If aliens were sending us signals, this is what they might look like

April 8 2022, by Matt Williams

Composite image of the SKA combining all elements in South Africa and Australia. Credit: SKAO

For more than 60 years, scientists have been searching the cosmos for possible signs of radio transmission that would indicate the existence of extraterrestrial intelligence (ETI). In that time, the technology and methods have matured considerably, but the greatest challenges remain. In addition to having never detected a radio signal of extraterrestrial origin, there is a wide range of possible forms that such a broadcast

could take.

In short, SETI researchers must assume what a signal would look like, but without the benefit of any known examples. Recently, an international team led by the University of California Berkeley and the SETI Institute developed a new machine learning tool that simulates what a message from extraterrestrial intelligence (ETI) might look like. It's known as Setigen, an open-source library that could be a game-changer for future SETI research.

The research team was led by Bryan Brzycki, an astronomy graduate student at UC Berkeley. He was joined by Andrew Siemion, the Director of the Berkeley SETI Research Center, and researchers from the SETI Institute, Breakthrough Listen, the Dunlap Institute for Astronomy & Astrophysics, the Institute of Space Sciences and Astronomy, International Center for Radio Astronomy Research (ICRAR), and the Goergen Institute for Data Science.

Since the 1960s, the most common method of SETI has involved searching the cosmos for radio signals that are artificial in origin. The first such experiment was Project Ozma (April to July 1960), led by famed Cornell astrophysicist Frank Drake (creator of the Drake Equation). This survey relied on the 25-meter dish at the National Radio Astronomy Observatory in Green Bank, West Virginia, to monitor Epsilon Eridani and Tau Ceti at frequencies of about 400 kHz around 1.42 GHz.

These searches have since expanded to cover larger areas of the night sky, wider frequency ranges, and greater signal diversity. As Brzycki explained to Universe Today via email:

"In the 1960s, the idea was to focus on a region around a well-known frequency where neutral hydrogen emits radiation in interstellar space,

1.42 GHz. Since this natural emission is prevalent throughout the galaxy, the idea is that any intelligent civilization would know about it, and potentially target this frequency for transmission to maximize the chance of detection. Since then, especially as technology has rapidly advanced, radio SETI has expanded along all axes of measurement.

"We now can take measurements across a bandwidth of multiple GHz instantaneously. As storage has improved, we can collect huge amounts of data, allowing higher resolution observations in both time and frequency directions. By the same token, we've done surveys of nearby stars and other direction in the galaxy, to maximize exposure to potentially interesting directions in the sky."

Another major change has been the incorporation of machine learning-based algorithms designed to find transmissions amid the radio background noise of the cosmos and correct for radio frequency interference (RFI). The algorithms employed in SETI surveys have fallen into one of two categories: those that measure voltage time-series data and those that measure time-frequency spectrogram data.
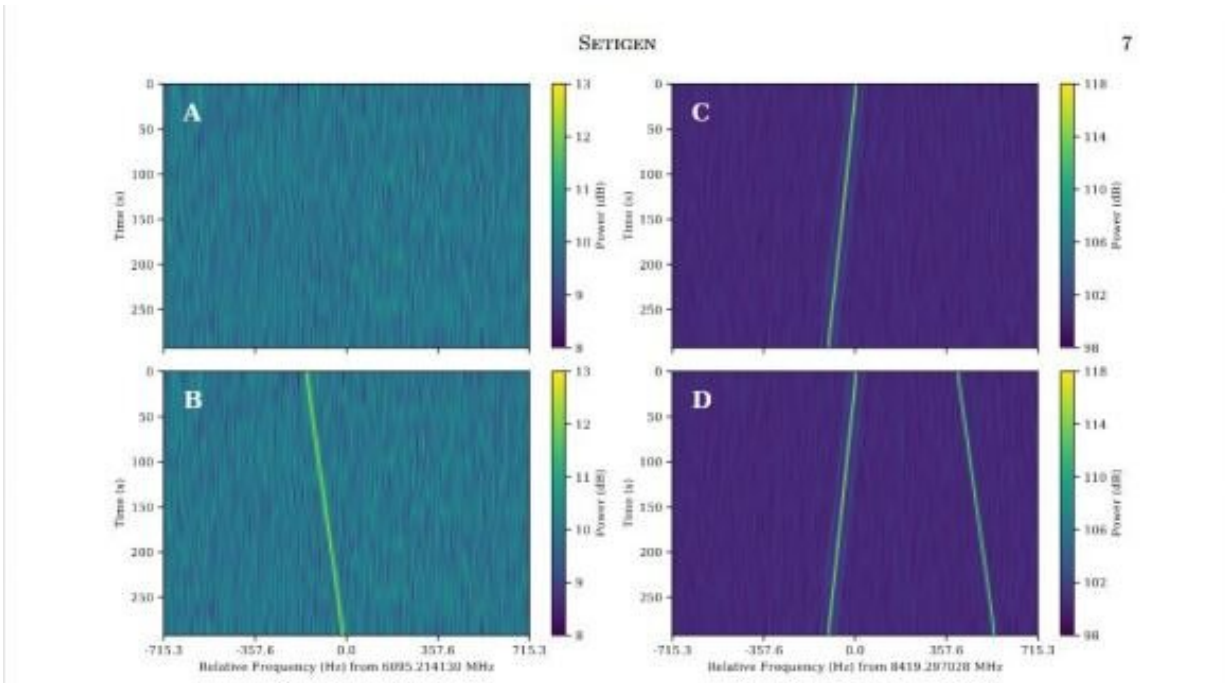
Radio spectrogram plots created from Setigen frames. Credit: Brzycki et al.

"The raw data collected by a radio antenna are voltage measurements; a radio wave induces a current in the antenna, which is read out and recorded as a voltage," said Brzycki. "A radio telescope is really just an antenna augmented by a parabolic dish to focus a larger area of light, increasing resolution and brightness. It turns out that intensity is proportional to voltage squared. Further, we care about the intensity as a function of frequency and time (the when and where of a potential signal)."

To get this, says Brzycki, astronomers start by employing algorithms that calculate the power of each frequency being observed towards the input time series data. In other words, the algorithm transforms radio signal data from a function of space and/or time into a function dependent on spatial frequency or temporal frequency—aka. a Fourier Transform

(FT). By squaring this, astronomers can measure the intensity of each frequency over the data-collection period.

"To get a full spectrogram, an array of intensity as a function of time and frequency, we take a section of the voltage-time series, get the FT, then repeat this process over the entire observation so that we can effectively stack a series of FT-data arrays on top of each other in the time direction," Brzycki added. "[O]nce you decide on a time resolution, we figure out the number of time samples needed and calculate the FT to see how much power lies in each frequency bin."

The primary search algorithm used by SETI researchers is known as the "incoherent tree deDoppler" algorithm, which shifts the spectrum of radio waves to correct for frequency drift and maximizes the signal-to-noise ratio of a signal. The most comprehensive SETI search program ever mounted, Breakthrough Listen, uses an open-source version of this algorithm known as TurboSETI, which has served as the backbone of many "technosignatures" searches (aka. signs of technological activity). As Brzycki explained, this method has some drawbacks: "The algorithm makes the assumption that a potential SETI signal is continuous with a high duty-cycle (meaning that it's almost always 'on'). Looking for a continuous sine-wave signal is a good first step since it's relatively easy and inexpensive in power for humans to produce and transmit such signals.

"Since TurboSETI is targeted for straight-line signals that are always 'on,' it can struggle to pick up alternate morphologies, like broadband and pulsed signals. Additional algorithms are being developed to try to detect these other kinds of signals, but as always, our algorithms are only as effective as the assumptions we make of the signals they are targeted for."

For SETI researchers, machine learning is a way of identifying

transmissions in raw radio frequency data and classifying multiple types of signals. The main issue, says Brzycki, is that the astronomical community doesn't have a dataset of ET signals, which makes supervised training difficult in the traditional sense. To this end, Brzycki and his colleagues developed a Python-based open-source library called Setigen that facilitates the production of synthetic radio observations.

"What Setigen does is facilitate the production of synthetic SETI signals, which can be used in entirely synthetic data, or added on top of real observational data to provide a more realistic noise and RFI background," said Brzycki. "This way, we can produce large datasets of synthetic signals to analyze the sensitivity of existing algorithms or to serve as a basis for machine learning training."

This library standardizes synthesis methods for search algorithm analysis, especially for existing radio observation data products like those used by Breakthrough Listen. "These come in both spectrogram and complex voltage (time series) formats, so having a method of producing mock data can be really useful for testing production code and developing new procedures," Brzycki added.

One of the 42 dishes in the Allen Telescope Array that searches for signals from space. Credit: Seth Shostak / SETI Institute.

Right now, algorithms for multi-beam observations are being developed using Setigen to produce mock signals. The library is also being constantly updated and improved as SETI research progresses. Brzycki and his colleagues also hope to add support for broad-band signal synthesis to aid search algorithms that target non-narrowband signals. More robust SETI surveys will be possible in the near future as next-generation radio telescopes become operational.

This includes Breakthrough Listen, which will be incorporating data from the MeerKAT array in South Africa. There's also the Square Kilometre Array (SKA), a massive radio telescope project that will

combine data from observatories in South Africa and Australia. These include the MeerKAT and Hydrogen Epoch of Reionization Array (HERA) in South Africa and the Australian SKA Pathfinder (ASKAP) and Murchison Widefield Array (MWA) in Australia.

Alas, there's is still the most limiting factor regarding SETI, which is our extremely limited frame of reference. When it comes right down to it, astronomers have no idea what an extraterrestrial signal would look like because we've never seen one before. This, paradoxically, makes it harder to tease out evidence of technosignatures amid the background noise of the cosmos. As such, astronomers are forced to take the "low-hanging fruit" approach, which means looking for technological activity as we know it.

However, by establishing parameters based on what is theoretically possible, scientists can narrow the search and increase the odds that they will find something someday. As Brzycki summarized:

"The only potential solution to this is some sort of unsupervised machine learning survey that minimizes our assumptions; work is being done on this front. Setigen certainly relies on this assumption—the synthetic signals one can produce are heuristic in nature, in that the user decides what they should look like.

"At the end of the day, the library provides a way of evaluating our existing algorithms and creating datasets of potential signals to develop new search methods, but the fundamental issues of where and when will always remain—the best we can do is to keep on looking."

At times like this, it is good to remind ourselves that the Fermi Paradox only needs to be resolved once. The moment we detect a radio transmission in the cosmos, we will know for certain that we are not alone in the Universe, that intelligent life can and does exist beyond

Earth, and is communicating using technologies we can detect.

More information: Bryan Brzycki et al, Setigen: Simulating Radio Technosignatures for SETI. arXiv:2203.09668v1 [astro-ph.IM], doi.org/10.48550/arXiv.2203.09668

Provided by Universe Today