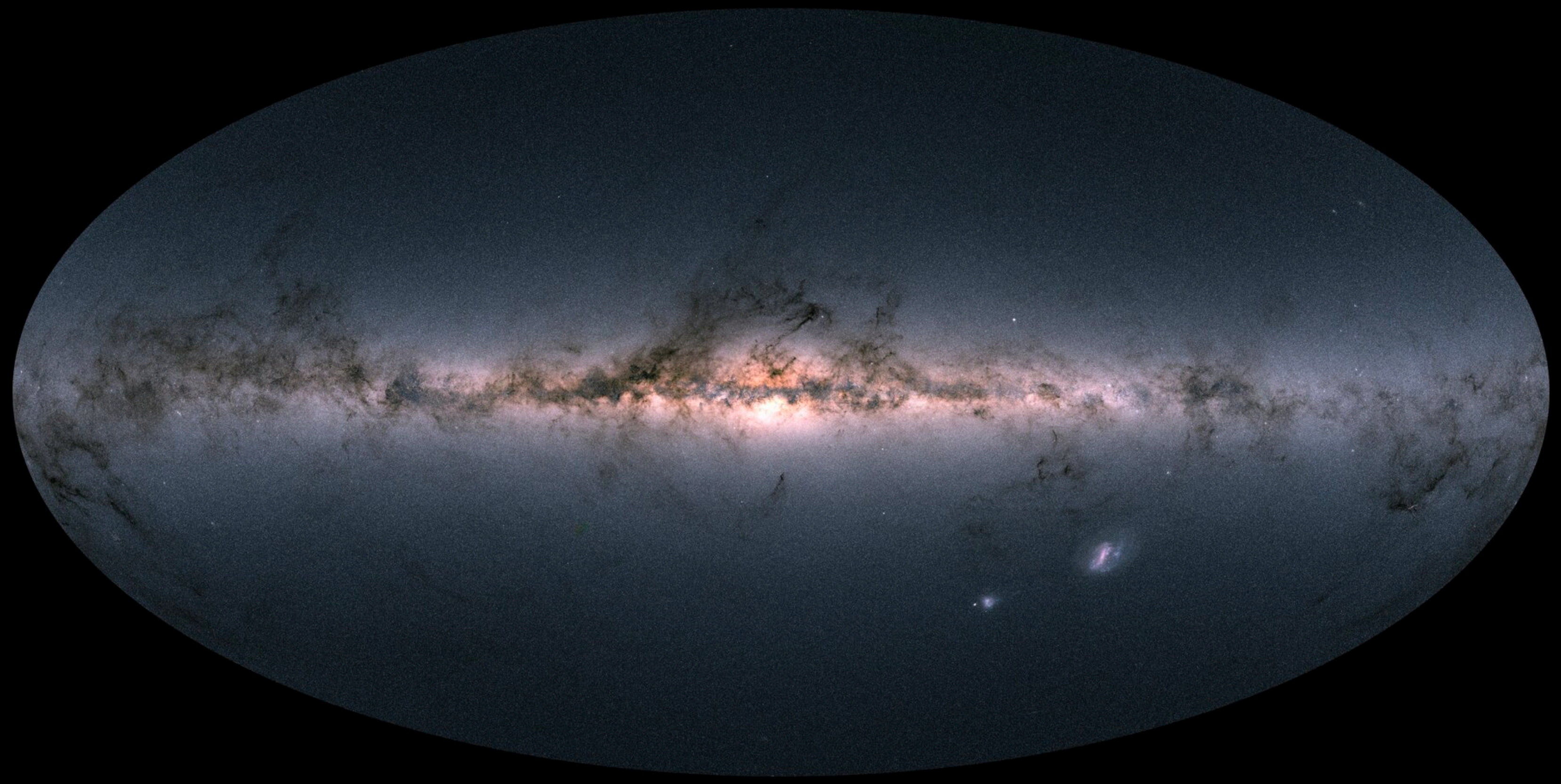
Scintillation Timescale Throughout the Milky Way (d=1 kpc, V=10 km/s, 6 GHz)
Scintillation Timescale Throughout the Milky Way (d=8 kpc, V=10 km/s, 6 GHz)
Scintillation Timescale Throughout the Milky Way (d=2 kpc, V=10 km/s, 6 GHz)
Scintillation Timescale Throughout the Milky Way (d=16 kpc, V=10 km/s, 6 GHz)
Scintillation Timescale Throughout the Milky Way (d=4 kpc, V=10 km/s, 6 GHz)
Scintillation Timescale Throughout the Milky Way (d=32 kpc, V=10 km/s, 6 GHz)
Sgr B2
b (deg)
l (deg)
log10[Δtd (s)]

Scattering intensity

- Ionosphere — weak $m_d \ll 1$
- IPM — mostly weak
- ISM — can be strong! $m_d \approx 1$

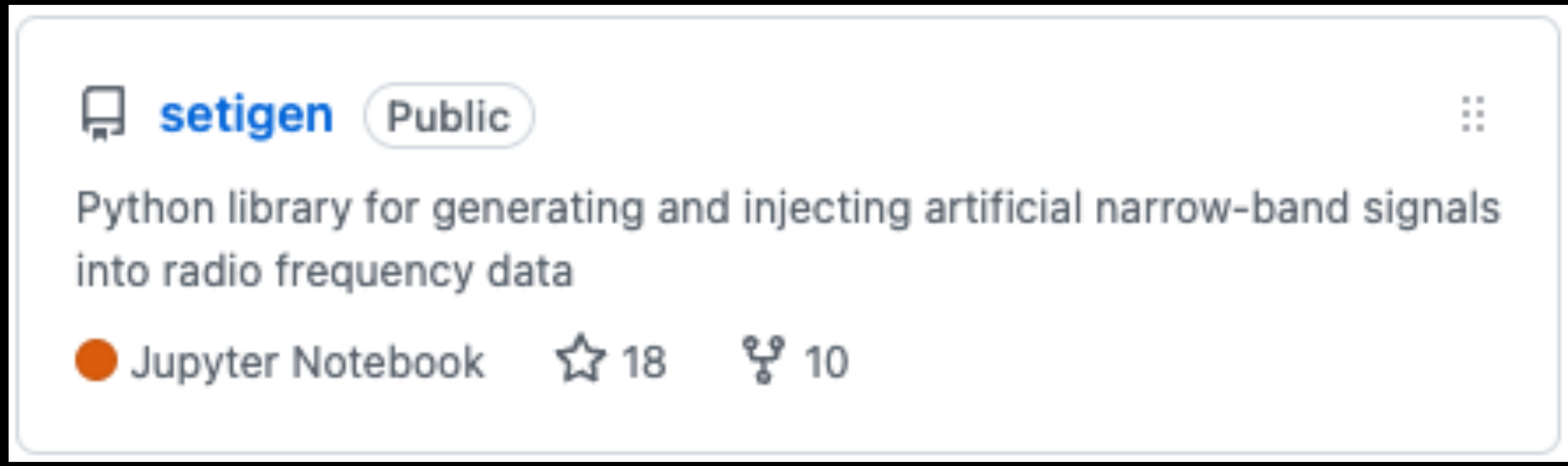


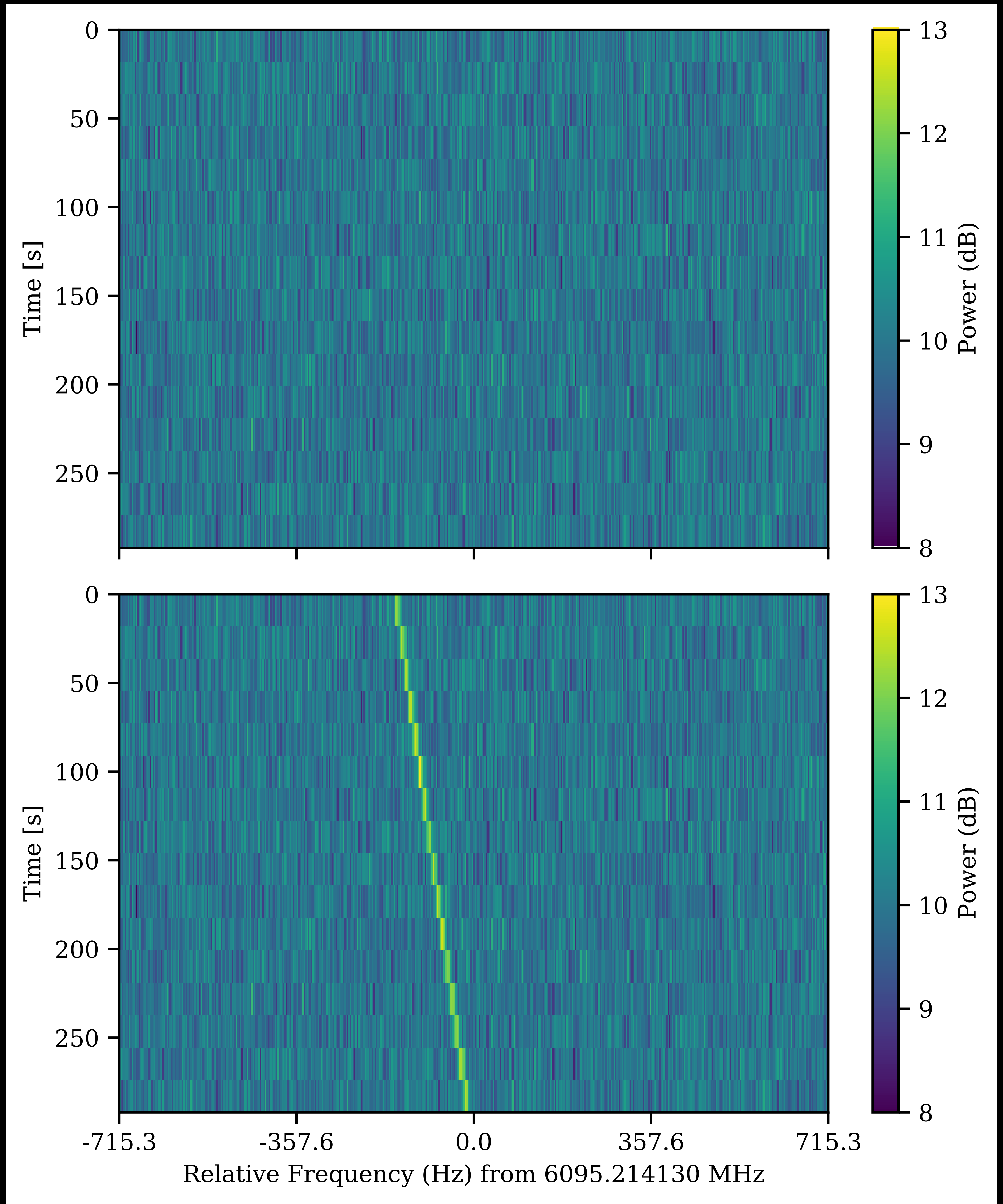




Setigen

- Python library for synthetic spectrogram and voltage data
- Specific focus on narrowband signal generation and injection

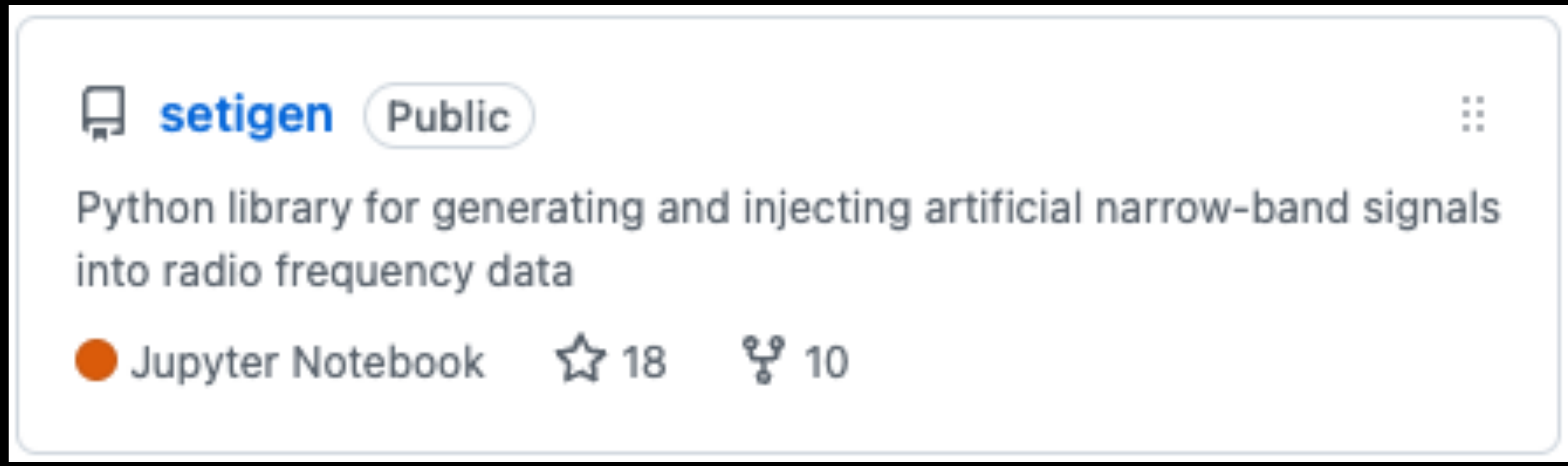


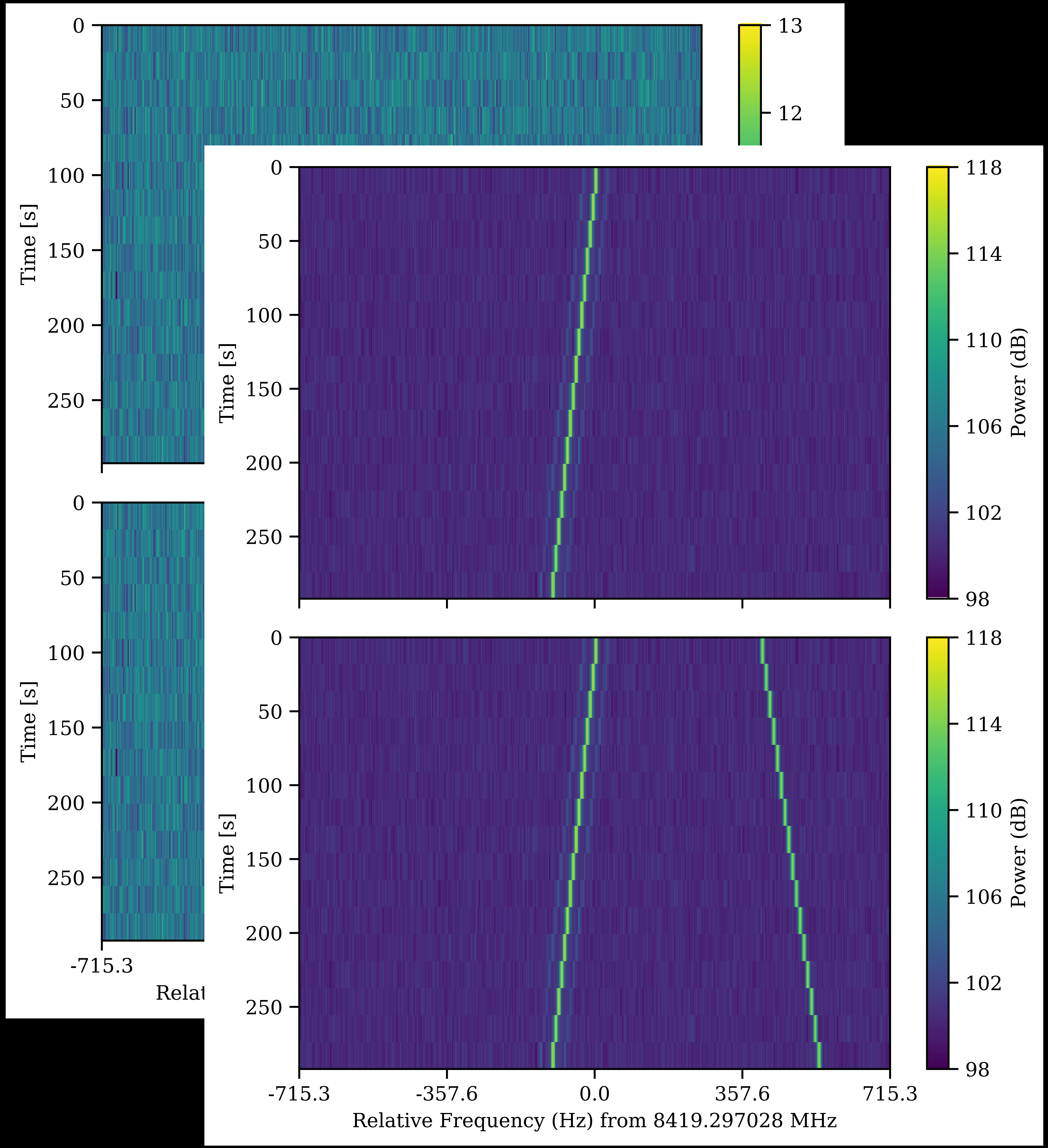


Brzycki et al. 2022, AJ

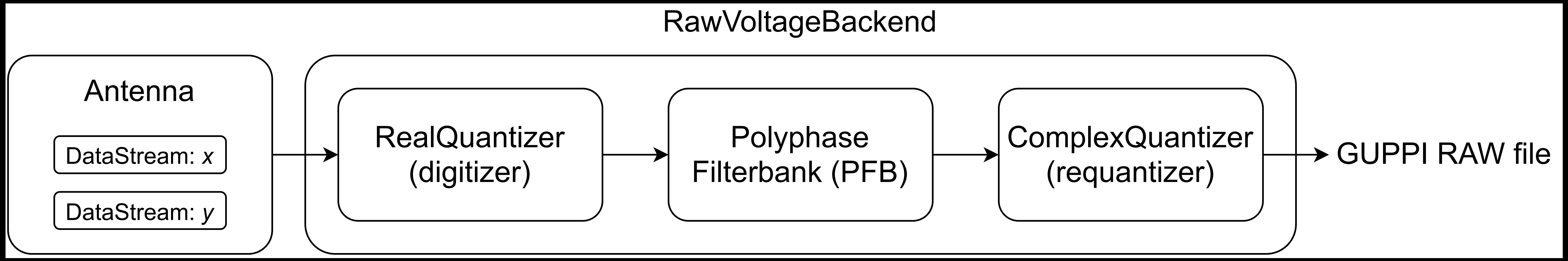
Setigen

- Python library for synthetic spectrogram and voltage data
- Specific focus on narrowband signal generation and injection



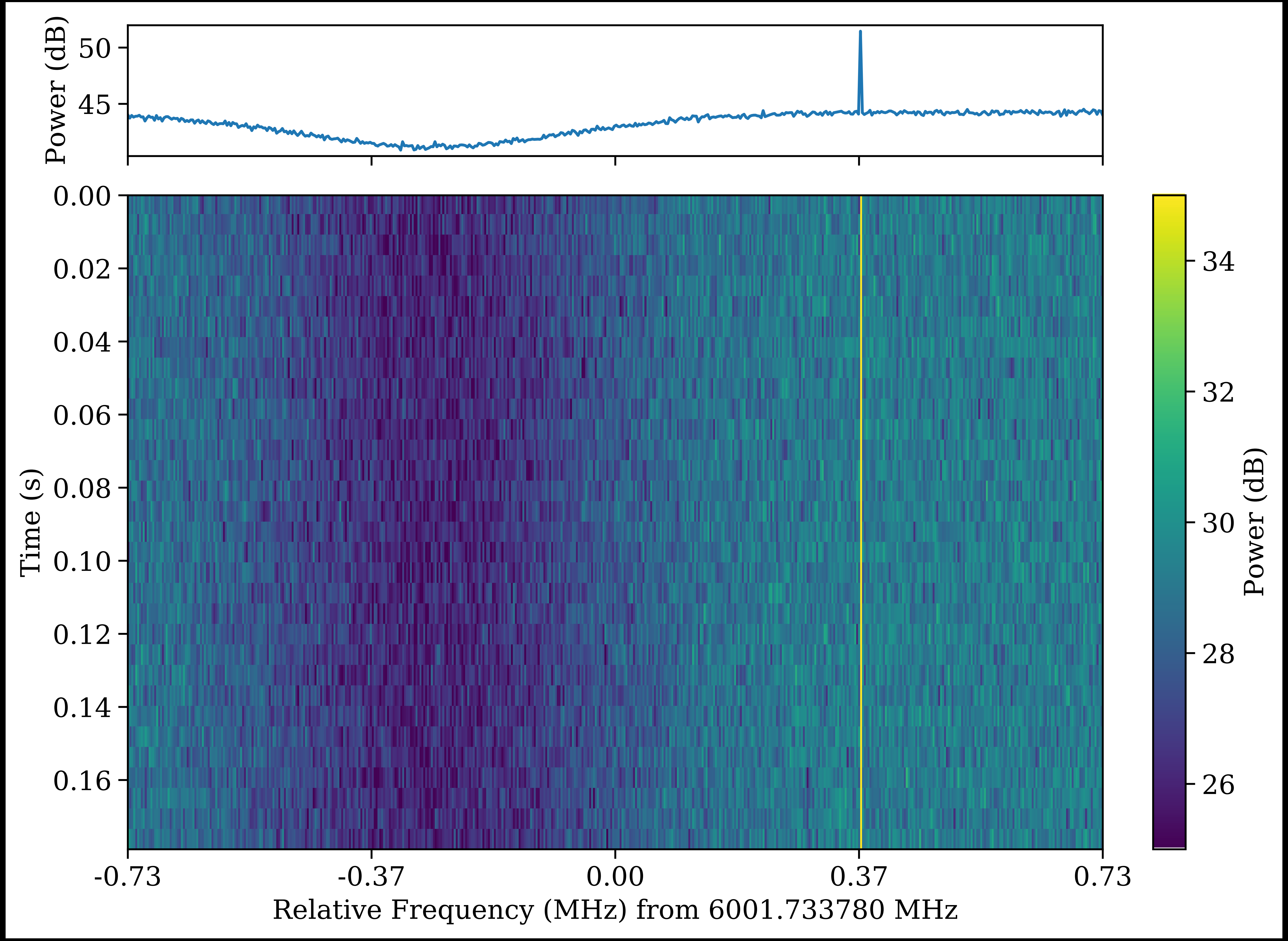


Brzycki et al. 2022, AJ



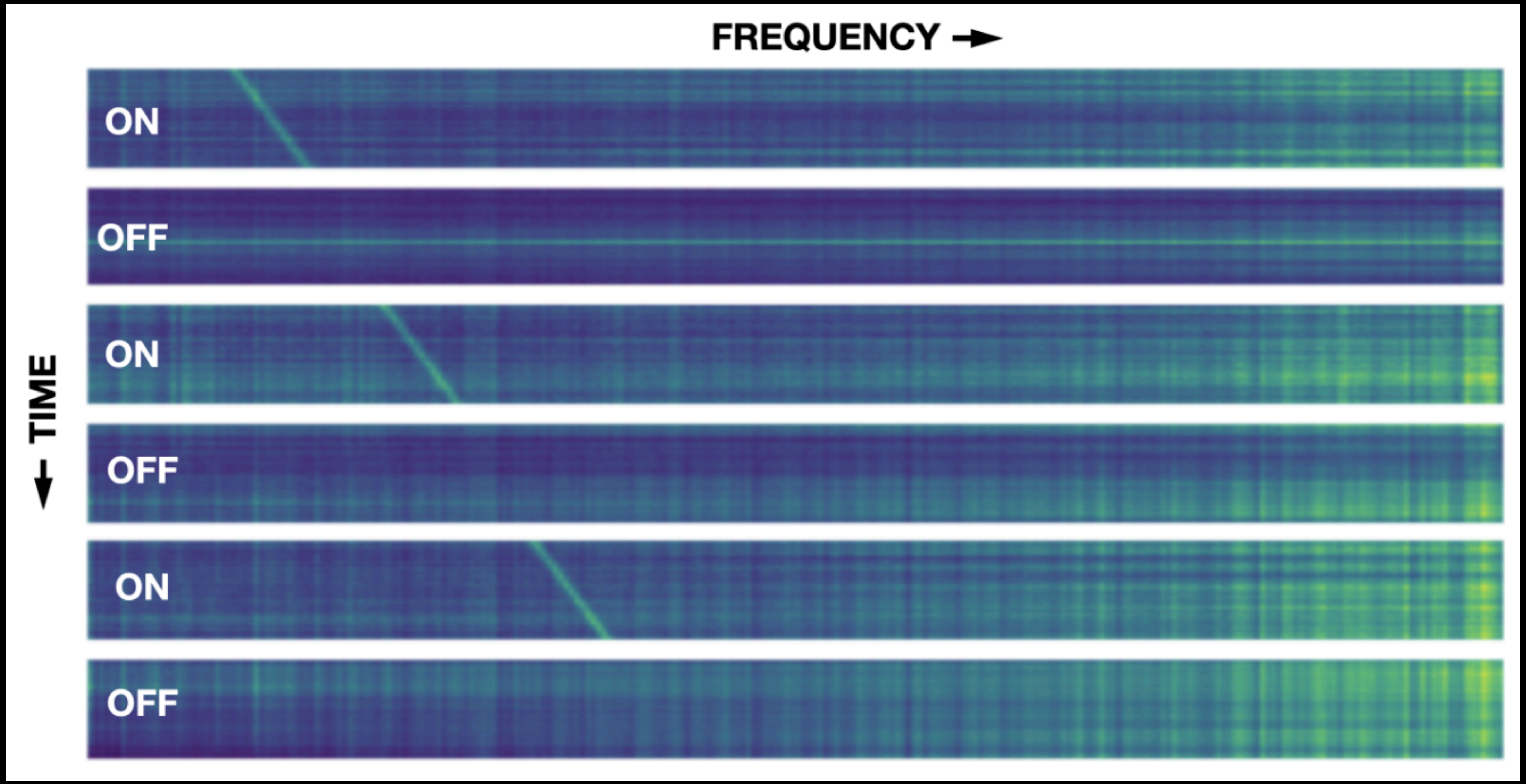
- Synthetic complex voltage data

- Simple models of backend components, such as a polyphase filterbank



Applications of Setigen beyond my research

- Injection — recovery testing
- ML dataset production (e.g. Kaggle)
- Multibeam search surveys
- Development of software for the Allen Telescope Array

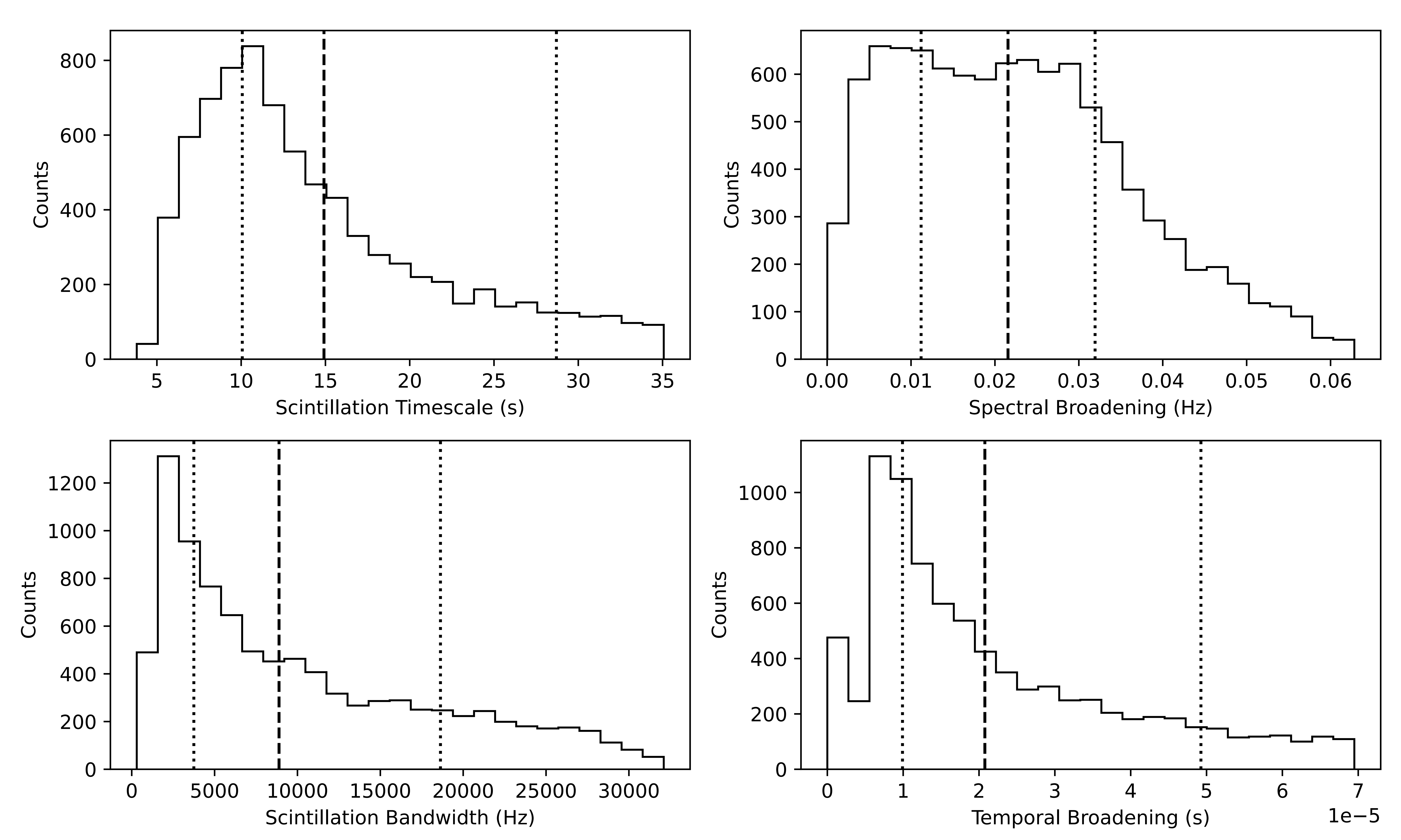


Breakthrough Listen x Kaggle 2021

Inter-quartile range

Median

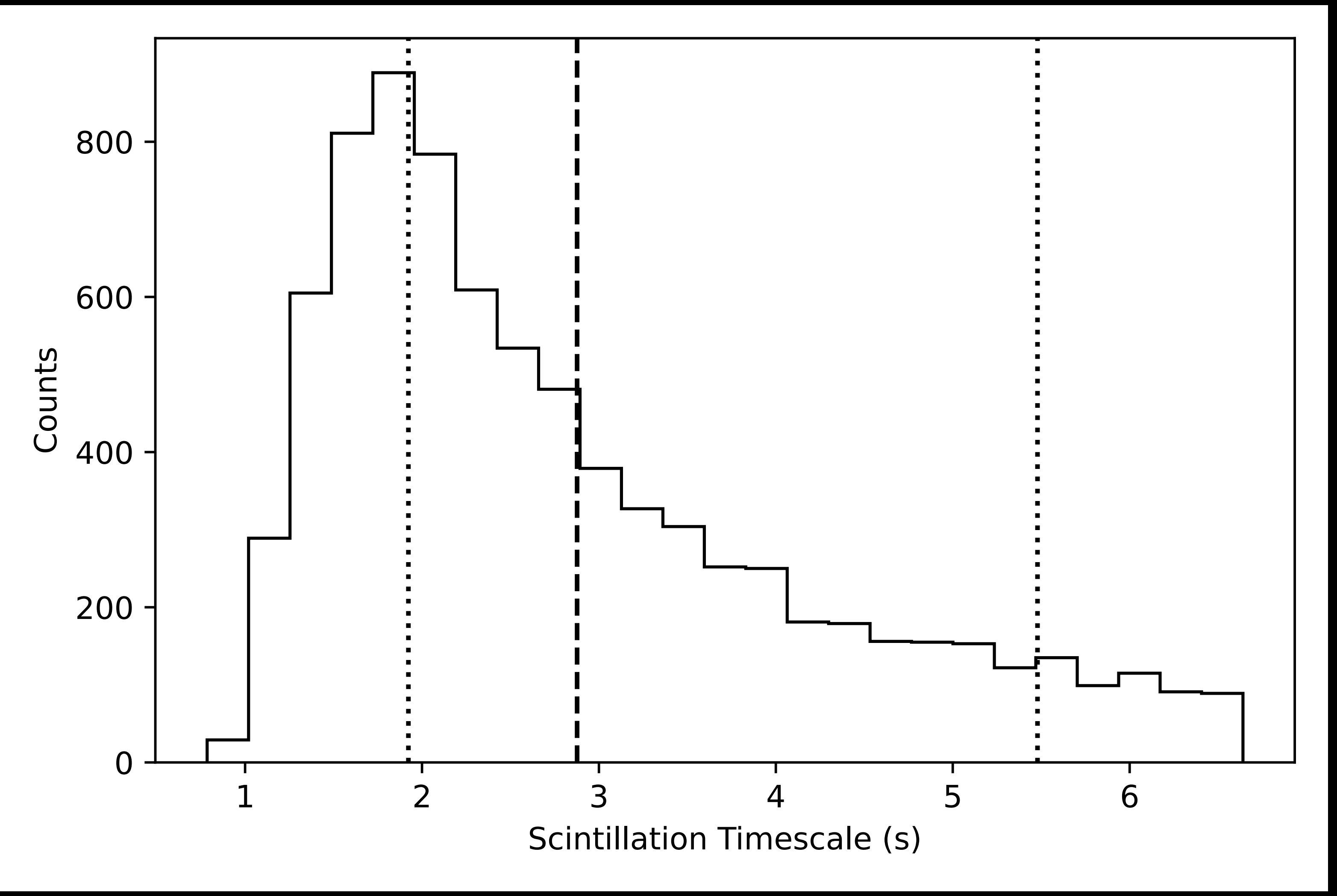
C-band (l, b) = (1, 0)

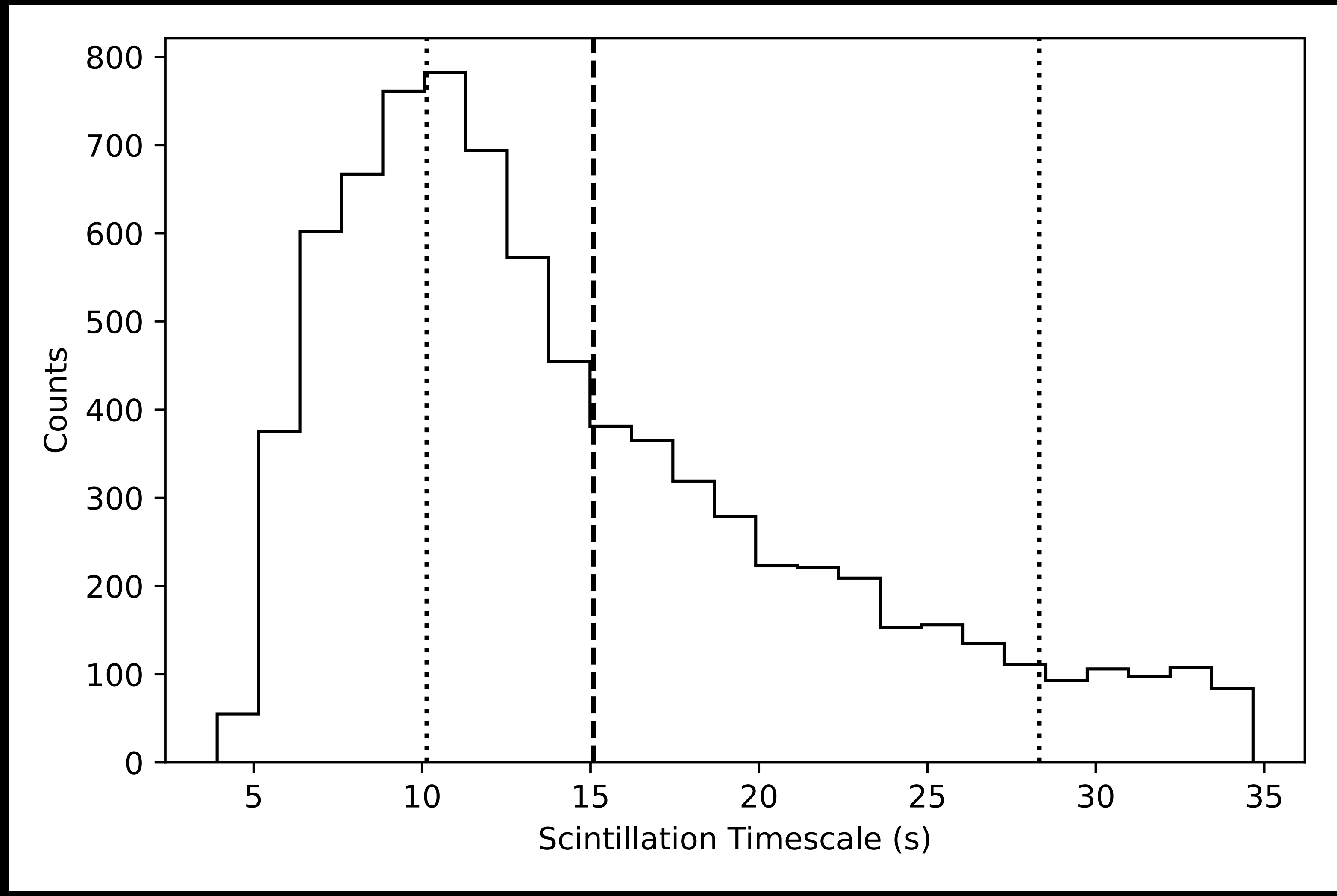


Monte Carlo-sampled timescales

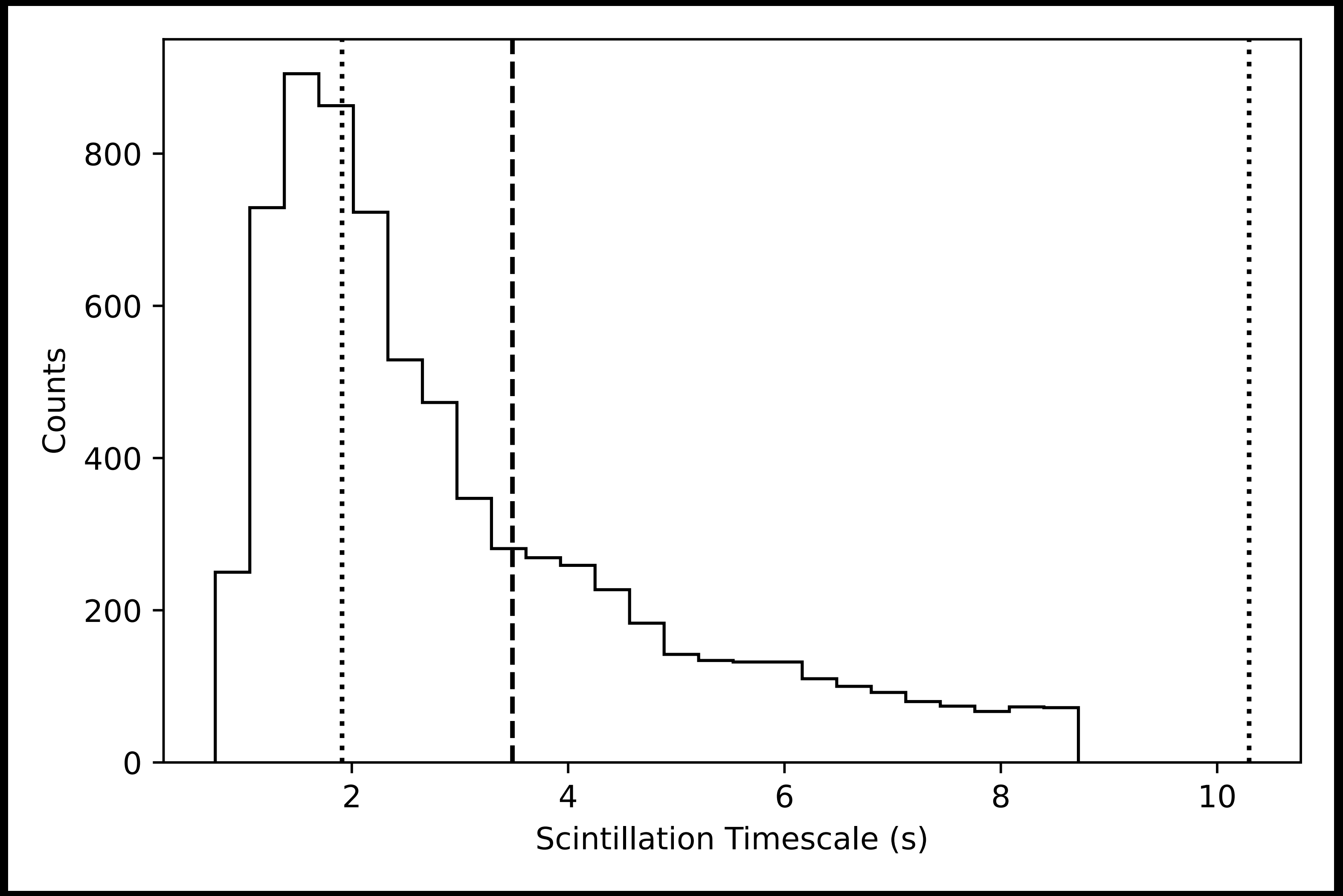
L band | C band

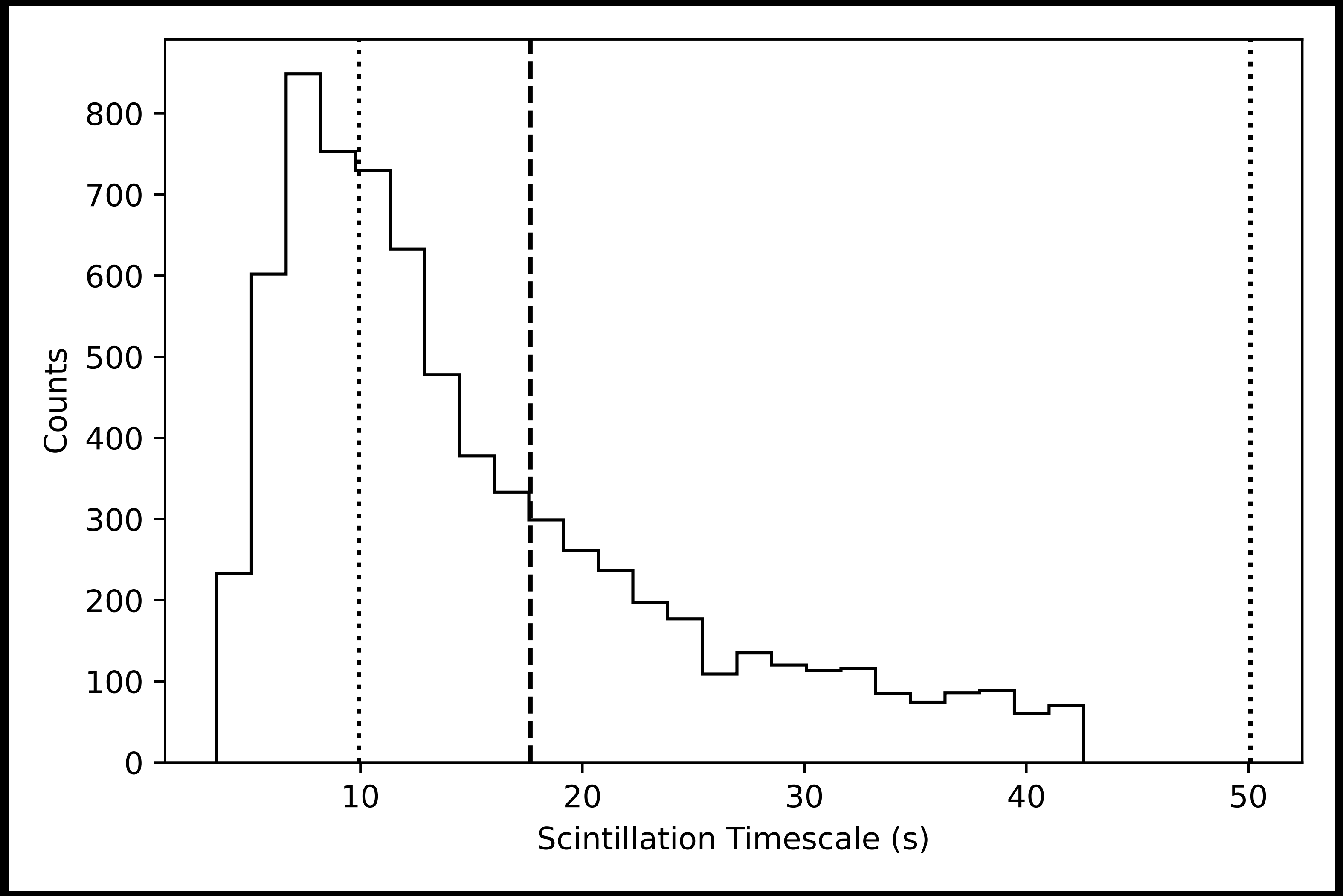
Density





Uniform





Distance sampling

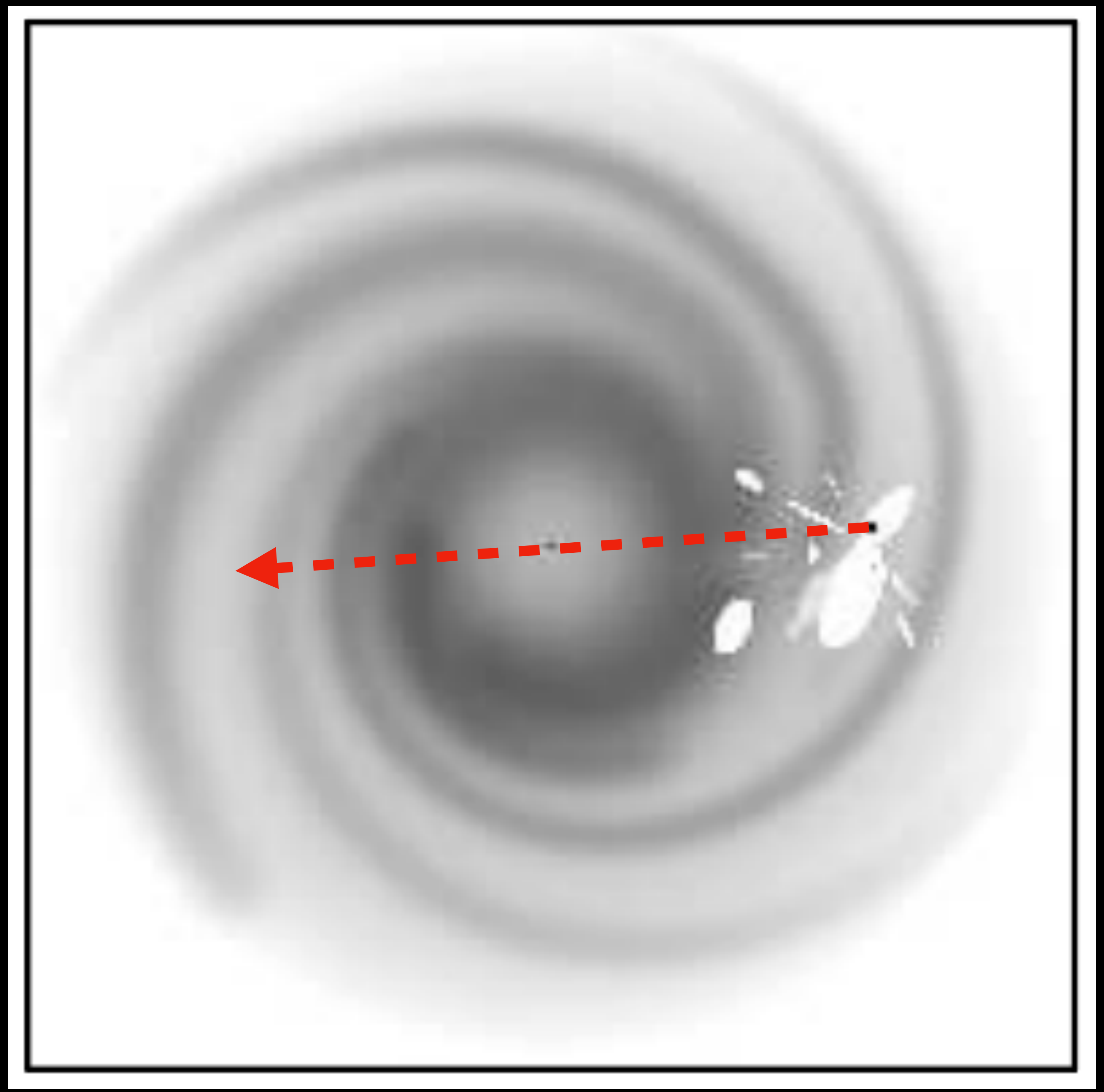
Statistics at different bands

(l, b) = (1, 0)

Band	Frequency (GHz)	Median (s)	IQR (s)	Mode (s)
LOFAR	0.110 — 0.240	0.22	0.14 — 0.41	0.14
L	1.1 — 1.9	2.9	1.9 — 5.6	1.9
S	1.8 — 2.8	4.8	3.3 — 9.0	3.1
C	3.95 — 8	15	10 — 28	11
X	8 — 11.6	28	19 — 52	16

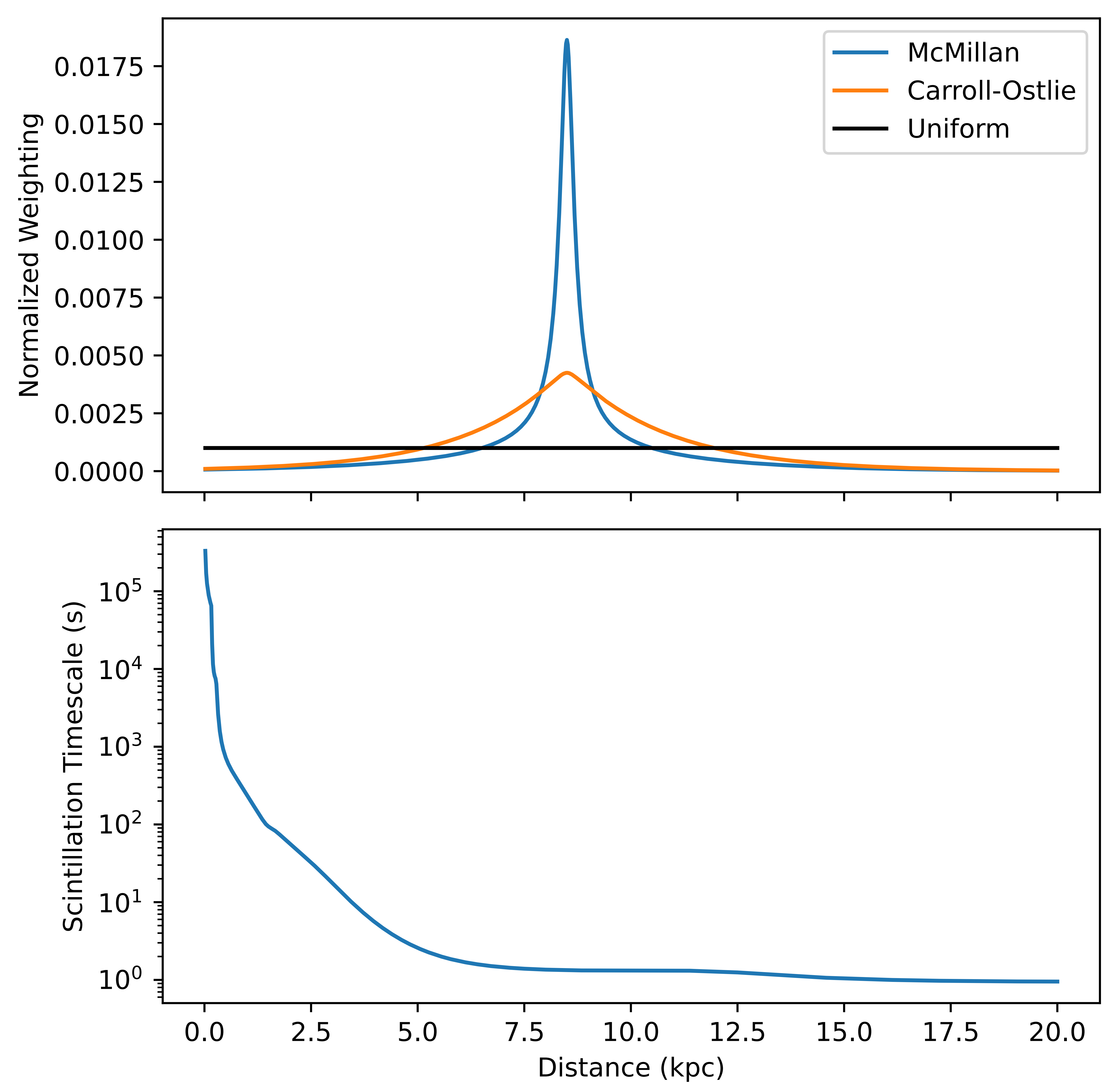
$$\Delta t_d \propto \nu^{6/5} v_T^{-1}$$

Density-based sampling



Cordes & Lazio 2002

(l, b) = (1, 0)



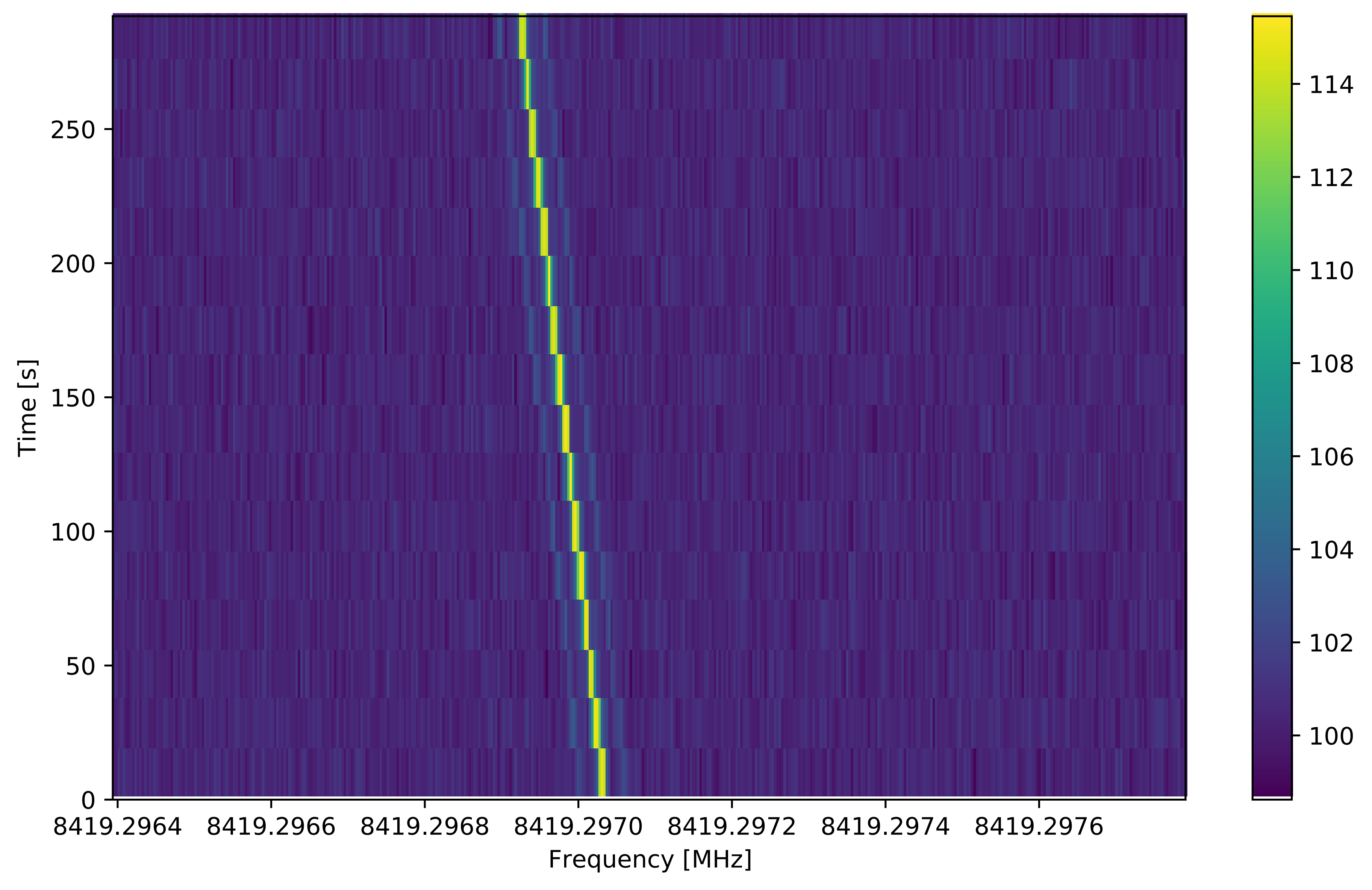
McMillan 2017, Gowanlock et al. 2011, Carroll & Ostlie 2007

What would strongly scattered signals look like?

- Temporal scintillation
- Spectral broadening
- Pulse broadening
- Spectral de-correlation

What would strongly scattered signals look like?

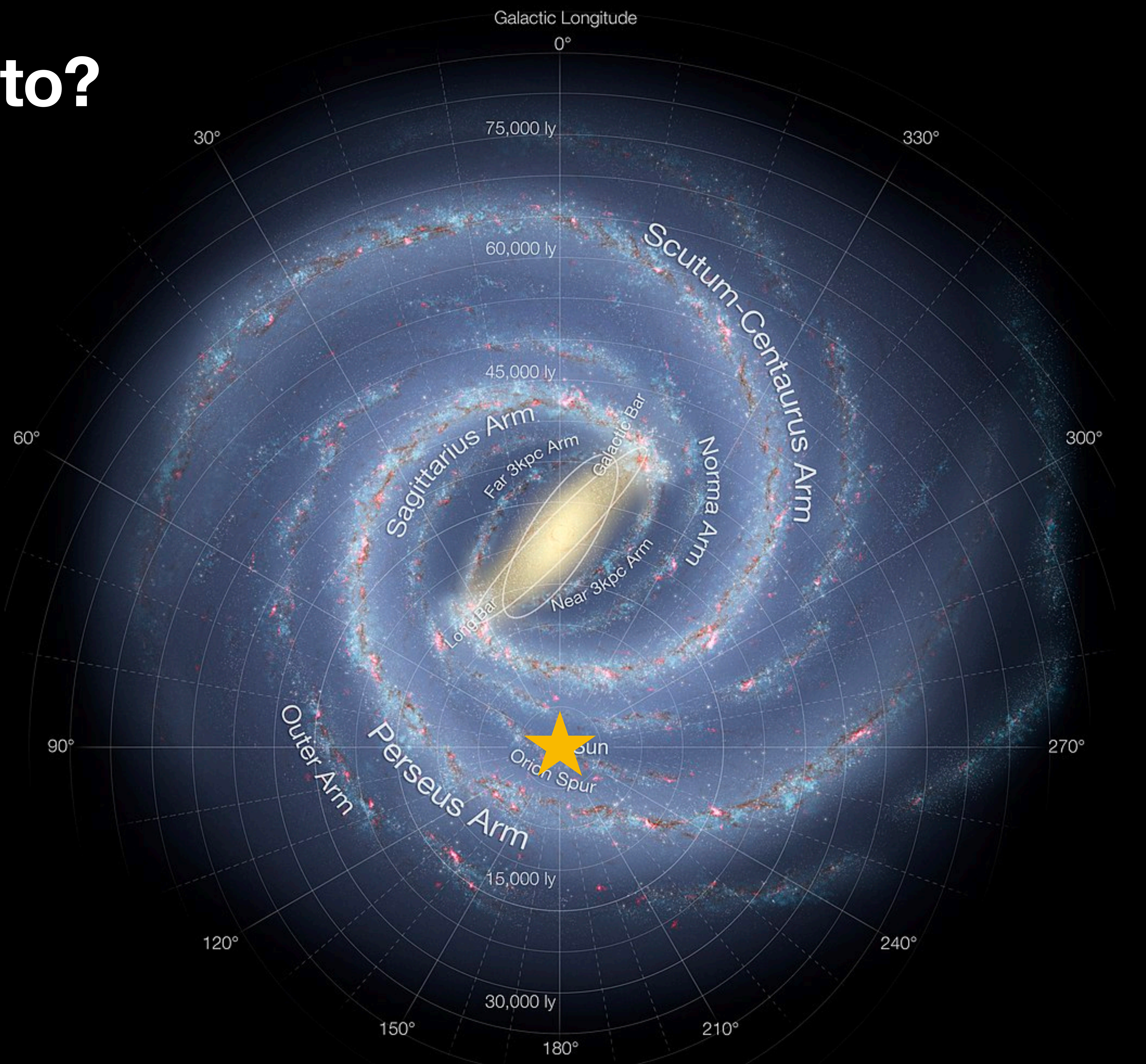
- Assuming a 100% duty-cycle narrowband transmitter



- Temporal scintillation
- ~~Spectral broadening~~
- ~~Pulse broadening~~
- ~~Spectral de-correlation~~

Why is this worth looking into?

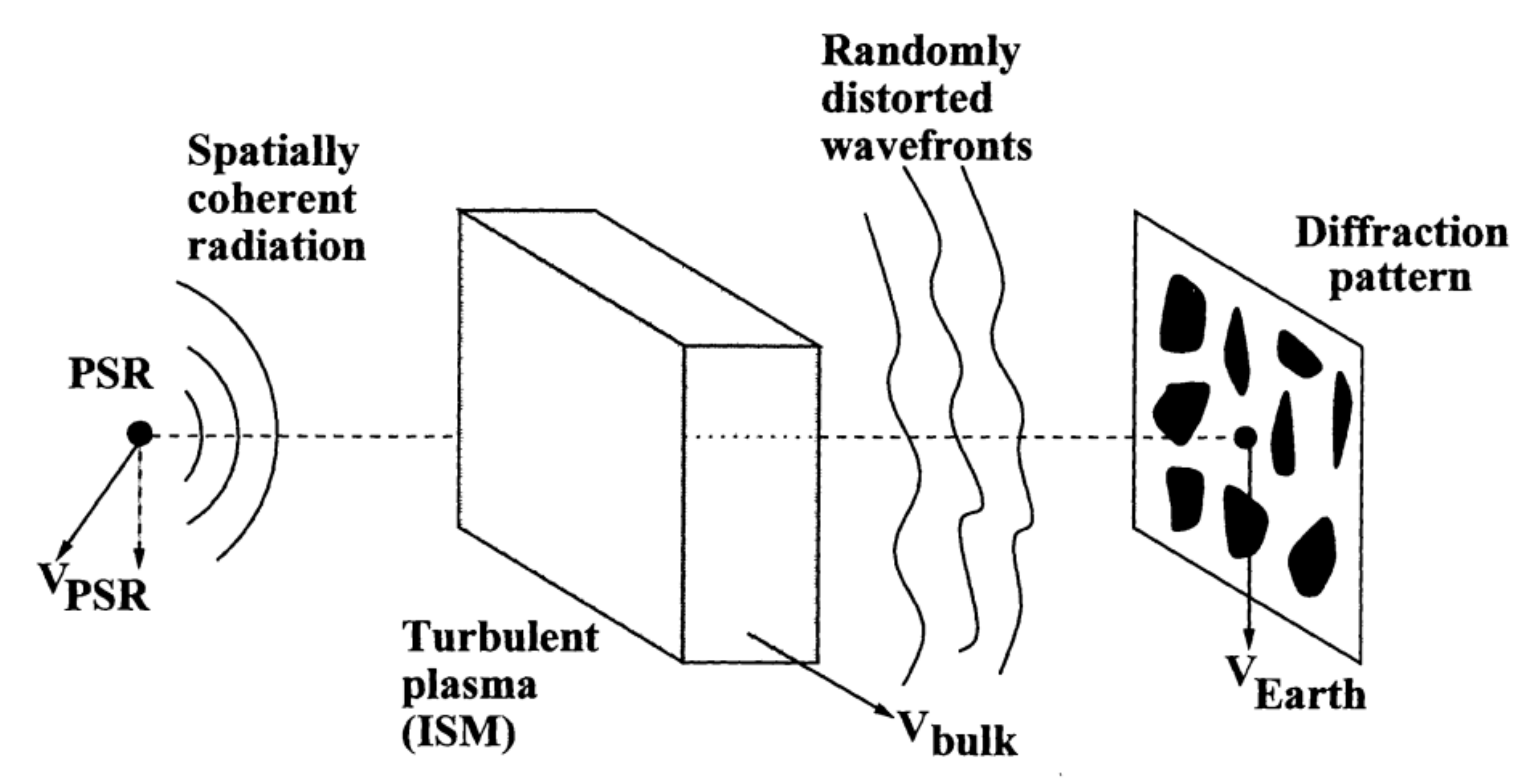
- Astrophysical modulation as a filter for technosignature candidates
- Looking towards the Galactic Center is well motivated by SETI
- Could provide a framework for using more of the actual signals during narrowband analysis



NASA/JPL

We focus on so-called diffractive scintillations

- Electron density fluctuations give rise to phase fluctuations
- Multi-path propagation
- Interference pattern with characteristic spatial and spectral scales
- Can lead to 100% intensity modulation on characteristic temporal scales Δt_d



Cordes 2002

Parameter space exploration of scattering parameters

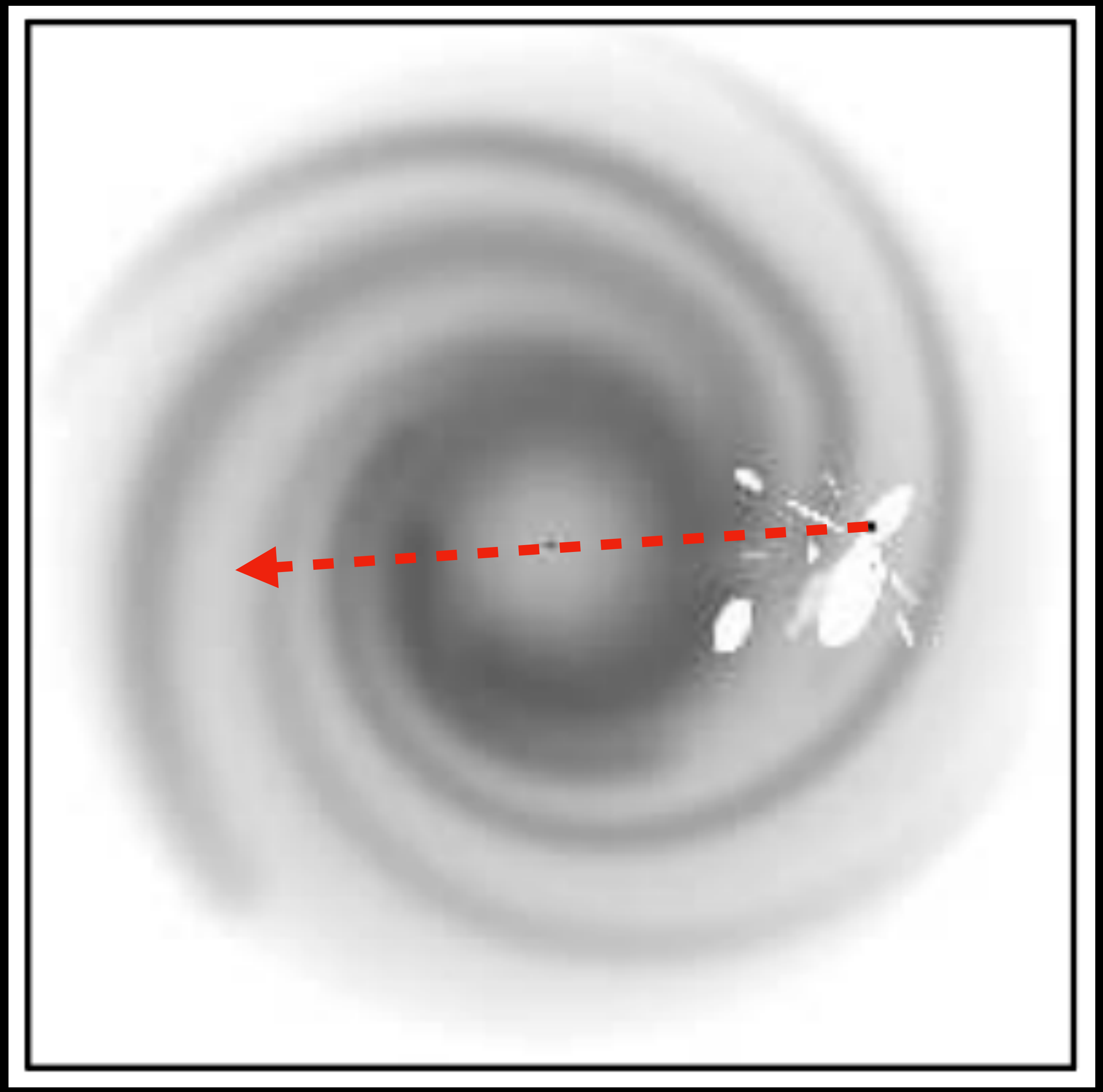
- A priori, we do not know:
  - Sky direction
  - Frequency
  - Distance
  - Transverse velocity

Monte Carlo sampling!

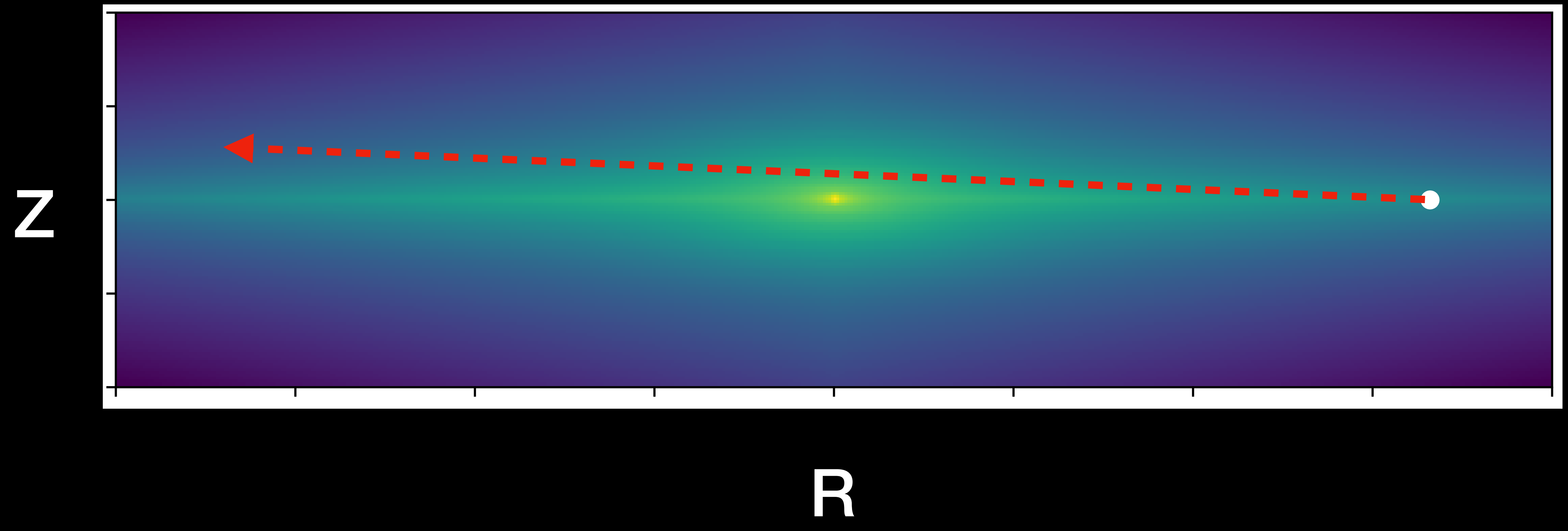
$$\Delta t_d \propto \nu^{6/5} v_T^{-1}$$

- Sky direction → Chosen parameter
- Frequency → Uniform sampling within chosen band
- Distance → Uniform or density based sampling
- Transverse velocity → Uniform sampling

Density-based sampling



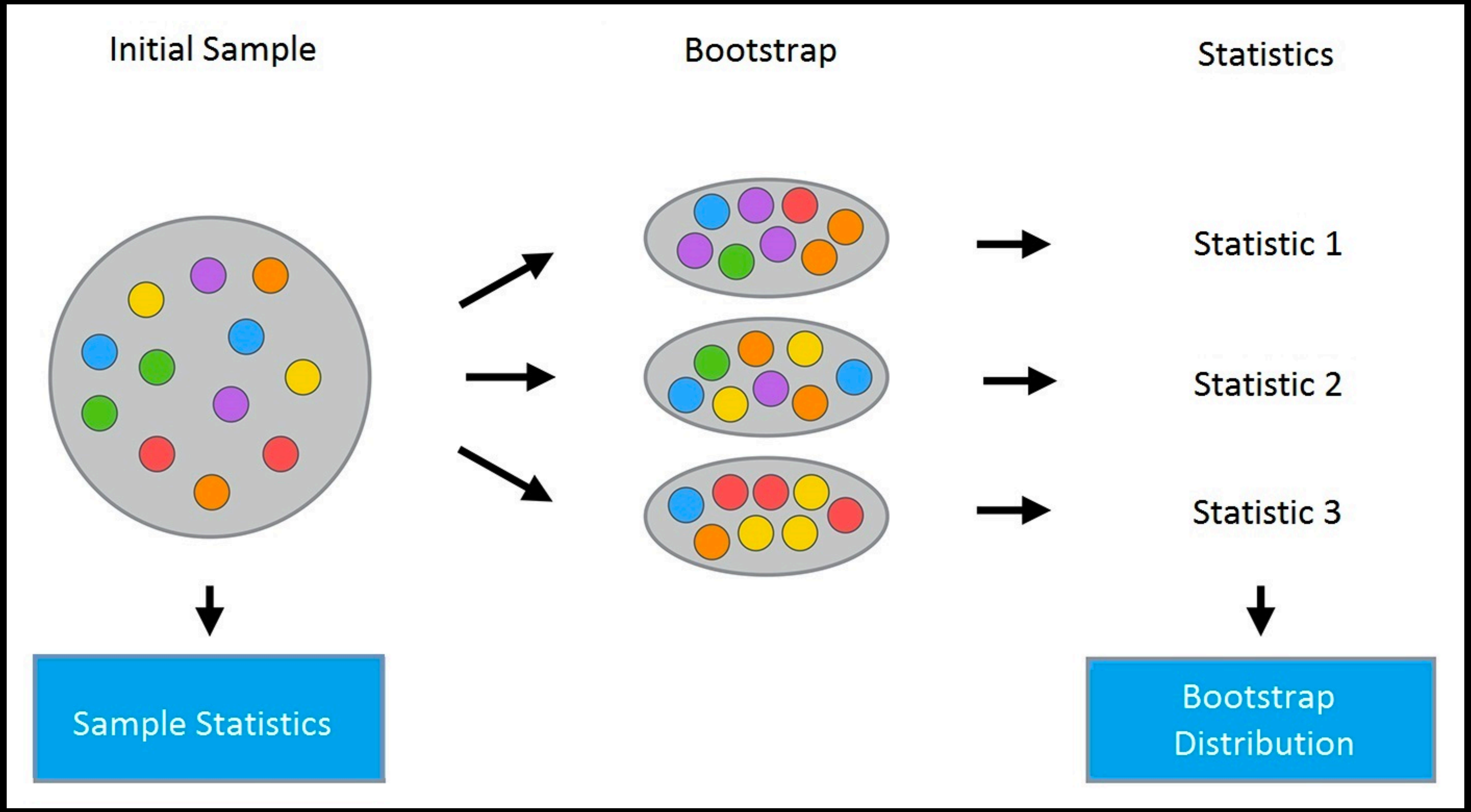
Cordes & Lazio 2002



McMillan 2017, Gowanlock et al. 2011, Carroll & Ostlie 2007

Low sample regime

- Spread of values around the asymptotic “truth”
- Both correlated and uncorrelated samples within the same observation
- How can we evaluate this?



Kaggle

Quick way to produce synthetic data with asymptotic statistics

- (Cario & Nelson 1996) The ARTA random process:
  - Matches a target intensity distribution
  - Matches a target autocorrelation structure (with custom asymptotic precision)

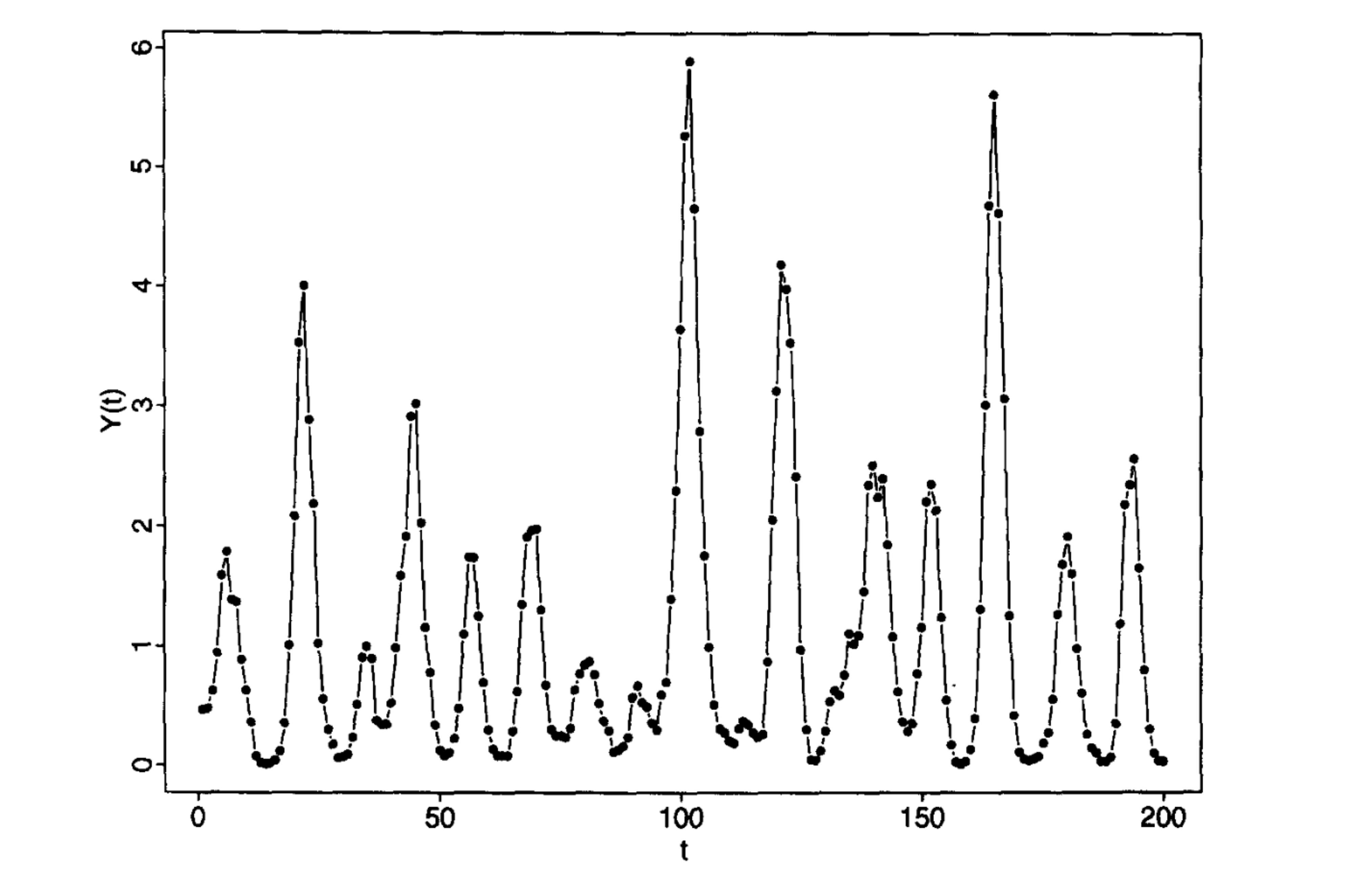
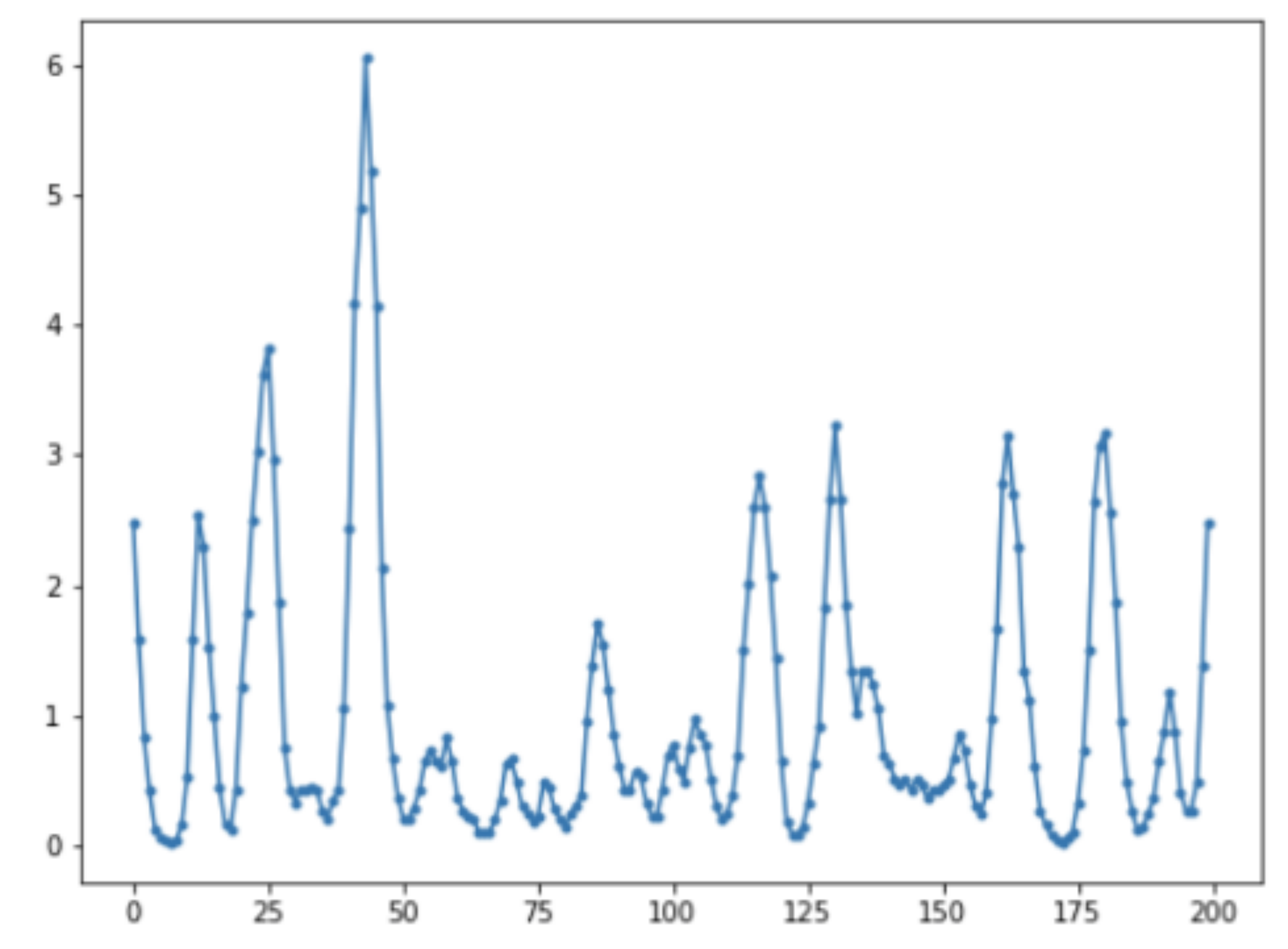


Fig. 1. Sample path of an ARTA process with exponential marginals and autocorrelations $\rho_1 = 0.9$ and $\rho_2 = 0.6$.

Cario & Nelson 1996



Our implementation