Radio Search for Extraterrestrial Intelligence
SETI is fun !

Geographically-spaced Synchronized Signal Detection System

New Hampshire Astronomical Society

Feb. 8, 2019

Concord, New Hampshire

Presented by Skip Crilly
Education & Public Outreach
Volunteer Science Ambassador
Green Bank Observatory
West Virginia

Rev. Feb. 8, 2019

Abstract

Radio Frequency Interference (RFI) is a confounding problem in radio SETI, as false positives are introduced into receiver signals. Various methods exist to attempt to excise suspected RFI, with a possibility that true positives are rejected, and that un-excised RFI remain as false positives. Uncertain far side-lobe antenna patterns add to the uncertainty. To ameliorate the RFI problem, a system having geographically-spaced simultaneous and synchronized reception has been implemented. A radio telescope at the Green Bank Observatory in Green Bank, West Virginia has been combined with a radio telescope of the Deep Space Exploration Society, near Haswell, Colorado to implement a spatial filter having a thrice-Moon-distance transmitter rejection. Approximately 135 hours of simultaneous synchronized pulse observations have been captured from November 2017 through February 2019. This presentation describes the problem, observation system, observed results and a proposed hypothesis to be subjected to attempts at refutation through further experimentation and RFI and ETI transmitter signal model development.

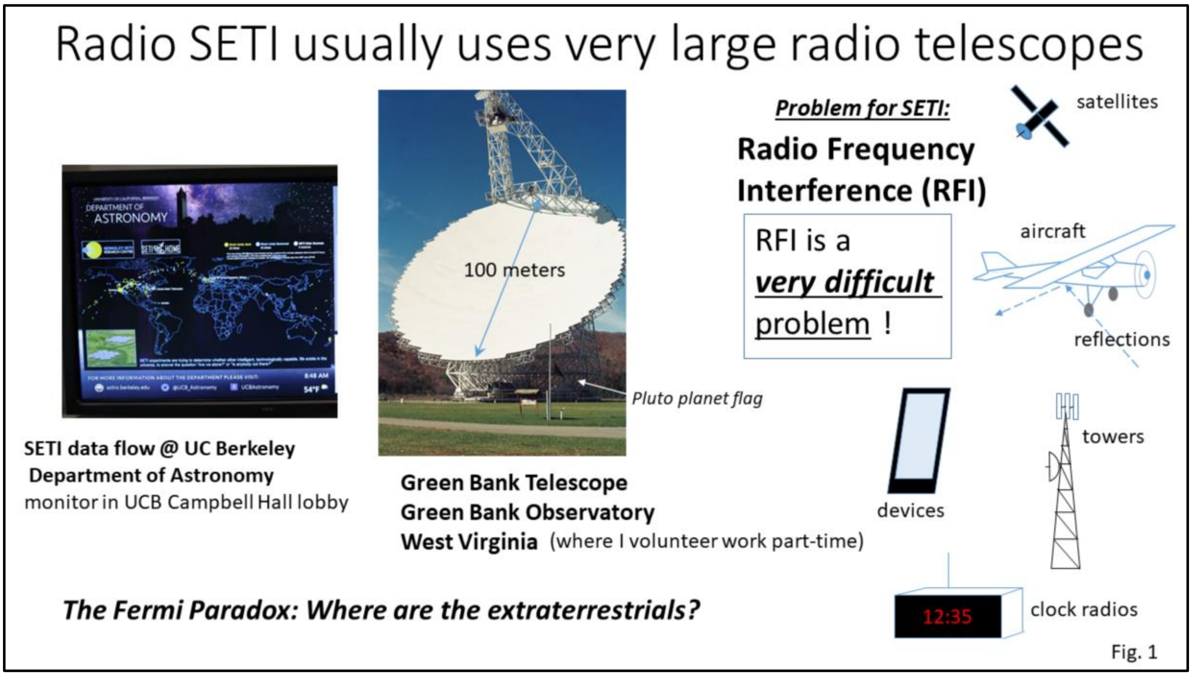


Fig. 1

Problem

Radio SETI has been plagued by an increasingly difficult problem – Radio Frequency Interference (RFI). As communication systems and computing devices have modernized, human-caused electromagnetic emissions and speculated extraterrestrial-caused radio signals have become difficult to differentiate.

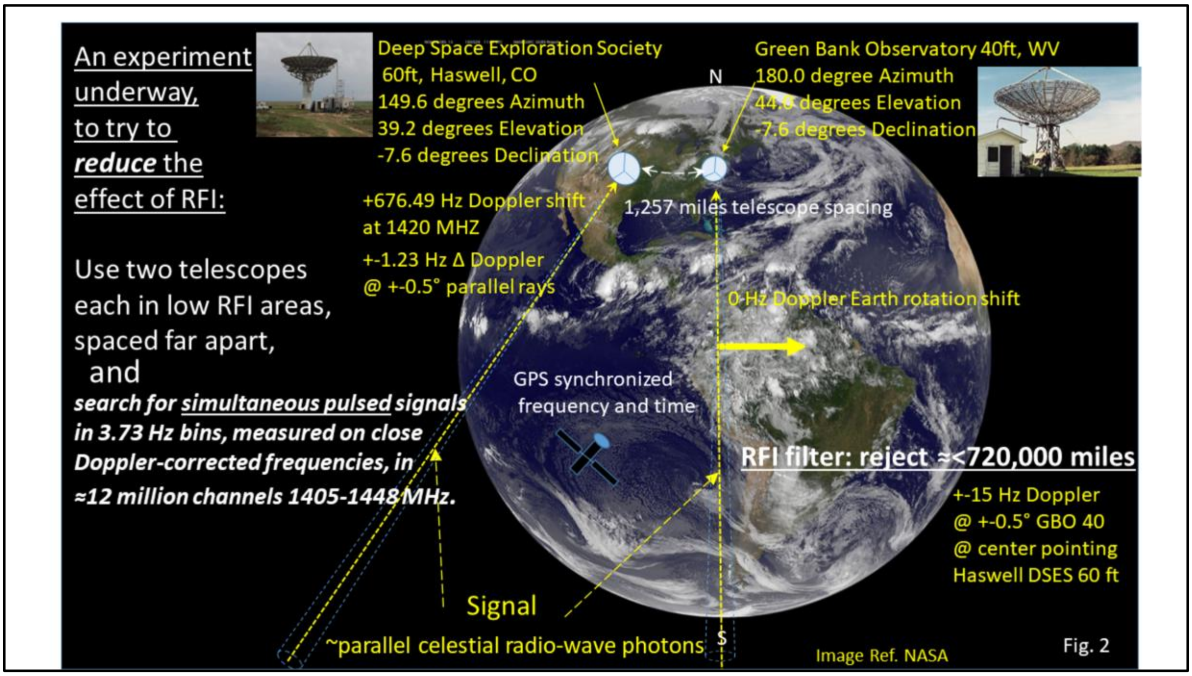
Antenna pointing direction is typically used to identify celestial transmitting sources. Issues arise:
1. In a terrestrial, airborne and satellite RFI environment, antenna pattern sidelobes change as antenna pointing is changed; highly sensitive telescopes can yield false positives, given the uncertainty of off-beam RFI and antenna patterns.
2. Multi-pixel feeds aid in directional differentiation, yet do not entirely solve the problem of multipath local RFI entering the sidelobes of the feed elements, resulting in outlying false positives.
3. Long term monitoring of identified candidate pointing directions requires time-consuming and careful follow-up, especially in the presence of sporadic RFI, and intermittent ETI transmissions.

Machine-based learning of human-sourced signals, and local RFI source identification, can help ameliorate the problem, through RFI excision. However, one is left with a quandary: How does one differentiate between a modern human communication system (RFI) and a typical celestial communication system (ETI), both designed to operate under a common Shannon's Law scenario?

Energy-efficient communication signals are speculated to be transmitted by celestial intentional transmitters, as high information transfer and low transmitter cost are apparent goals. Communication signals contributing to these goals are expected to span large radio frequency bandwidths, contain narrow bandwidth elements for detectability, and in their limit, be indistinguishable from random noise.

Modern communication systems transmit different information to closely spaced receive antennas, in systems using multiple transmit and/or receive antennas. Increased channel capacity results.

If energy-efficient and high capacity communications signals are in the limit indistinguishable from random noise, and multiple receive antennas receive different parts of a single transmitted information stream, how can a communicative and gregarious ETI make itself known? Relatively easily identifiable signals seem to be required, leading to the idea proposed in this paper.



Idea to help solve the RFI problem

Receiver spatial filtering may be assumed by the transmitter, leading to a method to make a transmitter known to a receiver, by intermittently transmitting spatial-**simultaneous** and time-**simultaneous** narrow bandwidth pulses, within an otherwise efficient and high capacity information stream. Additional characteristics of received signals, correlated to these simultaneous pulses, should ideally make the overall event likelihood low in noise and RFI, and support an ETI or very distant RFI transmitter hypothesis

Receiver antennas (Fig. 2) may be distantly spaced to search for simultaneous pulses of hypothetical celestial origin. For example, a 1,257 mile receive antenna spacing provides a 720,000 mile spatial filter, to counter local and near-space RFI. The spatial filtering distance is considerably farther than the distance to the two antennas' -3 dB beam overlap point, because Doppler shift within the antennas' beamwidth gives space-based-RFI signals a frequency difference potentially greater than an FFT bin size. For example, there is a +-15 Hz difference in Doppler-induced frequency shift for signals arriving across the Green Bank Forty Foot Telescope's -3 dB beamwidth, e.g. when the two electromagnetic signal rays and antenna spacing baseline form a triangle, and therefore indicate signals from close-in space transmitters.

Receiver bandwidth selection

The choice of a 3.73 Hz FFT bin bandwidth and 1405 to 1448 MHz results from the processing tradeoffs of on-the-fly narrowband pulse signal detection and storage using an SNR threshold. A celestial transmitter may transmit with different occupied bandwidths and pulse durations, to increase channel capacity. Signals that are guessed to be intended for ease of detectability likely have various pulse durations, to avoid a receiver-transmitter un-matched filter scenario. The 3.73 Hz bin bandwidth may be considered to be a matched filter to one set of potentially transmitted identification pulses.

The 1405 to 1448 MHz spectrum is chosen because it is a somewhat well-RFI-protected wide band of spectrum. Contiguous frequency coverage is prioritized in signal processing, over contiguous time, due to the need to identify CW and narrow bandwidth, potentially drifting, RFI sources within the chosen band. The pulse detection system operates at one-quarter duty cycle, at four second triggered intervals.

Signal Processing

Four 0.27 s duration contiguous time uniform windows are 2^25 point FFT-filtered and applied to a per-bin SNR threshold, before approximately 3 seconds of processing time, and a dwell time required to allow GPS-sync'd time triggering of the two telescopes' intermediate frequency analog-to-digital converters. Local oscillator frequency is synchronized using OCXOs at each site, each locked to a GPS signal. SNR thresholds are two-fold, at 11.2 dB per FFT bin and 12.0 dB 4x0.27s contiguous time average SNR, to reduce noise-induced false positives. Intentionally transmitted signals are expected to often partially straddle adjacent and/or alternate time windows, thus allowing the 12.0 dB threshold to select and store these pulses and partially reject noise-induced pulses. The Noise in SNR is measured in each 256 bin segment of the FFT output. ADC sampling is at 125 MSPS. Polarization is circular at Haswell and linear at Green Bank, resulting in a 3 dB maximum polarization match-to-mismatch, given random Poincare-Sphere transmitted polarization and site-differential Faraday rotation. Raw time domain data is not stored. The two telescopes' software systems do not communicate with each other, to avoid near-simultaneous software-event induced corruption of data.

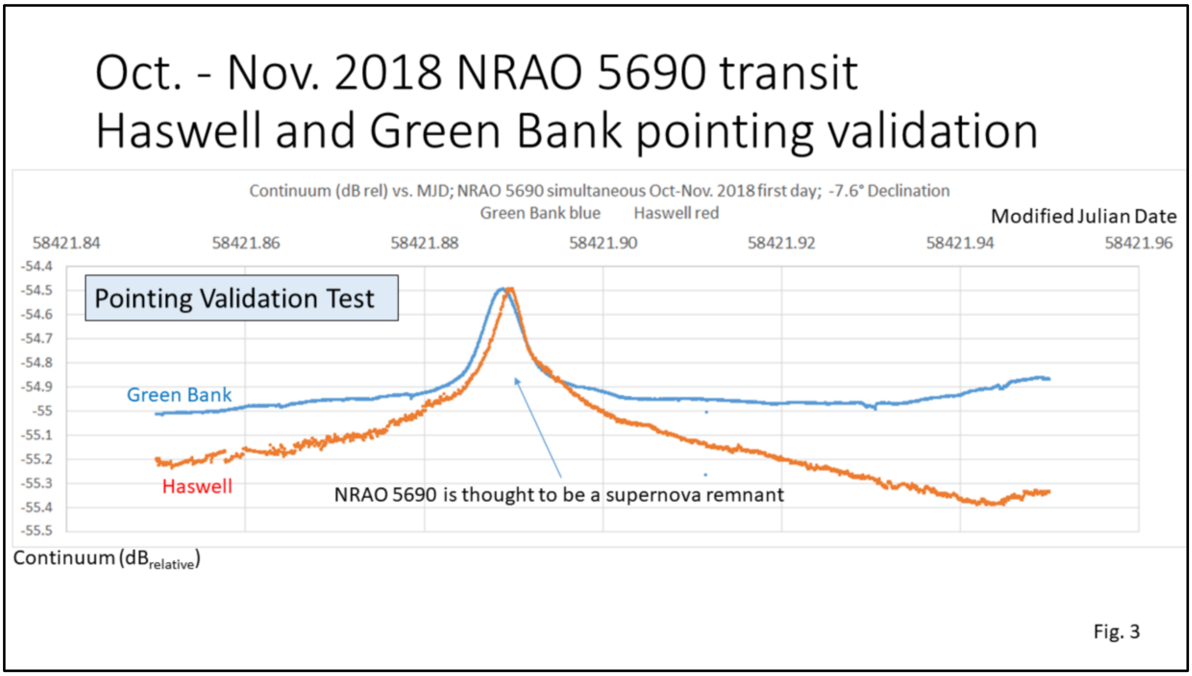
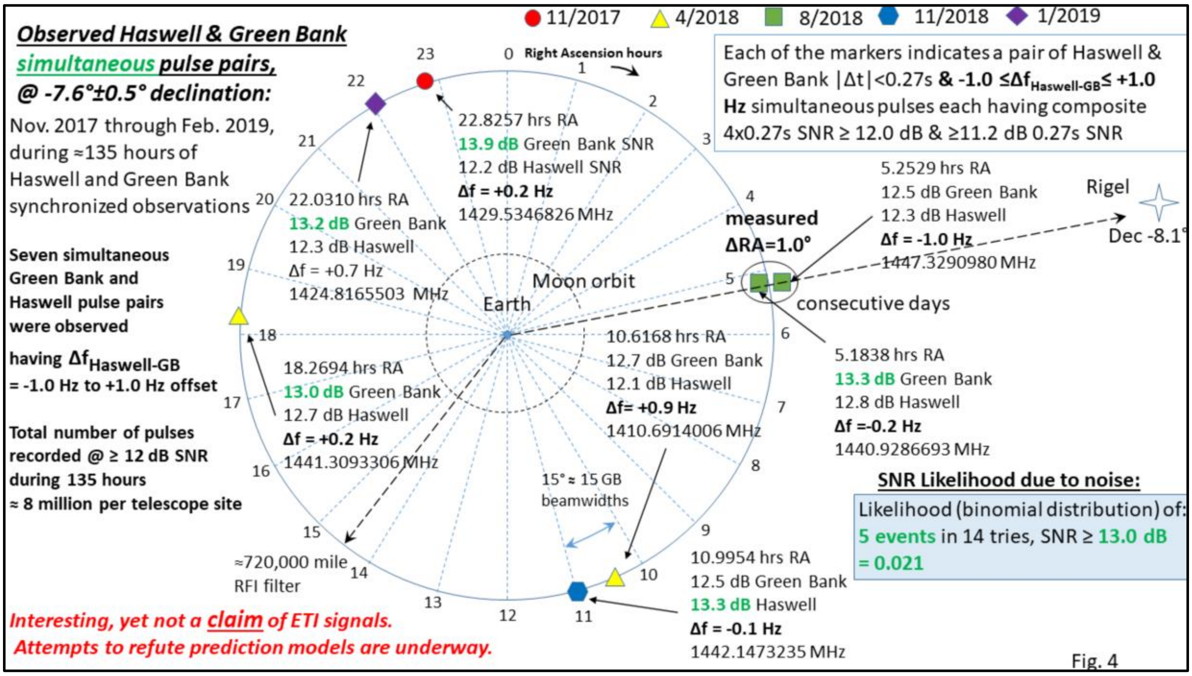


Fig. 3

Antenna Pointing Validation

Haswell and Green Bank telescope pointing was validated during each observation run. The October-November 2018 validation is shown in Fig. 3. The NRAO 5690 object's continuum, at 1405-1445 MHz and 1s frequency / time integration, is used to verify that antenna pointing matches reasonably well at the two distant telescope antennas. Antenna pointing does not change during observation runs.



Summary of simultaneous pulses observed during 135 hours

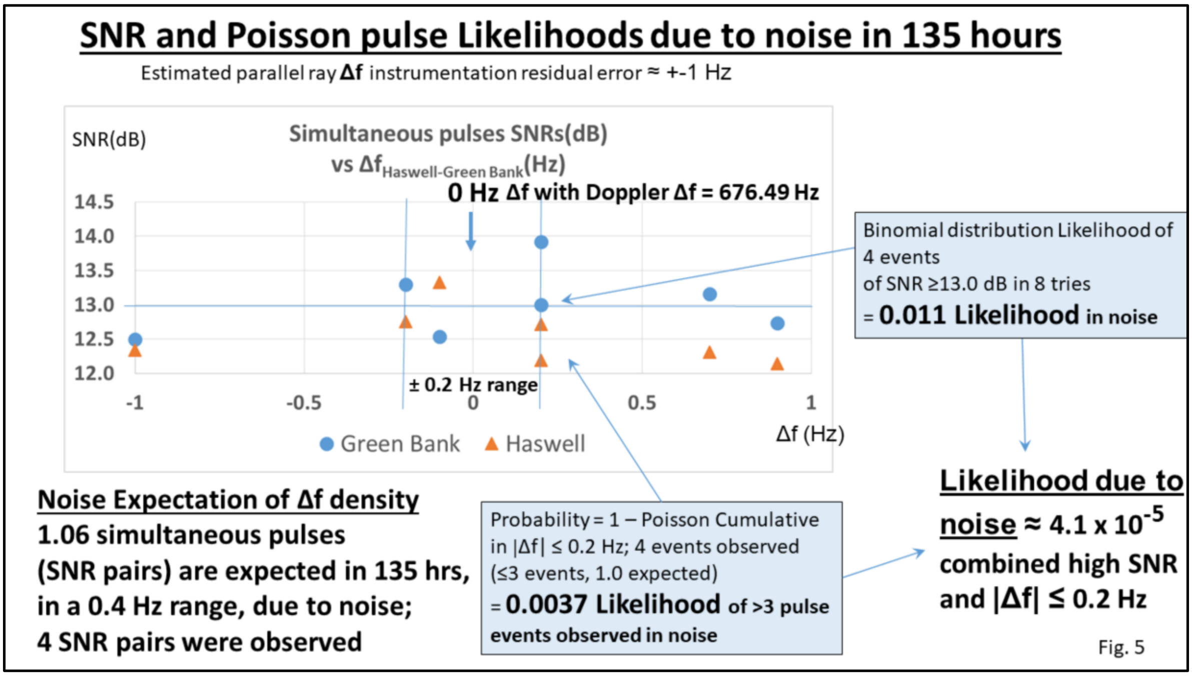
The plot in Fig. 4, of an Earth distant view, indicates the occurrences of simultaneous narrowband pulses, observed during 135 hours of observation in November 2017 to February 2019. The pointing Right Ascension of pulses are shown that have Haswell to Green Bank offset frequency in the -1.0 to +1.0 Hz range, after Doppler correction, in an FFT bin bandwidth of 3.73 Hz, 1405-1448 MHz, at -7.6° declination. The outer circle represents the approximate expected closest distance to a transmitting source that is subjected to the 1,257 mile antenna-spacing RFI spatial filter.

Two of the seven simultaneous pulses repeated on consecutive days within 1.0 degree of pointing, near the direction of Rigel.

The signal to noise ratios of five of the fourteen pulses (≥ 13.0 dB) appear to be anomalous, (binomial noise Pr. 0.021), considering a Rayleigh amplitude, exponential power, noise-caused hypothesis, after suspected RFI excision.

Five simultaneous pulses are expected in 135 hours, on average due to noise, in a 2.0 Hz offset range, based on an analysis of a large Doppler offset range, derived while assuming that almost all simultaneous pulses in a larger Doppler offset range are due to noise. Therefore, it is thought at least some of the seven simultaneous pulses observed are likely due to noise.

The repetition of a transmitting signal, observed at a single celestial pointing direction, is unexpected from a near-space transmitter, as the transmitter is affected by the Sun-Earth-Moon gravity wells and needs thrust to retain close to the same celestial pointing direction during the period of a day. Such a celestial-station-keeping thrusting system would not naturally be expected in a human-built transmitter. RFI modeling is underway to quantify this idea.



Anomalous population of simultaneous pulses in |Δf| ≤ 0.2 Hz (Fig. 5)

An anomalous population of four simultaneous pulses, within the set of seven simultaneous pulses, exhibited frequency offsets in a 0.4 Hz range, from -0.2 Hz to +0.2 Hz, with moderately high SNR (two un-correlated factor Pr. 4.1×10^{-5}).

This anomalous population may have its effect size analyzed for significance, using Cohen's d.

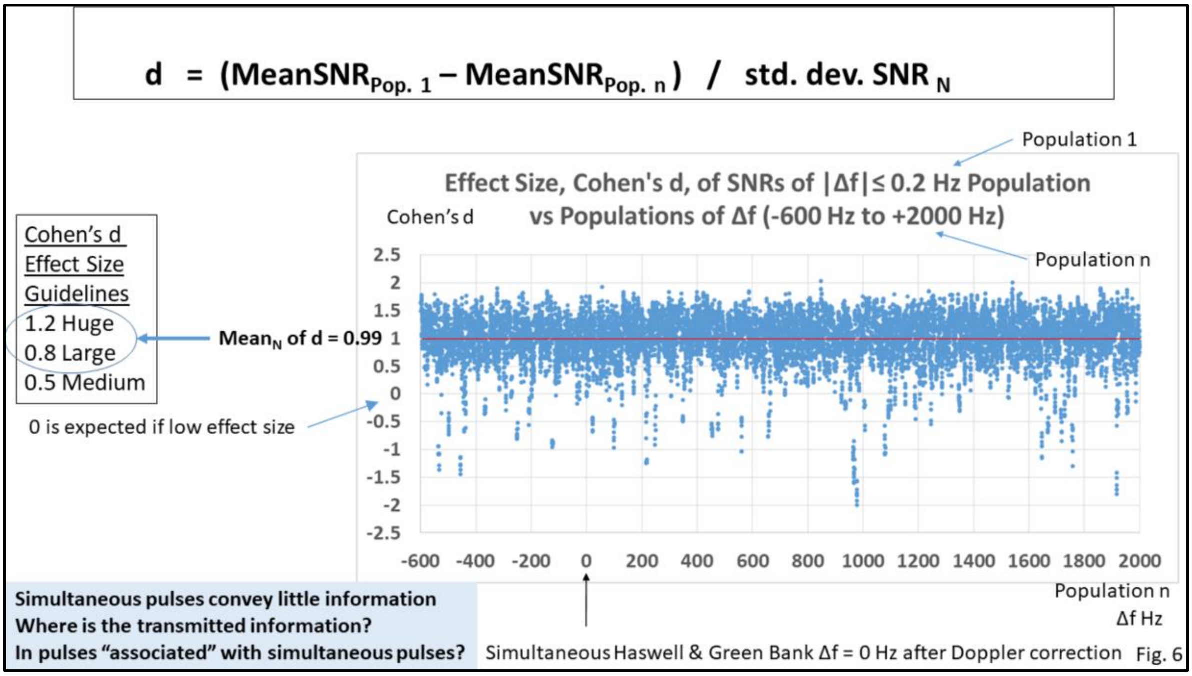


Fig. 6

Cohen's d analysis of anomalous simultaneous pulse population

Cohen's d (Fig. 6) may be used to compare the effect size of various populations of data. The SNRs of the anomalous ±0.2 Hz range population are compared to the same-size populations at other Doppler offsets, over the -600 to +2000 Hz Δf range. The guideline of Cohen's d indicates that the effect size of the selected population is midway between "large" and "huge".

A problem arises with simultaneous pulses in a communication system

If simultaneous pulses are present in a communication system, to provide detectability, the receiver might expect to observe additional close-time and/or close-frequency pulses that contain information. Intermittent simultaneous pulses, surmised to be for identification, are concise, and similar at spaced receiver locations, and therefore do not appear to carry a great amount of information. Other signals appear to be needed in a transmitter signal to increase channel capacity. To address this issue, hypothetical additional pulses potentially carry information and are referred to as "associated" pulses.

It is not necessary that associated transmitted pulses be present at multiple receiver locations, as a transmitter antenna system may use spatial filtering to increase channel capacity. If signal detectability has been achieved, in time and pointing direction, further transmitted signals are expected to ***not*** be simultaneously present, and may be readily received at high channel capacity. This concept is discussed further in hypothesis development.

Green Bank *associated* pulses

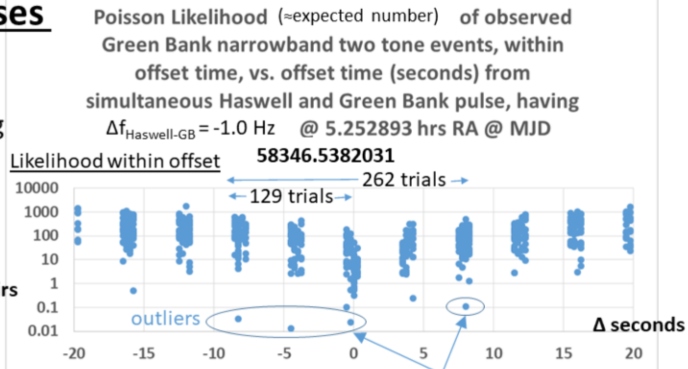
Within ± 8.25s of time of frequency offset = -1.0 Hz ***simultaneous*** 5.2529 hrs RA-pointing Haswell & Green Bank pulse

Likelihood Estimate

Tone spacing Rank in file	Δtime /file duration
	8s / 10,800s = 0.00074
	i.e. **129 trials of tone pairs**
339 Hz: 15/174,033	
607 Hz: 29/174,033	
2700.8 Hz: 139/174,033 =0.000799	**2794.0 Hz: 149/174,033 =0.000856**

Binomial distribution (3 seen,129 trials, pr. 0.000799) **=0.00016**

Binomial distribution (4 seen,262 trials, pr. 0.000856) **= 8.26x 10^{-5}**



Offset time		RF Freq. (¼ speed)		cSNR		Tone spacing (0.25 s per tone speed)		Event Likelihood	Event Likelihood within offset time
-8.25	s	1427.1206200	MHz	12.739	dB	607.2	Hz	0.000216	**0.0330**
-4.50	s	1429.6278708	MHz	12.596	dB	339.0	Hz	0.000120	**0.0128**
-0.50	s	1429.7780670	MHz	12.681	dB	5971.6	Hz	0.002120	0.1039
-0.25	s	1439.3661335	MHz	12.451	dB	2700.8	Hz	0.000959	**0.0240**
4.25	s	1430.8780819	MHz	12.682	dB	11153.5	Hz	0.003960	0.2416
8.00	s	1447.7421694	MHz	12	dB	2794.0	Hz	0.000992	0.1121

Fig. 7

7

Anomalous associated pulses are observed (Fig. 7)

A set of four two-tone pairs appear to exhibit the properties one might expect of associated pulses:

1. the observed four two-tone pulses are calculated to be rarely observed (Pr. 8.26x10^{-5}) in random noise,
2. due to a lack of symbol symmetry and/or repetition in the tones, the associated pulses appear to contain a moderate Shannon Entropy, indicating the potential of high information content,
3. the associated pulses occur close in time to the simultaneous pulse, indicating that a noise-caused hypothesis of simultaneous and associated tones may be refuted to a high likelihood,
4. the associated tones described above were not observed in Haswell data,
5. the rank of increasing-value-sorted tone spacing in a four hour duration file was used to estimate close tone likelihoods, after CW and narrowband RFI was excised, implying that un-excised broadband pulsed RFI is unlikely to adequately explain the near-simultaneity of the associated and simultaneous pulses. If multi-tone pulsed RFI is commonly present in the four hour data file, the measured rank of the close-spaced tones would indicate its presence, i.e. as a large number of close-tone outliers. These outliers were not observed.

A relationship between highly anomalous associated pulses and one of the two near-Rigel-pointing simultaneous pulses appears evident. Simultaneous ±1.0 Hz offset pulses occur randomly due to noise, on average, at approximately twenty-seven hour intervals, (97,200 seconds) while a highly unlikely multi-tone event (Pr. 8.26 x 10^{-5} in seventeen seconds) occurred at Green Bank, i.e. within ±8.25 s of the simultaneous Haswell-Green Bank pulse. The simultaneous and associated events appear to be time-correlated.

The four close-tone pairs' noise likelihood of 8.26 x 10^{-5} is calculated using the binomial distribution and a common event probability, based on the noise-likelihood of the 2794.0 Hz tone spacing. The expected composite multi-tone likelihood value is significantly lower than the calculated composite likelihood, due to the unlikely presence of the closer tone spacing pairs, at 607.2 and 339.0 Hz, given the probability as the 2794.0 Hz tone pair. During system validation, a comparison of post-RFI-excision tone spacing likelihood, in sky data, to the calculated probability of noise-induced Poisson close tone spacing, has been examined, with close match of sky noise to theory. The existence of bursting multi-tone RFI does not appear evident.

Further analysis is required, and potential equipment issues and assumptions need to be questioned, as a multi-tone transmit-receive mechanism using simultaneous and associated pulses appears to be present in the observed data, implying a powerful signal identification and RFI amelioration mechanism exists.

Green Bank associated pulses
re: RA 5.1838 hr ($\Delta f_{Haswell-GB}$ = -0.2 Hz)
at time of simultaneous pulse i.e. @ MJD 58345.53806130

Offset Time		RF Freq.		cSNR		FFT Δbins	Tone Spacing		Noise-caused Event Likelihood	Event Likelihood within offset frequency
0	s	1408.0909692	MHz	12.766	dB	2171903	8090969.2	Hz	2.493405525	
0	s	**1409.3006417**	MHz	12.106	dB	324719	1209672.5	Hz	0.372786514	
0	s	**1409.7767003**	MHz	12.217	dB	127791	476058.6	Hz	**0.14670765**	
0	s	**1410.0691803**	MHz	12.941	dB	78512	292480.0	Hz	**0.090133977**	
0	s	**1410.3755482**	MHz	**13.642**	dB	82240	306367.9	Hz	**0.094413825**	**0.003745399**
0	s	1414.4826546	MHz	12.381	dB	1102493	4107106.5	Hz	1.265692868	
0	s	1420.6385441	MHz	12.673	dB	1652459	6155889.5	Hz	1.897069253	
0	s	1421.4794949	MHz	12.301	dB	225741	840950.8	Hz	0.259156996	
0	s	1424.5024636	MHz	13.375	dB	811472	3022968.8	Hz	0.931592603	
0	s	1426.5113376	MHz	12.346	dB	539253	2008874.0	Hz	0.61907756	
0	s	1428.0339979	MHz	12.014	dB	408736	1522660.3	Hz	0.469240385	
0	s	1430.7576656	MHz	12.078	dB	731129	2723667.8	Hz	0.839356587	
0	s	1432.9897180	MHz	12.592	dB	599162	2232052.4	Hz	0.687854771	
0	s	**1440.9286693**	MHz	**13.293**	dB	2131096	7938951.3	Hz	2.446557947	**Simultaneous pulse**
0	s	1445.0470410	MHz	12.769	dB	1105517	4118371.8	Hz	1.269164505	
0	s	1447.0053740	MHz	18.865	dB	525686	1958333.0	Hz	0.603502264	

atypical

Typical tone spacing due to noise

13.642 dB rank
2,001/151,517
Pr. (noise) ≈ 0.013

Product multiplied by 3 because the four pulses could have been here

continuous CW RFI

Fig. 8

A second set of near-Rigel-pointing associated pulses are observed (Fig. 8)

A set of four anomalous close frequency spaced tones was observed **in the same 0.27 s interval** as the time of the 5.1838 hrs RA-pointing (near Rigel) simultaneous pulse. The combination of a high SNR and close tone spacing places the noise-caused likelihood of this event at approximately 0.00015.

There is a possibility that the Green Bank associated pulses are actually simultaneous at Haswell, yet below the SNR threshold of detection at Haswell. Four two-tone pairs at RA 5.2529 hrs, and the quad set at 5.1838 hrs, i.e. twelve anomalous pulses that do not appear in data captured at the Haswell receive site, leads to a thought that associated pulses may **not** be intentionally transmitted as simultaneous pulses.

Rician statistical modeling of simultaneous and associated pulses, considering RFI, SNR and polarization, is expected to refute, or support, simultaneous and associated pulses' hypotheses.

Bayesian Candidate-Model Inference

Candidate-Models are ***selectively*** a **null** hypothesis and the **alternate** hypotheses

Candidate models	Probability prior beliefs (Model m)	Explanation
A. Model Noise + telescopes	Probability **Model N** belief ≈ 1^- i.e. good confidence	e.g. a "good" null hypothesis. Does receiver noise and quiet sky explain data?
B. Model RFI + telescopes	Probability **Model R** belief ≈ 1^- to 0^+ depending on RFI knowledge	Does RFI into spaced antennas explain data?
C. Model ETI + telescopes	Probability **Model E** belief ≈ 0^+ to 1^- over time	Do the Drake Equation, Shannon's Law & observations explain data?
D. Models Unknown	Probability **Model U** belief ≈ 1^- to 0^+ over time	Unknown-caused – Where are the "gotchas"? Natural objects having low Doppler Spread?

Given the Data and Model, solve this 3 factor equation, for each Model m:

$$\Pr(\boldsymbol{Model\ m}\ \text{"}explains\text{"}\ given\ \boldsymbol{Data}) = \Pr(\boldsymbol{Data}\ \text{"explained"},\ given\ \boldsymbol{Model\ m}) \times \frac{\Pr(Model\ m\ belief)}{\Pr(Data\ valid)}$$

This Pr() is a Likelihood Function, developed for each Model m

Model Likelihood Functions
N: Rayleigh & Poisson stats
R: Find & model RFI
E: More observations
U: Find "gotchas"

Model Prior Belief Values
Pr. N ≈ 1 except "outliers"
Pr. R ≈ increasing to 1^- ?
Pr. E ≈ increasing to 1^- ?
Pr. U ≈ decreasing to 0^+?

Assume that Pr(Data Valid) ≈ **1** i.e. that Anomalous Data ("outliers" on the tail) are being observed. If Pr(Data Valid) << 1, then all known Models will ***appear*** to explain the Data.

Fig. 9

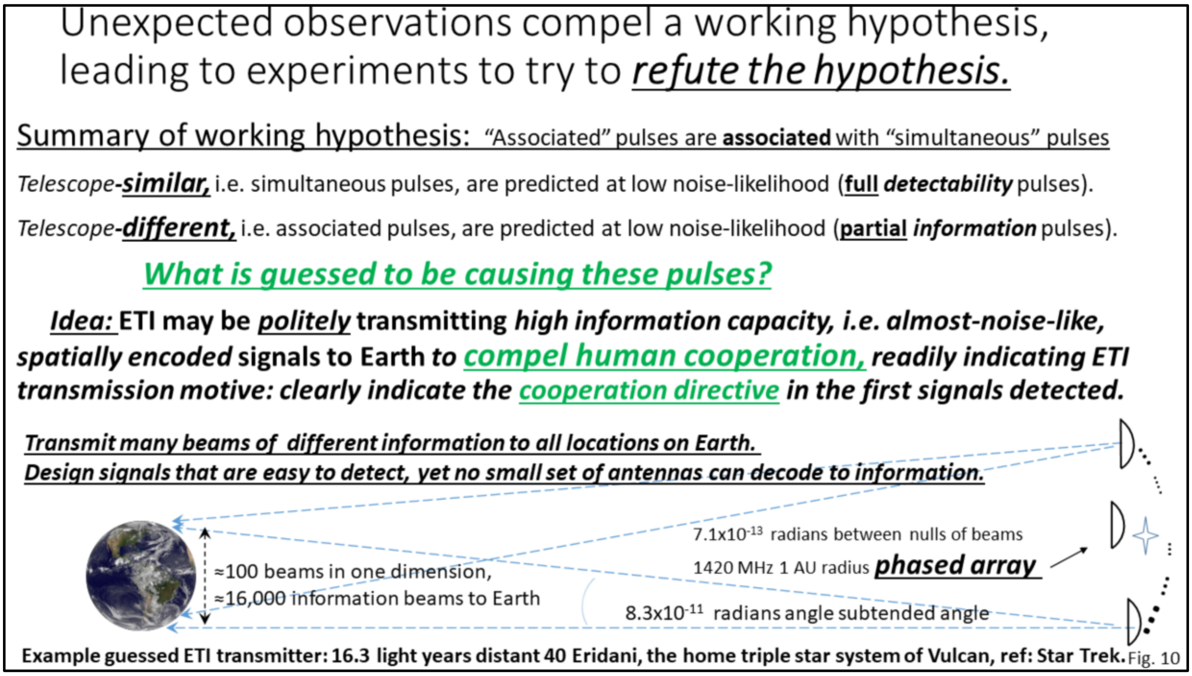
Anomalous data compels the analysis of multiple hypotheses

Bayesian Inference (Fig. 9) may be used to compare hypotheses, given that each Model, related to its hypothesis, may explain observed Data, to a varying degree. Observed Data is considered to include simultaneous pulses, pulses apparently associated in frequency and time with simultaneous pulses, and other associated pulses.

As observational work progresses, model development and follow-up observations may provide a refutation of various hypotheses. For example, the noise hypothesis has a high Model belief, based on our knowledge of communication systems. However, the noise hypothesis produces a Model having an apparent low probability in explaining observed anomalous Data. If the anomalous Data is considered valid, the three factor Bayesian Inference then yields a low probability that the Noise Model explains the observed Data.

Various receiver antenna spacing may be implemented. It is expected that multi-telescope observation of the same associated pulses should occur at close telescope spacing, if transmitted signals are celestially transmitted. The noise hypothesis may then be refuted to a statistical likelihood. Close antenna spacing, however, increases the probability that an RFI hypothesis, and its particular RFI Models, will explain the observed Data. Close antenna spacing and large antenna spacing may then be used together to selectively refute various hypotheses, to statistical significance.

Simultaneous narrow band pulse detection SETI systems provide a vehicle to refute RFI hypotheses, because the RFI models that include distant spaced telescopes, naturally filter RFI signals from the observed Data.



Reasoning that underlies an ETI hypothesis

A hypothesis is summarized in Fig. 10 that speculatively explains the reasoning behind the concept of simultaneous and associated pulses in a celestial communication system, in addition to the reasoning of ease of detectability and high celestial channel capacity.

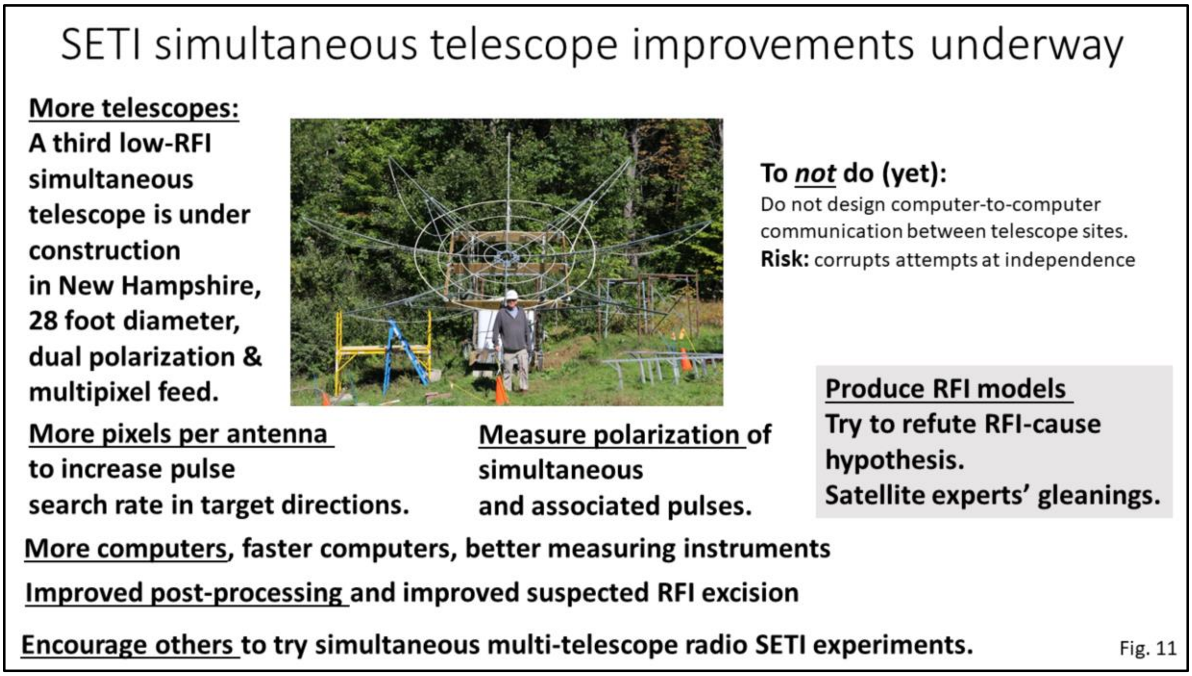
An altruistic and gregarious transmitting civilization might desire to convey useful information to other civilizations, without inducing ***information-fear*** among the potentially low-altruistic and warring polities of receiving civilizations.

The transmitter intentionally transmits an easily understood **"message within the message"**, at the signal layer, indicating that received signals cannot be decoded to information, without **cooperation** among the geographically-spaced polities of the receiving civilization.

After adequately refuting non-ETI hypotheses, each polity will immediately determine that the signals cannot be decoded to information with a small number of receive antennas, in a limited geographical area of the home planet. Cooperation will be thought to be important, and easy to reliably partake in, compared to planetary warring and attempted domination by one polity. Information-fear is ameliorated as the "message in a message" is gleaned similarly by the humans in all receiving polities.

The idea that a polity's controlled land and resources establishes the power of a polity, becomes insignificant in light of the power of a planetary civilization cooperating among all polities. Civilizations that learn to cooperate on their home planet will be able to benefit from information provided by advanced celestial civilizations. Civilizations that do not cooperate among polities on their home planet will be left without useful information they may likely need to survive.

Quantification and modeling of the expected parameters of simultaneous and associated pulses is underway, to aid in the refinement of further experimental procedures and refutation of various hypotheses, to statistical significance.



Improvements underway

Improvements (Fig. 11) are underway to enhance the statistical significance of observations that provide detail to the models needed in Bayesian Inference hypothesis selection.

Thank you

Steve Plock

Deep Space Exploration Society team

Green Bank Observatory team

family and friends

Thank You !!

Questions?