Title
Maximizing Signal Detection and Improving Radio Frequency Interference Identification in the Search for Radio Technosignatures

Author
Pinchuk, Pavlo

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Maximizing Signal Detection and Improving Radio Frequency Interference Identification in the Search for Radio Technosignatures

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Physics

by

Pavlo Pinchuk

2021

ABSTRACT OF THE DISSERTATION

Maximizing Signal Detection
and Improving Radio Frequency Interference Identification
in the Search for Radio Technosignatures

by

Pavlo Pinchuk
Doctor of Philosophy in Physics
University of California, Los Angeles, 2021
Professor Jean-Luc Margot, Chair

In this work, I describe significant advancements to the signal detection and Radio Frequency Interference (RFI) identification capabilities of modern radio technosignature detection algorithms. These improvements are presented alongside the results of the analysis of four annual UCLA radio technosignatures searches spanning 2016–2019. First, I describe the UCLA SETI Group's initial versions of the signal detection and RFI identification algorithms, which were able to detect approximately 850 000 candidate signals within a frequency range of 1.15–1.73 GHz over $\sim$2 hours of observations with the 100 m diameter Green Bank Telescope in 2016. Next, I describe an improved candidate signal detection algorithm that detected approximately 6 million signals in a 2017 search for technosignatures with identical observational parameters. Importantly, I show that the common practice of ignoring frequency space around candidate detections can reduce the number of signals detected by a factor of four or more and presents significant problems when estimating figures of merit or upper limits on the prevalence of technosignatures. I then present further improvements to these detection algorithms, which introduce the use of the topographic prominence

for detection purposes and nearly double the signal detection count of some previously analyzed data sets. I also describe improvements to direction-of-origin filter algorithms, which are designed to remove most of the signals attributable to RFI from the data. The updated algorithms ensure a unique link between signals observed in separate scans. Finally, I present a novel machine-learning-based RFI mitigation algorithm, which helps address a major remaining challenge in the search for radio technosignatures. Specifically, I describe the design and deployment of a Convolutional Neural Network (CNN) that can determine whether or not a signal detected in one scan is also present in another scan. This CNN-based filter outperforms both a baseline 2D correlation model as well as existing filters over a range of metrics and reduces the number of signals requiring visual inspection after the application of traditional filters by a factor of 6–16 in nominal situations.

The dissertation of Pavlo Pinchuk is approved.

Michael P. Fitzgerald

Troy A. Carter

David Saltzberg

Jean-Luc Margot, Committee Chair

University of California, Los Angeles

2021

To my parents . . .

for their tireless support, care, and nourishment,

and to Evie . . .

for her unfaltering love throughout this journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my parents for all of the long phone calls, the care packages, the cozy family holidays, and the support they gave me along this journey.

I would like to thank Evie for believing in me and giving me a reason to wake up every morning.

I would like to thank John for taking a leap of faith and moving to the great city of Los Angeles with me.

I would like to thank Sanjana for all the great discussions and adventures hunting for free food all over the UCLA campus.

I would like to thank Ashok for all of the much-needed coffee breaks and for working around my extrememly long-running programs on our shared server.

I would like to sincerely thank my advisor, Jean-Luc Margot, for his mentorship over the past five years. He has helped me overcome numerous difficulties, and I would have abandoned my studies long ago if it were not for his patience, care, and dedication to my success. I am forever grateful for the opportunity to grow as a person and a scholar under his guidance.

A special additional thank you to the members of my dissertation committee: David Saltzberg, Troy A. Carter, and Michael P. Fitzgerald. Their advice helped make this work the best it can be.

I gratefully acknowledge the American Astronomical Society and IOP Publishing for their permission to reprint articles originally published in The Astronomical Journal.

Chapter 1 is partially inspired by Pinchuk et al. [2019]. I thank and acknowledge the coauthors for their contributions: J.-L. Margot, A. H. Greenberg, T. Ayalde, C. Bloxham, A. Boddu, L. G. Chinchilla-Garcia, M. Cliffe, S. Gallagher, K. Hart, B. Hesford, I. Mizrahi, R. Pike, D. Rodger, B. Sayki, U. Schneck, A. Tan, Y. Xiao, and R. S. Lynch.

Chapter 2 is an adapted and partially modified version, including figures, of Margot et al. [2018]. I thank and acknowledge the coauthors for their contributions: J.-L. Margot, A. H. Green-

berg, A. Shinde, Y. Alladi, S. Prasad MN, M. O. Bowman, C. Fisher, S. Gyalay, W. McKibbin, B. Miles, D. Nguyen, C. Power, N. Ramani, R. Raviprasad, J. Santana, and R. S. Lynch.

Chapter 3 is an adapted and partially modified version, including figures, of Pinchuk et al. [2019]. I thank and acknowledge the coauthors for their contributions: J.-L. Margot, A. H. Greenberg, T. Ayalde, C. Bloxham, A. Boddu, L. G. Chinchilla-Garcia, M. Cliffe, S. Gallagher, K. Hart, B. Hesford, I. Mizrahi, R. Pike, D. Rodger, B. Sayki, U. Schneck, A. Tan, Y. Xiao, and R. S. Lynch.

Chapter 4 is an adapted and partially modified version, including figures, of Margot et al. [2021]. I thank and acknowledge the coauthors for their contributions: J.-L. Margot, R. Geil, S. Alexander, S. Arora, S. Biswas, J. Cebreros, S. P. Desai, B. Duclos, R. Dunne, K. K. L. Fu, S. Goel, J. Gonzales, A. Gonzalez, R. Jain, A. Lam, B. Lewis, R. Lewis, G. Li, M. MacDougall, C. Makarem, I. Manan, E. Molina, C. Nagib, K. Neville, C. O'Toole, V. Rockwell, Y. Rokushima, G. Romanek, C. Schmidgall, S. Seth, R. Shah, Y. Shimane, M. Singhal, A. Tokadjian, L. Villafana, Z. Wang, I. Yun, L.a Zhu, and R. S. Lynch.

Chapter 5 is an adapted and partially modified version, including figures, of Pinchuk and Margot [2021]. I thank and acknowledge J.-L. Margot for his contributions.

I gratefully acknowledge partial support from the Nathan P. Myhrvold Graduate Fellowship.

VITA

2013	Advanced Placement Scholar with Distinction
2013—2015	University of Colorado, Colorado Springs President's list
2016	B.S. Physics summa cum laude, *University of Colorado, Colorado Springs* Colorado Springs, Colorado
2016	Alyne and Leon Camp Fellow
2017	M.S. Physics, Department of Physics and Astronomy, *University of California, Los Angeles* Los Angeles, California
2017–2019	Graduate Teaching Assistant, Department of Physics and Astronomy, *University of California, Los Angeles* Los Angeles, California
2017–2021	Graduate Research Assistant, UCLA SETI Group, Department of Earth, Planetary, and Space Sciences, *University of California, Los Angeles* Los Angeles, California
2020	Nathan P. Myhrvold Graduate Fellow

PUBLICATIONS

J.-L. Margot, A. H. Greenberg, **P. Pinchuk**, A. Shinde, Y. Alladi, S. Prasad MN, M. O. Bowman, C. Fisher, S. Gyalay, W. McKibbin, B. Miles, D. Nguyen, C. Power, N. Ramani, R. Raviprasad, J. Santana, and R. S. Lynch. A Search for Technosignatures from 14 Planetary Systems in the Kepler Field with the Green Bank Telescope at 1.15—1.73 GHz. *The Astronomical Journal*, 155:209, 2018. doi: 10.3847/1538-3881/aabb03

P. Pinchuk, J.-L. Margot, A. H. Greenberg, T. Ayalde, C. Bloxham, A. Boddu, L. G. Chinchilla-Garcia, M. Cliffe, S. Gallagher, K. Hart, B. Hesford, I. Mizrahi, R. Pike, D. Rodger, B. Sayki, U. Schneck, A. Tan, Y. Xiao, and R. S. Lynch. A Search for Technosignatures from TRAPPIST-1, LHS 1140, and 10 Planetary Systems in the Kepler Field with the Green Bank Telescope at 1.15–1.73 GHz. *The Astronomical Journal*, 157(3):122, 2019. doi: 10.3847/1538-3881/ab0105

J.-L. Margot, **P. Pinchuk**, R. Geil, S. Alexander, S. Arora, S. Biswas, J. Cebreros, S. P. Desai, B. Duclos, R. Dunne, K. K. L. Fu, S. Goel, J. Gonzales, A. Gonzalez, R. Jain, A. Lam, B. Lewis, R. Lewis, G. Li, M. MacDougall, C. Makarem, I. Manan, E. Molina, C. Nagib, K. Neville, C. O'Toole, V. Rockwell, Y. Rokushima, G. Romanek, C. Schmidgall, S. Seth, R. Shah, Y. Shimane, M. Singhal, A. Tokadjian, L. Villafana, Z. Wang, I. Yun, L.a Zhu, and R. S. Lynch. A Search for Technosignatures Around 31 Sun-Like Stars with the Green Bank Telescope at 1.15—1.73 GHz. *The Astronomical Journal*, 161(2):55, 2021. doi: 10.3847/1538-3881/abcc77

P. Pinchuk, J.-L. Margot. A Machine-Learning-Based Direction-of-Origin Filter for the Identification of Radio Frequency Interference in the Search for Technosignatures. *submitted* (2021).

CHAPTER 1

Introduction

The question "Are we alone in the Universe?" is one of the most enduring and fundamental unanswered questions in modern science. The quest for a definitive answer to this question motivates the search for evidence of life in the Solar System and beyond. Over time, two primary strategies for the search have emerged. One focuses on biosignatures, which are defined as scientific evidence of past or present life. The other focuses on technosignatures, which are defined as any scientific evidence of the existence of extinct or extant technology. Given our present knowledge of astrobiology, it is impossible to reliably predict which strategy will succeed first. However, the search for technosignatures offers four distinct advantages compared to telescopic or robotic searches for biosignatures.

First, present-day human technology limits the targets for a biosignature search to planetary systems around a few dozen nearby stars, including our own Sun. In contrast, the number of available targets for a directed technosignature search is likely at least a billion times larger, given the abundance of potentially habitable exoplanets [Borucki, 2016] and our ability to detect technosignatures emitted thousands of light years away [e.g., Margot et al., 2018]. Second, the detection of a biosignature outside the Solar System, along with its interpretation, may remain ambiguous and controversial for many years, whereas a confirmed detection of an extraterrestrial technosignature of the type described in this work would offer a high level of certainty in interpretation. Third, the information content of some biosignatures may remain limited post-detection, while the content of an intentional extraterrestrial transmission could potentially provide an unparalleled advance in knowledge. Finally, a substantial search for technosignatures can be accomplished with only a

small (<1%) fraction of the budget currently allocated to the search for biosignatures, including missions like Mars 2020, Europa Clipper, and the James Webb Space Telescope.

The UCLA SETI Group has been conducting annual observations for radio technosigatures since 2016. This search builds on the legacy of technosignature searches performed in the period 1960–2010 [Tarter, 2001, Tarter et al., 2010, and references therein] and compares favorably to other recent searches in terms of end-to-end sensitivity, frequency drift rate coverage, and signal detection count per unit bandwidth per unit integration time. So far, the UCLA SETI Group has focused on detecting signals that are narrow in the frequency domain, which would provide compelling evidence of the existence of another civilization. A repeatedly observable signal that originates from an extrasolar source and is narrow (< 10 Hz) in the frequency domain is a technosignature because natural sources do not emit such narrowband signals. The narrowest reported natural emission spans about 500 Hz and corresponds to OH (1612 MHz) maser emission [Cohen et al., 1987].

Although no technosignatures have been discovered to date, the tools and algorithms used to analyze the radio baseband data have undergone multiple levels of improvements. In the following chapters, I describe significant advancements to the signal detection and Radio Frequency Interference (RFI) identification capabilities of modern radio technosignature detection algorithms. These improvements are presented alongside the results of the analysis of four annual UCLA radio technosignatures searches spanning 2016–2019.

In Chapter 2, I describe the UCLA SETI Group's initial versions of the signal detection and RFI identification algorithms. Specifically, I describe a search for technosignatures that is sensitive to Arecibo-class transmitters located within ∼420 ly of Earth and transmitters that are 1000 times more effective than Arecibo within ∼13 000 ly of Earth. The observations focused on 14 planetary systems in the Kepler field and used the L-band receiver (1.15–1.73 GHz) of the 100 m diameter Green Bank Telescope. The initial signal detection algorithms identified approximately 850 000 candidates, corresponding to a hit rate density of 1.691×10^{-3} hits per hour per hertz (see Section 4.5 for a definition of hit rate density). Most (99%) of these signals were automatically classified

as human-generated RFI. A large fraction (>99%) of the remaining candidate signals were also flagged as anthropogenic RFI because they have frequencies that overlap those used by global navigation satellite systems, satellite downlinks, or other interferers detected in heavily polluted regions of the spectrum.

In Chapter 3, I describe an improved candidate signal detection algorithm that detected approximately 6 million signals in a 2017 search for technosignatures with identical observational parameters. The observations focused primarily on planetary systems in the Kepler field, but also included scans of the recently discovered TRAPPIST-1 and LHS 1140 systems. The improved data processing pipeline yielded a hit rate density of 1.07×10^{-2} hits per hour per hertz, a significant improvement over the pipeline described in the previous chapter, and classified over 98% of the detected signals as anthropogenic Radio Frequency Interference (RFI). I discuss the problems associated with the common practice of ignoring frequency space around candidate detections in radio technosignature detection pipelines, which include inaccurate estimates of figures of merit and unreliable upper limits on the prevalence of technosignatures. The improved candidate signal detection algorithm mitigates these problems and improves the efficiency of the search. Specifically, the new algorithm increases the number of candidate detections by a factor of more than four compared to the detection algorithms described in Chapter 2.

In Chapter 4, I present further improvements to these detection algorithms and apply them to a search for technosignatures in regions surrounding 31 Sun-like stars near the plane of the Galaxy. Specifically, the improved candidate signal detection procedure relies on the topographic prominence of the signal power and nearly doubles the signal detection count of some previously analyzed data sets. The improved algorithms detected candidate signals at a hit rate density of 2.21×10^{-2} hits per hour per hertz, significantly outperforming the algorithms used previously by the UCLA SETI Group as well as those used by the Breakthrough Listen (BL) team. I also describe improvements to the direction-of-origin filters that remove most RFI. These improvements promote unique links between signals observed in separate scans. Additionally, I describe the results of a preliminary signal injection and recovery analysis to test the performance of the UCLA SETI

Group's pipeline. The search described in this chapter compares favorably to other recent searches in terms of end-to-end sensitivity, frequency drift rate coverage, and signal detection count per unit bandwidth per unit integration time.

In Chapter 5, I present a novel machine-learning-based RFI mitigation algorithm, which helps address a major remaining challenge in the search for radio technosignatures. Specifically, I describe the design and deployment of a Convolutional Neural Network (CNN) that can determine whether or not a signal detected in one scan is also present in another scan. CNNs offer a promising complement to existing filters because they can be trained to analyze dynamic spectra directly, instead of relying on inferred signal properties. I describe the compilation of several data sets consisting of labeled pairs of images of dynamic spectra, as well as the design and training of a CNN that can determine whether or not a signal detected in one scan is also present in another scan. This CNN-based DoO filter outperforms both a baseline 2D correlation model as well as existing DoO filters over a range of metrics. I show that the CNN reduces the number of signals requiring visual inspection after the application of traditional DoO filters by a factor of 6–16 in nominal situations. Integrating this ML-based DoO filter into existing radio technosignature search pipelines has the potential of providing accurate RFI identification in near-real-time.

Chapter 6 briefly summarizes the results presented in this work.

CHAPTER 2

A search for technosignatures from 14 planetary systems in the Kepler field with the Green Bank Telescope at 1.15–1.73 GHz

Analysis of Kepler mission data suggests that the Milky Way includes billions of Earth-like planets in the habitable zone of their host star. Current technology enables the detection of technosignatures emitted from a large fraction of the Galaxy. We describe a search for technosignatures that is sensitive to Arecibo-class transmitters located within ∼420 ly of Earth and transmitters that are 1000 times more effective than Arecibo within ∼13 000 ly of Earth. Our observations focused on 14 planetary systems in the Kepler field and used the L-band receiver (1.15–1.73 GHz) of the 100 m diameter Green Bank Telescope. Each source was observed for a total integration time of 5 minutes. We obtained power spectra at a frequency resolution of 3 Hz and examined narrowband signals with Doppler drift rates between ±9 Hz s^{-1}. We flagged any detection with a signal-to-noise ratio in excess of 10 as a candidate signal and identified approximately 850 000 candidates. Most (99%) of these candidate signals were automatically classified as human-generated radio-frequency interference (RFI). A large fraction (>99%) of the remaining candidate signals were also flagged as anthropogenic RFI because they have frequencies that overlap those used by global navigation satellite systems, satellite downlinks, or other interferers detected in heavily polluted regions of the spectrum. All 19 remaining candidate signals were scrutinized and none were attributable to an extraterrestrial source.

2.1 Introduction

Analysis of Kepler mission data suggests that the Milky Way includes billions of Earth-like planets in the habitable zone (HZ) of their host star [e.g., Borucki, 2016]. The possibility that intelligent and communicative life forms developed on one or more of these worlds behooves us to conduct a search for extraterrestrial intelligence. Here, we describe an L-band radio survey of 14 planetary systems selected from the Kepler mission field. Our analysis methods are generally similar to those used by Siemion et al. [2013], but our observations sample a different slice of the search volume. In addition, our analysis examines signals of lower signal-to-noise ratio (10 vs. 25) and larger range of Doppler drift rates (± 9 Hz s^{-1} vs. ± 2 Hz s^{-1}) than recent Breakthrough Listen results [Enriquez et al., 2017].

We define a "technosignature" as any measurable property or effect that provides scientific evidence of past or present technology, by analogy with "biosignatures," which provide evidence of past or present life. The detection of a technosignature such as an extraterrestrial signal with a time-frequency structure that cannot be produced by natural sources would provide compelling evidence of the existence of another civilization. A signal that is narrow (< 10 Hz) in the frequency domain is a technosignature because natural sources do not emit such narrowband signals. The narrowest reported natural emission spans about 550 Hz at L-band and corresponds to OH (1612 MHz) maser emission [Cohen et al., 1987]. A monochromatic signal that shifts by ± 10 Hz as a function of time according to a complex sequence in a manner similar to that used in transmitting the 1974 Arecibo message [The staff at the National Astronomy and Ionosphere Center, 1975] is another technosignature. This work focuses on detecting signals that are narrow in the frequency domain, and is sensitive to both of these examples. Our data are also amenable to searching for signals that are narrow in the time domain (e.g., pulses).

Our search is not predicated on the assumption of deliberate transmissions aimed at Earth. Earthlings, for instance, use high-power ($\sim 10^6$ W) transmissions to study asteroids that may pose an impact hazard [e.g., Naidu et al., 2016]. These transmissions use monochromatic, binary phase-

coded, or chirp signals, all of which would be recognized as technosignatures by alien civilizations. In most such observations, less than a millionth of the energy is absorbed and scattered by the asteroid, and the remainder propagates beyond the asteroid at the speed of light. Our search is agnostic about whether radio transmissions were intended for detection by a distant civilization (e.g., a beacon) or not (e.g., a radar or inter-planet telecommunication system).

Sections 2.2, 2.3, 2.4, and 2.5 describe the observations, analysis, discussion, and conclusions, respectively.

2.2 Observations

We selected 14 exoplanet host stars (Table 2.1) from the Kepler catalog. A majority of these stars host small habitable zone planets with radii $R_p < 2R_E$, where R_E is Earth's radius [Kane et al., 2016]. Although such planets may be advantageous for the development of extraterrestrial life forms, advanced civilizations may be capable of thriving in a variety of environments, and we do not restrict our search to small habitable zone planets. Indeed, because planets and habitable zone planets are common among most stars in the Galaxy [e.g., Petigura et al., 2013, Batalha, 2014], there is no compelling reason to search the Kepler field [e.g., Siemion et al., 2013] as opposed to other fields. There is, however, a possible increased probability of detecting technosignatures when observing planetary systems edge-on, because it has been shown that most planets have low orbital inclinations [$< 3°$; Fang and Margot, 2012], in which case interplanetary transmissions may be aligned with Earth.

Our observing sequence was inspired by a solution to the traveling salesperson problem, which minimized the time spent repositioning the telescope. Each target was observed twice in the following 4-scan sequence: target 1, target 2, target 1, target 2. The integration time for each scan was $\tau = 150$ s, yielding a total integration time of 5 minutes per target.

We conducted our observations with the 100 m diameter Green Bank Telescope (GBT) [Jewell and Prestage, 2004] on 2016 April 15, 16:00-18:00 Universal Time (UT). We recorded both linear

polarizations of the L-band receiver, which has a frequency range of 1.15–1.73 GHz. Over this frequency range, the full-width half maximum (FWHM) beam width of the telescope is 11 arcmin–7.3 arcmin. The aperture efficiency is $\sim$72%,[1] which provides an effective area of $\sim$5600 m^2 and telescope sensitivity of $\sim$2 K/Jy.[2] At elevations above $\sim$20 degrees, the system temperature is $\sim$20 K and the system equivalent flux density (SEFD) is $\sim$10 Jy. The L-band receiver is located at the Gregorian focus of the telescope, which was designed with an off-axis reflector to minimize stray radiation.

We used the GUPPI backend [DuPlain et al., 2008] in baseband recording mode and sampled 800 MHz of bandwidth from 1.1 to 1.9 GHz. The signal was channelized into 256 channels of 3.125 MHz bandwidth each. The raw voltages of the in-phase and quadrature channels were digitized with 8-bit quantization. GUPPI's baseband recording mode enables reduction of the data storage requirements by a factor of four with minimal signal degradation with an optimal four-level (two-bit) sampler [Kogan, 1998]. In this mode, the quantization thresholds are set to $-0.981599\ \sigma$, 0, $+0.981599\ \sigma$, where σ is the root-mean-square (rms) of the voltage and the quantized levels are set to ± 1 and ± 3.335875. The quantization efficiency, which is the ratio of signal power that is observed with the four-level sampler to the power that would be obtained with no quantization loss, is $\eta_Q = 0.8825$ [Kogan, 1998]. Eight computers handled the transfer of our data to 8 disk arrays at an aggregate rate of 800 MB/s or 2.88 TB/h.

Calibration procedures at the beginning of our observing window consisted of recording a monochromatic tone at 1501 MHz (center frequency + 1 MHz), performing a peak and focus procedure on a bright radio source near the Kepler field, and observing a bright pulsar near the Kepler field (PSR B2021+51) [Manchester et al., 2005].

[1] https://science.nrao.edu/facilities/gbt/proposing/GBTpg.pdf

[2] 1 Jy = 10^{-26} Wm^{-2}Hz^{-1}.

2.3 Analysis

2.3.1 Validations

We verified the validity of our data-processing pipeline by analyzing the monochromatic tone data and recovering the signal at the expected frequency. We also folded the pulsar data at the known pulsar period and recovered the characteristic pulse profile.

2.3.2 Data selection

We unpacked the data to 4-byte floating point values, computed Fourier transforms of the complex samples with the FFTW routine [Frigo and Johnson, 2005], and calculated the signal power (Stokes I) at each frequency bin. Signals with frequencies outside the range of the L-band receiver (1150–1730 MHz) were discarded. We used the GBT's notch filter to mitigate interference from a nearby aircraft surveillance radar system. Signals with frequencies within the 3 dB cutoff range of the notch filter (1200–1341.2 MHz) were also discarded.

2.3.3 Bandpass correction

Individual channels of the GUPPI instrument exhibited a mostly uniform bandpass response, with a few notable exceptions. We fit a 16-degree Chebyshev polynomial to the median bandpass response of well-behaved channels. We divided each power spectrum by this median response, which normalized signal levels across the entire bandpass.

2.3.4 Spectral analysis

We used a Fourier transform length of 2^{20}, corresponding to a time interval of $\Delta\tau = 0.336$ s and yielding a frequency resolution $\Delta f = 2.98$ Hz. This choice of transform length and frequency resolution was dictated by our desire to examine drift rates of order 10 Hz s^{-1} (Section 2.3.5). We stored about 450 consecutive power spectra, depending on the exact integration time of each

scan, in frequency–time arrays of 2^{20} columns and 432–451 rows (hereafter, "time-frequency diagrams", sometimes known as "spectrograms" or "spectral waterfalls"). The average noise power was subtracted and the array values were scaled to the standard deviation of the noise power.

2.3.5 Drift rate analysis

Because radio signals experience time-variable Doppler shifts due to the rotational and orbital motions of both emitters and observers, uncompensated signals smear in frequency space on a timescale $\Delta f / \dot{f}$, where Δf is the frequency resolution and $\dot{f}$ is the Doppler drift rate. The duration of a time series required to obtain a spectrum with resolution Δf is $1/\Delta f$, such that the maximum drift rate that is observable without smear is $\dot{f}_{\rm max} = \pm(\Delta f)^2$. The maximum Doppler drift rates due to Earth's rotational and orbital motions are $\sim$0.15 Hz s^{-1} and $\sim$0.034 Hz s^{-1}, respectively, at the maximum frequency of our observations and at Green Bank's latitude. The Doppler drift rates due to the emitters are unknown. We examined drift rates of up to $\dot{f}_{\rm max} = \pm(\Delta f)^2 = \pm 8.88$ Hz s^{-1}, corresponding to accelerations of $\pm$1.6 m s^{-2}, which admit a wide range of planetary radii, spin rates, orbital semi-major axes, and orbital periods. We applied a de-smearing procedure to compensate for the accelerations of both emitter and observer. Specifically, we implemented a computationally advantageous tree algorithm [Taylor, 1974, Siemion et al., 2013], which enabled examination of 512 Doppler drift rates from 0 to 8.88 Hz s^{-1} in linearly spaced increments of 0.0173 Hz s^{-1}. This algorithm reads the frequency–time arrays, then shifts and sums all powers corresponding to each of the 512 possible drift rate values, effectively enabling integration of the signal power over the entire scan duration without smear. Application of the algorithm with positive and negative Doppler drift rates resulted in two frequency–drift rate arrays of 2^{20} columns and 512 rows for each scan.

2.3.6 Candidate signal detection

We identified candidate signals with an iterative procedure. We searched for the element with the highest signal-to-noise ratio (S/N) in the frequency–drift rate arrays. The characteristics of this candidate signal (unique identifier, source name, scan number, scan start time, frequency at start of scan, drift rate, S/N, frequency resolution) were stored in a SQL database for subsequent analysis. Because a candidate signal would often be detected redundantly at multiple drift rate values adjacent to that with the highest S/N, we decided to keep only the instance with the highest S/N value. In order to do so, we blanked the frequency–drift rate arrays in a region of frequency extent $\pm\dot{f}_{\max}\tau$ centered on the frequency of the highest S/N candidate signal. We then repeated the procedure and searched for the element with the next highest S/N. All candidate signals with S/N > 10 were identified in this fashion and stored in the database. We counted 858 748 candidate signals, which amounts to $\sim$750 000 candidates per hour of on-source integration time in the useful frequency range of the GBT L-band receiver.

2.3.7 Rejection algorithms

A signal from a distant source at rest or in uniform motion with respect to the observer exhibits no time variation in the value of the Doppler shift. Signals from extraterrestrial sources, unless cleverly compensated for a specific location on Earth, experience a nonzero Doppler drift rate due in part to the rotational and orbital motions of Earth. For these reasons, we categorized all signals with zero Doppler drift rate as likely terrestrial and eliminated them from further consideration. About a quarter (231 181) of the candidate signals were flagged on this basis, leaving 627 567 candidates with nonzero Doppler drift rates.

To further distinguish between radio-frequency interference (RFI) and genuine extraterrestrial signals, we implemented two additional filters. First, we flagged any signal that was not detected in both scans of the same source. This filter can rule out many anthropogenic signals that temporarily enter the beam (e.g., satellite downlinks). Second, we flagged any signal that appeared in more

than one position on the sky. This filter can rule out many anthropogenic signals that are detectable through the antenna sidelobes. A logical AND was used to automatically flag candidate signals that remained for consideration after the rejection steps. Our rejection filters used the scan start times, frequencies, Doppler drift rates, and frequency resolutions stored in the SQL database to properly recognize signals from the same emitter observed at different times. These filters successfully flagged 617 410 of the remaining signals as likely anthropogenic, leaving 10 157 signals for further investigation.

Overall, our rejection filters automatically eliminated 99% of the initial detections as RFI.

2.3.8 Known interferers

Several regions of the spectrum exhibit an unusually high density of detections (Figure 2.1). Most of these high-signal-density regions can be attributed to known interferers. We discarded all candidate signals in these regions, which we defined by the frequency extents shown in Table 2.2. For signals generated by global navigation satellite systems, we used the signal modulation characteristics, typically binary phase-shift keying (BPSK), to delineate the frequency extent. For satellite downlinks, we used the Federal Communications Commission (FCC) table of frequency allocations. On average, the density of detections in the frequency regions ascribed to these satellites is $\sim$3 500 detections per MHz of bandwidth. In contrast, the density of signals in regions of the spectrum that exclude these interferers is $\sim$500 detections per MHz of bandwidth.

We flagged 9 663 candidate signals out of 10 157 as most likely due to these known interferers, leaving 494 signals for further consideration.

2.3.9 Additional interferers

After we excised the known interferers listed in Table 2.2, we identified hundreds to thousands of detections with common time-frequency characteristics in narrow regions of the spectrum spanning a total of $\sim$9 MHz (Table 2.3). Because these classes of signals were each detected in at least six

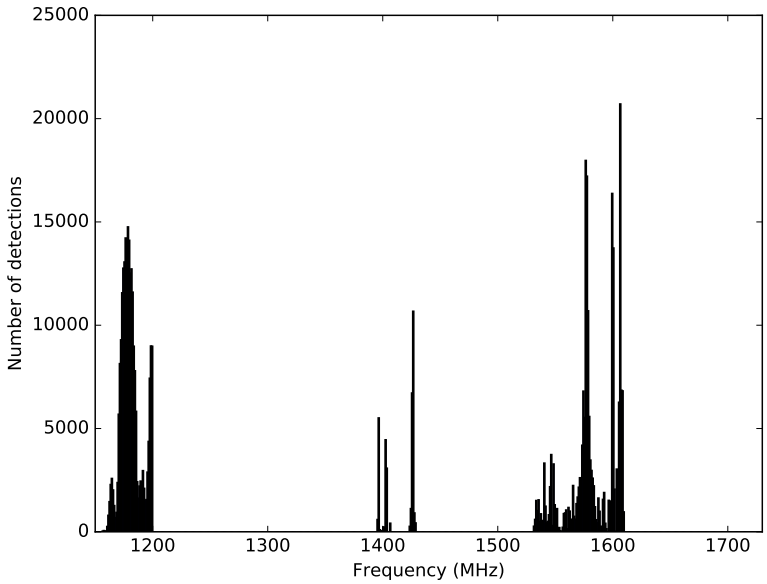

Figure 2.1: Number of detections as a function of frequency, prior to application of our rejection algorithms (Section 2.3.7). Most of the high-density regions are due to the known interferers listed in Table 2.2. The three clusters near 1400 MHz correspond to additional interferers described in Table 2.3.

distinct pointing directions, they are almost certainly anthropogenic, and we flagged them as RFI. We flagged 456 candidate signals out of 494 as most likely due to RFI, leaving 38 signals for further consideration.

We detected strong interference in the radio astronomy protected band (1400–1427 MHz). The interferers are visible in histograms of our detections in two 9-MHz-wide regions of the spectrum (Figure 2.2). We briefly describe the characteristics of these interferers below.

The interferers near 1396 MHz, 1400 MHz, 1403 MHz, and 1406 MHz are similar. They have a comb-like appearance in the frequency domain, with spectral features spaced every kilohertz. Each feature has a bandwidth of about 140 Hz and resembles the modulation of a double-sideband suppressed-carrier transmission. In addition to these features, strong carriers are observed at 1396.18 MHz, 1396.75 MHz, and 1402.94 MHz. Most of these interferers are in the radio as-

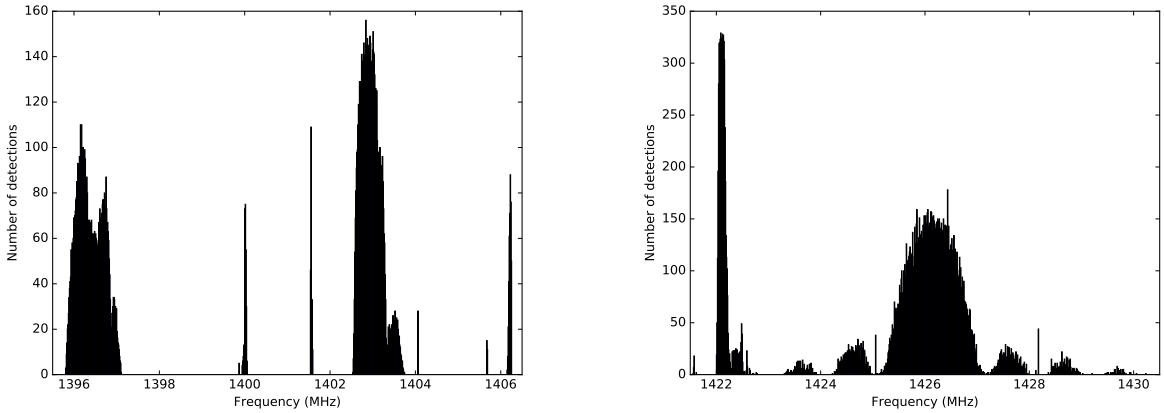

Figure 2.2: Number of detections as a function of frequency, prior to the application of our rejection algorithms (Section 2.3.7), in two 9 MHz-wide regions of the spectrum that partially overlap the 1400–1427 MHz radio astronomy protected band. Signal characteristics are listed in Table 2.3.

tronomy protected band.

The region of the spectrum near 1401.5 MHz is characterized by a number of narrow lines that seem to cluster near 10 discrete frequency regions. They have small (mostly $<$0.2 Hz s^{-1}) positive or negative Doppler drift rates. Some of these lines exhibit somewhat erratic behavior as a function of time, perhaps indicating an unstable oscillator. The small Doppler drift rates suggest a terrestrial source.

The interferer near 1422 MHz generates a broad ($\sim$270 kHz) region of increased noise power without distinct lines, making the identification of the Doppler drift rate difficult.

The region between 1423 MHz and 1429 MHz exhibits some interferers that are approximately 70 Hz wide and reminiscent of those described by Siemion et al. [2013, their figure 5]. Others are time-variable signals (12–15 s periodicity) that are approximately 70 Hz wide. These characteristics are similar to those of some Air Route Surveillance Radars (ARSR) that track aircraft in all azimuthal directions at 5 rotations per minute, except that these systems are designed to operate between 1250 MHz and 1350 MHz. If due to ARSR, it is unclear whether the interference is caused

at the source or generated by intermodulation products in the GBT receiver system. The sinc-type appearance of the histogram (Figure 2.2) is reminiscent of the power spectrum of a BPSK waveform with 1 MHz bandwidth and compressed pulse width of 1 μs. Such waveforms are used to provide radar imaging with 150 m resolution [e.g., Margot et al., 2000]. Some ARSR systems use non-linear frequency modulation (NLFM) to provide a nominal range resolution of 116 m. Some NLFM schemes exhibit a sinc-like power spectrum. In addition to the periodic interferers, strong carriers with variable drift rates are observed near 1425.05 MHz and 1428.18 MHz.

Figure 2.3 shows a graphical summary of the frequency regions that were excluded from our analysis.

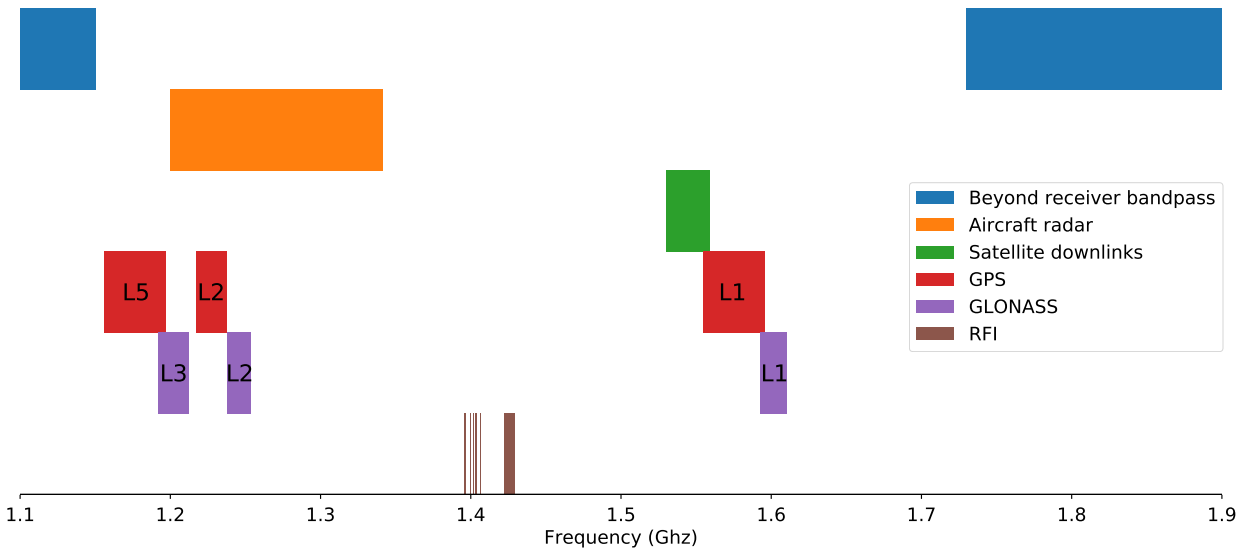

Figure 2.3: Color-coded summary of frequency regions that were excluded from our analysis.

2.3.10 Evaluation of remaining candidates

The 38 remaining signals represent 19 pairs of scans. The characteristics of the first scans are shown in Table 2.4. We generated frequency-time diagrams for all 19 pairs. We then searched our

database for signals at similar frequencies and compared their time-frequency diagrams to those of our top candidates. This process revealed that all 19 candidates are detected in more than one direction on the sky, ruling out the possibility of an extraterrestrial signal. Examination of the time-frequency diagrams revealed groups of signals that can be attributed to the same source of RFI. We provide a brief description of the signals below, along with a few examples of time-frequency diagrams. To our knowledge, these interferers have not been reported in the literature, although they may of course have been detected in other searches.

The candidate signal near 1151 MHz (Figure 2.4) is a monochromatic signal with a substantial Doppler drift rate that is observed in at least eight distinct directions on the sky.

The candidate signal near 1375 MHz has a complex time-frequency structure that is observed in at least one other direction on the sky.

The candidate signal near 1414 MHz (Figure 2.5) exhibits both high S/N and somewhat erratic frequency behavior. The rejection filter logic likely failed because the Doppler behavior of this signal is erratic.

The candidate signal near 1444 MHz is a low S/N monochromatic signal that is observed in at least four other directions on the sky.

Both candidates near 1453 MHz exhibit a set of intermittent narrow lines that are observed in multiple directions on the sky.

All five candidates near 1457 MHz are due to anthropogenic RFI and share the same characteristics, i.e., a monochromatic signal superimposed on a broadband, elevated noise level.

The candidate signal near 1461 MHz is a single line that is observed in at least two directions on the sky. Moreover, signals with similar lines that are offset in frequency by almost exactly 20 kHz and 30 kHz are observed in other directions on the sky.

The candidate signal near 1467 MHz (Figure 2.6) exhibits a set of monochromatic lines. The strongest one exceeded our S/N threshold. It is observed in at least four directions on the sky.

The candidate signal near 1472.254 MHz is detected in at least three directions on the sky.

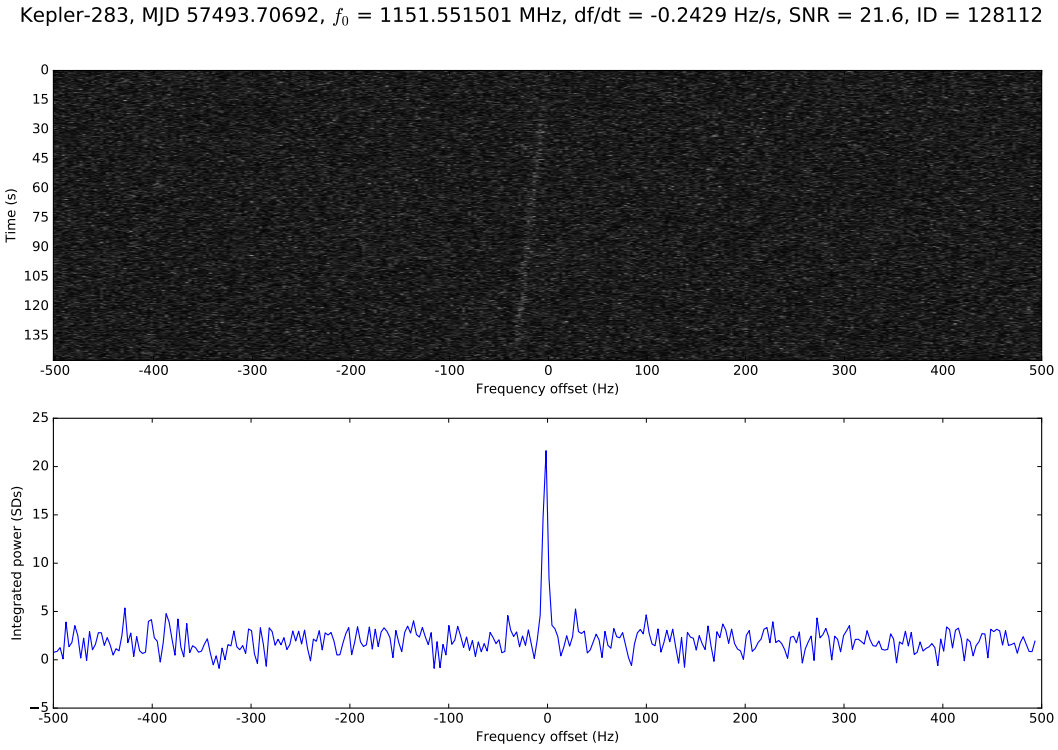

Figure 2.4: Time-frequency characteristics of the first sample candidate signal. (Top) Time-frequency diagrams showing consecutive power spectra; (Bottom) Integrated (i.e., shifted and summed) power spectra.

The candidates near 1472.262 MHz correspond to the same signal that is observed in at least two distinct directions.

The candidates near 1485 MHz and 1501 MHz are monochromatic signals that are detected in at least two and six directions on the sky, respectively.

The candidate signal near 1623 MHz has a complex time-frequency structure that is observed in several directions on the sky. An approximately linear ramp with slope -150 Hz s^{-1} appears then disappears with on and off durations of 5 to 10 s.

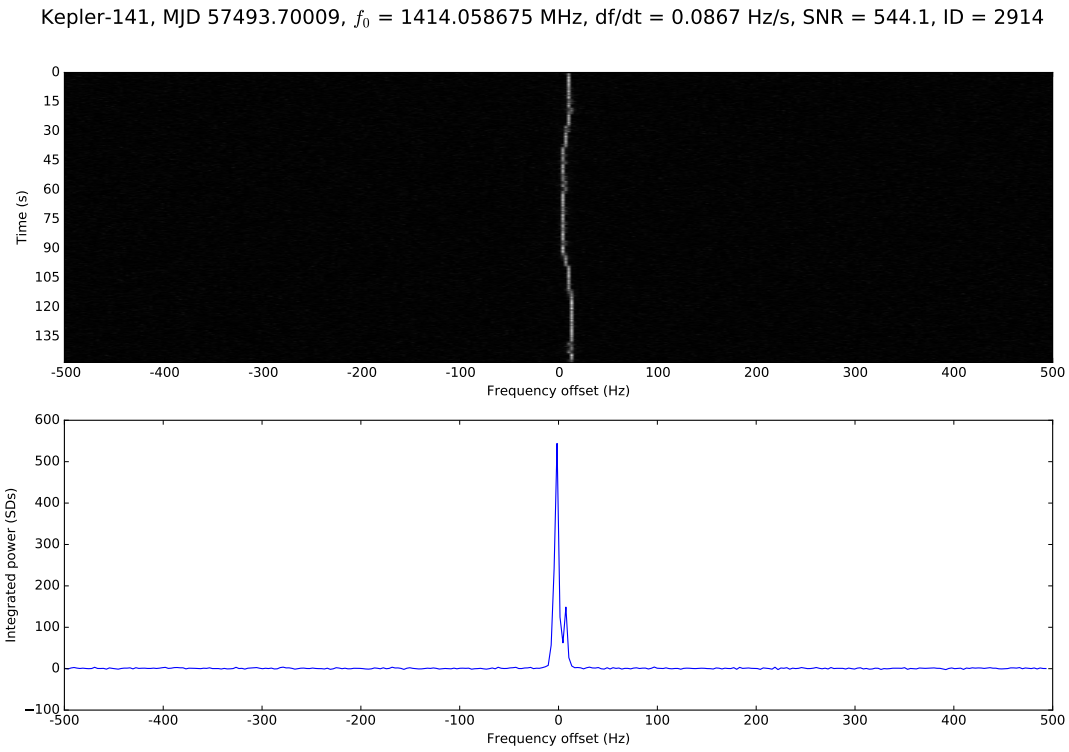

Figure 2.5: Time-frequency characteristics of the second sample candidate signal. (Top) Time-frequency diagrams showing consecutive power spectra; (Bottom) Integrated (i.e., shifted and summed) power spectra.

2.4 Discussion

2.4.1 Search volume

Although no extraterrestrial signals were identified in our analysis to date, we emphasize that our study encompassed only a small fraction of the search volume. The fraction of the sky that was covered in our search is 14 times the solid angle of the GBT beam. At the center frequency of our search, each beam has a solid angle of 0.015 deg^2. Considering all 14 sources, we covered 0.21 deg^2 or about 5 ppm of the entire sky. Our observations lasted a total of 5 minutes on each source, which is about 10 ppm of a terrestrial year. After the elimination of polluted bands, our use-

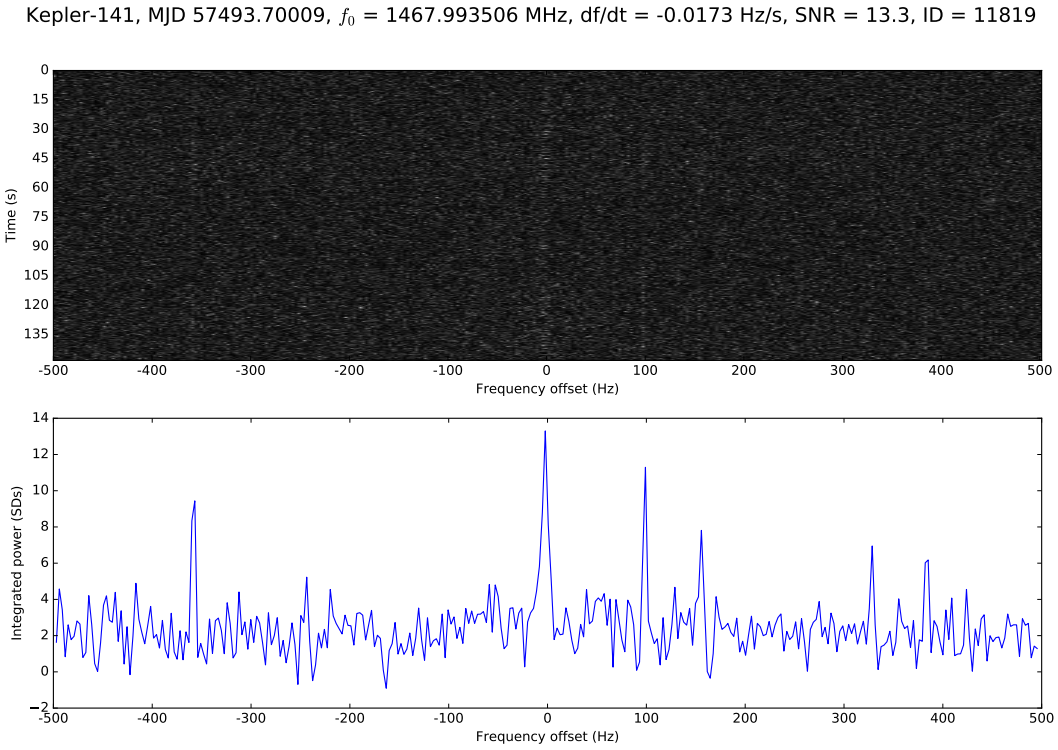

Figure 2.6: Time-frequency characteristics of the third sample candidate signal. (Top) Time-frequency diagrams showing consecutive power spectra; (Bottom) Integrated (i.e., shifted and summed) power spectra.

ful bandwidth spanned almost exactly 300 MHz, which is a small fraction of the electromagnetic spectrum available for transmission.

We computed the Drake figure of merit [Drake, 1984, Enriquez et al., 2017] that corresponds to our search parameters:

$$\mathrm{DFM} = \frac{\Delta f_{\mathrm{tot}} \Omega}{F_{\mathrm{det}}^{3/2}}, \tag{2.1}$$

where Δf_{tot} = 300 MHz is the total bandwidth examined, Ω = 0.21 deg^2 is the total area of the sky covered, and F_{det} = 10 Jy is the minimum flux density required for a detection (Section 2.4.3). In these units, we find DFM $\simeq 2 \times 10^6$. Referring to the values provided in Enriquez et al. [2017], our search is about 150 times larger than that of Horowitz and Sagan [1993] and 10 times larger

than that of Gray and Mooley [2017]. It amounts to about 2% and 12% of the recent large surveys by Enriquez et al. [2017] and Harp et al. [2016], respectively.

2.4.2 Existence limits

Attempting to place existence limits on the basis of SETI observations is a difficult exercise. Certainly, we can place limits only on the kinds of signals that we are looking for, not actual limits on the presence of civilizations. For instance, Enriquez et al. [2017] attempted to place a limit on the number of 100%-duty cycle transmitters (e.g., a radio beacon) and suggested that fewer than 0.1% of the stellar systems within 50 pc possess such transmitters. However, beacons operating at frequencies lower than 1.1 GHz, larger than 1.9 GHz, or in the 1.2–1.34 GHz range would be undetected in their (and our) search, which makes general claims about the number of beacons unreliable. In this spirit, we describe the types of signals that are detectable with our search, but we do not attempt to make inferences about the prevalence of radio beacons in the Galaxy.

2.4.3 Sensitivity

For the detection of narrowband signals above a floor with noise fluctuations, the S/N can be expressed as $P_r/\Delta P_{\rm noise}$, where P_r is the received power and $\Delta P_{\rm noise}$ is the standard deviation of the receiver noise given by

$$\Delta P_{\rm noise} = \frac{k_B T_{\rm sys} \Delta f}{\sqrt{\Delta f \tau}}, \tag{2.2}$$

with k_B Boltzmann's constant, $T_{\rm sys}$ the system temperature, Δf the frequency resolution, and τ the integration time [e.g., Ostro, 1993, Naidu et al., 2016]. The power received by a transmitter of power P_t and antenna gain G_t located at distance r is

$$P_r = \frac{P_t G_t}{4\pi r^2} \frac{A_e}{2}, \tag{2.3}$$

where A_e is the effective area of the receiving station and the factor of 1/2 accounts for reception by a single-polarization feed. If powers received in both polarizations are added incoherently as in this work, $\mathrm{S/N} = \sqrt{n_{\rm pol}} \times P_r/\Delta P_{\rm noise}$, with $n_{\rm pol} = 2$. The S/N is proportional to the

product of factors that relate to the transmitter-receiver distance, transmitter performance, receiver performance, quantization efficiency, and data-taking and data-analysis choices, as shown in this expression:

$$\mathrm{S/N} = \left(\frac{1}{4\pi r^2}\right)(P_t G_t)\left(\frac{A_e}{2k_B T_{\mathrm{sys}}}\right)(\eta_Q)\left(\frac{n_{\mathrm{pol}}\,\tau}{\Delta f}\right)^{1/2}. \tag{2.4}$$

The second factor is known as the effective isotropic radiated power (EIRP) and the third factor is the inverse of the system-equivalent flux density (SEFD), so we can rewrite the S/N as

$$\mathrm{S/N} = 9\left(\frac{100\ \mathrm{ly}}{r}\right)^2\left(\frac{\mathrm{EIRP}}{10^{13}\ \mathrm{W}}\right)\left(\frac{10\ \mathrm{Jy}}{\mathrm{SEFD}}\right)\left(\frac{\eta_Q}{1}\right)\left(\frac{n_{\mathrm{pol}}}{1}\right)^{1/2}\left(\frac{\tau}{1\ \mathrm{s}}\right)^{1/2}\left(\frac{1\ \mathrm{Hz}}{\Delta f}\right)^{1/2}. \tag{2.5}$$

The S/N is maximized when the frequency resolution of the data matches the bandwidth of the signal.

We used the Arecibo planetary radar as a prototype transmitter with P_t = 10^6 W, G_t = 73.4 dB, and EIRP = 2.2 $\times 10^{13}$ W. The nominal receiver is the GBT with A_e = 5600 m^2, T_{sys} = 20 K, and SEFD = 10 Jy. The quantization efficiency of the four-level sampler is $\eta_Q = 0.8825$. Our data-taking and data-analysis choices correspond to $n_{\mathrm{pol}} = 2$, τ = 150 s, and Δf = 3 Hz. With these values, we find $\Delta P_{\mathrm{noise}} = 3.9 \times 10^{-23}$ W and $P_r = 3.1 \times 10^{-22}$ W for a transmitter located at a distance of 420 ly (128 pc) from Earth. Such transmissions would be at the limit of detection with our minimum S/N threshold of 10.[3]

The target stars with known distances in our sample are located at distances that exceed 420 ly, such that a more powerful transmitter, a more sensitive receiver, longer integration times, or narrower frequency resolutions would be needed for detection. Transmitters with considerably larger EIRPs than that of Arecibo may be available to other civilizations. A transmitter with 1000 times Arecibo's EIRP would be detectable in our search from distances of up to 13 250 ly. Current technology enables the detection of technosignatures emitted from a large fraction of the Galaxy (Figure 2.7). In such a vast search volume, there are billions of targets accessible to a search for technosignatures. In contrast, the search for biosignatures will be limited in the foreseeable future to a few targets in the Solar System and to a few hundred planetary systems around nearby stars.

[3] An S/N threshold of 10 with our search parameters and $\eta_Q = 1$ corresponds to detection of flux densities of 10 Jy.

One can use Figure 2.7 to evaluate the detectability of the Arecibo planetary radar by other civilizations. From a S/N standpoint, an increase by a factor of 10 in EIRP is equivalent to an increase in effective area by a factor of 10. In other words, the colored lines in Figure 2.7 also represent the detectability of Arecibo by remote antennas with effective areas that are 1-1000 times larger than those of the GBT or Arecibo, assuming similar system temperatures and search parameters ($n_{\rm pol} = 2$, τ = 150 s, and Δf = 3 Hz). Equation (2.5) can be used to evaluate detectability with other transmitter, receiver, or search parameters. Typical planetary radar transmissions have a duty cycle of approximately 50%, and the maximum tracking duration for celestial sources at Arecibo is $\sim$2.5 h. For observations of main belt asteroids and more distant bodies at opposition, the pointing direction over the entire tracking duration changes by less than 2 arcmin, i.e., less than the width of the beam for S-band (2.38 GHz) planetary radar transmissions.

2.5 Conclusions

We described the results of a search for narrowband signals from extraterrestrial sources using two hours of GBT telescope time in 2016. We identified 858 748 candidate signals. Our rejection filters automatically eliminated 99% of the candidates, leaving 10 157 candidate signals for further inspection. Almost all of the remaining signals were ruled out because they were attributable to anthropogenic RFI, leaving 19 pairs of candidate signals. All of these candidates were observed in more than one direction on the sky, thereby ruling them out as extraterrestrial signals.

Table 2.1: Target host stars listed in order of observations. Distances in parsecs are from the NASA Exoplanet Archive. Habitable zone categories 1 and 2 refer to small ($R_p < 2R_E$) planets in the conservative and optimistic habitable zones described by Kane et al. [2016], respectively. In multi-planet systems, only the lowest category is listed.

Host star	Distance (ly)	HZ category
Kepler-399	NA	
Kepler-186	561^{+42}_{-33}	Cat. 1
Kepler-452	NA	Cat. 2
Kepler-141	NA	
Kepler-283	NA	Cat. 1
Kepler-22	620	
Kepler-296	737^{+91}_{-59}	Cat. 1
Kepler-407	NA	
Kepler-174	NA	Cat. 2
Kepler-62	1200	Cat. 1
Kepler-439	2260^{+215}_{-124}	
Kepler-438	473^{+65}_{-75}	
Kepler-440	851^{+52}_{-150}	Cat. 2
Kepler-442	1115^{+62}_{-72}	Cat. 1

Table 2.2: Spectral regions exhibiting a high density of detections per unit frequency. The number of detections are reported both prior to and after application of our rejection filters (Section 2.3.7). The density column shows the number of pre-filter detections per MHz. Because some bands overlap, the totals are not the arithmetic sums of the table entries.

Spectral region (MHz)	Width (MHz)	Pre-filter detections	Post-filter detections	Density (# per MHz)	Identification
1554.96 – 1595.88	40.92	107 955	2657	2638	GPS L1
1155.99 – 1196.91	40.92	206 093	5289	5036	GPS L5
1592.9525 – 1610.485	17.5325	80 908	1110	4614	GLONASS L1
1192.02 – 1212.48	20.46	37 792	518	1847	GLONASS L3
1530 – 1559	29	28 814	89	994	Satellite downlinks
Total	129.385	459 543	9663	3552	

Table 2.3: Characteristics of likely anthropogenic interferers. The number of detections are reported both prior to and after the application of our rejection filters (Section 2.3.7). Doppler drift rate and S/N statistics are computed on the pre-filter detections. Strong interferers are detected in the radio astronomy protected band (1400–1427 MHz).

Spectral region (MHz)	Width (MHz)	Pre-filter detections	Post-filter detections	Median drift rate (Hz s^{-1})	S/N (min/median/max)
1395.810 – 1397.097	1.287	6210	46	0.451	10.0/15.9/1087.2
1399.872 – 1400.054	0.182	287	0	0.434	10.0/12.5/26.1
1401.547 – 1401.599	0.052	220	2	-0.017	10.0/19.0/1423.6
1402.536 – 1403.750	1.214	7536	76	0.434	10.0/19.5/451.8
1406.145 – 1406.250	0.105	418	4	0.434	10.0/12.6/25.6
1422.000 – 1422.270	0.270	5150	2	N/A	10.0/17.4/84.1
1423.308 – 1428.971	5.663	20 134	326	0.121	10.0/21.9/7642.2
Total	8.773	39 955	456		

Table 2.4: Characteristics of top candidates listed in increasing order of frequency. Epoch, frequency, and drift rate refer to the beginning of the first scan with units of modified Julian date (MJD), Hz, and Hz s^{-1}, respectively. S/N refers to the integrated power over the scan duration after correcting for the corresponding Doppler drift rate.

ID	Source	Epoch (MJD)	Frequency (Hz)	Drift rate (Hz s^{-1})	S/N
128112	Kepler-283	57493.70692	1151551501.442434	-0.2429	21.6
141584	Kepler-442	57493.74497	1375489655.604034	-8.8124	16.7
2914	Kepler-141	57493.70009	1414058674.868273	0.0867	544.0
93976	Kepler-399	57493.68869	1444533390.553847	0.0173	14.2
36563	Kepler-186	57493.69100	1453857276.541974	0.0173	16.0
52175	Kepler-22	57493.70934	1453895304.341606	0.0173	19.1
81528	Kepler-296	57493.71608	1457414449.371766	0.3123	18.9
108574	Kepler-407	57493.71818	1457414503.015998	0.3469	10.7
81536	Kepler-296	57493.71608	1457442281.787187	0.2949	14.5
108572	Kepler-407	57493.71818	1457453120.902177	0.3643	12.2
81533	Kepler-296	57493.71608	1457490984.788880	0.3123	16.7
108848	Kepler-407	57493.71818	1461771970.293017	-0.0173	17.6
11819	Kepler-141	57493.70009	1467993506.067759	-0.0173	13.3
123438	Kepler-438	57493.73606	1472254848.842477	0.0173	16.4
72937	Kepler-283	57493.70692	1472262660.038624	-0.0347	13.8
24383	Kepler-174	57493.72476	1472262722.623560	-0.0347	75.4
153968	Kepler-440	57493.74274	1485092589.943495	-0.0173	10.4
956243	Kepler-442	57493.74497	1501557959.611854	0.1041	10.2
1129207	Kepler-439	57493.73383	1623514653.815893	-0.1735	22.0

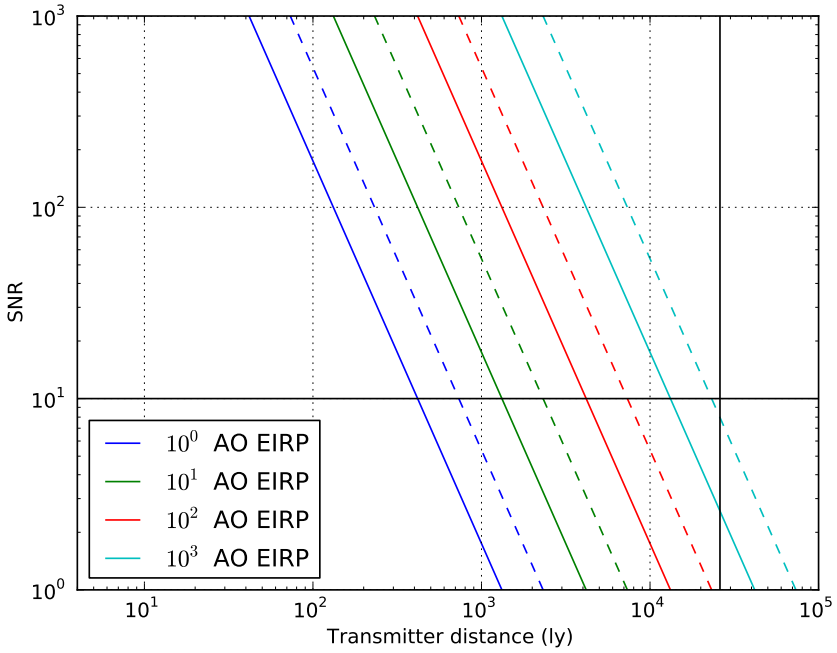

Figure 2.7: S/N of detections as a function of transmitter distance from Earth, assuming search parameters identical to those used in this work ($n_{\mathrm{pol}} = 2$, τ = 150 s, and Δf = 3 Hz). Colored lines represent transmitter EIRPs that are equivalent to 1–1000 times that of the Arecibo Observatory (AO) planetary radar. Solid and dashed colored lines represent reception with the GBT (SEFD = 10 Jy) and AO (SEFD = 3.2 Jy), respectively. The black horizontal line represents the threshold for detection used in this work. The black vertical line represents the distance to the center of the Galaxy.

CHAPTER 3

A search for technosignatures from TRAPPIST-1, LHS 1140, and 10 planetary systems in the Kepler field with the Green Bank Telescope at 1.15–1.73 GHz

As part of our ongoing search for technosignatures, we collected over three terabytes of data in May 2017 with the L-band receiver (1.15–1.73 GHz) of the 100 m diameter Green Bank Telescope. These observations focused primarily on planetary systems in the Kepler field, but also included scans of the recently discovered TRAPPIST-1 and LHS 1140 systems. We present the results of our search for narrowband signals in this data set with techniques that are generally similar to those described by Margot et al. [2018]. Our improved data processing pipeline classified over 98% of the $\sim$ 6 million detected signals as anthropogenic Radio Frequency Interference (RFI). Of the remaining candidates, 30 were detected outside of densely populated frequency regions attributable to RFI. These candidates were carefully examined and determined to be of terrestrial origin. We discuss the problems associated with the common practice of ignoring frequency space around candidate detections in radio technosignature detection pipelines. These problems include inaccurate estimates of figures of merit and unreliable upper limits on the prevalence of technosignatures. We present an algorithm that mitigates these problems and improves the efficiency of the search. Specifically, our new algorithm increases the number of candidate detections by a factor of more than four compared to Margot et al. [2018].

3.1 Introduction

In this chapter, we present a search for radio technosignatures using the L-band receiver of the 100 m diameter Green Bank Telescope (GBT). We scanned a total of 12 sources within a 2 hour observational window with the goal of detecting narrowband emissions. Narrowband ($\sim$ 10 Hz) signals are diagnostic of engineered emitters because the narrowest known sources of natural emission span a larger ($\sim$ 500 Hz) range of frequencies at L band [e.g., Cohen et al., 1987]. Our search builds on the legacy of technosignature searches performed in the period 1960–2010 [Tarter, 2001, Tarter et al., 2010, and references therein] and complements some recent efforts [Siemion et al., 2013, Harp et al., 2016, Enriquez et al., 2017, Gray and Mooley, 2017, Margot et al., 2018], but differs primarily in the choice of sample of the vast parameter space still left to search. In addition, we are generally sensitive to a wider range of signal drift rates (± 8.86 Hz s^{-1}) than the cited works, and to signals with lower signal-to-noise-ratio (S/N$>$10) than a recent large survey [Enriquez et al., 2017, S/N$>$25].

We present a substantial improvement to the signal detection algorithms used in the search pipelines of Siemion et al. [2013], Enriquez et al. [2017], Margot et al. [2018]. In these pipelines, a wide (several hundred Hertz) window surrounding the frequency of each candidate detection is removed from further consideration, whether this candidate appears in a primary [Siemion et al., 2013, Margot et al., 2018] or secondary [Enriquez et al., 2017] scan.

Consequently, other legitimate signals within that window are never analyzed, which implies that both figures of merit for the completeness of the search and upper limits on the prevalence of technosignatures reported in these works are inaccurate. By adjusting the way in which the frequency space surrounding each candidate signal is handled, we no longer need to discard legitimate signals and are able to analyze a larger fraction of the frequency space. This advance is appreciable. We estimate that this improvement alone increases our detection count by a factor of more than four as compared to Margot et al. [2018]. The increase in the detection count of other search pipelines after implementation of a similar improvement would likely be substantial.

Our data acquisition techniques, a brief overview of which is presented in Section 3.2, are generally similar to those presented by Margot et al. [2018]. Section 3.3 explains our data analysis procedures, including major changes and improvements to our data processing pipeline. The results of our search are presented in Section 3.4, followed by a discussion and conclusions in Sections 3.5 and 3.6, respectively.

3.2 Data Acquisition

3.2.1 Sources

For our observations, we selected 10 sources (Table 3.1) from the Kepler catalog based on habitability criteria presented by Kane et al. [2016]. These criteria take into account the size and location of exoplanets with respect to the host stars. Small planets ($R_p < 2R_E$, where R_p and R_E are the radii of the planet and Earth, respectively) in the conservative and optimistic habitable zones fall under Habitable Zone (HZ) categories 1 and 2, respectively. Conversely, planets with any radius in the conservative and optimistic habitable zones fall under HZ categories 3 and 4, respectively.

For our observations, we selected all seven host stars reported by Kane et al. [2016] with confirmed planets in HZ categories 1 and 2 (Table 3.1). We supplemented these with three more host stars with confirmed planets in HZ categories 3 and 4 (Table 3.1). In addition to the 10 sources in the Kepler field, we observed the recently discovered TRAPPIST-1 and LHS 1140 systems. The TRAPPIST-1 system hosts seven Earth-sized, temperate exoplanets orbiting an ultra-cool dwarf star [Gillon et al., 2017]. The benign equilibrium temperatures of some of the planets in the system make the prospect of liquid water on their surfaces, and thus the possibility of life, plausible. The LHS 1140 system harbors only one known planet, which orbits its M dwarf host star within the habitable zone [Dittmann et al., 2017].

3.2.2 Observations

We conducted our observations with the 100 m diameter Green Bank Telescope (GBT) on May 4, 2017, 15:00 – 17:00 Universal time (UT). We recorded both linear polarizations of the L-band receiver using the GUPPI backend in its baseband recording mode [DuPlain et al., 2008]. GUPPI was configured to channelize 800 MHz of recorded bandwidth into 256 channels of 3.125 MHz each. We verified telescope pointing accuracy with standard procedures and tested signal integrity by injecting a monochromatic tone near the receiver frontend.

We observed all our targets in pairs in order to facilitate the detection and removal of signals of terrestrial origin (Section 3.3.4). The sources were paired by approximately minimizing the slew time of the telescope for the duration of the observing block and then taking consecutive pairs from the solution. Pairings were adjusted to eliminate any ambiguity in the direction of origin of detected signals. Specifically, we required angular separations larger than 1° (several times the ~ 9 arcmin beamwidth of the GBT at the center frequency of our observations) between pair members. The pairs are listed consecutively in Table 3.1 (i.e., the first pair is Kepler-442 and Kepler-440, and so on). Each pair was observed for a total of ~ 330 s using an "on" - "off" - "on" - "off" sequence, where "on" represents a scan of the first source in the pair, and "off" represents a scan of its partner.

3.3 Analysis

3.3.1 Data Pre-Processing

After unpacking the data, we computed 2^{20}-point Fourier transforms of the digitized raw voltages, yielding a frequency resolution of $\Delta\nu = 2.98$ Hz. We chose this frequency resolution because it is small enough to provide unambiguous detections of technosignatures and large enough to examine Doppler frequency drift rates of up to ~ 10 Hz s^{-1} (Section 3.3.2).

We modeled the bandpass response of GUPPI's 256 channels by fitting a 16-degree Chebyshev

polynomial to the median bandpass response of all channels within the operating range of the GBT L-band receiver (1.15 — 1.73 GHz), excluding channels that overlap the frequency range (1200 – 1341.2 MHz) of a notch filter designed to mitigate radio frequency interference (RFI) from aircraft radar. Our experience indicates that this process allows us to apply bandpass corrections that yield the expected flat baselines. After applying the bandpass correction to all 256 channels, we stored consecutive power spectra of length 2^{20} as rows in a time-frequency array and normalized the result to zero mean and unit variance of the noise power. We call the graphical representation of such arrays time-frequency diagrams, though they are also often referred to as "spectrograms", "spectral waterfalls", or "waterfall plots".

3.3.2 Doppler De-Smearing

Due to the orbital and rotational motions of both the emitter and the receiver, we expect extraterrestrial technosignatures to drift in frequency space [e.g., Siemion et al., 2013, Margot et al., 2018]. A de-smearing algorithm is required to avoid spreading the power of a given signal over multiple channels. Since the Doppler drift rates due to the emitters are unknown, we examined 1023 linearly spaced drift rates in the range ± 8.86 Hz s^{-1}, with a step of $\Delta \dot{f} = 0.0173$ Hz s^{-1}. To accomplish this, we made use of a computationally advantageous Doppler de-smearing algorithm [Taylor, 1974, Siemion et al., 2013] which computes an array containing de-smeared power spectra, where each de-smeared spectrum represents a time integration of the consecutive power spectra after correcting for a given Doppler drift rate. A single pass of this algorithm computes 512 power spectra for all drift rates in the range 0 to 8.86 Hz s^{-1} . To obtain power spectra at negative drift rates, we applied the algorithm a second time and reversed the search direction.

In a previous analysis, Margot et al. [2018] performed a search for technosignatures on each of the resulting arrays individually. As a result, it was possible for a candidate signal to be detected twice; once with the correct drift rate, and once with a spurious drift rate of the opposite sign (Figure 3.1). This duplication increased the false detection count and was occasionally problematic in subsequent stages of the data-processing pipeline. To avoid this problem, we concatenated the

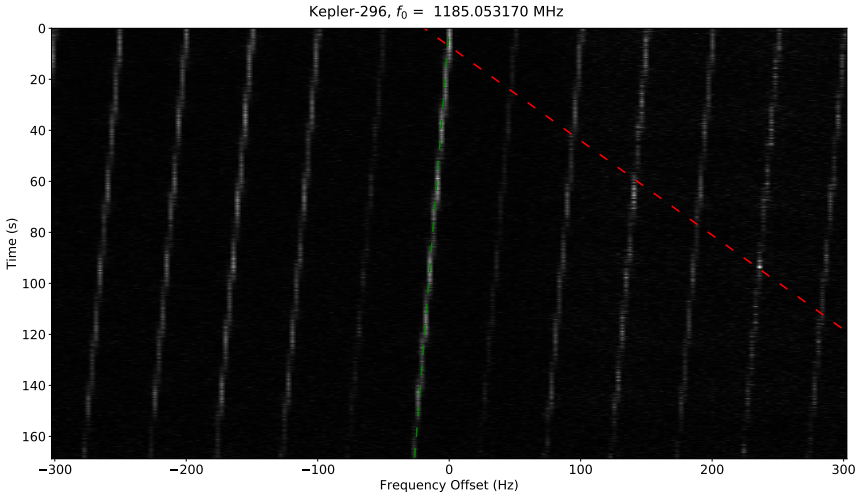

Figure 3.1: Time-Frequency diagram of a signal that would result in multiple detections in an earlier version of our data processing pipeline. In this diagram, pixel intensity represents signal power. The green dashed line shows the correct drift rate of the signal, while the red dashed line indicates a drift rate along which signal values integrate to a summed power that exceeds our detection threshold, resulting in an additional, spurious detection. Our improved pipeline eliminates the possibility of spurious detections for signals of this nature.

outputs of both applications of the de-smearing algorithm into a single array prior to subsequent analysis. This array contained all 1023 possible drift rates (one duplicate calculation at 0 Hz s^{-1} was removed).

3.3.3 Candidate Signal Detection

The output of the Doppler de-smearing algorithm contains the integrated power spectra of the scan at various drift rates, making it ideal for identifying promising candidates for extraterrestrial technosignatures (hereafter, "candidate signals"). We performed this search iteratively by identifying the signal with the highest S/N and storing its characteristics in a structured query language (SQL) database before moving on to the next signal. Drift rates that are similar to the drift rate with maximum S/N often yield integrated powers with large S/N values as well, and it is important to avoid

redundant detections of the same signal. Siemion et al. [2013], Margot et al. [2018] avoided the redundancy by discarding all detections within a frequency range Δf centered on the start frequency of the highest-S/N signal. For the analysis at 3 Hz resolution, they defined the frequency range to be $\Delta f = 2\dot{f}_{max}\tau$, where $\dot{f}_{max} = 8.86$ Hz s^{-1} and τ is the duration of the scan (typically around 150 s). This choice guaranteed that duplicate detections for signals with the highest detectable drift rates were not recorded. However, this procedure also removed all other valid candidates within a $\sim$ 3000 Hz window of every detected candidate signal. Moreover, due to the iterative nature of the search, high S/N signals were always detected first, which prevented lower-S/N signals in their vicinity from being detected. This procedure removed many legitimate candidate signals in the vicinity of higher-S/N signals, ultimately leaving large regions of the spectrum unexamined. This issue results in two unacceptable consequences. First, it yields incorrect calculations of figures of merit, because such calculations rely on an accurate measure of the bandwidth examined during a particular search. Second, it renders attempts to place upper limits on the abundance of technosignature sources unreliable because the pipeline eliminates the very signals it purports to detect. Problems of this nature affect the results of several studies, including those of Siemion et al. [2013], Enriquez et al. [2017], Margot et al. [2018]. To properly quantify the impact of these discarded signals on estimates of the prevalence of technosignatures, a proper injection and recovery study must be conducted. This study is beyond the scope of our work.

We designed a novel procedure to alleviate these shortcomings. In order to avoid redundant detections, we simply require that no new detection contain any subset of the points in time-frequency space belonging to any other already-detected signal. In other words, we require that none of our detections "cross" in time-frequency space. In this context, the frequency range used to discard redundant detections is no longer a constant, but rather a function of the drift rate of a new potential candidate and the bandwidth of the already-detected signal. If $\dot{f}_0$ is the drift rate of a known candidate, and $\dot{f}$ is the drift rate of the potential candidate signal, then the potential candidate is

marked redundant if its frequency f at the start of the scan satisfies

$$\begin{aligned} f_0 - \Delta f_b < f < f_0 + (\dot{f} - \dot{f}_0)\tau + \Delta f_b \qquad &\text{if} \quad \dot{f} > \dot{f}_0 \\ f_0 + \Delta f_b > f > f_0 + (\dot{f} - \dot{f}_0)\tau - \Delta f_b \qquad &\text{if} \quad \dot{f} < \dot{f}_0 \end{aligned} \tag{3.1}$$

where f_0 is the frequency of a known candidate signal, τ is the scan duration, and Δf_b is half of the signal bandwidth. Because we do not want new detections that contain *any* part of an already-detected signal, we must account for its non-zero bandwidth by extending the frequency range as in Equation 3.1. For implementation details of this procedure, including the estimation of bandwidth, see Appendix 3.A.

One drawback of this method is that potential technosignatures may be discarded if they "cross" a stronger signal of terrestrial origin in time-frequency space. There is no reason to assume *a priori* that a valid candidate signal would not exhibit this behavior. However, superimposed signals are, by their very nature, difficult to detect. Other detection pipelines, including those of Siemion et al. [2013], Enriquez et al. [2017], Margot et al. [2018] are also blind to such signals.

We identified all candidate signals with S/N > 10 in channels within the operating range of the GBT L-band receiver (1.15 — 1.73 GHz), excluding channels that overlap the GBT notch filter at 1200 – 1341.2 MHz. A total of 5,840,149 candidate signals were detected.

3.3.4 Doppler and Direction-of-Origin Filters

After all candidate signals were identified, we applied several filter procedures in order to distinguish anthropogenic signals from promising technosignature candidates. The overall purpose of the algorithms is generally similar to those described by Margot et al. [2018].

Our filter procedures (hereafter, "filters") are designed to search the SQL database and flag the most promising candidates. The first filter flagged all candidate signals with non-zero Doppler drift rates. Signals with zero Doppler drift rates, defined here as signals that drift across less than one frequency channel over the course of a scan, are of no interest to us because the corresponding emitters will generally not be in motion with respect to the receiver, which suggests they are

terrestrial in nature. The second filter flagged a signal as a promising candidate if it is persistent, i.e., it is detected in both scans of its source. This filter removes any anthropogenic signals that may have temporarily entered the beam during one of the scans. The third filter marks a signal as a technosignature candidate if its direction of origin is unique, i.e, it is not detected in the scans corresponding to other sources. This filter eliminates many anthropogenic signals that are detectable through the antenna sidelobes. For implementation details, see Appendix 3.B. If both a candidate signal and its corresponding signal in the other scan of the source were flagged by all three filters, then the candidate signal was marked as a high-interest signal. Of the 5,840,149 total detections, $5\,743\,209$ ($> 98\%$) were discarded as a result of these filters.

3.3.5 Frequency Filters

A majority of the signals detected in our search have frequencies in operating bands of known interferers, such as Global Navigation Satellite Systems or Aircraft Radar. Figure 3.2 depicts our detection count superimposed onto frequency bands of known anthropogenic RFI. The interferer labeled simply as "RFI" is particularly interesting, as it overlaps the radio astronomy protected band near 1420 MHz. The time-frequency structure of this RFI is similar to that described by Siemion et al. [2013] and Margot et al. [2018], who attributed the likely origin of the RFI to intermodulation products of Air Route Surveillance Radars (ARSR). Table 3.2 describes the properties of the regions of most prominent detected anthropogenic RFI. These regions are reminiscent of some of Harp et al. [2016, Table 2]'s "permanent RFI bands." All candidate signals detected within these regions were removed from consideration because of their likely anthropogenic nature.

After removing candidates with frequencies in the operating bands of some well-known anthropogenic interferers (Table 3.2), we observed that regions of high signal density remained. An example of such a region is shown in Figure 3.3. Note the presence of many narrow clusters of high-signal-density regions. Although it is possible that a valid technosignature could be found within one of these regions, it would likely be difficult to detect and establish its validity given the vast amount of strong, ambient RFI. With this challenge in mind, we developed a method to

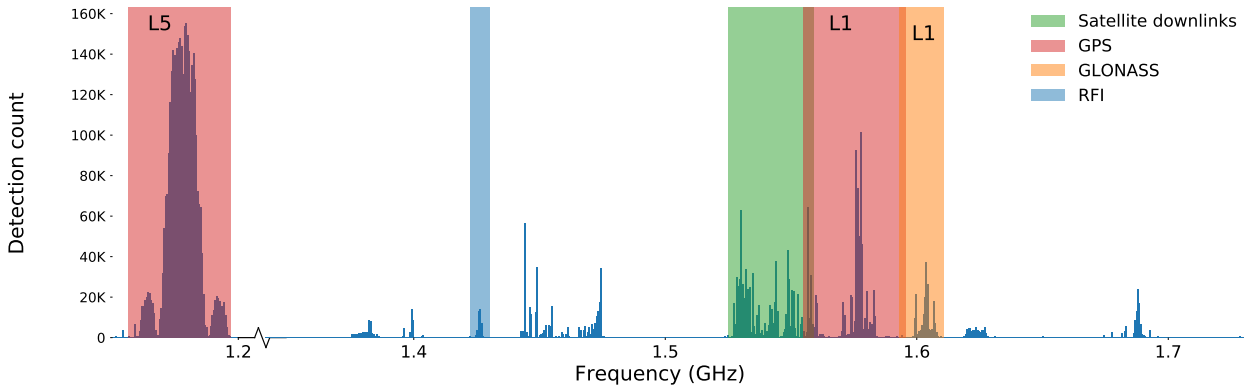

Figure 3.2: Candidate signal count as a function of frequency, superimposed on operating bands of known interferers. Note that the majority of detections occur in congested bands. The region marked simply as "RFI" contains signals from an interferer, likely ARSR, whose intermodulation products overlap the radio astronomy protected band [Margot et al., 2018]. Frequencies overlapping the GBT notch filter at 1200 – 1341.2 MHz are excluded.

discriminate signals found within densely populated frequency regions from signals found in quieter parts of the spectrum. Specifically, we measured the signal density in 1-kHz-wide frequency bins and discarded candidates within that bin if the signal density exceeded a pre-defined threshold value. For this work, we chose a threshold value of 1000 signals MHz^{-1}, which corresponds to a sharp drop-off in histograms of signal density. For further details, see Appendix 3.C. A sample result of this procedure is shown in Figure 3.3, where signals left after our density thresholding procedure are shown in red. More than 96% of the 581 433 signals found outside of the well-known interferer operating bands listed in Table 3.2 were discarded using this procedure.

3.4 Results

Signals that remained after application of our Doppler, direction-of-origin, and frequency filters were marked as final technosignature candidates. Thirty such signals remained, and their properties are given in Tables 3.3 and 3.4. Further examination of the final 30 candidates revealed that 13 of

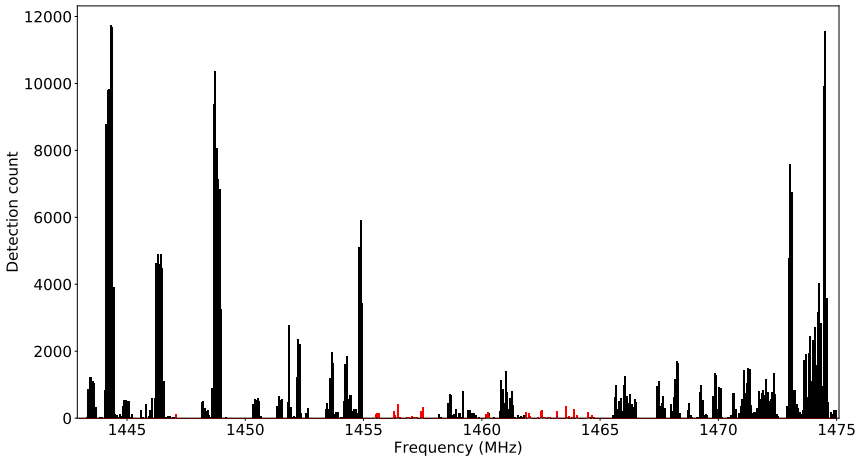

Figure 3.3: Example of a region of high signal density detected outside of known interferer operating bands. Signals left after application of our signal density thresholding procedure are shown in red. Most of the remaining signals are eliminated by the Doppler and direction-of-origin filters, which are applied independently.

them are anthropogenic because they are also present in the "off" scan of the source. They were not correctly identified by our filters for a variety of reasons, which we summarize into categories below.

Category 'a' refers to signals with S/N$<$10 in the "off" scan. Because the "off" scan detections were not recorded in the database, it was not possible for our filters to flag this category of signals as RFI.

Category 'b' refers to signals whose drift rates between the "on" and "off" scans differed by more than our allowed tolerance, which we set to $\pm\Delta\dot{f} = \pm 0.0173$ Hz s^{-1} (Appendix 3.B). Because the "off" scan detections were not correctly paired to the "on" scan detections, it was not possible for our filters to flag this category of signals as RFI.

Category 'c' refers to signals that 'cross' a signal of a higher S/N in scan of a different source. As discussed before, our signal detection methods are currently blind to such signals.

Category 'd' refers to broad signals for which it is difficult to accurately determine a drift

rate. Such signals are naturally difficult to 'pair', since the integrated power may peak at different frequencies within the bandwidth of the signal for different scans.

Category 'e' refers to signals that exhibit non-linear behavior in frequency as a function of time. Our pipeline is not currently well-equipped to handle such signals because the Doppler de-smearing algorithm can only detect linear drifts in frequency vs. time.

All categories described above pinpoint potential areas for improvement for our current pipeline.

Of the remaining 17 signals, we believe that 15 are anthropogenic. These signals appear in three sources: Kepler-296 (3), LHS 1140 (5), and TRAPPIST-1 (7). In each source, the signals exhibit similar modulation properties and are detected at similar frequencies but at different drift rates, implying that they are unlikely to be of extraterrestrial origin. These signals are labeled as Category 'f' in Tables 3.3 and 3.4. The time-frequency diagrams of the "on" - "off" - "on" scans for a sample signal from each of the three sources are shown in Figures 3.4, 3.5, and 3.6.

The remaining two candidates are shown in Figures 3.7 and 3.8. Note that the drift rate for both candidates is close to zero (Table 3.4). Additionally, both candidates are found within 100 Hz of other candidates flagged by our pipeline as RFI. As a result, we cannot conclude that these signals, nor any of the other signals found in this search, are of extraterrestrial origin.

3.5 Discussion

3.5.1 Drake Figure of Merit

The Drake Figure of Merit [Drake, 1984] is often used to compare the parameter space examined by different searches and is given by

$$DFM = \frac{\Delta\nu_{tot}\Omega}{F_{\rm det}^{3/2}} \tag{3.2}$$

where $\Delta\nu_{tot}$ is the total bandwidth observed, Ω is the total angular sky coverage, and $F_{\rm det}^{3/2}$ is the minimum flux density required for a detection. For our search, $F_{\rm det}^{3/2} = 9.4$Jy and $\Omega = N \times$

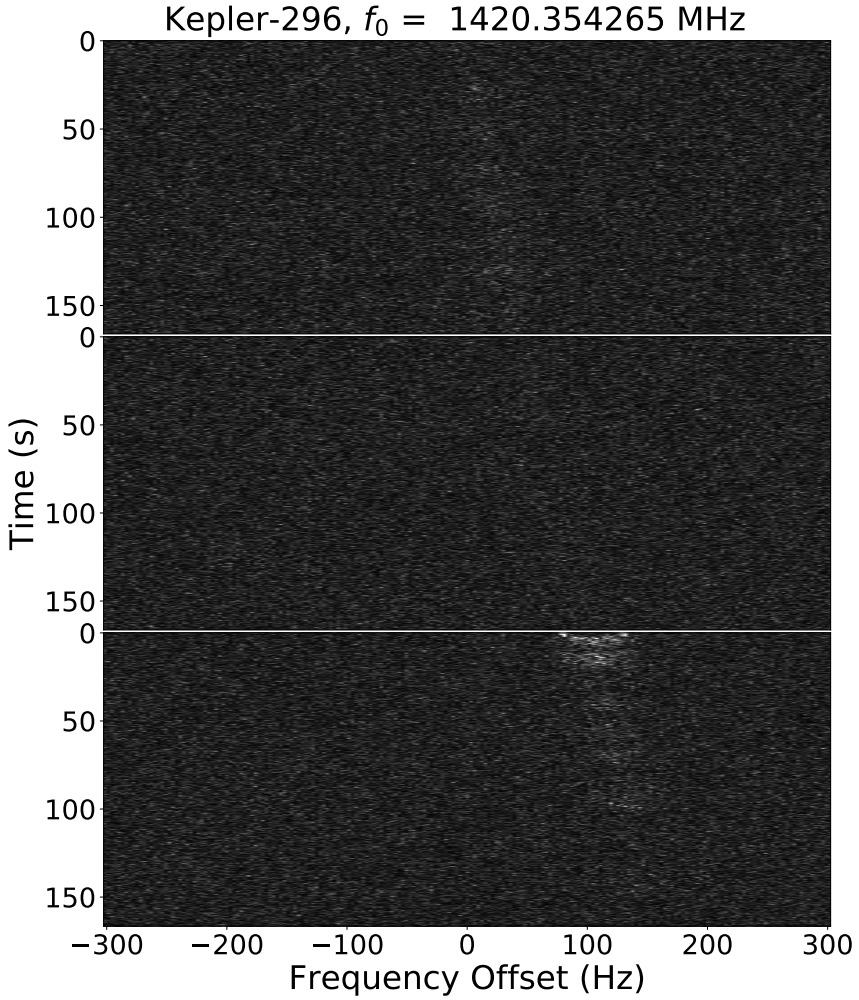

Figure 3.4: Time-Frequency diagrams for the "on" - "off" - "on" scans for a sample category 'f' candidate. This signal, found during the Kepler-296 scans, has a time-frequency structure reminiscent of the ARSR structure.

0.015 deg^2, where $N = 12$ is the number of individual sky pointings [Margot et al., 2018]. To compute $\Delta\nu_{tot}$, we take the bandwidth of the GBT L-band receiver (580 MHz) and subtract the bandwidth of the GBT notch filter (141.2 MHz), the bandwidth discarded due to known interferers (Table 3.2; 137.167 MHz), and the total bandwidth discarded during our density thresholding procedure (Section 3.3.5; 37.101 MHz). Using the resulting bandwidth $\Delta\nu_{tot} = 264.532$ MHz, we find that the DFM associated with this search is 1.6×10^6. This number amounts to about 1.7%

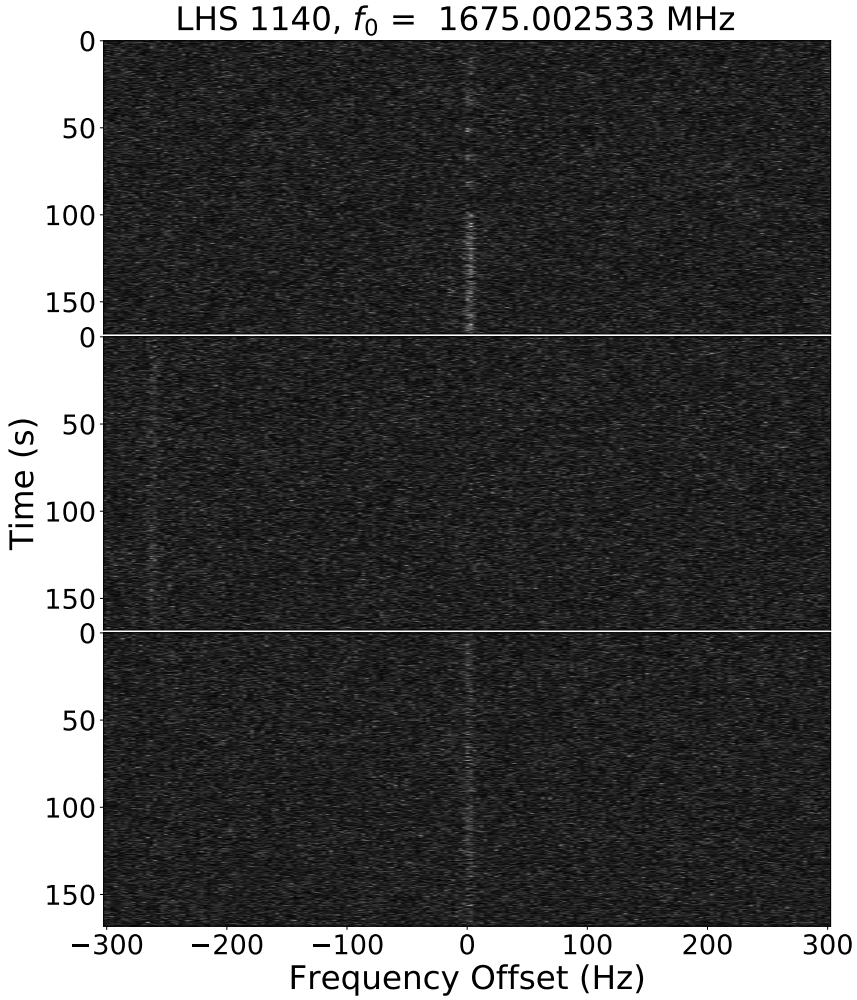

Figure 3.5: Time-Frequency diagrams for the “on” - “off” - “on” scans for a sample category ‘f’ candidate from the scans of LHS 1140.

and 10% of the recent large surveys presented by Enriquez et al. [2017] and Harp et al. [2016], respectively.

3.5.2 Increase of Candidate Detection Efficiency

The candidate detection procedures presented by Siemion et al. [2013], Enriquez et al. [2017], Margot et al. [2018] leave substantial regions of the spectrum unexamined, as pointed out in Section 3.3.3. This deficiency leads to an overestimation of the Drake Figure of Merit (DFM) associ-

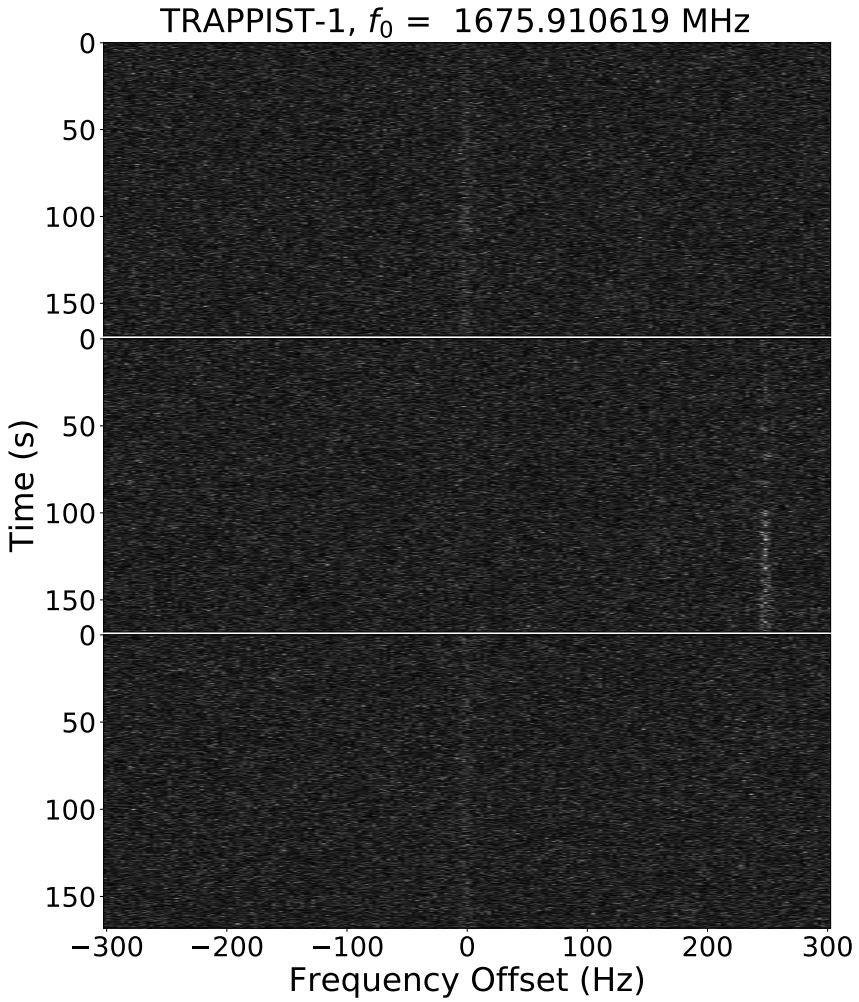

Figure 3.6: Time-Frequency diagrams for the “on” - “off” - “on” scans for a sample category ‘f’ candidate from the scans of TRAPPIST1.

ated with these searches, since this number is directly proportional to the total bandwidth examined (Equation 3.2).

We can calculate the magnitude of this overestimation for the work of [Margot et al., 2018] by noting that a signal with a drift rate $\dot{f}_i$ occupies no more than $\dot{f}_i\tau$ Hz of bandwidth, where τ is the scan duration. However, a window $f_w \approx 3000$ Hz was discarded around every detection, leaving $f_w - \dot{f}_i\tau$ Hz unexamined around every candidate signal. We can calculate the approximate fraction

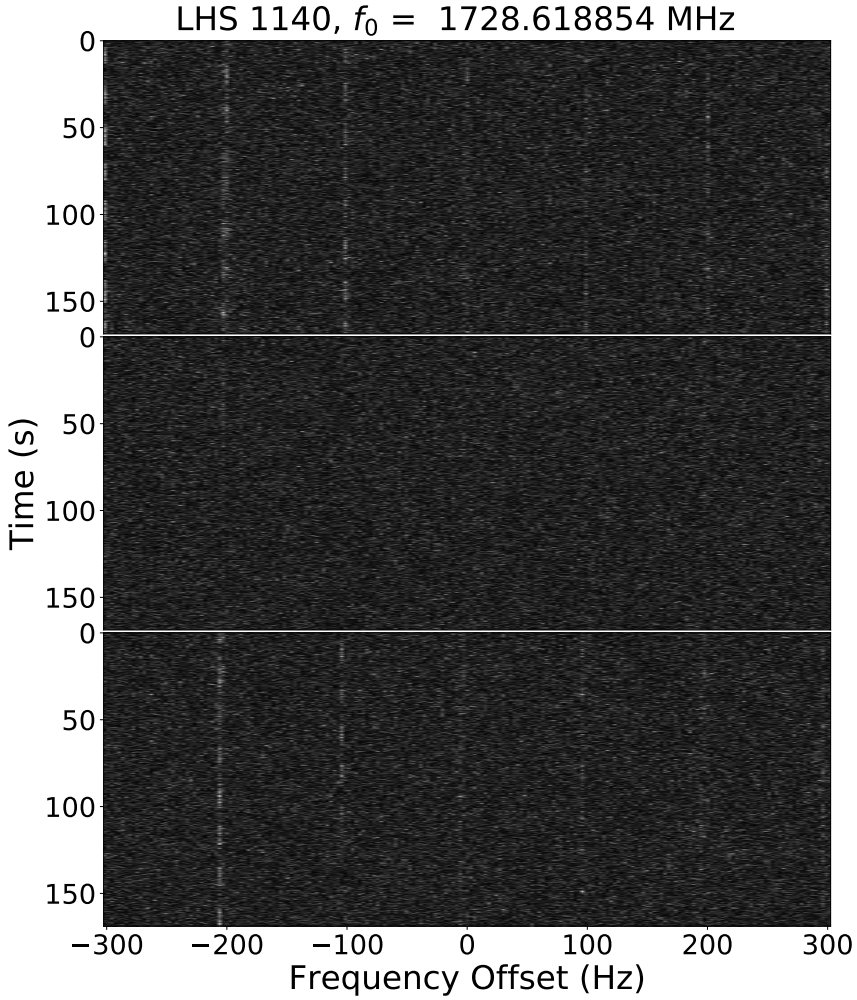

Figure 3.7: Time-Frequency diagrams for the "on" - "off" - "on" scans for the first promising technosignature candidates. The candidate signal is at the center of the diagram. This signal has a near-zero Doppler drift rate, and is in the vicinity of other signals that were discarded as anthropogenic RFI. For this reason, we cannot conclude that this signals is an extraterrestrial technosignature.

F by which the DFM was overestimated using

$$F = \frac{N\Delta f_{tot}}{N\Delta f_{tot} - \sum_i (f_w - \dot{f}_i \tau)}, \tag{3.3}$$

where Δf_{tot} is the total bandwidth that was searched (300 MHz), and N is the total number of scans. Using the database of detections found by Margot et al. [2018], we find $F = 1.048$, meaning that their reported DFM was overestimated by approximately 5%. Enriquez et al. [2017] eliminate

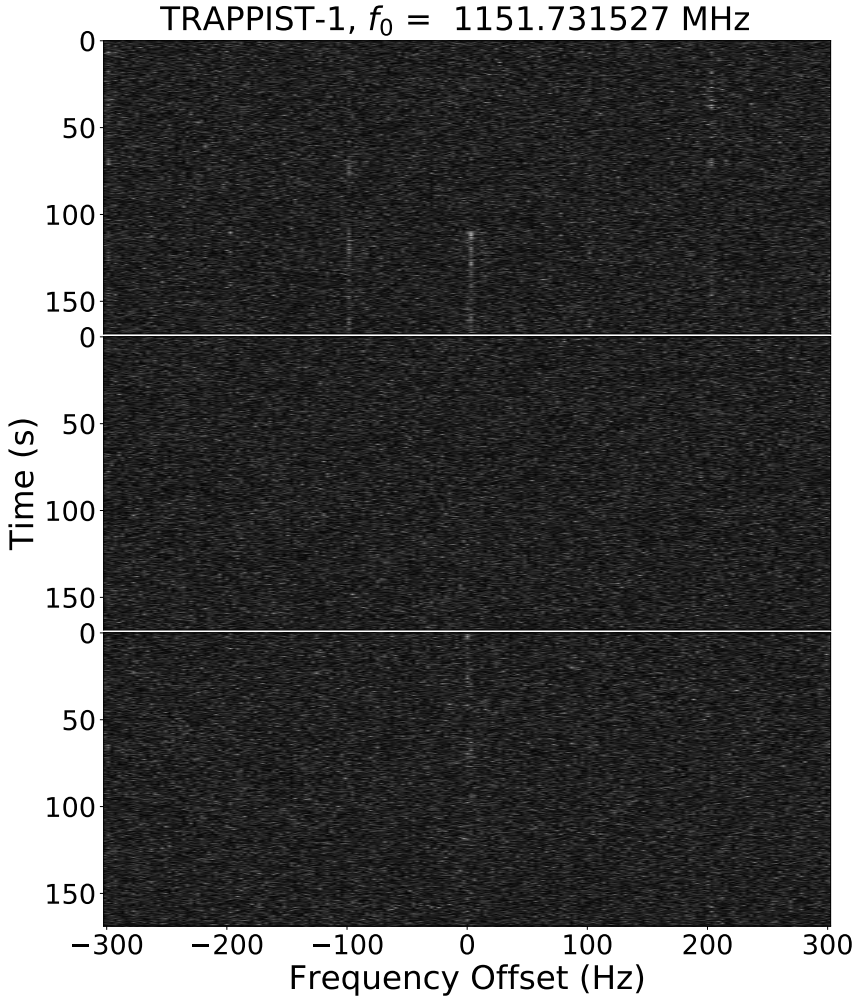

Figure 3.8: Time-Frequency diagrams for the "on" - "off" - "on" scans for the second promising technosignature candidates. The candidate signal is at the center of the diagram. This signal has a near-zero Doppler drift rate, and is in the vicinity of other signals that were discarded as anthropogenic RFI. For this reason, we cannot conclude that this signals is an extraterrestrial technosignature.

$f_w = 1200$ Hz of frequency space in their "on" scans centered around detections in their "off" scans. In the absence of complete information about their detections, we are unable to compute the magnitude by which their DFM was overestimated.

Our current procedure greatly improves on this situation by only requiring that detections do

not overlap in time-frequency space. This improvement allows us to search the space around detected signals instead of discarding frequency windows that are 1200-3000 Hz wide. As a result, our calculation of the DFM (Section 3.5.1) no longer suffers from overestimation problems.

We can get a better idea of the level of improvement to our data processing pipeline by comparing the signal counts recovered by the candidate detection method used by Margot et al. [2018] and the method that is used in this work (Figure 3.9). Evidently, our improved detection procedure recovers more than four times the number of detections reported by Margot et al. [2018]. Furthermore, 91 283 of the 97 083 signals that passed our Doppler and Direction-of-Origin filters (Section 3.3.4) were only detected as a result of this improvement. Eight of our final 30 candidates (Tables 3.3 and 3.4) would not have been detected with the detection algorithms of Margot et al. [2018].

The practice of blanking frequency space around candidate detections suggests that a large number of candidates, which could include valid technosignatures, are removed from consideration when using the detection algorithms of Siemion et al. [2013], Enriquez et al. [2017], Margot et al. [2018]. This practice makes the calculation of existence upper limits unreliable, because these pipelines remove from consideration many of the signals that they are designed to detect.

We have reprocessed the 2016 data discussed by Margot et al. [2018] with our improved algorithms (Appendix 3.D). We found no evidence of extraterrestrial technosignatures.

3.5.3 Existence Limits

Considering that radio SETI detection pipelines typically eliminate substantial fractions of the spectrum (e.g., Harp et al., 2016, Table 2; Enriquez et al., 2017, 1200-1350 MHz; Margot et al., 2018, Table 2; this work, Table 2) and further eliminate a fraction of the signals that they are designed to detect [Siemion et al., 2013, Enriquez et al., 2017, Margot et al., 2018], it is difficult to make general and robust statements about the prevalence of narrowband emitters in the Galaxy. One such claim by Enriquez et al. [2017] has been shown to be questionable [Margot et al., 2018].

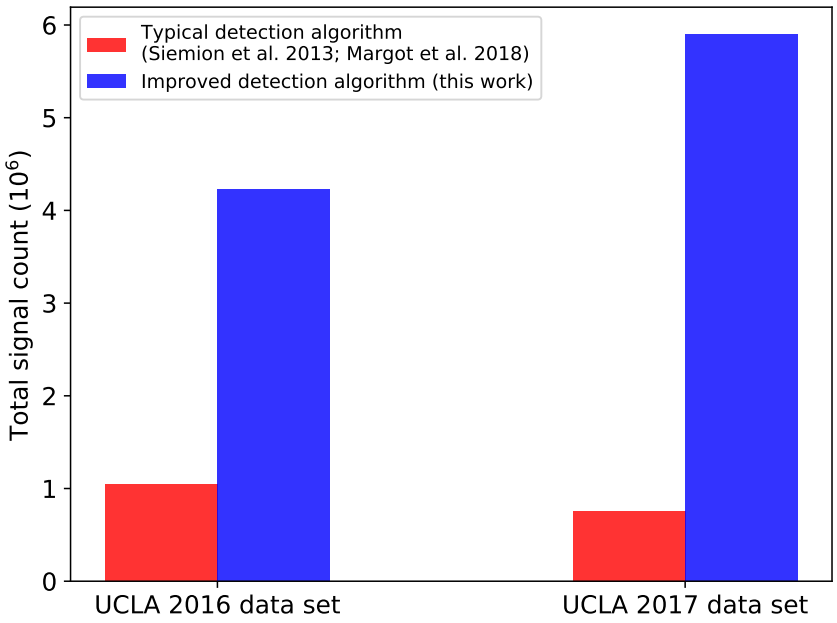

Figure 3.9: Detection counts obtained with the algorithm of Siemion et al. [2013], Margot et al. [2018] (red) and that presented in this work (blue). The removal of many legitimate detections by typical algorithms suggests that claims of existence limits based on the results of these algorithms and others like it [e.g., Enriquez et al., 2017] are questionable.

Injection of artificial signals in the data would demonstrate that a fraction of detectable and legitimate signals are not identified by existing pipelines. Until this fraction is properly quantified, it is wise to refrain from making overly confident claims about the prevalence of radio emitters in the galaxy.

3.5.4 Sensitivity

Margot et al. [2018] provide a detailed analysis of the sensitivity of a search performed with the 100 m GBT at a frequency resolution of $\Delta\nu = 2.98$ Hz sensitive to flux densities of 10 Jy. The results of that calculation [Margot et al., 2018, Figure 5] are generally applicable here because our search parameters are identical except for a slightly better sensitivity of 9.4 Jy. We estimate that civilizations located near the closest of our observed sources (TRAPPIST-1; ~ 40 ly) would

require a transmitter with only a small fraction ($< 1\%$) of the effective isotropic radiated power (EIRP) of the Arecibo planetary radar to be detectable in our search. Transmitters located as far as our most-distant observed source (Kepler-452; ~ 1800 ly) require approximately 18 times the Arecibo EIRP.

3.6 Conclusions

We described the results of a search for technosignatures using two hours of GBT telescope time in 2017. We identified 5,840,149 candidate signals, 98% of which were automatically eliminated by our rejection filters. Of the signals that remained, 30 were found outside of densely populated frequency regions and required further inspection. None of the remaining candidates were attributable to extraterrestrial technosignatures.

We found that quiet parts of the radio spectrum remain unexamined in the radio technosignature search pipelines of Siemion et al. [2013], Enriquez et al. [2017], Margot et al. [2018]. This problem results in inflated estimates of figures of merit and unreliable upper limits on the prevalence of technosignatures. To address this problem, we implemented a new procedure that increased the candidate detection efficiency by a factor of four or more compared to Margot et al. [2018].

3.A Candidate Signal Detection and Bandwidth Estimation

To identify promising candidate signals in the drift-rate-by-frequency array output by the Doppler de-smearing algorithm, we applied an iterative procedure that seeks out candidate signals with S/N exceeding 10. We began by arranging the 1023×2^{20} array so that the drift rates decrease linearly down the rows of the array (i.e., the first row contains the integrated power with drift rate 8.8644 Hz s^{-1}, the second row contains the integrated power with drift rate 8.8471 Hz s^{-1}, and so on). We then searched for the highest S/N signal in this array, and noted its drift rate $\dot{f}_0$ and frequency f_0 at the start of the scan.

We measured the bandwidth of candidate signals with the Python *SciPy* routine `peak_widths`. These routines measure the width from the central peak of the signal to the first point on either side above a pre-defined threshold value, using linear interpolation when necessary. We initially measured the bandwidth as the full width at half maximum (FWHM) of each signal. However, due to the peculiar nature of many of the detected signals this threshold value proved to be ineffective. We instead set the bandwidth measurement threshold at five times the standard deviation of the noise, which corresponds to half of our signal detection threshold (Figure 3.10).

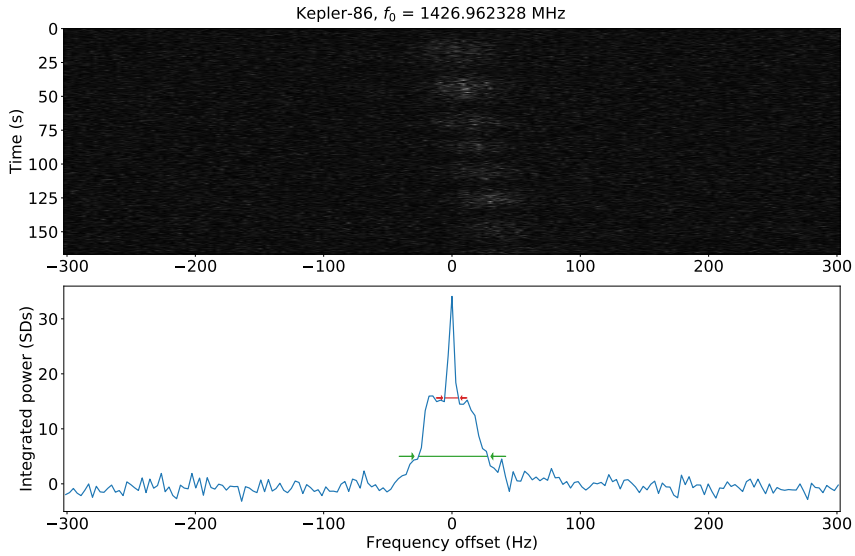

Figure 3.10: Sample result of the bandwidth estimation procedure. (Top) Time-Frequency diagram of signal. (Bottom) The power of the signal integrated with the best-fit drift rate of 0.2429 Hz s^{-1}, in units of standard deviations of the noise (σ). The estimated bandwidth measured at FWHM is shown by the red arrows and proved to be ineffective for the purpose of avoiding duplicate detections (e.g., signal outside this bandwidth exceeds our detection threshold of 10σ). The green arrows show the bandwidth measured at 5σ, which yields a more robust estimate of the bandwidth and which we used in our implementation.

By default, we used a search window of 200 frequency bins ($\sim$ 600 Hz) on either side of each signal to measure the bandwidth. For narrowband signals, this range is more than enough to

ensure that the entirety of a signal fits within the search window. If a detected signal is close to the edge of the channel and the required frequency window is not available on one side of the signal, we utilize the entire available frequency range on that side of the signal, and the full 600 Hz on the other. The *SciPy* routines conveniently provide the frequency coordinates of the left and right intersection points at the specified threshold, which we used to calculate the width of the signal to the left (Δf_{b_l}) and right (Δf_{b_r}) of the center frequency.

Signals with large bandwidths required additional care. If the median integrated power of the signal within the initial 1200 Hz window exceeded 5σ, we labeled the signal as broadband and increased the search window to 200000 frequency bins, or ~ 600 kHz, on either side of the main signal. We applied a Savitzky-Golay filter [Savitzky and Golay, 1964] with a window of 1001 frequency bins (~ 3000 Hz) and a polynomial order of 3 to the integrated spectra within the search window. This filter reduces noise by fitting a polynomial to all the points within the specified window and replacing the central point with the corresponding fit value. This filter was chosen for its computational advantages and simplicity. We then applied the bandwidth estimation procedure described above to the filtered points and stored the result in our database.

We used the bandwidth measurements in Equation 3.1, as follows:

$$\begin{aligned} f_0 - \Delta f_{b_l} < f < f_0 + (\dot{f} - \dot{f}_0)\tau + \Delta f_{b_r} \qquad &\text{if} \quad \dot{f} > \dot{f}_0 \\ f_0 + \Delta f_{b_r} > f > f_0 + (\dot{f} - \dot{f}_0)\tau - \Delta f_{b_l} \qquad &\text{if} \quad \dot{f} < \dot{f}_0 \end{aligned} \tag{3.4}$$

However, we chose to establish a minimum bandwidth to account for uncertainty in the determination of the drift rate of a signal (for example, if a signal's drift rate is time variable or is not a perfect multiple of the drift rate step $\Delta\dot{f}$). Specifically, we redefined Δf_{b_l} as min(Δf_{b_l}, $3\Delta\nu$) and Δf_{b_r} as min(Δf_{b_r}, $3\Delta\nu$), which amounts to a minimum of ~ 10 Hz on either side of each signal, for a minimum of 20 Hz per signal.

We removed redundant detections with frequencies given by Equation 3.4 by applying a mask to the drift-rate-by-frequency array. This mask is defined by drawing two lines through the array. The first is a vertical line at the frequency of the detected candidate signal. The second is a line

with a slope of $-\Delta\dot{f}/\Delta\nu$ crossing through the frequency and drift rate point corresponding to the detected signal (Figure 3.11). The boundaries defined by these two lines comprise the mask for zero bandwidth signals. In order to account for the finite bandwidth of a signal, the boundaries are shifted to the left and right so that the total width of the mask at the detection point matches the measured bandwidth of the signal (signified by the red arrows in Figure 3.11).

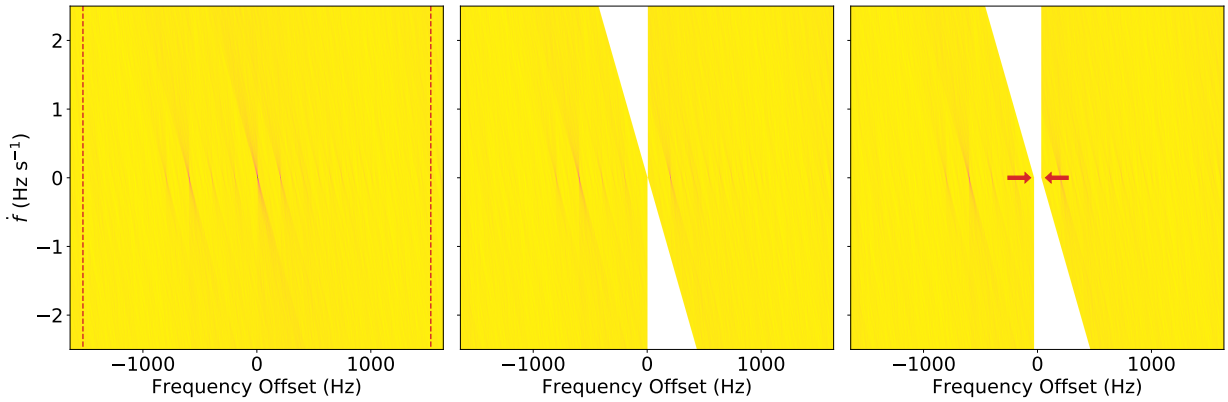

Figure 3.11: (Left) A portion of the original array output by the Doppler de-smearing algorithm for a sample scan (LHS 1140). The plot is centered on $f = 1546.879303$ MHz. The intensity of the plot represents the integrated power at a given drift rate and frequency. Signals of interest are represented as local maxima. The minimum integrated power in this array exceeds our 10 S/N detection threshold, therefore all signals shown are detectable by our pipeline. The signal with maximum S/N in this portion of the array is located at the center frequency. The area between the dashed vertical lines represents the slice of the array that would have been discarded after a signal detection with the approach of Siemion et al. [2013] and Margot et al. [2018], i.e., only one of the signals would have been reported. (Middle) Same array after application of our new masking procedure, assuming that the measured bandwidth of the signal is 0. With our new procedure, valid signals in the vicinity of the strongest signal are not discarded. (Right) Same array after application of our new masking procedure, assuming that the measured bandwidth of the signal is 60 Hz. This number was chosen for visualization purposes, and does not represent the true measured bandwidth (~ 3 Hz) of the center peak.

3.B Doppler and Direction-of-Origin Filters

To distinguish anthropogenic signals from potential extraterrestrial technosignatures, we invoked several filter procedures within our database to flag promising technosignature candidates. The first filter flags candidate signals with non-zero Doppler drift rates.

The direction-of-origin filters require signals from different scans to be compared and possibly paired. We pair two signals if they have similar drift rates and compatible frequencies, i.e., the frequencies at the beginning of each scan are within some tolerance of a linear relationship with a slope equal or nearly equal to the drift rate. We quantify these tests as follows. Consider a signal with start time t_0, start frequency f_0, and drift rate $\dot{f}_0$, and another signal from a different scan with corresponding values t, f, $\dot{f}$. We define $\Delta t = t - t_0$ and require that f falls in the interval $[f_-, f_+]$ for pairing, where

$$f_\pm = (\dot{f}_0 \pm 2\Delta\dot{f})\Delta t \pm \Delta\nu. \tag{3.5}$$

In this work, the values $\Delta\dot{f}$ and $\Delta\nu$ are given by 0.0173 Hz $s^{-1}$and 2.98 Hz, respectively. To account for uncertainty in the drift rate determination, we allow for a drift rate difference of $\Delta\dot{f}$. We thus query the database for all candidate signals with a frequency at the start of the scan in the range $[f_-, f_+]$ with a drift rate $\dot{f}_0 - \Delta\dot{f} \leq \dot{f} \leq \dot{f}_0 + \Delta\dot{f}$ Hz s^{-1}. Two signals are considered paired if the following condition holds:

$$\min_i |(f - f_0) + \dot{f}_i \Delta t| \leq \Delta\nu \tag{3.6}$$

where $\dot{f}_i \in \{\dot{f}_0, \dot{f}_0 \pm \Delta\dot{f}, \dot{f}, \dot{f} \pm \Delta\dot{f}\}$.

To determine whether a signal is persistent, i.e., whether it is detected in both scans of its source, we apply the above procedure to each candidate and all candidates detected in the second scan of the same source. If a match is found, both signals are flagged. To determine whether a signal's direction of origin is unique, i.e, to ensure that it is not detected in the scans corresponding to other sources, we apply the above procedure to each candidate and all candidates detected in either scan of all other sources. All matches are discarded from consideration. The candidate

signals that remain are flagged as having a single direction of origin.

3.C Signal Density Thresholding

In order to remove candidate signals that were likely to be anthropogenic RFI, we developed a procedure to filter signals based on the density of nearby detections. We began by dividing the 1.15 – 1.73 GHz range into 1-kHz-wide frequency bins. For each bin, we measured the signal density by counting the number of detections within a window centered on the bin. For these calculations, we excluded the regions listen in Table 3.2. We tested four different window sizes: 1 kHz, 10 kHz, 100 kHz, and 1 MHz. For each window size, we plotted a histogram of signal counts. We found that a 1 MHz window resulted in a distinctive transition between small and large signal densities at approximately 1000 signals MHz^{-1} (Figure 3.12), and we used this threshold and window size to filter out regions of high signal density.

3.D Re-analysis of 2016 Data

Margot et al. [2018] presented the results of a search for technosignatures around 14 planetary systems in the Kepler field conducted on April 15, 2016, 16:00 - 18:00 universal time (UT) with the GBT. They discarded a frequency region of width $\pm\dot{f}_{max}\tau \simeq 3000$ Hz around each detection, where $\dot{f}_{max} = 8.86$ Hz s^{-1} is the maximum drift rate detectable by their search and $\tau \approx 150$ s is the integration time of the scan. This process or a variant of it is also implemented in the radio SETI detection pipelines of Siemion et al. [2013] and Enriquez et al. [2017]. As a result, a substantial portion of the spectrum remains unexamined in these searches, and the Drake figure of merit associated with these searches is overestimated (Section 3.5).

To remedy this situation, we have re-analyzed the data obtained by Margot et al. [2018] using the candidate detection procedure described in this work. We have found a total of 4 228 085 signals as compared with 1 046 144 reported previously. Our Doppler and direction-of-origin rejec-

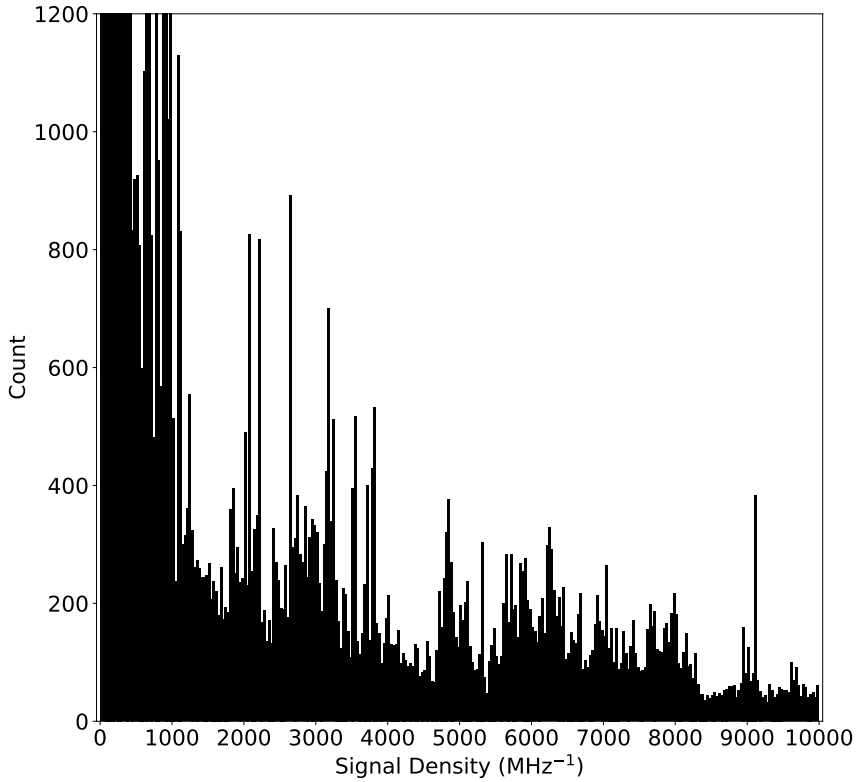

Figure 3.12: Signal densities of 1kHz regions in the 1.15–1.73 GHz range. The plot is clipped at a max of 1200 signals MHz^{-1}. Note the sharp dropoff at approximately 1000 signals MHz^{-1}.

tion algorithms (Section 3.3.4, Appendix 3.B) automatically labeled more than 99% of the detected signals as anthropogenic RFI. After removing all remaining signals found within operating bands of the interferers described by Margot et al. [2018, Tables 2 and 3], we were left with 18 technosignature candidates. Seven of these had been identified by Margot et al. [2018] and attributed to anthropogenic RFI. The properties of the remaining 11 technosignature candidates are given in Table 3.5. These signals were further scrutinized and categorized according to the procedure described in Section 3.4. All were found to be attributable to anthropogenic RFI.

We note that 3 of the 19 final technosignature candidates described by Margot et al. [2018] were found to be part of a broadband RFI signal and were removed via Equation 3.1. The remaining 9 signals were correctly labeled as RFI by our Doppler and direction-of-origin rejection algorithms

or frequency-based filters. This enhancement in classification performance was only possible because of the improvements to the candidate signal detection algorithms presented in this work and because the raw data were preserved for re-analysis.

Table 3.1: Target host stars listed in order of observation. Distances in light years (ly) were obtained from the NASA Exoplanet Archive. Habitable Zone categories are described by Kane et al. [2016]. Categories 1 and 2 refer to small ($R_p < 2R_E$) planets in the conservative and optimistic habitable zones, respectively, while categories 3 and 4 refer to planets of any radius in the conservative and optimistic habitable zones, respectively. TRAPPIST-1 and LHS 1140 were categorized on the basis of orbital radii from the NASA Exoplanet Archive and HZ boundaries as calculated with the algorithm of [Kopparapu et al., 2013].

Host Star	Distance (ly)	HZ Category
Kepler-442	1115^{+62}_{-72}	1, 2
Kepler-440	851^{+52}_{-150}	2
Kepler-174	1210.04^{+63}_{-56}	3, 4
Kepler-62	1200	1, 2
Kepler-296	737^{+91}_{-59}	1, 2
Kepler-86	1128.5^{+44}_{-42}	4
Kepler-22	620	4
Kepler-283	1477.49^{+67}_{-74}	1, 2
Kepler-452	1787.34^{+395}_{-243}	2
Kepler-186	561^{+42}_{-33}	1, 2
TRAPPIST-1	39.5 ± 1.3	1, 2
LHS 1140	40.67 ± 1.37	1

Table 3.2: Spectral regions exhibiting a high density of detections per unit frequency. Known anthropogenic interferers are listed in the 'Identification' column.

Frequency Region (MHz)	Total detection count	Density (# per MHz)	Identification
1155.99 - 1196.91	3 579 122	87 466	GPS L5
1192.02 - 1212.48	96 233	4703	GLONASS L3
1422.320 - 1429.992	41 757	5443	ARSR products?
1525 - 1559	809 877	23 820	Satellite downlinks
1554.96 - 1595.88	718 711	17 564	GPS L1
1592.9525 - 1610.485	184 207	10 507	GLONASS L1

Table 3.3: Characteristics of Top 15 Candidates from Kepler sources. Properties are listed for the first scan of the source only. For a description of the 'Category' column, see Section 3.4.

Source	Epoch (MJD)	Frequency (Hz)	Drift Rate Hz s^{-1}	S/N	Category
Kepler-86	57877.66340278	1151559391.617775	0.0867	11.2	a
		1457412120.699883	0.4510	19.7	b
Kepler-174	57877.65087963	1501593533.158302	-0.0520	33.8	b
Kepler-186	57877.68348380	1457412961.125374	0.5551	18.6	c
		1457488715.648651	0.1908	15.6	a
		1457489272.952080	0.1735	12.9	a
		1693601790.070534	0.1041	189.0	b
Kepler-296	57877.66091435	1420354264.974594	0.2429	11.2	f
		1420476266.741753	0.2255	15.4	f
		1431607288.122177	0.2082	10.5	f
		1435340115.427971	0.1561	26.6	a
Kepler-440	57877.64343750	1457473185.658455	0.3296	26.2	b
Kepler-442	57877.64098380	1436369919.776917	0.3123	11.3	a
		1457510107.755661	0.2949	22.4	d
		1501766377.687454	-0.0520	17.8	e

Table 3.4: Characteristics of Top 15 Candidates from LHS 1140 and TRAPPIST-1. Properties are listed for the first scan of the source only. For a description of the 'Category' column, see Section 3.4.

Source	Epoch (MJD)	Frequency (Hz)	Drift Rate Hz s^{-1}	S/N	Category
LHS 1140	57877.69550926	1675002533.197403	0.01735	28.0	f
		1676205691.695213	0.01735	23.5	f
		1676211157.441139	-0.01735	24.0	f
		1677041041.851044	-0.01735	12.9	f
		1678640750.050545	0.03469	17.0	f
		1728618854.284286	-0.01735	10.0	
TRAPPIST-1	57877.69228009	1151731526.851654	0.01735	16.2	
		1463927415.013313	0.03469	92.4	a
		1675910618.901253	-0.01735	14.0	f
		1676249122.619629	-0.01735	13.1	f
		1676298260.688782	-0.01735	14.6	f
		1677018946.409225	-0.01735	14.1	f
		1678110888.600349	0.03469	12.0	f
		1678460314.869881	0.01735	13.6	f
		1678716921.806335	0.01735	15.5	f

Table 3.5: Characteristics of Top 11 Candidates from 2016 Search. Properties are listed for the first scan of the source only. For a description of the 'Category' column, see Section 3.4.

Source	Epoch (MJD)	Frequency (Hz)	Drift Rate Hz s^{-1}	S/N	Category
Kepler-22	57493.70935185	1456613901.257515	0.52042	14.3	a
Kepler-296	57493.71608796	1454808911.681175	0.05204	10.3	a
		1454877239.465714	0.01735	14.0	a
		1457440766.692162	0.31225	12.9	a
		1457443791.627884	0.31225	11.9	b
Kepler-399	57493.68870370	1414058676.362038	0.01735	764.4	b
Kepler-407	57493.71819444	1620381447.672844	-6.40113	14.5	d
Kepler-440	57493.74275463	1457486632.466316	0.41633	21.4	b
		1457486835.122108	0.39899	23.6	b
Kepler-442	57493.74497685	1308946093.916893	-0.01735	11.3	e
		1376548311.114311	-0.24286	540.9	b

CHAPTER 4

A Search for Technosignatures Around 31 Sun-like Stars with the Green Bank Telescope at 1.15–1.73 GHz

We conducted a search for technosignatures in April of 2018 and 2019 with the L-band receiver (1.15 – 1.73 GHz) of the 100 m diameter Green Bank Telescope. These observations focused on regions surrounding 31 Sun-like stars near the plane of the Galaxy. We present the results of our search for narrowband signals in this data set as well as improvements to our data processing pipeline. Specifically, we applied an improved candidate signal detection procedure that relies on the topographic prominence of the signal power, which nearly doubles the signal detection count of some previously analyzed data sets. We also improved the direction-of-origin filters that remove most radio frequency interference (RFI) to ensure that they uniquely link signals observed in separate scans. We performed a preliminary signal injection and recovery analysis to test the performance of our pipeline. We found that our pipeline recovers 93% of the injected signals over the usable frequency range of the receiver and 98% if we exclude regions with dense RFI. In this analysis, 99.73% of the recovered signals were correctly classified as technosignature candidates. Our improved data processing pipeline classified over 99.84% of the $\sim$26 million signals detected in our data as RFI. Of the remaining candidates, 4539 were detected outside of known RFI frequency regions. The remaining candidates were visually inspected and verified to be of anthropogenic nature. Our search compares favorably to other recent searches in terms of end-to-end sensitivity, frequency drift rate coverage, and signal detection count per unit bandwidth per unit integration time.

4.1 Introduction

We describe a search for radio technosignatures with the L-band receiver of the 100 m diameter Green Bank Telescope (GBT). We used a total of 4 hours of GBT time in 2018 and 2019 to observe the regions around 31 Sun-like stars near the plane of the Galaxy. We have so far prioritized the detection of narrowband ($\sim$ 10 Hz) signals because they are diagnostic of engineered emitters [e.g., Tarter, 2001].

Our search builds on the legacy of technosignature searches performed in the period 1960–2010 [Tarter, 2001, Tarter et al., 2010, and references therein] and previous searches conducted by our group [Margot et al., 2018, Pinchuk et al., 2019]. Other recent searches include work conducted by Siemion et al. [2013], Harp et al. [2016], Enriquez et al. [2017], Gray and Mooley [2017], Price et al. [2020].

Our choice of search parameters has key advantages compared to the Breakthrough Listen (BL) searches described by Enriquez et al. [2017] and Price et al. [2020], which contend with much larger data volumes. Specifically, our search provides roughly uniform detection sensitivity over the entire range of frequency drift rates (± 8.86 Hz s^{-1}) whereas the BL searches suffer a substantial loss in sensitivity due to the spreading of signal power across up to 13–26 frequency resolution cells. In addition, we cover a range of frequency drift rates that is 2–4 wider than the BL searches with a time resolution that is 51 times better.

Our search algorithms are distinct from the BL algorithms in that they alleviate the necessity of discarding $\sim$kHz wide regions of frequency space around every detected signal. We abandoned this practice in previous work [Pinchuk et al., 2019]. In this work, we further refined our algorithm by implementing a candidate signal detection procedure that relies on the concept of prominence and by removing the requirement to compute the bandwidth of candidate signals. Our new approach, combined with better end-to-end sensitivity and drift rate coverage, enables a hit rate density or signal detection count per unit bandwidth per unit integration time that is $\sim$300 times larger than that of the BL search described by Price et al. [2020].

A key measure of the robustness and efficiency of a data processing pipeline is provided by the technique of signal injection and recovery [e.g., Christiansen et al., 2013], whereby artificial signals are injected into the raw data and the fraction of signals recovered by the pipeline is quantified. Despite the importance of this metric, we are not aware of an existing tool to quantify the recovery rates of data-processing pipelines in radio technosignature searches. We make a first step towards the implementation of this tool and show that our current pipeline detects 93% of the injected signals over the usable frequency range of the receiver and 98% if we exclude regions with dense RFI. In addition, our pipeline correctly flagged $99.73\%$of the detected signals as technosignature candidates. Although our current implementation requires additional work to fully capture the end-to-end pipeline efficiency, it can already illuminate imperfections in our and other groups' pipelines and be used to calibrate claims about the prevalence of other civilizations [e.g., Enriquez et al., 2017].

The chapter is organized as follows. Our data acquisition and analysis techniques are presented in Sections 4.2 and 4.3, respectively. Our preliminary signal injection and recovery analysis is described in Section 4.4. The main results of our search are outlined in Section 4.5. In Section 4.6, we describe certain advantages of our search, including dechirping efficiency, drift rate coverage, data archival practices, candidate detection algorithm, and hit rate density. We also discuss limits on the prevalence of other civilizations, search metrics such as the Drake Figure of Merit, and re-analysis of previous data with our latest algorithms. We close with our conclusions in Section 4.7.

4.2 Data Acquisition and Pre-Processing

Our data acquisition techniques are generally similar to those used by Margot et al. [2018] and Pinchuk et al. [2019]. Here, we give a brief overview and refer the reader to these other works for additional details.

4.2.1 Observations

We selected 31 Sun-like stars (spectral type G, luminosity class V) with median galactic latitude of 0.85° (Tables 4.1 and 4.2) because their properties are similar to the only star currently known to harbor a planet with life. We observed these stars with the GBT during two 2-hour sessions separated by approximately one year. During each observing session, we recorded both linear polarizations of the L-band receiver with the GUPPI backend in its baseband recording mode [DuPlain et al., 2008], which yields 2-bit raw voltage data after requantization with an optimal four-level sampler [Kogan, 1998]. The center frequency was set to 1.5 GHz and we sampled 800 MHz of bandwidth between 1.1 and 1.9 GHz, which GUPPI channelized into 256 channels of 3.125 MHz each. We validated the data acquisition and analysis processes at the beginning of each observing session by injecting a monochromatic tone near the receiver frontend and recovering it at the expected frequency in the processed data.

We observed all our targets in pairs in order to facilitate the detection and removal of signals of terrestrial origin (Section 4.3.2). The sources were paired in a way that approximately minimized telescope time overhead, i.e., the sum of the times spent repositioning the telescope. Pairings were adjusted to avoid pair members that were too close to one another on the plane of the sky with the goal of eliminating any possible ambiguity in the direction of origin of detected signals. Specifically, we required angular separations larger than 1° between pair members, i.e., several times the ~ 8.4 arcmin beamwidth of the GBT at 1.5 GHz.

Each pair was observed twice in a 4-scan sequence: A, B, A, B. The integration time for each scan was 150 s, yielding a total integration time of 5 minutes per target. CoRoT 102810550 and CoRoT 110777727 were each observed for an additional two scans. With 66 scans of 150 s duration each, our total integration time amounts to 2.75 hr.

4.2.2 Sensitivity

Margot et al. [2018] calculated the sensitivity of a search for narrowband signals performed with

the 100 m GBT. Assuming a System Equivalent Flux Density (SEFD) of 10 Jy, integration time of 150 s, and frequency resolution of 3 Hz, they found that sources with flux densities of 10 Jy can be detected with a signal-to-noise ratio (S/N) of 10. The results of that calculation are directly applicable here because our search parameters are identical to that study. Specifically, our search is sensitive to transmitters with the effective isotropic radiated power (EIRP) of the Arecibo planetary radar transmitter (2.2×10^{13} W) located 420 ly from Earth [Margot et al., 2018, Figure 5]. Transmitters located as far as the most distant source (CoRoT 102963038; $\sim 10,407$ ly) and with $\gtrsim 600$ times the Arecibo EIRP can also be detected in this search. Although we selected Sun-like stars as primary targets, our search is obviously sensitive to other emitters located within the beam of the telescope. A search of the Gaia DR2 catalog [Gaia Collaboration, 2016, 2018] inspired by Wlodarczyk-Sroka et al. [2020] reveals that there are 15,031 known stars with measured parallaxes within the half-power beamwidths associated with our 31 primary sources. The median and mean distances to these sources are 2088 and 7197 ly, respectively.

4.2.3 Computation of Power Spectra

After unpacking the digitized raw voltages from 2-bit to 4-byte floating point values, we computed consecutive power spectra with 2^{20}-point Fourier transforms, yielding a frequency resolution of $\Delta f = 2.98$ Hz. We chose this frequency resolution because it is small enough to provide unambiguous detections of narrowband ($<$10 Hz) technosignatures and large enough to examine Doppler frequency drift rates of up to nearly ± 10 Hz s^{-1} (Section 4.2.4). We processed all channels within the operating range of the GBT L-band receiver (1.15–1.73 GHz), excluding channels that overlap the frequency range (1200–1341.2 MHz) of a notch filter designed to mitigate radio frequency interference (RFI) from a nearby aircraft detection radar. Although Enriquez et al. [2017] and Price et al. [2020] used the L-band receiver over a larger frequency range (1.1–1.9 GHz), we used a narrower range because we observed serious degradation of the bandpass response beyond the nominal operating range of the receiver.

In order to correct for the bandpass response of GUPPI's 256 channels, we fit a 16-degree

Chebyshev polynomial to the median bandpass response of a subset of the processed channels that did not include strong RFI and that were not close to the cutoff frequencies of filters located upstream of the GUPPI backend. After applying the bandpass correction to all channels, we stored the consecutive power spectra as rows in time-frequency arrays (a.k.a time–frequency diagrams, spectrograms, spectral waterfalls, waterfall plots, or dynamic spectra) and normalized the power to zero mean and unit standard deviation of the noise power. The normalized power values reflect the S/N at each time and frequency bin.

4.2.4 Doppler Dechirping

Due to the orbital and rotational motions of both the emitter and the receiver, we expect extraterrestrial technosignatures to drift in frequency space [e.g., Siemion et al., 2013, Margot et al., 2018, Pinchuk et al., 2019]. To integrate the signal power over the scan duration while compensating for Doppler drifts in signal frequency, we used incoherent sums of power spectra, where each individual spectrum was shifted in frequency space by a judicious amount prior to summation. This technique is known as incoherent dechirping. Coherent dechirping algorithms exist [Korpela, 2012] but are computationally expensive and seldom used.

Because the Doppler drift rates due to the emitters are unknown, we examined 1023 linearly spaced drift rates in increments of $\Delta\dot{f} = 0.0173$ Hz s^{-1}over the range ± 8.86 Hz s^{-1}. To accomplish this task, we used a computationally advantageous tree algorithm [Taylor, 1974, Siemion et al., 2013], which operates on the dynamic spectra and yields time integrations of the consecutive power spectra after correcting approximately for each trial Doppler drift rate. The algorithm requires input spectra with a number of rows equal to a power of two, and we zero-padded the dynamic spectra with approximately 65 rows to obtain 512 rows. The output of this algorithm, which was run once for negative drift rates and once for positive drift rates, is stored in 1023×2^{20} drift-rate-by-frequency arrays that are ideal for identifying candidate signals, i.e., radio signals that exceed a certain detection threshold (Section 4.3.1). We quantify the sensitivity penalty associated with the use of the tree algorithm in Section 4.6.1.

4.3 Data Analysis

4.3.1 Candidate Signal Detection

We performed an iterative search for candidate signals on the drift-rate-by-frequency arrays obtained with the incoherent dechirping algorithm. Specifically, we identified the signal with the highest integrated S/N and stored its characteristics in a structured query language (SQL) database, then identified and recorded the signal with the second highest S/N, and so on. Redundant detections can occur when signals in the vicinity of a candidate signal have large integrated power along similar drift rates. Different data processing pipelines tackle these redundant detections in different ways. Siemion et al. [2013], Enriquez et al. [2017], Margot et al. [2018] and Price et al. [2020] discarded all detections within $\sim$kHz-wide regions of frequency space around every candidate signal detection. This method leaves large portions of the observed frequency space unexamined and biases the results towards high S/N signals because signals with lower S/N in their vicinity are discarded. Importantly, this method complicates attempts to place upper limits on the abundance of technosignature sources because the pipeline eliminates the very signals it purports to detect (Sections 4.6.4 and 4.6.5).

Pinchuk et al. [2019] introduced a novel procedure to alleviate these shortcomings. In order to avoid redundant detections, they imposed the restriction that two signals cannot cross in in the time-frequency domain of the scan and discarded detections only in a small frequency region around every candidate signal detection. The extent of this region was set equal to the bandwidth of the candidate signal measured at the 5σ power level, where σ is one standard deviation of the noise. Unfortunately, this bandwidth calculation can cause complications in some situations. For example, small noise fluctuations may result in an unequal number of candidate signals detected in two scans of a source. In the $\sim$400 Hz region of the spectrum shown on Figure 4.1, two signals (± 100Hz) are detected in the first scan but only one signal (-100Hz) is detected in the second scan. This incompleteness is detrimental to our direction-of-origin filters (Section 4.3.2), which rely on accurate signal detection across all scans. Moreover, discarding the region corresponding

to a bandwidth measured at 5σ prevents the detection of at least five other signals per scan in this example (Figure 4.1).

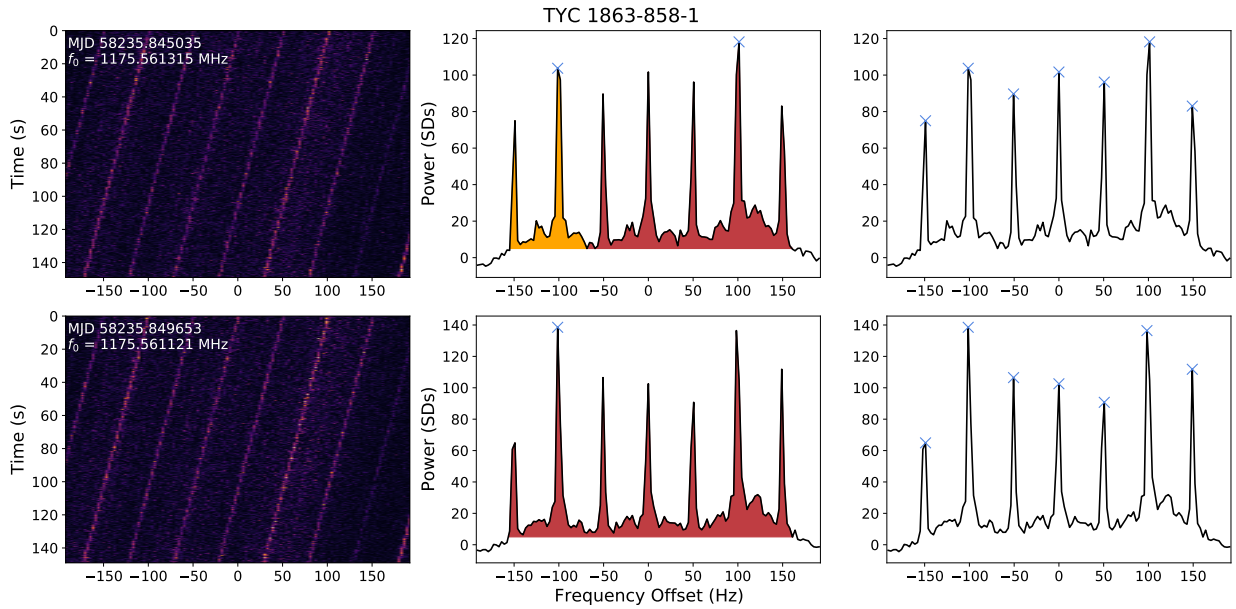

Figure 4.1: Comparison of signal detection procedures illustrated on a ∼400 Hz region for scans 1 (top) and 2 (bottom) of TYC 1863-858-1. (Left) Dynamic spectra, where pixel intensity represents signal power. (Middle) Integrated power spectra with blue crosses marking the signals that are detected with the procedure described by Pinchuk et al. [2019]. In the first scan, the strongest signal (+100 Hz) is detected and the corresponding 5σ bandwidth is shown in red. The second strongest signal (-100 Hz) is then detected and the corresponding 5σ bandwidth is shown in orange. In the second scan, only the strongest signal, which is now at -100 Hz, is detected. (Right) Integrated power spectra with blue crosses marking the signals that are detected with the procedure described in this work.

In this work, we improve on the procedure presented by Pinchuk et al. [2019] in two important ways. First, we identify candidate signal detections on the basis of the topography-inspired concept of prominence. The prominence of a signal is defined as the vertical distance between the peak and its lowest contour line, as implemented in the numerical computing package *SciPy* [Virtanen

et al., 2020]. Because our integrated spectra are one-dimensional, we take the larger of a peak's two 'bases' as a replacement for the lowest contour line. The high-frequency (low-frequency) base is defined as the minimum power in the frequency region starting on the high (low) frequency side of the peak and ending +500 Hz (-500 Hz) away or at the frequency location of the nearest peak with higher (lower) frequency and larger power, whichever results in the smallest frequency interval. While the $\pm$500 Hz limits are not essential to compute prominences, they do speed up the calculations. Second, we remove the bandwidth-dependence of Pinchuk et al. [2019]'s algorithm. Instead, we apply a local maximum filter to the drift-rate-by-frequency arrays in order to remove any points that are not a maximum in their local 3x3 neighborhood. We find that this filter in conjunction with the prominence-based candidate signal detection identifies the signals of interest without introducing redundant detections.

Signals are considered candidate detections if their prominence meets two criteria: (1) their prominence exceeds 10σ, where σ is one standard deviation of the noise in the integrated spectrum, (2) their prominence exceeds a fraction f of their integrated power. For this analysis, we settled on $f = 75\%$. The second requirement is necessary because power fluctuations superimposed on strong broadband signals that approach or exceed the 10σ detection threshold can yield prominences that exceed 10σ. With this second requirement, a signal with a prominence of 10σ above a 3.0-σ baseline would be marked as a detection, but the same signal above a 3.5-σ baseline would not. Figure 4.2 describes the detection space.

As a result of these candidate signal detection improvements, we now detect 1.23–1.75 and $\sim$12 times as many signals as we did with the data processing pipelines of Pinchuk et al. [2019] and Margot et al. [2018], respectively (Figure 4.3). We compare this signal detection performance to that of other searches in Section 4.6.4.

Once a signal with frequency f_0 and drift rate $\dot{f}_0$ is detected with the criteria described in this section, we follow the procedure outlined by Pinchuk et al. [2019]. Specifically, we eliminate any other candidate signal with frequency f and drift rate $\dot{f}$ if the following inequalities hold true at

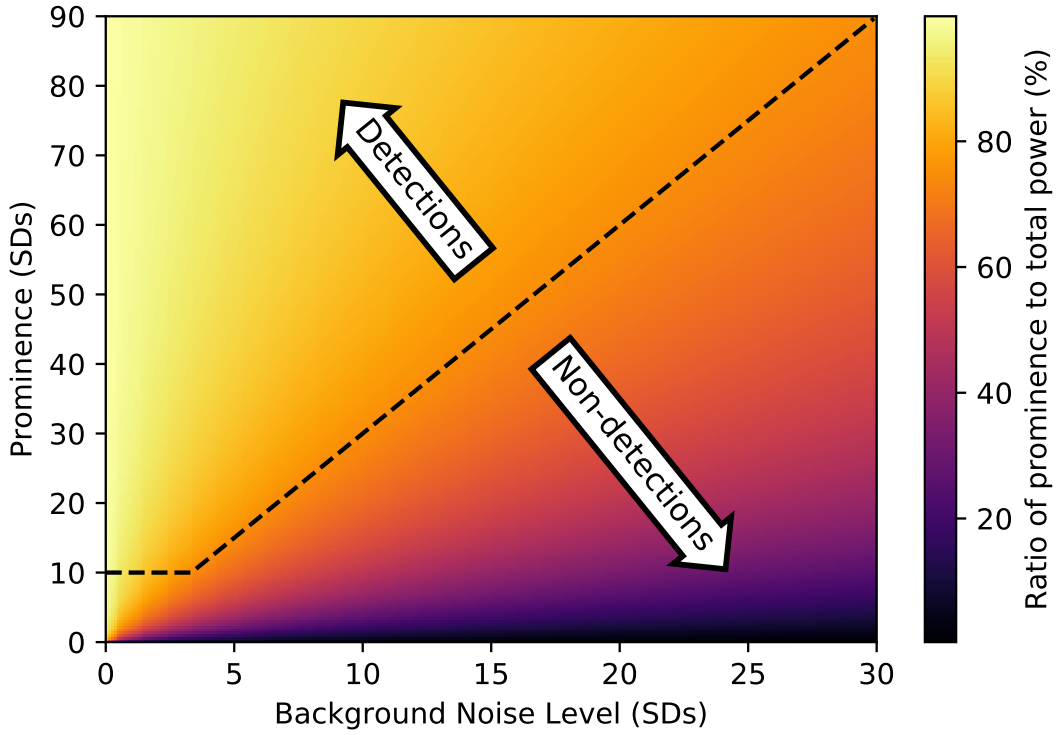

Figure 4.2: Illustration of detection criteria. Signals above the dashed black line line are marked as detections by our pipeline.

the start of the scan:

$$f_0 < f < f_0 + (\dot{f} - \dot{f}_0)\tau \qquad \text{for} \quad \dot{f} > \dot{f}_0 \tag{4.1}$$
$$f_0 > f > f_0 + (\dot{f} - \dot{f}_0)\tau \qquad \text{for} \quad \dot{f} < \dot{f}_0$$

where τ is the scan duration. Our candidate detection procedure was applied iteratively until all candidate signals with prominence $\geq 10\sigma$ were identified. Occasionally, signals with prominences $\geq 10\sigma$ but S/N < 10 get recorded in the database. This condition tends to occur primarily in regions with dense RFI where the baseline subtraction is imperfect. For this reason, we flagged all signals with S/N < 10 and did not consider them to be valid candidates.

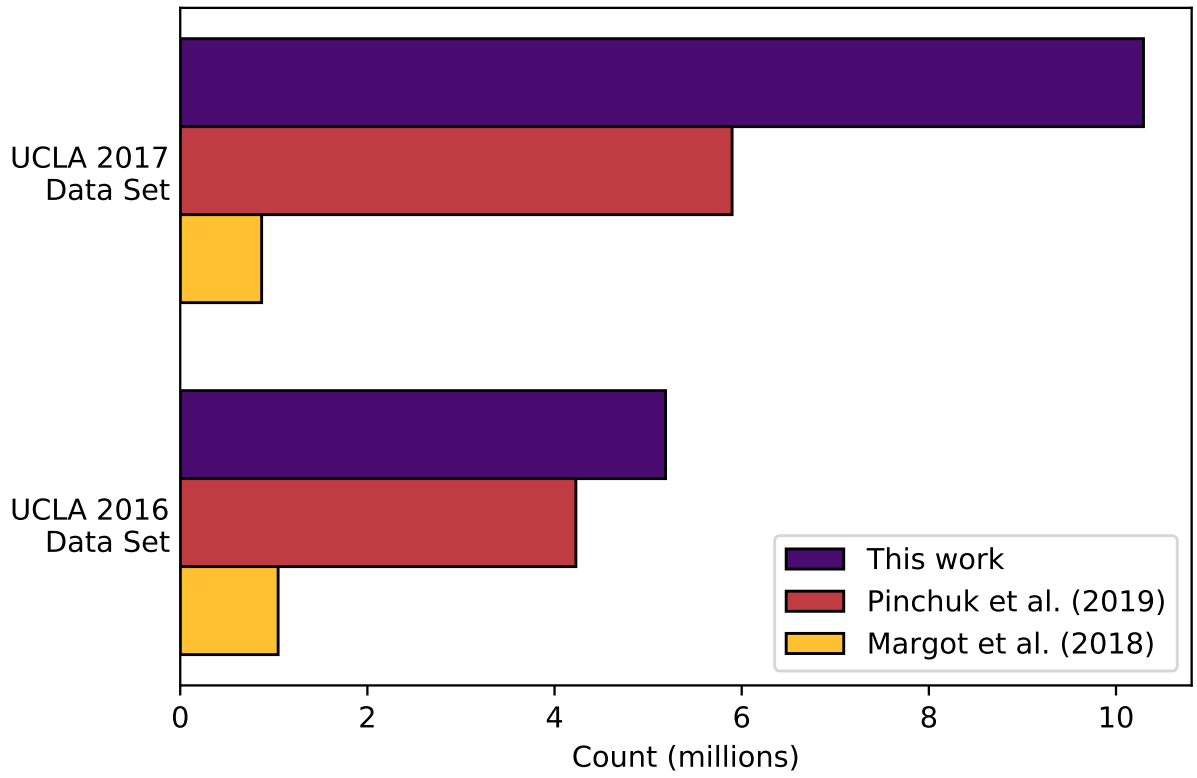

Figure 4.3: Detection counts obtained with the algorithms presented by Margot et al. [2018], Pinchuk et al. [2019], and this work. Our current pipeline detects 1.23–1.75 as many signals as Pinchuk et al. [2019]'s pipeline and ~ 12 times as many signals as Margot et al. [2018]'s pipeline.

4.3.2 Doppler and Direction-of-Origin Filters

After identifying all candidate signals, we applied a Doppler filter and improved variants of our direction-of-origin filters [Margot et al., 2018, Pinchuk et al., 2019] to detect and discard anthropogenic signals in the data.

We began by applying a Doppler filter, which is designed to remove all signals with zero Doppler drift rate, defined here as signals that drift less than one frequency resolution cell (Δf = 2.98 Hz) over the duration of a scan (τ = 150 s). The signals removed by this filter are of no interest to us because the corresponding emitters exhibit no line-of-sight acceleration with respect

to the receiver, suggesting that they are terrestrial in nature.

Next, we applied two direction-of-origin filters, which are designed to remove any signal that is either not persistent (i.e., not detected in both scans of its source) or detected in multiple directions on the sky (i.e., also detected in scans corresponding to other sources). Because the largest possible sidelobe gain is approximately -30 dB compared to the main lobe gain, signals detected in multiple directions on the sky are almost certainly detected through antenna sidelobes. The second filter is highly effective at removing such signals.

As explained by Pinchuk et al. [2019], the direction-of-origin filters compare signals from different scans and flag them according to the observed relationships. For example, if a signal from a scan of source A is paired with a signal from a scan of source B, then both signals are removed because they are detected in multiple directions of the sky. In our previous implementation of these filters, two signals were considered a pair if their drift rates were similar and their frequencies at the beginning of each scan were within a predetermined tolerance of a straight line with a slope corresponding to the drift rate. With this definition, it was possible for *multiple* signals in one scan to be paired with a *single* signal from a different scan, which is undesirable. For example, a valid technosignature candidate from one of the scans of source 'A' could be labeled as RFI because it was paired with a signal from one of the scans of source 'B', even if the signal in the scan of source 'B' was already paired with a different (RFI) signal from the scan of source 'A'.

We have redesigned our filter implementation to keep a record of all signals that are paired during filter execution. We use this record to impose the restriction that each signal is allowed to pair with only one other signal in each scan. Additionally, we implemented an improved pairing procedure that is loosely based on the Gale – Shapley algorithm [Gale and Shapley, 1962] designed to solve the stable matching problem. Our improved procedure operates as follows. We define the propagated frequency difference $\Delta F(f_i, f_j)$ of two signals from different scans to be

$$\Delta F(f_i, f_j) = \left| (f_i - f_j) + \overline{\dot{f}_{ij}} \Delta t_{ij} \right| \tag{4.2}$$

where f_i and f_j are the start frequencies of the two signals, $\overline{\dot{f}_{ij}} = (\dot{f}_i + \dot{f}_j)/2$ is the average of the

two signal drift rates, and $\Delta t_{ij} = t_j - t_i$ is the time difference between the two scans. Our updated algorithm iterates over all remaining unpaired candidate signals and updates the pairings until $\Delta F(f_i, f_j)$ is minimized for all signal pairs. In rare cases when the minimum value of $\Delta F(f_i, f_j)$ is not unique, multiple pairings are allowed, but no inference about the anthropogenic nature of the signals is made on the basis of these pairings alone.

To ensure that paired signals likely originated from the same emitter, we impose two additional requirements on all signal pairs. First, we require that

$$f_{ij,-} \leq f_j \leq f_{ij,+} \quad \text{or} \quad f_{ji,-} \leq f_i \leq f_{ji,+} \tag{4.3}$$

where $f_{ij,\pm} = f_i + (\dot{f}_i \pm \Delta\dot{f})\Delta t_{ij} \pm \Delta f$ represent the propagated frequency bounds and $\Delta\dot{f}$, Δf are the drift rate and frequency resolution, given by 0.0173 Hz $s^{-1}$and 2.98 Hz, respectively. This condition places an upper limit on $\Delta F(f_i, f_j)$ and we reject signal pairs whose propagated frequency differences exceed this bound. Second, we require that

$$\left|\dot{f}_i - \dot{f}_j\right| \leq 2\Delta\dot{f}, \tag{4.4}$$

and we reject signal pairs that do not satisfy this criterion. In tandem, these requirements reduce the possibility of pairing two unrelated signals.

To determine if a signal is persistent (first filter), we apply the pairing procedure to candidate signals detected in both scans of a source. Those signals left without a partner are deemed to originate from transient sources and are labeled as RFI. To determine whether a signal is detected in multiple directions of the sky (second filter), we apply the pairing procedure to signals from scans of different sources. In this case, all resulting pairs are attributed to RFI and discarded. Candidate signals remaining after the application of these procedures are marked for further inspection.

4.3.3 Frequency Filters

A majority of the candidate signals detected in our search are found in operating bands of known interferers. Table 4.3 describes the frequency ranges and signal counts associated with the most

prominent anthropogenic RFI detected in our data. Candidate signals detected within these frequency regions (except the ARSR products region) were removed from consideration because of their likely anthropogenic nature. The combined 2017 and 2018 signal detection counts in the excluded RFI regions (156,327/MHz) are considerably higher than outside of these regions (20,654/MHz) or in the 1400–1427 MHz radio astronomy protected band (6,949/MHz). The protected band is regrettably polluted, possibly as a result of intermodulation products generated at the telescope [Margot et al., 2018].

The useful bandwidth of our observations $\Delta f_{\mathrm{tot}} = 309.3$ MHz is computed by taking the operational bandwidth of the GBT L-band receiver (580 MHz) and subtracting the bandwidth of the GBT notch filter (141.2 MHz) and the total bandwidth discarded due to known interferers (Table 4.3; 129.5 MHz).

4.4 Preliminary Signal Injection and Recovery Analysis

A signal injection and recovery analysis consists of injecting artificial signals into the raw data and quantifying the fraction of signals that are properly recovered by the pipeline [e.g., Christiansen et al., 2013]. Although a rigorous injection analysis is beyond the scope of this paper, we performed a preliminary examination by injecting narrowband (2.98 Hz) signals into the dynamic spectra before applying the incoherent dechirping (Section 4.2.4), candidate detection (Section 4.3.1), and Doppler and direction-of-origin filtering (Section 4.3.2) procedures.

4.4.1 Generation and Injection of Artificial Signals

We selected 10 000 starting frequencies from a uniform distribution over the operating region of the GBT L-band receiver (1.15–1.73 GHz), excluding the frequency region of the GBT notch filter (1.2–1.3412 GHz). For each starting frequency, we also randomly selected a frequency drift rate from the discrete set $\{k \times \Delta\dot{f} \ : \ k \in \mathbb{Z}, -510 \leq k \leq 510\}$, with $\Delta\dot{f}$ = 0.0173 Hz s^{-1}. Each signal was randomly assigned to one of the sources and was injected into the first scan of this

source. A corresponding partner signal was injected into the second scan of this source. The starting frequency of the partner signal was obtained by linearly extrapolating the frequency of the signal in the first scan, i.e., by adding the product of the artificial drift rate and the known time difference between the two scans. The drift rate of the partner signal was set equal to the that of the original signal plus an increment randomly chosen from the set $\{-\Delta\dot{f}, 0, \Delta\dot{f}\}$.

We injected half of the signals at our detection threshold (10σ) to test the limits of our pipeline's detection capabilities. The remaining signals were injected at an S/N of 20 to test our sensitivity to stronger signals. A total of 20,000 signals were injected into the April 27, 2018 data. A full list of the injected signal properties is available as supplemental online material.

4.4.2 Recovery and Classification of Injected Signals

After injecting the signals into the dynamic spectra, we applied our candidate detection procedure (Section 4.3.1) and stored the output in a SQL database. Signals were considered properly recovered if their properties matched those of the injected signals within ± 2 Hz in frequency, $\pm\Delta\dot{f}$ in drift rate, and ± 0.1 in S/N. We found that our procedure recovered 18,528 (92.64%) of the injected signals. Outside of the regions with dense RFI described in Table 4.3, our pipeline performs better, with a recovery rate of 97.66%. We observe no significant difference in the recovery rate as a function of drift rate or scan number (Figure 4.5), but we do notice a $\sim$3% increase in the recovery rate for signals with larger S/N.

We found that most of the signals missed by our pipeline were injected in regions of known RFI (Figure 4.5). This pattern is a consequence of two known limitations of our candidate detection procedure. First, our algorithms only detect the signal with highest S/N when two signals intersect in time-frequency space [Pinchuk et al., 2019]. Second, signals with a low prominence superimposed on an elevated noise baseline are discarded (Section 4.3.1). High-density RFI regions such as the ones listed in Table 4.3 are conducive to both of these conditions, thereby reducing the recovery rate. A cursory analysis suggests that $\sim$70–80% of the non-detections are due to the

intersecting condition.

In order to quantify the performance of our Doppler and direction-of-origin filters (Section 4.3.2), we applied our filters to the entire set of detected signals, including the detections resulting from injected signals. To distinguish the performance of these filters from that of our detection algorithm, we removed 414 of the injected signals that were detected in only one scan of a source. Furthermore, we removed 21 signals that were injected with a Doppler drift rate of 0 (i.e., stationary with respect to the observer). Of the remaining 18,093 injected signals, 18,044 (99.73%) were flagged as promising technosignature candidates by our Doppler and direction-of-origin filters.

4.4.3 Performance of Data Processing Pipeline

The preliminary injection and recovery analysis described in this section identified some important limitations of our radio technosignature detection pipeline. Our detection algorithm, which is an improvement over those of Margot et al. [2018] and Pinchuk et al. [2019] (Figure 4.3) and outperforms those of Enriquez et al. [2017] and Price et al. [2020] (Section 4.6.4), experiences degraded performance in regions with dense RFI. In these regions, it is more likely for a technosignature candidate to intersect a strong RFI signal (Figure 4.4), thereby escaping detection by our pipeline. This limitation could be overcome by using the recorded drift rates and starting frequencies of two signals within a scan to determine whether the signals are predicted to intersect each other in the other scan of the source. If an intersection condition were detected, the known signal could be blanked or replaced with noise and a new detection procedure could be run to identify previously undetected signals. In the presence of strong RFI, a fraction of the injected signals escape detection because their prominence is below our detection threshold (i.e., prominence $< f \times$ integrated power, with f=75%). In some situations, a valid technosignature could also be removed if it were detected in a frequency region corresponding to a broadband signal. It may be possible to overcome this limitation in the future by including a comparison of the properties of the narrowband signal (e.g., drift rate, modulation, etc.) to those of the underlying broadband signal.

Our improved Doppler and direction-of-origin filters performed exceptionally well, only mislabeling 49 of the 18,093 injected signals. The signals that were incorrectly flagged were paired with an RFI signal of similar drift rate in a scan of a different source. This issue can be mitigated by expanding the signal matching criteria to include signal properties other than starting frequency and drift rate, such as bandwidth or gain ratio.

The results presented in this section provide important insights into the detection capabilities of our current data processing pipeline. In particular, they demonstrate that our pipeline still misses some of the narrowband signals that it is designed to detect. These results are also useful to identify specific areas in need of improvement.

4.4.4 Limitations of Current Signal Injection and Recovery Analysis

The analysis presented in this section is preliminary because it injects signals into the dynamic spectra and not into the raw data. Therefore, the current implementation does not consider certain data processing steps such as correcting for the bandpass channel response (Section 4.2.3), calculating the noise statistics and normalizing the power spectra to zero mean and unit variance, or applying the incoherent dechirping procedure (Section 4.2.4).

In future work, we will implement the capability to inject signals in the raw data. This improved implementation will allow us to quantify the detection performance of the entire pipeline. We anticipate that it will also be helpful in revealing additional areas for improvement.

4.5 Results

We applied the methods described in Section 4.3 to the data described in Section 4.2.1. We detected a total of 26,631,913 candidate signals over both 2018 and 2019 observation epochs. We used the total integration time of 2.75 hr and useful bandwidth of $\Delta f_{\text{tot}} = 309.3$ MHz to compute a signal detection count per unit bandwidth per unit integration time. In BL parlance, our detections are

referred to as "hits" [Enriquez et al., 2017, Price et al., 2020] and the hit rate density of this search is 3.1 $\times 10^{-2}$ hits per hour per hertz. In comparison, the L-band component of Price et al. [2020]'s search with the same telescope and S/N threshold resulted in 37.14 million hits in 506.5 hr over a useful bandwidth of 660 MHz, or a hit rate density of 1.1 $\times 10^{-4}$ hits per hour per hertz, almost 300 times smaller than ours. We discuss possible reasons for this large differential in Section 4.6.4.

A complete table of signal properties of the detected candidates is available online.[1] Our Doppler and direction-of-origin filters flagged 26,588,893 (99.84%) signals as anthropogenic RFI. A majority of the remaining 43,020 signals were detected within operating regions of known interferers (Table 4.3). Candidate signals remaining within these frequency regions were attributed to RFI and removed from consideration.

The remaining 4 539 signals were deemed most promising technosignature candidates. Visual inspection of all of these candidates revealed that they are attributable to RFI. Figure 4.6 shows an example of a promising signal that was ultimately attributed to RFI.

The vast majority of the most promising candidates were eliminated because they were detected in multiple directions on the sky. These signals escaped automatic RFI classification by our filters for one or more of the following reasons, which are generally similar to the "Categories" described by Pinchuk et al. [2019, Section 4]:

1. The S/N values of corresponding signals in scans of other sources were below the detection threshold of 10. This difficulty could perhaps be circumvented in the future by conducting an additional search for lower S/N signals at nearby frequencies.

2. The drift rate of the signal differed from those of corresponding signals in scans of other sources by more than our allowed tolerance ($\pm\Delta\dot{f} = \pm 0.0173$ Hz s^{-1}).

3. The signal was not detected in scans of other sources because it intersected another signal of a higher S/N.

[1] https://doi.org/10.5068/D1937J

4. The signal bandwidth exceeded 10 Hz, making it difficult to accurately determine a drift rate and therefore link the signal with corresponding signals in scans of other sources.

All of these difficulties could likely be overcome by a direction-of-origin filter that examines the time-frequency data directly instead of relying on estimated signal properties such as starting frequency and drift rate. We are in the process of implementing machine learning tools for this purpose.

Because automatic classification and visual inspection attributed all of our candidate signals to RFI, we did not detect a technosignature in this sample. We are preserving the raw data in order to enable reprocessing of the data with improved algorithms in the future, including searches for additional types of technosignatures.

4.6 Discussion

4.6.1 Dechirping Efficiency

Over sufficiently short ($\sim$5 min) scan durations, monochromatic signals emitted on extraterrestrial platforms are well approximated by linear chirp waveforms ($f(t) = f(t_0) + \dot{f}(t - t_0)$). Most radio technosignature detection algorithms rely on incoherent dechirping, i.e., incoherent sums of power spectra, to integrate the signal power over the scan duration (Section 4.2.4). In the context of incoherent sums, the magnitude of the maximum drift rate that can be considered without loss in sensitivity is given by

$$\dot{f}_{\max} = \frac{\Delta f}{\Delta T}, \tag{4.5}$$

where Δf is the adopted spectral resolution and ΔT is the accumulation time corresponding to one row in the dynamic spectra. If the drift rate of a signal exceeds this maximum drift rate ($\dot{f} > \dot{f}_{\max}$), the signal frequency drift exceeds Δf during ΔT, and power is smeared over multiple frequency channels, resulting in reduced sensitivity.

In this work ($\Delta f = 2.98$ Hz; $\Delta T = 1/\Delta f = 0.34$ s), the maximum sensitivity can be

obtained up to frequency drift rates of $\dot{f}_{\mathrm{max,UCLA}} = 8.88$ Hz s^{-1}. BL investigators Enriquez et al. [2017] and Price et al. [2020] used $\Delta f = 2.79$ Hz and $\Delta T = 51/\Delta f = 18.25$ s, which yields $\dot{f}_{\mathrm{max,BL}} = 0.15$ Hz s^{-1}. However, these authors conducted searches for signals with drift rates larger than 0.15 Hz s^{-1}, resulting in reduced sensitivity for >90% of the drift rates that they considered. For instance, at the largest drift rate considered by Price et al. [2020], the frequency drifts by 4 Hz s^{-1} × 18.25 s = 73 Hz (26 channels) during ΔT, and only ∼4% of the signal power is recovered in each frequency channel. We express this loss of signal power with a detection efficiency in the range 0–100% and refer to it as a dechirping efficiency.

To confirm the performance of the data processing pipelines, we conducted numerical experiments with both our algorithms and BL's turboSETI package [Enriquez et al., 2017]. For the purpose of these simulations, we created noise-free, constant-power dynamic spectra of linear chirp waveforms with the frequency and time resolutions appropriate for the UCLA and BL searches. By considering only integral pixel locations, we simulated frequency drift rates that are exact multiples of the elemental drift rates considered by our respective tree algorithms (0.0173 Hz s^{-1} for the UCLA searches, 0.0096 Hz s^{-1} for the BL searches). We ran the respective tree algorithms on the simulated spectra and recorded the power recovered at each drift rate as a function of total signal power (Figure 4.7, Left). The experiments show that, at the nominal frequency resolutions of ∼3 Hz, dechirping efficiencies of 100% are possible in our and other searches with $\dot{f} \leq \dot{f}_{\mathrm{max}}$, whereas dechirping efficiencies rapidly degrade to values as low as 4% in the BL searches with $\dot{f} > \dot{f}_{\mathrm{max,BL}}$

In this experiment, a perfect algorithm would recover 100% of the signal power, as long as $\dot{f} \leq \dot{f}_{\mathrm{max}}$. The tree algorithm (Section 4.2.4) is not perfect in that it reuses pre-computed sums to achieve $N \log N$ computational cost. As a result, the tree algorithm shifts every spectrum by an amount that is not always optimal. In other words, it is unable to perfectly dechirp most linear chirp waveforms. In our simulations of the UCLA pipeline, we do observe 100% of the power recovered for several drift rates (Figure 4.7, Left). On average, the pipeline recovers 72.4% ± 6.8% of the signal power. In the worst-case scenario, the fraction of power recovered is 60%. The

tree algorithm's dechirped waveform of this worst-case scenario reveals that 60% of the frequency bins are shifted to the correct locations and 40% are shifted to incorrect locations (Figure 4.7, Right). We quantified the dechirping efficiencies associated with the use of the tree algorithm for a variety of array dimensions (Table 4.4).

We computed a rough estimate of the mean dechirping efficiency in the search of Price et al. [2020] for the nominal frequency resolution of $\sim$3 Hz and a uniform distribution of candidate signals as a function of drift rate. We assumed a generous 100% efficiency between 0 and 0.15 Hz s^{-1} and the 1/x trend observed in Figure 4.7 between 0.15 Hz s^{-1} and 4 Hz s^{-1}. We found a mean efficiency of 16.5%. A weighted mean of the efficiency based on the exact distribution of signals as a function of drift rate would provide a more accurate and likely larger value.

We describe two alternate, partial solutions to the loss of sensitivity sustained during incoherent dechirping. The first is to reduce the frequency resolution of the dynamic spectra, thereby increasing the range of drift rates that can be explored without spreading power across multiple channels [e.g., Siemion et al., 2013]. However, this solution still results in loss of sensitivity. For narrowband signals, each doubling of the frequency resolution results in a $\sqrt{2}$ decrease in sensitivity. To reach the maximum drift rates of ± 4 Hz s^{-1}considered by Price et al. [2020], one would have to apply 4–5 doublings, resulting in frequency resolutions of 45–90 Hz and sensitivity to narrowband signals of 18–25% of the nominal value. Another, related approach, would be to use a drift-rate-dependent boxcar average of the integrated spectra to recover the power that has been spread over multiple channels, e.g., by averaging 26 channels at the maximum drift rates of ± 4 Hz s^{-1}considered by Price et al. [2020]. Doing so would degrade the frequency resolution to values up to 73 Hz and the sensitivity to narrowband signals to 20% of the nominal value.

4.6.2 Extreme Drift Rates

In a recent study[2] of the expected drift rates of a large class of bodies, including exoplanets with highly eccentric orbits and small semimajor axes, Sheikh et al. [2019] recommended searching drift rates as large as $\dot{f}/f_{obs} = 200$ nHz At the center frequency of our observations (1.5 GHz), this corresponds to a drift rate of 300 Hz s^{-1}. Our data archival policy (Section 4.6.3) would enable reprocessing of the data with parameters that are more conducive to large drift rates. For example, we could reprocess our data with Fourier transforms of length 2^{17}. This choice would increase our frequency resolution 8-fold to 24 Hz and would allow us to search for drift rates up to $\sim$ 570 Hz s^{-1} without incurring any sensitivity loss due to signal smearing over multiple frequency channels. In contrast, BL archive products include dynamic spectra but do not include most of the raw voltage data [Lebofsky et al., 2019, Enriquez et al., 2017, Price et al., 2020], making it impractical to conduct a search with archival products at drift rates larger than $\sim$1 Hz s^{-1} with adequate sensitivity (Figure 4.7).

4.6.3 Data Requantization and Preservation

Our choice of data recording parameters is largely driven by our dedication to preserve the raw voltage data recorded during our observations. We prefer to archive the raw data as opposed to derived data products such as dynamic power spectra, for four reasons. First, the raw 2-bit data require less storage space than the 32-bit dynamic spectra. Second, the dynamic power spectra can be easily regenerated from the raw data, but the reverse is not true, because phase information is lost in the process of computing power spectra. Third, there are large penalties associated with preserving incoherent averages of individual power spectra. Enriquez et al. [2017] and Price et al. [2020] average 51 consecutive spectra to keep the archival volume manageable, which degrades the sensitivity of the search by factors of up to $\sim$25 (Section 4.6.1) and the time resolution by a

[2] Sheikh et al. [2019] incorrectly delineated the search parameters of Enriquez et al. [2017] in their Figure 1. The maximum frequency excursion considered in that search is 600 Hz, not 12,000 Hz.

factor of 51 (Figure 4.8). As a result, the BL dynamic spectra would not be useful in confirming or interpreting a signal with 1 Hz modulation, for instance. Fourth, the only way to preserve the ability to conduct novel or improved data analysis with maximum sensitivity and resolution is to preserve the raw data. However, there are penalties associated with storing raw data in 2-bit format as opposed to 8-bit format [e.g., Price et al., 2020].

For this work and previous analyses [Margot et al., 2018, Pinchuk et al., 2019], we selected a data-taking mode that yields 2-bit raw voltage data after requantization with an optimal four-level sampler [Kogan, 1998]. The quantization efficiency, which is the ratio of signal power that is observed with the optimal four-level sampler to the power that would be obtained with no quantization loss, is 0.8825. Price et al. [2020] noted that a consequence of this requantization is that the S/N threshold used in this work (10) would need to be lowered by approximately 12% to detect the same number of candidate signals as 8-bit quantized data. While we agree with this statement, the S/N threshold of radio technosignature searches is somewhat arbitrary, and our choice compares favorably to that of other surveys (Table 4.5). Should the need ever arises to detect weaker signals, we would simply re-analyze our data with a lower S/N threshold. In addition, the sensitivity enabled by our decision to minimize the accumulation time when computing dynamic spectra (Section 4.6.1) offsets the losses due to quantization efficiency compared to pipelines with longer accumulation times. Specifically, if we apply the 0.8825 quantization efficiency to the results illustrated in Figure 4.7, we find that our overall sensitivity surpasses BL's sensitivity for any drift rate larger than 0.153 Hz s^{-1} and surpasses it by a factor of at least 5 for any drift rate larger than 1.11 Hz s^{-1} .

4.6.4 Candidate Signal Detection Count

Our results indicate a hit rate density of 3.1 $\times 10^{-2}$ hits per hour per hertz whereas Price et al. [2020] obtained a considerably lower value of 1.1 $\times 10^{-4}$ hits per hour per hertz with the same telescope and S/N threshold (Section 4.5). We investigate possible causes for this factor of $\sim$300 difference. First, our observing cadence involves two scans of 150 s each per source, whereas

Price et al. [2020] used three scans of 300 s each per source. The difference in integration time could perhaps be invoked to explain a factor of up to 3 difference in hit rate density, although a larger number of signals ought to be detectable with BL's longer scan durations. Second, our usable frequency range extends over 309.3 MHz, whereas Price et al. [2020] used a superset of that range extending over 660 MHz. A non-uniform distribution of dense RFI across the spectrum could perhaps be invoked to explain a factor of up to $\sim$2 difference in hit rate density. Third, we examine a range of drift rates that is twice as large as the range used by Price et al. [2020], which may explain a factor of $\sim$2 difference in hit rate density if the distribution of hits as a function of drift rate is roughly uniform. These small factors cannot explain the two orders of magnitude difference in hit rate density, which must be related to more fundamental effects. We surmise that the two most important factors are the difference in the effective sensitivities of our searches due to different dechirping efficiencies (Section 4.6.1) and the algorithmic difference in the identification of candidate signals or hits.

As detailed by Pinchuk et al. [2019], the candidate signal detection procedures used in several previous radio technosignature searches [e.g., Siemion et al., 2013, Enriquez et al., 2017, Margot et al., 2018, Price et al., 2020] unnecessarily remove kHz-wide regions of frequency space around every signal detection. This practice complicates attempts to place upper limits on the existence of technosignatures, because the algorithms discard many signals that are legitimate technosignature candidates. In addition, this practice leads to slight overestimates of search metrics, such as the Drake Figure of Merit (DFM) [Pinchuk et al., 2019]. Here, we quantify the number of signals that are unnecessarily discarded by algorithms that remove $\sim$kHz wide frequency regions around every detection.

To perform this comparison, we used the database of signals detected during the April 27, 2018 observations and we replicated the procedure described by Enriquez et al. [2017] and Price et al. [2020]. The "blanking" procedure used by Price et al. [2020] specifies "Only the signal with the highest S/N within a window ... $\pm$ 600 Hz is recorded as a hit." To replicate this step, we sorted the signals detected in each scan in decreasing order of S/N and iterated over the sorted lists. At every

iteration, we kept the signal with the largest remaining S/N value and eliminated all other signals within $\pm$ 600 Hz. The next step described by Price et al. [2020] combines hits that fall within a certain frequency range into groups as long as the signal is detected in every scan of the source. We replicated this step by grouping signals that were present in both scans of each source according to Price et al. [2020]'s prescription for frequency range. The third step of the procedure described by Price et al. [2020] reads: "Additionally, any set of hits for which there is at least one hit in the OFF observations within $\pm$ 600 Hz of the hit frequency from the first ON observation would be discarded." This elimination seems wasteful because the presence of OFF-scan signals with drift rates that are unrelated to the ON-scan drift rate results in elimination. To replicate this step, we removed all groups of signals for which one or both of the two OFF-scans contained an unrelated signal within $\pm$ 600 Hz of the detection in the first ON-scan. To determine whether signals were unrelated, we placed the following condition on drift rate:

$$|\dot{f} - \dot{f}_0| > 2\Delta\dot{f}, \tag{4.6}$$

where $\dot{f}$ is the drift rate of the OFF-scan signal, $\dot{f}_0$ is the drift rate of the ON-scan signal, and $\Delta\dot{f} = 0.0173$ Hz s^{-1} is the drift rate resolution. Our direction-of-origin filters (Section 4.3.2) also remove signals if they are found in multiple directions on the sky, but only in conjunction with careful analysis of the drift rates of both signals. Specifically, our filters only remove the two signals if their drift rates are within a tolerance of $2\Delta\dot{f}$. For the purpose of this blanking analysis, we kept signals that satisfy this criterion because both pipelines remove them during subsequent filtering. We found that our pipeline detected 10,113,551 signals whereas our pipeline with a blanking algorithm modeled after the descriptions given by Enriquez et al. [2017] and Price et al. [2020] detected only 1,054,144 signals. In other words, our pipeline detects $\sim$10 times as many signals as the BL-like pipeline over the same frequency range, with a corresponding increase in hit rate density.

To summarize, we found that our algorithmic approach to signal identification explains the largest fraction of the factor of $\sim$300 difference in hit rate density between our and the BL searches, likely followed by our better overall sensitivity for $>$90% of the frequency drift rates examined by

the BL pipeline (Section 4.6.1), likely followed by our shorter integration times and consideration of a wider range of drift rates. It is also possible that the limited dynamic range of our 2-bit voltage data makes our search susceptible to spurious detections at harmonics of strong RFI signals (Danny Price, pers. comm.). We are planning to quantify the importance of this effect in the future.

The differential in hit rate density has implications for the validity of existence limit estimates and figure of merit calculations described by Enriquez et al. [2017] and Price et al. [2020].

4.6.5 Existence Limits

We describe three issues that affect recent claims about the prevalence of transmitters in the Galaxy [Enriquez et al., 2017, Price et al., 2020, Wlodarczyk-Sroka et al., 2020].

First, the range of Doppler drift rates considered in these searches has been limited (± 2 Hz $s^{-1}$and ± 4 Hz s^{-1}), whereas transmitters may be located in a variety of settings with line-of-sight accelerations that would only be detectable at larger drift rates [e.g., Sheikh et al., 2019].

Second, these claims invoke transmitters with certain EIRP values that are calculated on the basis of the nominal sensitivity to non-drifting signals. However, the sensitivity to signals drifting in frequency is demonstrably degraded (Section 4.6.1) with the incoherent dechirping method used in these searches. The published EIRP values could be erroneous by factors of up to 25 for these searches, depending on the drift rate of the putative signal.

Third, our preliminary candidate signal injection and recovery analysis (Section 4.4) reinforces the concerns voiced by Margot et al. [2018] and Pinchuk et al. [2019] about Enriquez et al. [2017]'s claims. Pinchuk et al. [2019] argued that an injection and recovery analysis would demonstrate that a fraction of detectable and legitimate signals are not identified by existing pipelines, thereby requiring corrections to the claims. We have shown that our current pipeline misses $\sim 7\%$ of the signals injected into the dynamic spectra (Section 4.4). We surmise that the BL pipelines used by Enriquez et al. [2017] and Price et al. [2020] miss a substantially larger fraction of signals that they are meant to detect because of reduced sensitivity (Section 4.6.1), time resolution (Section 4.6.3),

and detection counts (Section 4.6.4) compared to our pipeline.

In light of these issues, published claims about the prevalence of transmitters in the Galaxy [e.g., Enriquez et al., 2017, Price et al., 2020, Wlodarczyk-Sroka et al., 2020] almost certainly need revision. As mentioned in Section 4.4.4, we are planning improvements to our signal injection and recovery analysis. Until this refined analysis is complete, we will not be in a position to make reliable inferences about the prevalence of radio beacons in the Galaxy.

4.6.6 Drake Figure of Merit

The Drake Figure of Merit [Drake, 1984] is a metric that can be used to compare some of the dimensions of the parameter space examined by different radio technosignature searches. It is expressed as

$$\mathrm{DFM} = \frac{\Delta f_{\mathrm{tot}} \Omega}{F_{\mathrm{det}}^{3/2}}, \tag{4.7}$$

where Δf_{tot} is the total bandwidth observed, Ω is the total angular sky coverage, and F_{det} is the minimum detectable flux. Assuming unit quantization and dechirping efficiencies, our search with S/N threshold of 10 is sensitive to sources with flux densities of 10 Jy and above [Margot et al., 2018]. For consistency with earlier calculations [Enriquez et al., 2017, Price et al., 2020], we have assumed that the bandwidth of the transmitted signal is 1 Hz, resulting in a minimum detectable flux $F_{\mathrm{det}} = 10^{-25}$ W m^{-2}. The sky coverage of this search is $\Omega = 31 \times 0.015$ deg^2 $= 0.465$ deg^2, i.e., 11 ppm of the entire sky. The useful bandwidth is $\Delta f_{\mathrm{tot}} = 309.3$ MHz (Section 4.3.3). We used these parameters to calculate the DFM associated with this search and found $\mathrm{DFM} = 1.11 \times 10^{32}$, where we have used units of GHz m^3 W$^{-3/2}$ for compatibility with Horowitz and Sagan [1993]. We reanalyzed our 2016 and 2017 data sets (Section 4.6.8) and recomputed DFM values of 5.00×10^{31} and 4.71×10^{31} for these data sets, respectively, with an aggregate DFM for our 2016–2019 searches of 2.08×10^{32}. However, we regard these values and all previously published DFM values with skepticism.

The DFM values published in recent works do not provide accurate estimates of search volume

or performance for a few reasons. First, the DFM relies on minimum detectable flux, but authors have ignored factors that can tremendously affect overall search sensitivity, such as quantization efficiency ($\sim$88% for 2-bit sampling) or dechirping efficiency (60–100% with the tree algorithm and the parameters of this search, and as low as $\sim$4% in the recent BL search described by Price et al. [2020]). Second, it does not take account of the range of drift rates considered in a search, which is clearly an important dimension of the search volume. Third, it ignores the quality of the signal detection algorithms, such that two surveys may have the same DFM even though their data processing pipelines detect substantially different numbers of signals (e.g., the blanking of kHz-wide regions of frequency space described in Section 4.6.4). For these reasons, we believe that DFM values calculated by authors of recent searches, including our own, are questionable indicators of actual search volume or performance. Horowitz and Sagan [1993] expressed additional concerns, stating that the DFM "probably does justice to none of the searches; it is a measure of the odds of success, assuming a homogeneous and isotropic distribution of civilizations transmitting weak signals at random frequencies."

In Section 4.6.1, we showed that the dechirping efficiency degrades rapidly for frequency drift rates larger than $\dot{f}_{\max}$ (Figure 4.7). As a result, the minimum detectable flux for non-drifting signals, which has been used by Enriquez et al. [2017] and Price et al. [2020] in their DFM estimates (Equation 4.7), is not representative of the minimum detectable flux of signals with $>$90% of the drift rates that they considered, which can be up to 25 times larger. Given the presence of this flux to the 3/2 power in the denominator of the DFM, we believe that the DFMs of these searches have been inadvertently but considerably overestimated. Other figures of merit, such as Enriquez et al. [2017]'s 'Continuous Waveform Transmitter Rate Figure of Merit' (CWTFM), are also affected by this problem.

We can use our estimates of the mean dechirping efficiencies to quantify plausible errors in DFM estimates. In Section 4.6.1, we computed a rough estimate of 16.5% for the mean efficiency of the BL search conducted by Price et al. [2020], suggesting that the DFM of their search has been overestimated by a factor of $\sim$15. This value may be revised down once a more accurate estimate

of the mean dechirping efficiency becomes available. For the UCLA searches conducted between 2016 and 2019, the mean dechirping efficiency is 72.4% and the quantization efficiency is 88.25%, resulting in an overall efficiency of 64% and DFM overestimation by a factor of $\sim$2.

4.6.7 Other Estimates of Search Volume

The range of drift rates considered in a search program obviously affects the probability of success of detecting a technosignature. For instance, a search restricted to drift rates smaller than $\dot{f}_{\mathrm{max,BL}} = 0.15$ Hz s^{-1} could fail to detect the signal from an emitter on an Earth-like planet. The frequency drift rate dimension of the search volume does not appear to have been fully appreciated in the literature. It is distinct from the "modulation" dimension described by Tarter et al. [2010], who focused on "complex ... broadband signals". It also appears to be distinct from the "modulation" dimension of Wright et al. [2018], who contemplated drift rates on the order of the "Earth's barycentric acceleration", i.e., 0.03 Hz s^{-1} at the center frequency of our observations. It is also absent from the CWTFM used by Enriquez et al. [2017], Price et al. [2020], Wlodarczyk-Sroka et al. [2020]. The development of an improved figure of merit for radio technosignature searches is beyond the scope of this work. However, we recommend that improved figures of merit include the range of line-of-sight accelerations between emitter and receiver as a dimension of the search volume as well as explicit guidelines regarding the treatment of quantization and dechirping efficiencies.

4.6.8 Re-analysis of 2016 and 2017 Data

Margot et al. [2018] presented the results of a search for technosignatures around 14 planetary systems in the Kepler field conducted on April 15, 2016, 16:00 - 18:00 universal time (UT) with the GBT. Pinchuk et al. [2019] presented the results of a similar search conducted on May 4, 2017, 15:00 - 17:00 universal time (UT) that included 10 planetary systems in the Kepler field but also included scans of TRAPPIST-1 and LHS 1140.

We reprocessed these data with our updated algorithms and detected a total of 13,750,469 candidate signals over the 2016 and 2017 epochs of observation. Tables of signal properties of the detected candidates are available online for both the 2016 [Margot et al., 2020a] and 2017 [Margot et al., 2020b] data sets. We found that 13,696,445 (99.61%) signals were automatically flagged as anthropogenic RFI and 54,024 signals were labeled as promising. Candidate signals found within operating regions of known interferers (Table 4.3) were attributed to RFI and removed from consideration. Visual inspection of all of the remaining 4,257 candidate signals revealed that they are attributable to RFI. With this improved analysis, we confirm the initial results that no technosignatures were detected in the data obtained in 2016 [Margot et al., 2018] and 2017 [Pinchuk et al., 2019].

4.7 Conclusions

We described the results of a search for technosignatures that used 4 hours of GBT time in 2018 and 2019. We identified 26,631,913 candidate signals, 99.84% of which were automatically classified as RFI by rejection filters. Of the signals that remained, $4\,539$ were found outside of known RFI frequency bands and were visually inspected. All of these were attributable to RFI and none were identified as a technosignature.

We presented significant improvements to our signal detection and direction-of-origin filter algorithms. We tested the signal recovery of the updated procedures with a preliminary signal injection and recovery analysis, which showed that our pipeline detects $\sim 93\%$ of the injected signals overall. This recovery rate increases to $\sim 98\%$ outside of known RFI frequency bands. In addition, our pipeline correctly identified 99.73% of the artificial signals as technosignatures. This signal injection and recovery analysis provides an important tool for quantifying the signal recovery rate of a radio technosignature data processing pipeline. Planned improvements to this tool will further illuminate imperfections in our and other groups' pipelines and point to additional areas for improvement.

Our search represents only a modest fraction of the BL searches described by Enriquez et al. [2017] and Price et al. [2020] in terms of number of targets and data volume. However, our search strategy has advantages compared to these searches in terms of sensitivity (up to 25 times better), frequency drift rate coverage (2–4 times larger), and signal detection count per unit bandwidth per unit integration time ($\sim$300 times larger).

We described limitations of recent Drake Figure of Merit (DFM) calculations in assessing the probability of success of different search programs. These calculations have ignored important factors such as quantization and dechirping efficiencies. In addition, the DFM does not take account of the range of drift rates considered in a search nor the quality of signal detection algorithms. As a result, we suggest that recent DFM calculations are questionable indicators of actual search volume or performance. We recommend that improved metrics include the range of line-of-sight accelerations between emitter and receiver as a dimension of the search volume as well as explicit guidelines regarding the treatment of quantization and dechirping efficiencies.

Table 4.1: Target host stars listed in order of observation (2018). Successive pairs are separated by a blank line. Spectral types and parallax measurements were obtained from the SIMBAD database [Wenger et al., 2000]. Distances in light years (ly) were calculated from the parallax measurements. The Modified Julian Date (MJD) refers to the beginning of the first scan.

Host Star	Spectral Type	Parallax (mas)	Distance (ly)	MJD of Scan 1
April 27, 2018 20:00 – 22:00 UT				
TYC 1863-858-1	G0V	1.9547 ± 0.036	1669 ± 31	58235.84503472
TYC 1868-281-1	G2V	3.8622 ± 0.046	844 ± 10	58235.84737269
HD 249936	G2V	1.9515 ± 0.0412	1671 ± 35	58235.85430556
TYC 1864-1748-1	G2V	3.0350 ± 0.043	1075 ± 15	58235.85671296
HIP 28216	G2V	1.2479 ± 0.0909	2614 ± 190	58235.86393519
HD 252080	G5V	5.7621 ± 0.0511	566 ± 5	58235.86640046
HD 251551	G2V	4.5105 ± 0.0755	723 ± 12	58235.87356481
HD 252993	G0V	6.9544 ± 0.0401	469 ± 3	58235.87591435
TYC 742-1679-1	G5V	8.4009 ± 0.0360	388 ± 2	58235.88324074
HD 255705	G5V	6.8001 ± 0.0570	480 ± 4	58235.88562500
HD 254085	G0V	6.5338 ± 0.0967	499 ± 7	58235.89270833
HD 256380	G8V	2.3058 ± 0.0398	1415 ± 24	58235.89505787
TYC 739-1501-1	G2V	...	...	58235.90204861
HD 256736	G2V	6.1808 ± 0.0802	528 ± 7	58235.90435185
TYC 739-1210-1[a]	G5V	9.9809 ± 0.0424	327 ± 1	58235.91119213

(a) The source paired with TYC 739-1210-1 was observed only once and not analyzed.

Table 4.2: Target host stars listed in order of observation (2019). Successive pairs are separated by a blank line. Spectral types and parallax measurements were obtained from the SIMBAD database [Wenger et al., 2000]. Distances in light years (ly) were calculated from the parallax measurements. The Modified Julian Date (MJD) refers to the beginning of the first scan.

Host Star	Spectral Type	Parallax (mas)	Distance (ly)	MJD of Scan 1
April 26, 2019 22:00 – 24:00 UT				
TYC 148-515-1	G5V	5.4947 ± 0.0416	594 ± 4	58599.92175926
CoRoT 102810550	G2V	1.1333 ± 0.0247	2878 ± 63	58599.92392361
CoRoT 102830606	G2V	2.2193 ± 0.0357	1470 ± 24	58599.93030093
TYC 149-362-1	G5V	1.1606 ± 0.0581	2810 ± 141	58599.93254630
TYC 149-532-1	G2V	7.2260 ± 0.0380	451 ± 2	58599.93917824
CoRoT 102827664	G4V	2.3394 ± 0.0274	1394 ± 16	58599.94144676
CoRoT 102936925	G4V	1.0454 ± 0.0223	3120 ± 67	58599.94826389
CoRoT 110695685	G4V	1.5743 ± 0.0541	2072 ± 71	58599.95049769
CoRoT 110864307	G2V	1.5985 ± 0.0442	2040 ± 56	58599.95706019
CoRoT 102951397	G2V	1.1909 ± 0.0249	2739 ± 57	58599.95931713
CoRoT 102963038	G3V	0.3134 ± 0.0256	10407 ± 850	58599.96596065
HD 50388	G8V	7.3465 ± 0.0598	444 ± 4	58599.96820602
TYC 4805-3328-1	G5V	2.5383 ± 0.0455	1285 ± 23	58599.97480324
CoRoT 110777727	G1V	1.5407 ± 0.0457	2117 ± 63	58599.97699074
CoRoT 110776963	G4V	2.5207 ± 0.0436	1294 ± 22	58599.98373843
TYC 4814-248-1	G2V	2.9668 ± 0.0426	1099 ± 16	58599.98597222

Table 4.3: Definitions of operating regions of known anthropogenic interferers and associated signal counts. The column labeled "Post-filter Count" lists the number of signals remaining after application of our Doppler and direction-of-origin filters. The time-frequency structure of the RFI labeled as "ARSR products" is similar to that described by Siemion et al. [2013], Margot et al. [2018], and Pinchuk et al. [2019]. These products are likely intermodulation products of Air Route Surveillance Radars (ARSR).

Frequency Band (MHz)	Detection Count	Post-filter Count	Identification
1155.99 – 1196.91	11,937,074	15,034	GPS L5
1192.02 – 1212.48	135,769	276	GLONASS L3
1422.32 – 1429.99	190,530	2 945	ARSR products
1525 – 1559	8,258,612	341	Satellite downlinks
1554.96 – 1595.88	5,016,951	19,621	GPS L1
1592.95 – 1610.48	933,813	3,569	GLONASS L1

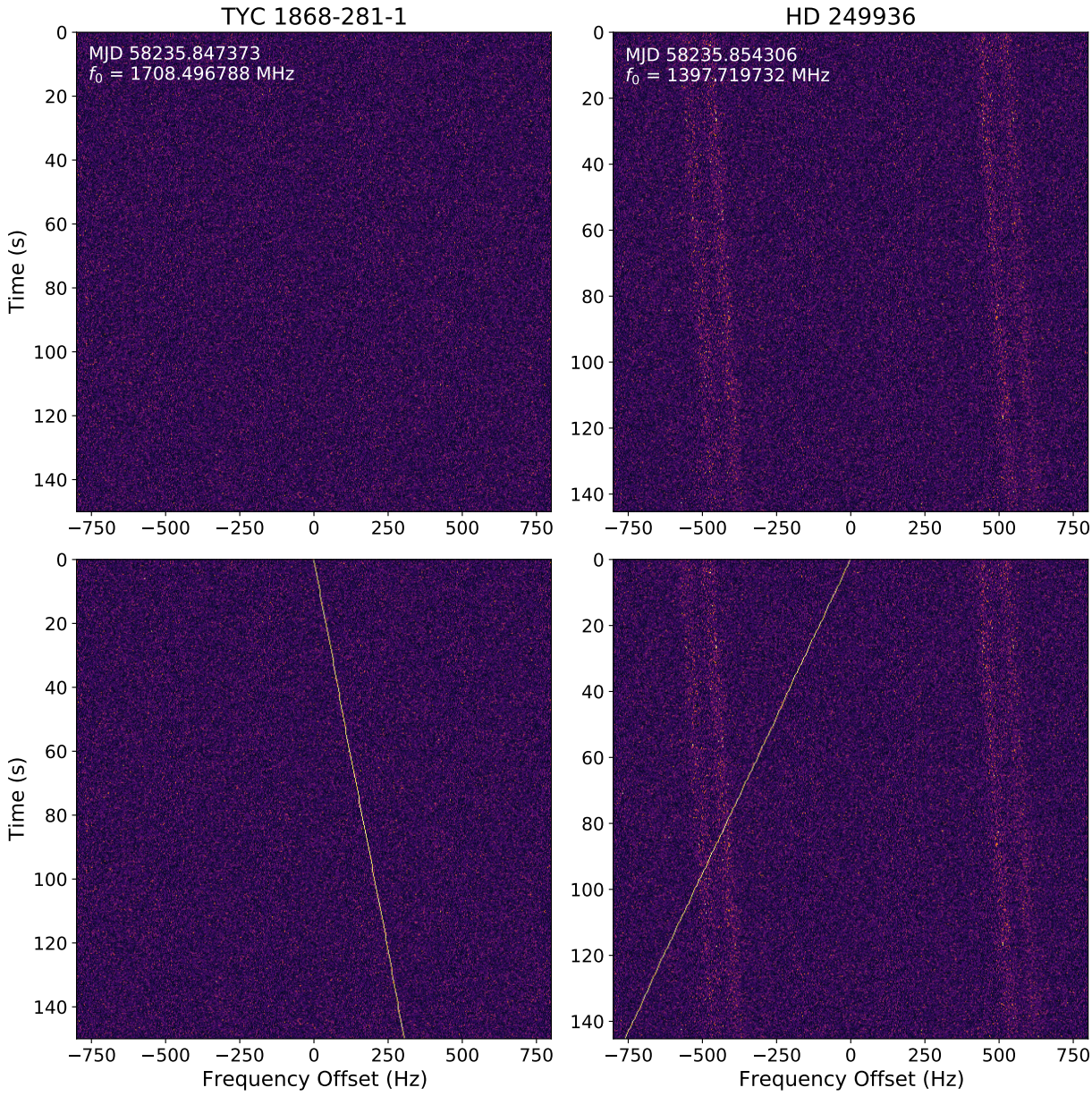

Figure 4.4: (Top) Time-Frequency diagram before signal injection. (Bottom) Time-Frequency diagram after signal injection. The injected signal S/N was increased twentyfold to facilitate visualization. The bottom left panel shows a signal that was successfully recovered by our data processing pipeline. The injected signal in the bottom-right panel crosses a stronger RFI signal and was missed by our detection algorithm.

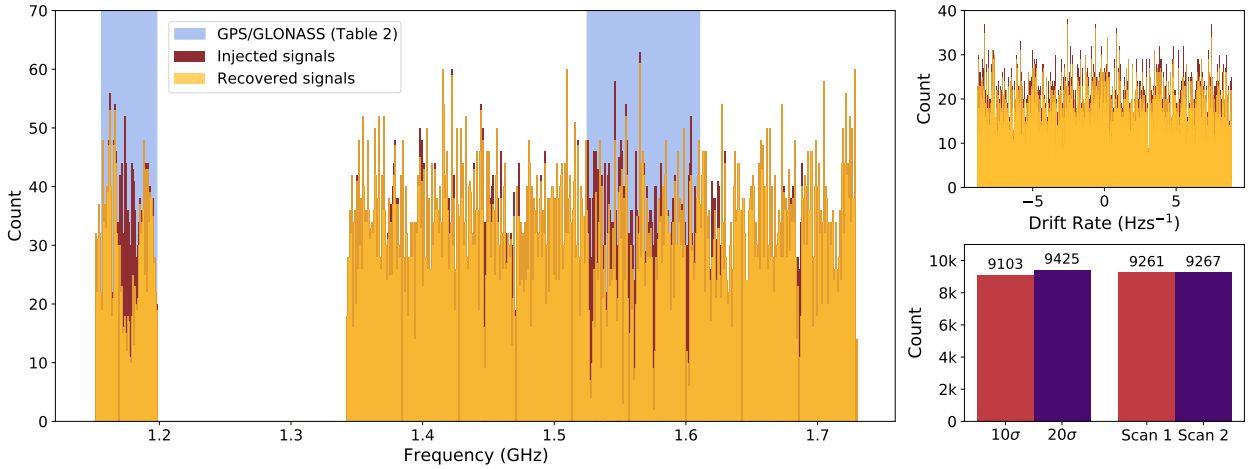

Figure 4.5: (Left) Frequency distribution of injected and recovered signals. The number of signals recovered within known RFI regions (such as GPS or GLONASS) is substantially lower than in other regions. (Top right) Drift rate distribution of injected and recovered signals. (Bottom right) Signals recovery counts as a function of S/N and scan number. We observe no significant difference in the recovery rate as a function of drift rate or scan number, but we do notice a $\sim$3% increase in the recovery rate of signals with larger S/N.

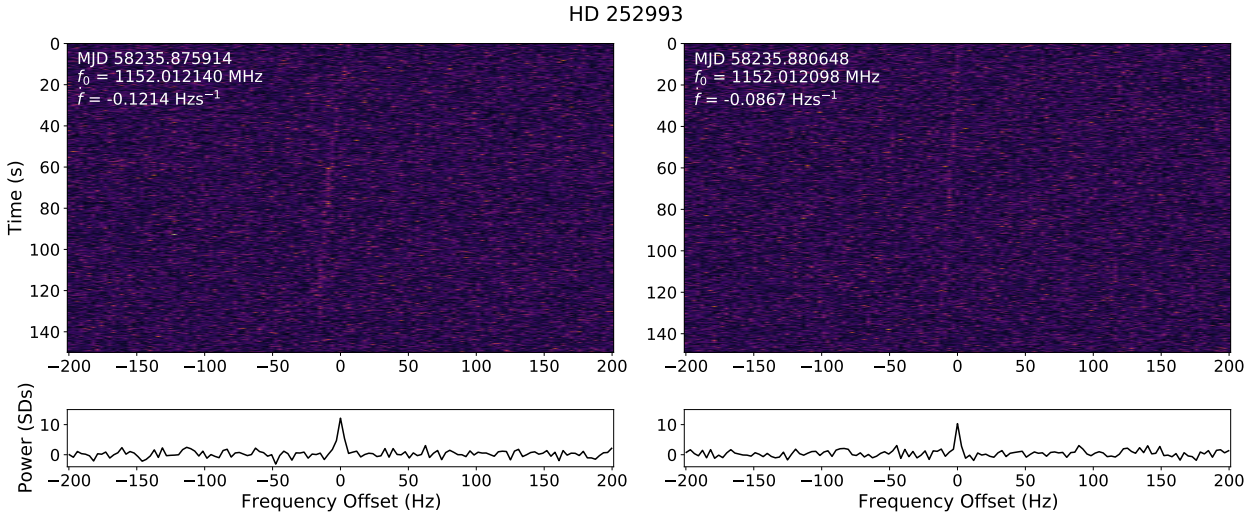

Figure 4.6: Dynamic spectra (top) and integrated power spectrum (bottom) of a final candidate signal that appears in scan 1 (left) and scan 2 (right) of HD 252993. Although this signal exhibits many of the desirable properties of a technosignature (e.g., narrowband, non-zero Doppler drift rate, persistence), it was ultimately rejected because it was visually confirmed to appear in multiple directions on the sky.

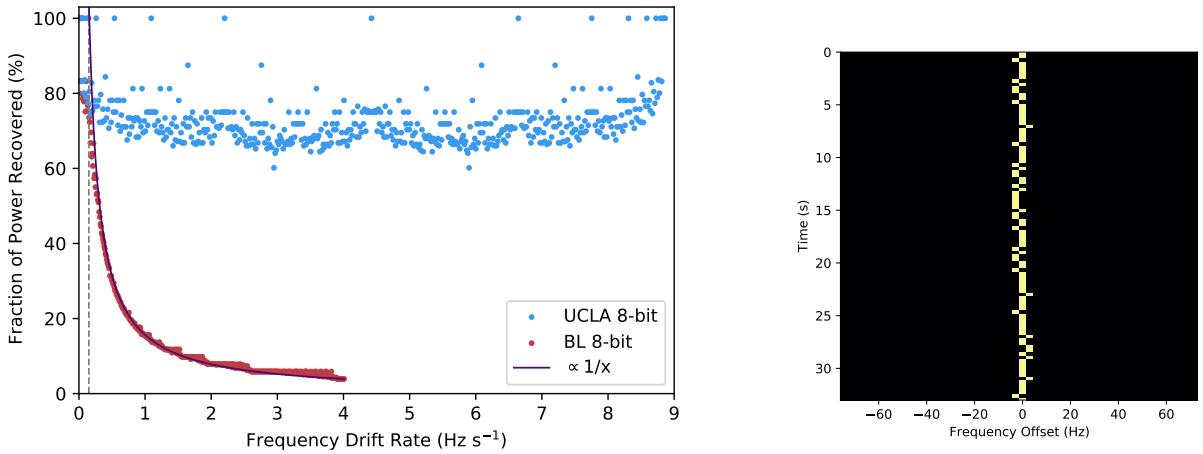

Figure 4.7: (Left) Dechirping efficiencies of the UCLA (blue) and BL (red) data processing pipelines as a function of Doppler frequency drift rate at the nominal frequency resolutions of $\sim$3 Hz. Our choices of data-taking and data-processing parameters result in a fairly uniform efficiency (72.4% $\pm$ 6.8%) across the full range of drift rates considered, with values below 100% due to imperfections of the tree algorithm (see text). The BL choices result in considerably reduced detection efficiency beyond $\dot{f}_{\mathrm{max,BL}} = 0.15$ Hz s^{-1}(dashed vertical line), with values as low as 4% due to smearing of the signal power across multiple frequency bins. The performance at frequencies beyond $\dot{f}_{\mathrm{max}}$ is well approximated by a $1/x$ function (purple line), consistent with the inverse bandwidth dependence of the amplitude of a linear chirp power spectrum. (Right) Dynamic spectrum of a linear chirp waveform dechirped imperfectly by the tree algorithm. In this worst-case scenario for $\dot{f} \leq \dot{f}_{\mathrm{max}}$, only 60% of the spectra are shifted by the correct amounts and only 60% of the power is recovered in the appropriate frequency channel. Only the first 100 rows ($\sim$30 s) are shown.

Table 4.4: Dechirping efficiencies resulting from incoherent dechirping of power spectra with a computationally advantageous but approximate tree algorithm (Section 4.2.4).

Rows	Min (%)	Max (%)	Mean (%)	Median (%)	STD (%)
4	100.00	100.00	100.00	100.00	0.00
8	75.00	100.00	93.75	100.00	11.57
16	75.00	100.00	90.62	93.75	10.70
32	68.75	100.00	85.16	81.25	11.20
64	68.75	100.00	81.64	78.12	9.83
128	64.06	100.00	77.93	75.00	8.88
256	64.06	100.00	75.17	73.44	7.68
512	60.16	100.00	72.42	71.09	6.84
1024	60.16	100.00	70.08	69.14	6.10
2048	56.84	100.00	67.92	66.60	5.53
4096	56.84	100.00	66.01	64.94	5.06

Table 4.5: S/N thresholds used in recent searches for radio technosignatures.

Reference	S/N
Gray and Mooley [2017]	7
Harp et al. [2016]	9/6.5
UCLA SETI searches	10
Price et al. [2020]	10
Enriquez et al. [2017]	25
Siemion et al. [2013]	25

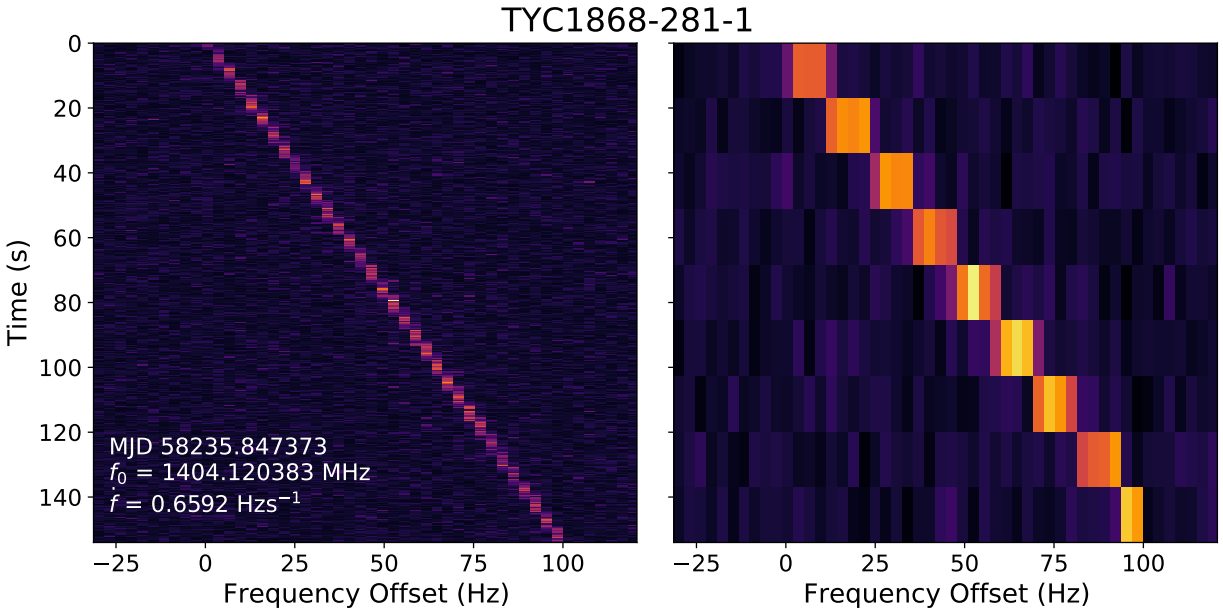

Figure 4.8: Representative dynamic spectra of a signal shown with the nominal time resolution of $\sim$1/3 Hz = 0.33 s (left) and with the degraded time resolution resulting from time-averaging 51 consecutive spectra (right).

CHAPTER 5

A Machine-Learning-Based Direction-of-Origin Filter for the Identification of Radio Frequency Interference in the Search for Technosignatures

Radio frequency interference (RFI) mitigation remains a major challenge in the search for radio technosignatures. Typical mitigation strategies include a direction-of-origin (DoO) filter, where a signal is classified as RFI if it is detected in multiple directions on the sky. These classifications generally rely on estimates of signal properties, such as frequency and frequency drift rate. Convolutional neural networks (CNNs) offer a promising complement to existing filters because they can be trained to analyze dynamic spectra directly, instead of relying on inferred signal properties. In this work, we compiled several data sets consisting of labeled pairs of images of dynamic spectra, and we designed and trained a CNN that can determine whether or not a signal detected in one scan is also present in another scan. This CNN-based DoO filter outperforms both a baseline 2D correlation model as well as existing DoO filters over a range of metrics, with precision and recall values of 99.15% and 97.81%, respectively. We found that the CNN reduces the number of signals requiring visual inspection after the application of traditional DoO filters by a factor of 6-16 in nominal situations.

5.1 Introduction

Radio technosignature searches have increased dramatically both in scope and complexity since the early days of the search for extraterrestrial intelligence [Drake, 1965, Tarter, 2001, Tarter et al., 2010, Drake, 2011, and references therein]. In the past three years alone, the UCLA SETI Group's radio technosignature detection algorithms have undergone multiple levels of improvements, increasing the total number of detections in a typical 2-hour observing window by more than an order of magnitude [Pinchuk et al., 2019, Margot et al., 2021]. Our current pipeline detects 200 times more signals per unit bandwidth per unit integration time than recent Breakthrough Listen (BL) searches [Enriquez et al., 2017, Price et al., 2020, Gajjar et al., 2021]. We have also improved our radio frequency interference (RFI) excision algorithms, yielding RFI classification accuracies of $> 99\%$ on data sets with millions of candidate signals. Other groups are making progress along these lines as well. For instance, the custom hardware system that enabled the early 90s NASA High Resolution Microwave Survey was migrated to a software platform whose RFI excision capabilities have continued to evolve [Harp et al., 2016]. Traas et al. [2021] reported the results from a search of 28 targets selected from the *TESS* Input Catalog and also described an improvement to the BL RFI excision technique.

Despite these advancements, RFI remains the biggest challenge to the search for technosignatures. Pinchuk et al. [2019] and Margot et al. [2021] described several pitfalls of current RFI identification algorithms that rely on inferred signal properties, such as estimates of frequency and frequency drift rate. They suggested that these hurdles might be overcome by an algorithm that instead examines the structure of candidate signals in time-frequency space. Because the time-frequency structure of a signal resembles an image, we can readily apply modern computer vision techniques to this problem, as also suggested by Cox et al. [2018], Zhang et al. [2018], Harp et al. [2019], and Brzycki et al. [2020].

The last decade (2010-2020) has seen considerable advances in the field of Convolutional Neural Networks (CNNs). In 2012, Krizhevsky et al. [2012] introduced the AlexNet architecture,

which won the ImageNet ILSVRC challenge [Russakovsky et al., 2015] the same year. This architecture achieved a top-five error rate of 17%, which represents the percentage of test images (256×256 pixels) for which the network's top five predictions, chosen from a total of 1000 classes, did not include the correct answer. This was an unprecedented accomplishment at the time. An explosion of CNN architectures followed in subsequent years, each performing better than the last [Simonyan and Zisserman, 2014, Szegedy et al., 2015, He et al., 2016, Chollet, 2017]. Modern CNN architectures have achieved top-five error rates of 3% or less [e.g., Tan and Le, 2019].

Machine learning has permeated both the workforce and the research industry, often leading to large improvements in challenging classification problems. In particular, astronomers have already applied CNNs to push the boundaries of astronomical data analysis. For example, Schawinski et al. [2017] trained a Generative Adversarial Network composed of two CNNs (one to classify samples and another to generate them during training) to recover features such as galaxy morphology from low–signal–to–noise and low angular resolution images. Shallue and Vanderburg [2018] trained a deep CNN to predict whether a signal found in *Kepler* data was a transiting exoplanet or a false positive, allowing them to detect and validate a five–planet resonant chain around Kepler–80 and a new, eighth planet around Kepler–90. Zhang et al. [2018] detected 72 new pulses from the repeating fast radio burst FRB 121102 using a CNN trained on radio astronomy data obtained with the Green Bank Telescope. For a more detailed overview of machine learning and CNNs applied to astronomy, see Baron [2019] and references therein.

CNN applications have been explored in the context of radio technosignature searches. Cox et al. [2018] and Harp et al. [2019] both generated a labeled set of synthetic candidate signals from a small (< 10) number of RFI classes to train a CNN for RFI classification. Although this approach provides a relatively simple way to obtain a labeled training set, the synthetic signals may not be representative enough of actual signals and may therefore introduce a bias during model training. Zhang et al. [2018] avoided this problem by using self-supervised learning to train their network. Specifically, their CNN was trained to predict the future time-frequency structure of a signal given the time-frequency structure from a past subset of the total observation. This method allowed the

observations to act as both the training set and the training labels for the model. However, in order to apply the network for RFI excision, the similarity of the predicted signals and the observed signals must be evaluated. Because this task is not trivial, the RFI classification performance of the network may suffer. Brzycki et al. [2020] explored the application of CNNs to technosignature candidate signal detection. Specifically, the authors trained a CNN to detect up to two signals in a frequency span of $\sim$1400 Hz. Although a potential improvement over the detection algorithms of Enriquez et al. [2017] and Price et al. [2020], this CNN cannot yet compete with the detection algorithms described by Margot et al. [2021], which can detect hundreds of signals within the same frequency range.

In this work, we describe an application of CNNs to the excision of RFI in technosignature data. The article is organized as follows. Our motivation and approach to this problem are presented in Section 5.2. Our data compilation procedure is detailed in Section 5.3. In Section 5.4, we describe our approach to CNN model selection and hyperparameter tuning. We also describe a non-ML baseline model that we use as a point of comparison to the trained CNN. Section 5.5 summarizes our results, including the final model performance on the test set as well as on archival data. In Section 5.6, we describe several failure modes of our trained network and offer avenues for future improvements. We present conclusions in Section 5.7.

5.2 Motivation and Approach

Modern radio technosignature programs detect millions of signals per survey [e.g., Siemion et al., 2013, Harp et al., 2016, Enriquez et al., 2017, Margot et al., 2018, Pinchuk et al., 2019, Price et al., 2020, Margot et al., 2021, Gajjar et al., 2021]. These signals must be carefully analyzed to determine whether or not they are of anthropogenic nature. The standard approach to perform this analysis is the direction-of-origin (DoO) filter. This filter labels a signal as RFI if it is not persistent in one direction on the sky or if it is detected in multiple directions on the sky. Theoretically, this filter is powerful enough to remove all RFI signals that are detected in multiple scans. In practice,

this filter often fails on a small subset of signals, but even failure rates as low as 1% can be costly because visual inspection of the remaining signals may be necessary. For instance, the filter failure rates in the searches of Pinchuk et al. [2019] and Margot et al. [2021] were 1.66% and 0.162%, respectively, requiring further examination of 96,940 and 43,020 signals, respectively.

The main pitfall of the DoO filter is the accuracy with which a unique signal can be linked across multiple scans. This "signal pairing" is required for both the persistence-test (present in all scans of the source) and the uniqueness-test (absent in scans of other sources) portions of the filter. Different surveys implement this pairing functionality in various ways. For example, Enriquez et al. [2017] and Price et al. [2020] consider two signals to be from a common origin if the frequency at which the latter signal is detected is within a generous tolerance of ± 600 Hz of the detection frequency of the first signal, even if the corresponding frequency drift rates are unrelated. Although this approach speeds up the analysis by discarding a large portion of the candidate signals, it is problematic because it may eliminate valid technosignatures. A more rigorous approach was adopted by Pinchuk et al. [2019] and Margot et al. [2021]. In both of these searches, two signals were paired only if their frequency drift rates and frequencies extrapolated to a common epoch are within a small tolerance. More robust versions of this filter could include tests of other signal properties, such as signal bandwidth or off-axis gain ratio.

In all four searches described in the preceding paragraph, the filter was applied to estimates of signal properties produced by a computer program on the basis of the time-frequency structure of each signal. Therefore, the efficiency of the filter relies heavily on the accuracy of the derived signal properties. When the estimates of these signal properties are imprecise or incorrect, or when the underlying assumption of a linear drift rate is violated, the filter classification fails. Pinchuk et al. [2019] detailed five different signal types for which their DoO filter exhibited a degraded performance. Importantly, this limitation can likely be overcome by an algorithm that examines the time-frequency structure of each signal directly.

In this work, our approach is to train a CNN to pair signals by directly examining the corresponding dynamic spectra. The trained network is then used to examine the data as follows. For

each signal detected in the survey, we extract a portion of the dynamic spectrum centered on the time-frequency location of the signal as the first input to the network. This dynamic spectrum is guaranteed to contain a signal, because the minimum detection threshold in typical SETI searches is set at $\geq$ ten times the standard deviation of the noise. Using an estimate of the drift rate of the signal in this dynamic spectrum, we extrapolate the expected detection frequency to the starting epoch of a different (typically subsequent) scan. We then extract a portion of the dynamic spectrum of the second scan to use as the second input to our network. This portion has the same dimensions as those of the first input and is centered on the expected detection frequency. The output of the network provides an assessment of whether or not the second dynamic spectrum contains the same signal as the first dynamic spectrum. The CNN will be trained to perform this task with a large ($\sim$1 million) labeled training set, such that the CNN can make this assessment by recognizing patterns in the images, as opposed to calculating and comparing estimates of signal properties like frequency and drift rate.

In what follows, we will use the terms "first" or "top" image to refer to the dynamic spectrum of the first scan, which must contain a signal of interest. Likewise, we will use the terms "second" or "bottom" image to refer to the dynamic spectrum of the second scan, which may or may not contain the same signal.

5.3 Data Preparation

5.3.1 Observations

We compiled our data set from the observations presented by Pinchuk et al. [2019]. Those observations were conducted on 2017 May 4, 15:00 – 17:00 Universal time (UT) with the 100 m diameter Green Bank Telescope (GBT). Both linear polarizations of the *L*-band receiver were recorded with the GUPPI back end in its baseband mode [DuPlain et al., 2008]. GUPPI was configured to channelize 800 MHz of recorded bandwidth into 256 channels of 3.125 MHz each.

The observations primarily consisted of sources from the Kepler field, but also included scans of TRAPPIST-1 and LHS 1140. A total of twelve sources were scanned. A full list of targets and their properties can be found in Table 1 of Pinchuk et al. [2019]. This article also includes details relating to the formation of the dynamic spectra, which have a time resolution of 0.336 s and frequency resolution $\Delta\nu$ of 2.98 Hz.

A total of 10,293,618 signals were detected in the data, 8,592,771 of which have a signal-to-noise ratio (S/N) ≥ 10. Because we had a large number of signals to choose from, we carefully pruned the data according to principles described in Section 5.3.3 in order to obtain the best possible training candidates.

5.3.2 Definition of Data Sets

In order to successfully train our CNN, we need to set aside several small portions of our data that we can use to evaluate the model performance during and after training. Typically, it is recommended to set aside 10–20% of the training data as a "validation" set that is used to evaluate important metrics like precision, recall, and a model cost function or loss [Géron, 2019]. These metrics can then be used to tune model hyperparameters (Section 5.4.3) or identify problems like overfitting, which occurs when a neural network simply learns to reproduce the labels of the training data and therefore generalizes poorly to any other data. Standard ML practices suggest that another 10–20% of the training data should be set aside as a "test" set that is only ever used to evaluate the performance of the final model. This evaluation is important in order to obtain an accurate estimate of how well the model generalizes to data that it has never seen before.

When the training data is representative of the data that the model will see in a production environment, the training, validation, and test sets are enough to successfully train a CNN from start to finish. However, because we needed to label a large training set in an automatic manner (Section 5.3.4.1), our training data consists of two images taken from a single scan, whereas the production data consists of two images from entirely separate scans. This difference between

training and production data makes our application atypical and requires adjustments to standard ML practices. In particular, it is possible for the network to perform well on the training data but poorly on the production data. This situation occurs when there is a "data mismatch" and often requires some manipulation of the training set in order to better match the production data. Importantly, this condition must be detected and addressed before the model is put into production. If the validation set is generated as a subset of the training set, then data mismatch is impossible to detect. On the other hand, if the validation data is comprised only of data that the model will see in production, it is not possible to discern whether poor model performance is attributable to data mismatch or to model training issues, such as overfitting, in the absence of other information. The solution to this problem is to create an additional data set, the "train–dev" set, which is a subset of 10–20% of the training data and is used to monitor the performance of the model during training and detect problems like overfitting. With the "train-dev" set on hand, we can compile validation and test sets that match the data that the model will see in production. These two data sets are hand-labeled, as is standard in most ML applications, and are therefore much smaller than the training and train–dev sets. The validation and test sets provide a useful way to select an appropriate CNN architecture, tune hyperparameters, and measure the model's generalization to new data, among other uses.

We compiled a subset of signals from the hand-labeled validation and training sets to evaluate the performance of the UCLA SETI Group DoO filter. In order to facilitate a fair comparison to our CNN, we only kept the labeled image pairs that corresponded to different scans of the same source, and we only applied the persistence portion of the DoO filter to these signals (the signal pairing logic is identical for both components of the filter). This is important because the DoO filter examines many different scans from a single observing session to look for the presence of a given signal elsewhere on the sky. However, when determining if the signal is persistent in its detection direction, the filter only examines the two scans of the source containing the signal of interest. The latter is consistent with a standard application of the CNN trained in this work and therefore offers the best comparison between the performance of the existing DoO filter and the

CNN.

We also made use of a small hand-labeled data set of image pairs to optimize and evaluate a baseline model (Section 5.4.1). This model, which does not rely on ML techniques, provides a mechanism to test the performance improvement due to the ML application. Table 5.1 summarizes the data sets utilized in this work.

Table 5.1: Data set name, size, and usage for all data sets presented in this work. Sections 5.3.4 and 5.3.5 detail the data compilation strategy, including the choice of data set size. The parentheses in the "Usage" column specify a particular use case for the data set. The final column details whether or not the data set was labeled by hand.

Name	Size	Usage	Hand-labeled?
Training	1,000,000	Training and evaluating the CNN	No
Train-dev	100,000	Evaluating the CNN (Overfitting)	No
Validation	1,156	Evaluating the CNN (Data mismatch)	Yes
Test	1,272	Evaluating the CNN (Final results)	Yes
UCLA DoO test	1,238	Evaluating the UCLA direction-of-origin filter	Yes
Baseline	524	Optimizing and evaluating the baseline model	Yes

Note that the validation and test sets are much smaller than the training and train–dev sets because the former had to be analyzed and labeled by hand whereas the latter were labeled automatically, as described in Section 5.3.4.

5.3.3 Data Selection Filters

We began by examining the distribution of drift rates of the 8,592,771 signals detected in 2017 (Figure 5.1, Left). We observed that the vast majority ($> 98\%$) of detected signals have drift rates $|\dot{f}| \leq 2$ Hz s^{-1}. Moreover, by examining the dynamic spectra of signals with drift rates $|\dot{f}| > 2$ Hz s^{-1}, we observed a lack of the narrowband characteristics that are often chosen as one possible

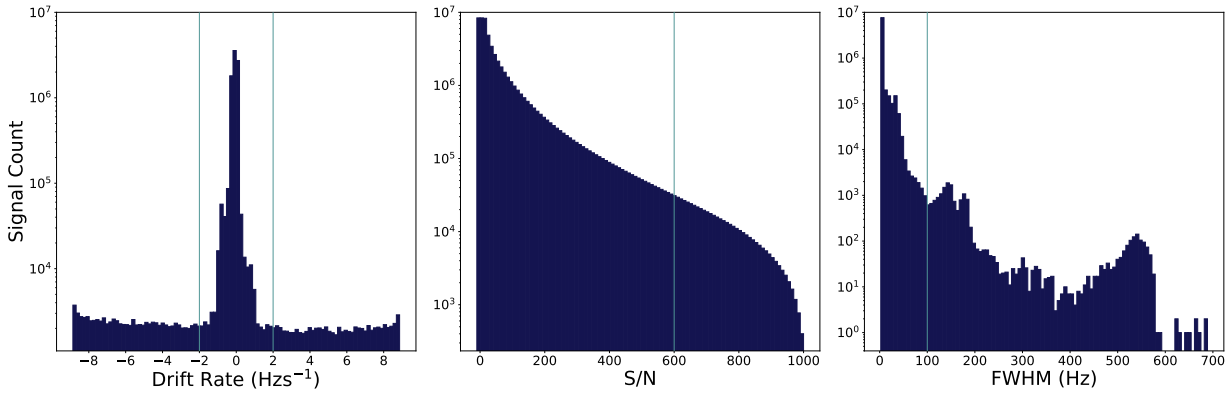

Figure 5.1: Histograms of the (Left) drift rates, (Middle) S/N, and (Right) bandwidths as measured by a FWHM metric of detected signals.

diagnostic of extraterrestrial engineered emitters. For these reasons, we excluded 168,684 (1.96%) signals with drift rates $|\dot{f}| > 2$ Hz s^{-1} from the training set.

We performed a similar cut on the basis of the S/N of the detected signals. Instead of significantly altering the S/N distribution, however, we opted to remove any signals with extremely large S/N. Based on the cumulative distribution of S/N shown in the middle panel of Figure 5.1 we chose to discard any signals with a S/N > 600. This threshold was large enough to preserve $> 99\%$ of the remaining signals but also remove any extreme S/N outliers. As a result of this filter, we removed 58,399 (0.69%) signals with extreme S/N from the training set.

Next, we examined the bandwidth of the remaining signals, as quantified by a full width at half maximum (FWHM) metric. The bandwidth of the training signals is especially important because we needed our training-set signals to fit within a 225×225 image. This image size was chosen to satisfy a few considerations. First, the number of lines cannot exceed half the number of lines in the dynamic spectra, which is ∼490 for the 2017 data at 2.98 Hz frequency resolution. Second, the number of columns must be an odd number so that signals can be perfectly centered in the image. Third, the input size for models like ResNet52 are images of size 224×224 pixels [He et al., 2016] and have been shown to be manageable with modern CNN architectures. Our chosen image size corresponds to an upper limit of ∼670 Hz on the bandwidth of the signals. However, because

we are only interested in narrowband signals ($\leq$ 10 Hz) in this work, we can set the bandwidth threshold much lower than the theoretical upper limit. The distribution shown in the right panel of Figure 5.1 suggests a threshold value of 100 Hz. As a result of this filter, we removed 15,044 (0.18%) signals with large bandwidths from the training set.

Finally, we discarded 6447 signals that were detected close enough to the edge of their 3.125 MHz wide channel such that the signal overlapped with a neighboring channel. Although we could "stitch" neighboring channels together to fully recover these signals, given the vast number of signals left to choose from, we decided to simply remove this tiny fraction of signals from consideration for the purpose of building the training set. When we evaluate actual data with the CNN, we do combine two channels when necessary to correctly represent signals located near the channel edges.

Overall, the filters above discarded 248,594 of the 8,592,771 detected signals, leaving 8,344,177 (97.11%) available to use for our training set. If we assume that the distributions of drift rates, S/N, and bandwidths remain relatively constant over time and the RFI observed with the 2017 antenna pointing directions is representative of other directions, then the network trained on our pruned data set should be applicable to $\sim$97% of the detected signals in future searches with similar parameters. If the RFI environment changes so much in time or in space that it severely alters the distributions of signal properties, this percentage value could change. Although the RFI environment may evolve in time and space, we were able to verify that these filters still captured $\sim$97% of the detected signals in searches conducted in 2018 and 2019 by Margot et al. [2021] with different antenna pointing directions. Specifically, 97.4% (9,845,561 out of 10,113,551) and 97.2% (16,048,515 out of 16,518,362) of the signals detected in 2018 and 2019, respectively, passed the data selection filters described above.

The range of signals selected for the labeled training set translates into a finite domain of applicability for the CNN. Because the CNN was not designed for signals with S/N < 10, frequency drift rates $|\dot{f}| > 2$ Hz s^{-1} , or bandwidths > 100 Hz, it may not perform well when applied to such signals. Note that the applicability of the training set to new data sets and the CNN's decision

accuracy are two different concepts. We evaluate the latter in Section 5.5.2.

5.3.4 Generation of the Training and Train–dev Set

Labeling a sizable training set by hand is time-consuming. To bypass this limitation, we developed a strategy to synthetically generate our labeled data set. Because we are interested in training a neural network to supplement our DoO filters by detecting whether or a not a signal is present in two separate images, we need a training set that consists of pairs of images labeled with a binary flag indicating the persistence of a signal across both scans.

5.3.4.1 Creation of Image Pairs

In order to simulate a pair of scans containing the same signal, we split the image representing a single scan into two parts along the time dimension. We then evaluated whether the signal was detected in both the top and bottom parts. If the signal was detected in both, we labeled the pair as a positive sample. Otherwise, we labeled the pair as a negative sample to signify that the signal was present in only one of the two parts. In practice, the detection decisions are implemented by computing the ratio of signal powers in the top and bottom parts.

The power ratio calculations rely on the simplifying assumption that the total integrated power associated with each signal is distributed evenly throughout the duration of the scan. In other words, we expect the signal in each half of the spectrum to contribute equally to the total power. We calculated the signal power detected in each half of the scan and recorded these values as power ratios (P_{top}, P_{bottom}), where the denominator is half of the total signal power in the scan. The signal power was calculated by summing $2n + 1$ pixels at each timestep along a line with a slope equal to the drift rate of the signal, where $n = \left\lfloor 1.5 \times \frac{B_{\mathrm{FWHM}}}{2\Delta\nu} \right\rfloor$ pixels on either side of the line, B_{FWHM} is the bandwidth of the signal measured in Hz, $\Delta\nu$ = 2.98 Hz is the frequency resolution of the data, and $\lfloor\ \rfloor$ is the floor operator. The drift rate of the signal was assumed to be the same in both halves of the spectrum. To calculate the signal power in the the top half of the

dynamic spectrum, we started at the pixel corresponding to the detection frequency of the signal. To calculate the power of the signal in the bottom half, we started at the center frequency obtained by linearly extrapolating the signal detection frequency to the appropriate time. The total signal power was obtained by summing the two halves. By comparing the ratio of powers in the top and bottom halves of each scan to a suitable threshold, we were able to assign an appropriate label to each signal. Section 5.3.4.2 describes the selection of the threshold.

This approach allowed us to label a large amount of signals in a short period of time. We chose to compile the training and train–dev data sets from a pool of 1,100,000 total signals, which is a random selection among the 8,344,177 signals that meet certain criteria described below. For reference, the popular MNIST handwritten digit dataset [LeCun et al., 1998] as well as the CIFAR–10 [Krizhevsky et al., 2009] multiclass image dataset both contain 60,000 samples each. Both the training and train–dev set contain an equal ratio of positive and negative samples. We set aside 100,000 signals for the train–dev set, leaving 1 million signals to be used as the training set.

In order to be accepted into the training set or train–dev data set, signals had to satisfy several criteria. Most importantly, the top image in a pair, which mimics the first scan in an actual observing sequence, must always contain a signal. Moreover, the primary signal must always be centered in the top image and nearly centered in the bottom image, i.e., the signal must start in or near the middle of the frequency array in the topmost time bin. This requirement affects the construction and processing of the images, which are described in Section 5.3.4.3. In particular, we allowed a small tolerance on the location of the signal in the bottom image, but the signal in the top image must always start at column 113 (if counting from 1) in the first row of the 225×225 images. Both of these criteria can easily be met in production, because one can apply these cropping steps to detected signals with known starting frequencies. For signals whose bandwidth spans several pixels, the starting frequency is defined as the starting frequency reported by the detection algorithm, which is where most of the power is detected [Margot et al., 2021].

5.3.4.2 Selection of Suitable Signals

We began by examining the S/N of each signal in the top half of the scan only. In order to satisfy the underlying assumption that there is definitely a signal in the center of the first image (represented in the training set by the top half of the scan), we required a minimum top-half S/N of at least 6 (Figure 5.2), which corresponds to a $\sim$1 in a billion false detection rate. We validated our choice of threshold by examining a sample of signals below the cutoff value. We found that most of these signals are faint and difficult to detect visually, while the rest are not present at all. On the contrary, signals above this threshold are clearly visible in the dynamic spectra. Appendix 5.A illustrates these two cases.

As we performed our final selection, we needed to allow for variations in the S/N of the top and bottom portions of the positive signals, since the top and bottom portions represent two separate scans and we have empirically observed that the S/N can change substantially between scans. To do so, we compared the integrated power values from the top and bottom halves of each signal. Specifically, we examined the distribution of the ratio of bottom to top integrated powers (i.e., $P_{\text{bottom}}/P_{\text{top}}$). We found that approximately 50% of the signals have a ratio between 0.75 and 1.25 (Figure 5.3), so we randomly selected 550,000 signals from this region to represent our positive samples (i.e., a signal is detected in both scans/images). Additionally, we selected 50,000 signals with a ratio of 0.2 or lower to partially represent our set of negative samples (see blue line in Figure 5.3). To ensure the absence of a signal from the bottom half, we verified that none of these samples had any signals with a prominence value greater than 3 times the standard deviation of the noise in the bottom half (see Margot et al. [2021], Section 3.1 for an in-depth discussion of the prominence calculation). Appendix 5.A depicts a sample of these signals. Ideally, *all* negative samples would be obtained with this method, but there were not enough of these signals to provide the necessary negative samples. Since this category is grossly under-represented in the data, we used data augmentation to create more negative samples. Specifically, the remaining 500,000 negative samples were obtained by taking samples from the region with a $P_{\text{bottom}}/P_{\text{top}}$ ratio between 0.75 and 1.25 and altering them in four different ways to remove any signals present

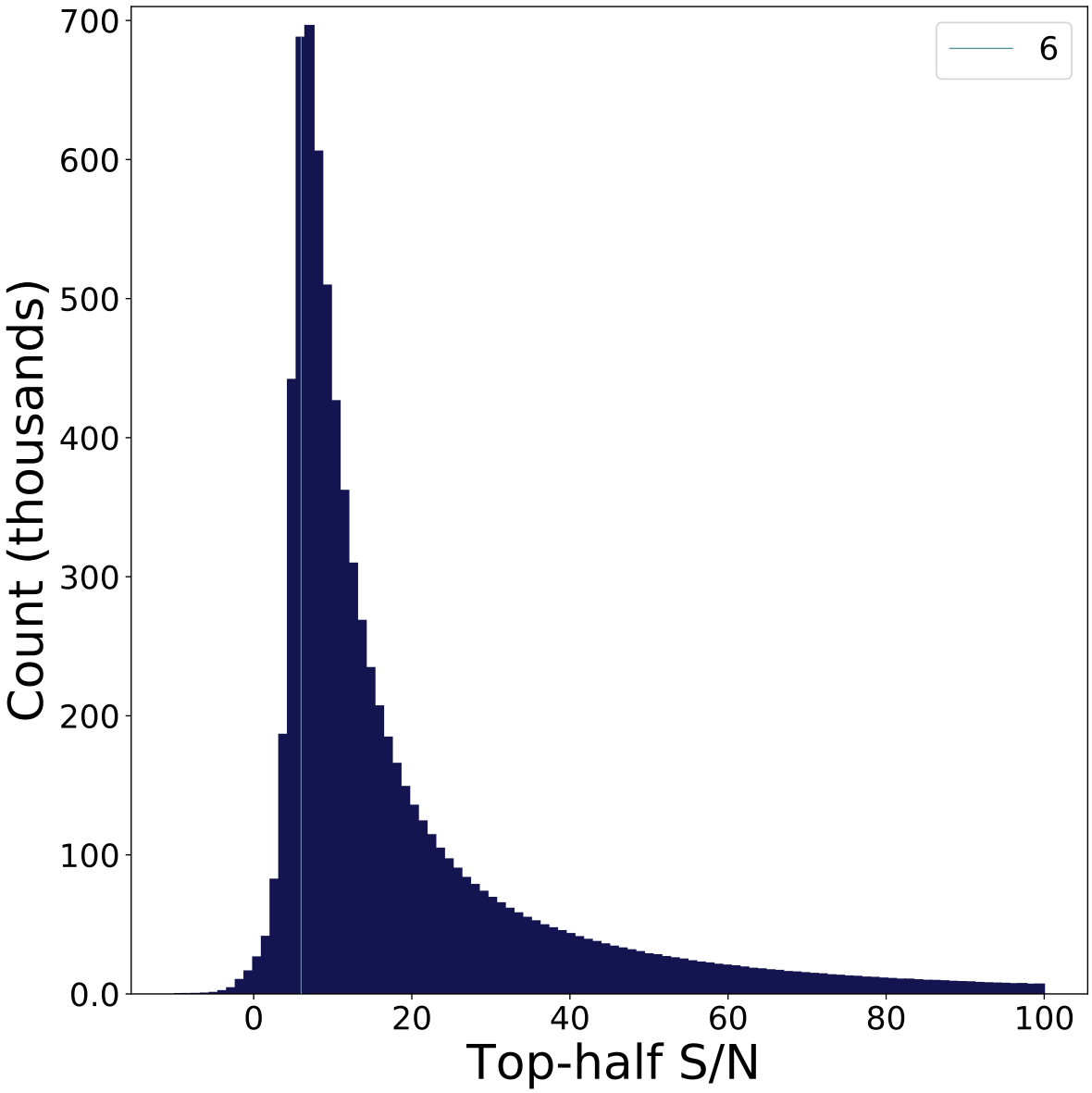

Figure 5.2: Histogram of the S/N detected in the top half of each scan. The blue vertical line shows the S/N cutoff value of 6 used to remove signals with low power in the top half of the scan.

in the bottom half. This process is described below.

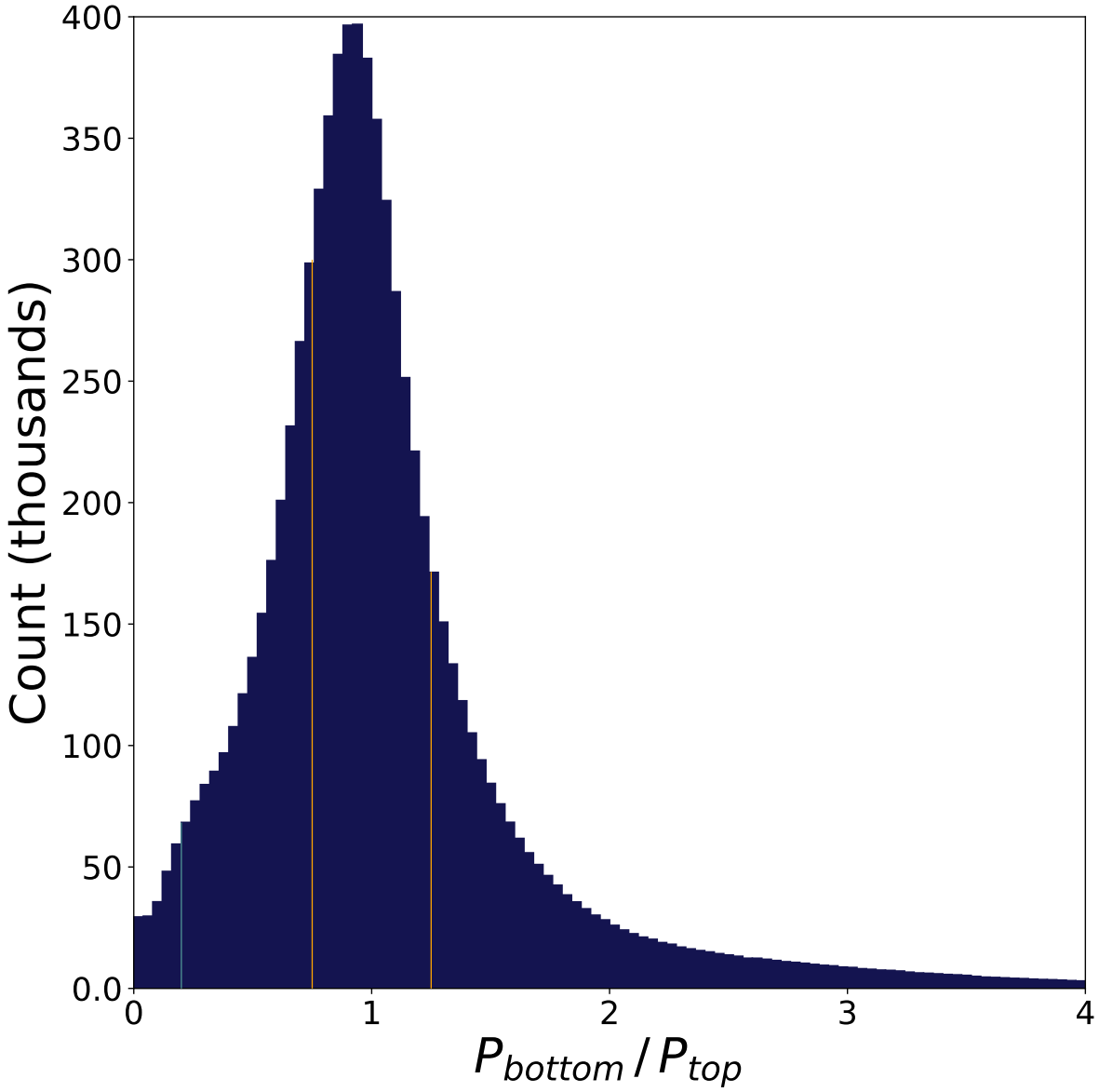

Figure 5.3: Distribution of the ratio of integrated powers, or $P_{\mathrm{bottom}}/P_{\mathrm{top}}$ ratio. The orange vertical lines delimit the lower and upper bounds that we used to select 500,000 positive samples (0.75 and 1.25, respectively). The blue vertical line is plotted at a ratio of 0.2. We selected 50,000 signals below this value to represent a portion of our negative samples.

5.3.4.3 Processing of Selected Signals

After selecting the signals for our training set, we applied some processing to ensure that signals in the positive and negative categories were representative of their respective labels.

We first considered the 550,000 signals from the positive category. When we applied our machine learning algorithm to real data, we obtained the bottom image by extracting a portion of the spectrum from the second scan centered on the frequency value calculated by extrapolating the frequency detected in the first scan. If the same signal is present in both scans, and if the signal's drift in time-frequency space is approximately linear[1], and if the drift rate estimate is approximately correct, this method will ensure that the signal also appears in the bottom image, but it does not guarantee that the signal will be perfectly *centered* in the bottom image. For instance, a small discrepancy between the actual and estimated drift rates can result in an offset between predicted and actual frequency values. To simulate this scenario in our training set, we shifted all of the signals in the bottom image of our samples by $\pm$0-5 pixels (0-15 Hz). The exact shift for each image was randomly selected from a distribution of shift values described in Section 5.4.1.

The 550,000 signals for the negative category were compiled using five distinct procedures. The first 50,000 signals were selected from the distribution shown in Figure 5.3 with a $P_{\text{bottom}}/P_{\text{top}}$ ratio of < 0.2. For each of these signals, we verified the lack of any signals with a prominence value greater than 3 times the standard deviation of the noise in the bottom half of the scan. The next 125,000 negative samples were obtained by selecting unused signals from the "positive" range ($0.75 \leq P_{\text{bottom}}/P_{\text{top}} \leq 1.25$) and shifting the signal in the bottom image by $\pm$6-10 pixels (18-30 Hz). By doing so, we forced the algorithm to learn that a positive detection requires the bottom signal to be detected in close proximity to the extrapolated frequency, which is calculated on the basis of signal properties in the top image. We obtained another 125,000 negative samples by once again selecting unused signals from the "positive" range and replacing the bottom signal with an

[1]The assumption of linearity is reasonable for a source analyzed at *L* band with a frequency resolution of $\sim$2.98 Hz, a scan duration of 150 s, and a line-of-sight jerk below $2.311e^{-5}\text{ms}^{-3}$. For reference, the maximum line-of-sight jerk for Earth's spin and orbit are $2.453e^{-6}\text{ms}^{-3}$ and $1.181e^{-9}\text{ms}^{-3}$, respectively.

unrelated signal (also sampled from the "positive" range). This group of negative samples forced the algorithm to compare signal properties and *not* pair two unrelated signals that may have been detected at similar frequencies in two different scans. Another 125,000 negative samples were obtained by selecting leftover signals from the "positive" range and replacing the bottom image with noise. The noise was generated by sampling values from a χ^2 distribution with four degrees of freedom that was fit to the bottom image after removing any power values belonging to any signals detected in the spectrum. The signals were removed by obtaining the database records of all signals detected within the relevant portion of the spectrum and discarding any power values within 2 times the measured bandwidth along the linear drift rate of each signal. The final 125,000 samples were obtained similarly, but instead of replacing the entire bottom image with noise, only the power values belonging to the signal in the bottom image were replaced with samples values from a χ^2 distribution that was fit to the bottom image with the same procedure as above.

An example product of each of the above procedures is shown in Figure 5.4.

Before finalizing the training and train–dev sets, we examined the drift rate distribution of the 1.1 million signals selected with the process described above. This distribution is biased towards signals with negative drift rates (Figure 5.5; left). This bias is expected from most low- and medium-Earth-orbit satellites, such as Global Positioning System (GPS) satellites, which orbit in a prograde fashion with respect to the telescope. In order to avoid inadvertently introducing this bias into our model, we selected 364,184 signals with a negative drift rate using a stratified split [Géron, 2019] on the signal drift rates, and horizontally flipped the images corresponding to these signals. The resulting drift rate distribution exhibited a significantly reduced bias between -0.5 and 0 Hz s^{-1} at the expense of a slight bias between -2 and -1 Hz s^{-1} (Figure 5.5; right).

At the end of the compilation process, our training set consisted of 550,000 positive samples and 550,000 negative samples. The 1.1 million samples were separated into 1 million training samples and 100,000 train–dev samples. The samples were separated using a stratified split on the bandwidth of the signals. Each set contained an equal amount of positive and negative samples.

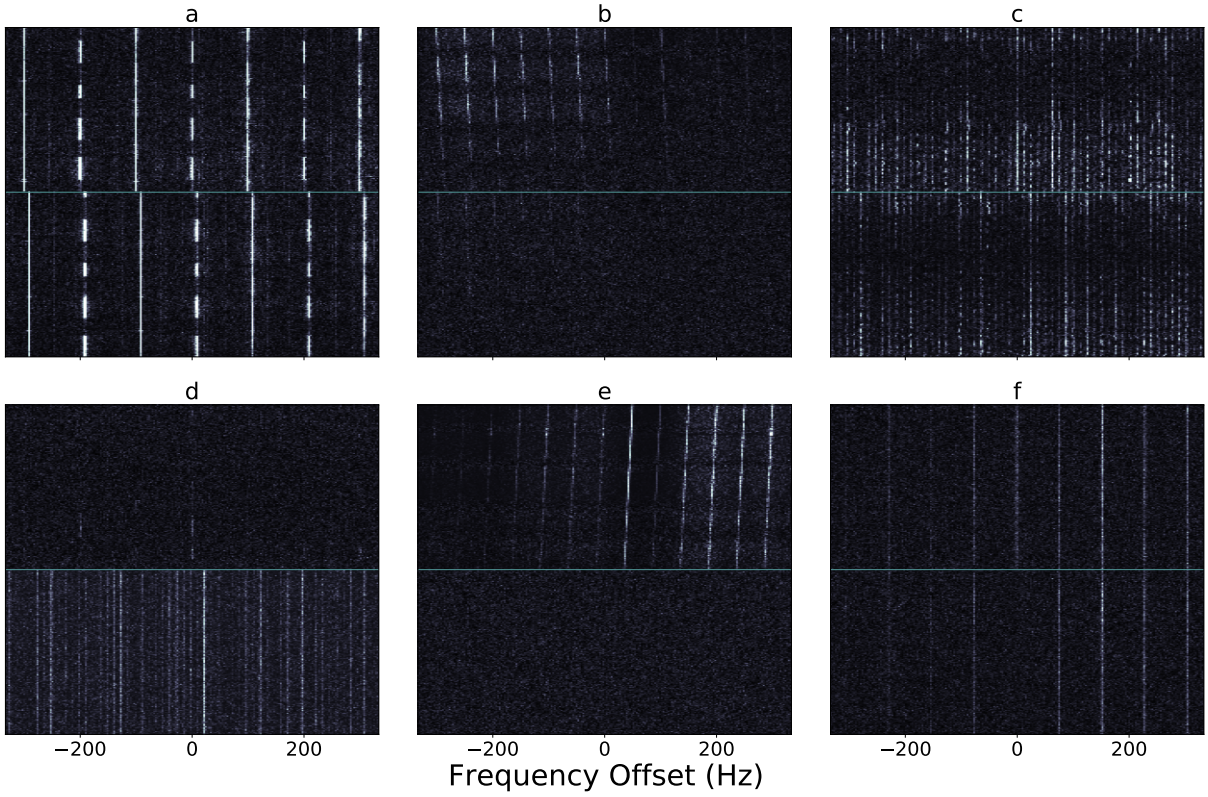

Figure 5.4: Sample signals used in the ML labeled training set. (a) Signal from the positive category, shifted 3 pixels (∼9 Hz) to the right. (b) Sample signal from the negative category with a $P_{\mathrm{bottom}}/P_{\mathrm{top}}$ ratio of < 0.2. (c) Sample signal from the negative category, shifted 8 pixels (∼24 Hz) to the right. (d) Sample signal from the negative category with an unrelated signal in the bottom image. (e) Sample signal from the negative category with a bottom image consisting completely of simulated noise. (g) Sample signal from the negative category with the primary (center) signal replaced by noise in the bottom image.

5.3.5 Creation of Validation, Test, and Baseline Model Data Sets

We compiled a small set of 1,156 hand-labeled images, where the top and bottom images are extracted from two separate scans. These images are a true representation of the samples that the network will see during production, so we use them as our validation set. By comparing the

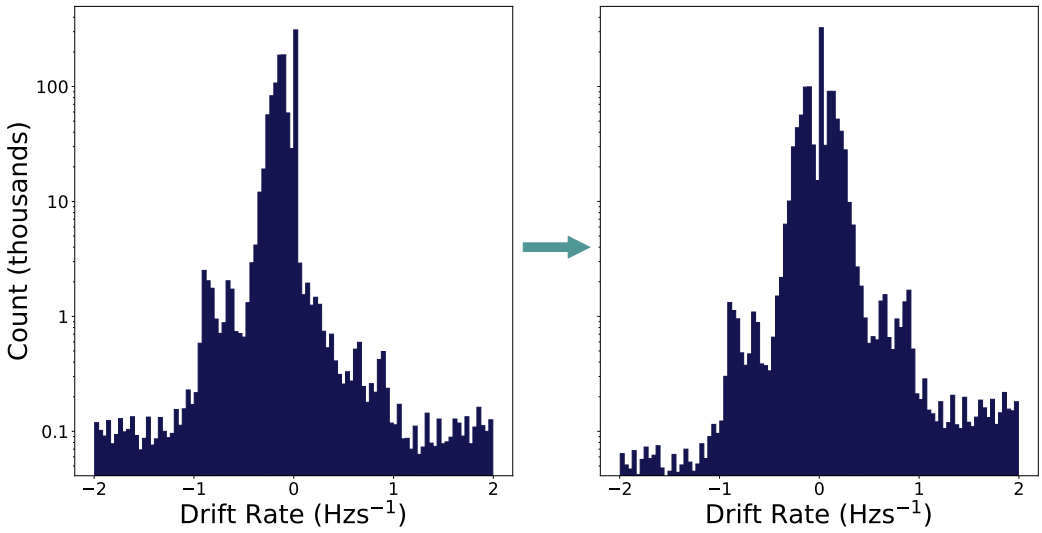

Figure 5.5: (Left) Drift rate distribution of the 1.1 million signals selected to be part of our training set. Note a significant bias towards signals with a negative drift rate value. (Right) Drift rate distribution of the same set of signals after applying a horizontal flip to 364,184 negative drift rate signals. The bias that affects ∼685,000 signals between -0.5 and 0 Hz s^{-1} is almost entirely removed at the expense of a slight bias introduced between -2 and -1 Hz s^{-1}that affects ∼2500 signals.

model performance on the train–dev and the hand-labeled validation set, we can assess whether or not there is a mismatch between the training set and the real test data. We also compiled a set of 1,272 hand-labeled images to serve as the test set. These samples contain signals from two different scans, as would be the case in the production environment. Finally, we selected 524 hand-labeled samples, where each sample contains a signal in both scans, to optimize and evaluate our baseline model (Section 5.4.1). This data set has no samples in common with any of the 4 data sets described above.

5.4 Models

5.4.1 Baseline Model

We devised a simple correlation-based model to serve as a benchmark or baseline for our results. We began by selecting a baseline data set as described in Section 5.3.5. We then calculated the 2D correlation coefficient between the two signals in each data sample, using all available time steps in a region of frequency width $2w + 1$ centered on each signal. The correlation coefficient is given by

$$\rho(A, B) = \frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{A_i - \mu_A}{\sigma_A} \right) \left(\frac{B_i - \mu_B}{\sigma_B} \right), \tag{5.1}$$

where A_i and B_i are the individual pixels of image A and image B, μ_j and σ_j are the mean and standard deviation of the pixels under consideration in image j, $j \in \{A, B\}$, respectively, and N is the total number of pixels compared. To ensure that our results were not influenced by poor localization of the signals in the images, we shifted the bottom image by ± 3 pixels in the frequency dimension and computed $\rho(A, B)$ in each case. We report the maximum correlation score from the set of seven resulting values. We tested both a large ($w = 15$ pixels, $\sim$50 Hz) and small ($w = 3$ pixels, $\sim$10 Hz) window size, and found that the latter gave the best results in terms of model precision and recall.

After computing the correlation values, we selected a threshold value in order to assign a label for each set of images. The label is positive (i.e., "True", 1) if the signals in the images are strongly correlated, or negative (i.e., "False", 0) if the signals in the images are unrelated. Typically, this threshold is chosen by finding the best trade-off between precision and recall [Géron, 2019]. Precision is defined as the ratio of the true positive count (i.e., label=prediction=1) to the sum of the true positive and false positive counts (i.e., prediction=1). In other words, when a model with a precision value of 1 predicts that an image pair belongs to the positive class, it is always correct. On the other hand, recall is defined as the ratio of the true positive count to the sum of the true positive and false negative counts (i.e., label=1). A model with a recall value of 1 will always correctly classify all the positive samples. A perfect model would have both recall and precision

values of 1. In practice, there is always a trade-off between the two metrics.

In our application, precision is more important than recall because a larger precision value minimizes the number of false positives. False positives represent valid candidate technosignature signals that were only detected in one image (or direction of the sky), yet were still classified as RFI. For this reason we chose our threshold as the correlation value that yielded a precision $\geq 95\%$. At this threshold (0.0551), the recall was 33.7% (Figure 5.6). In other words, the baseline model only detects $\sim$1/3 of the RFI in the data, but it does so with 95% precision.

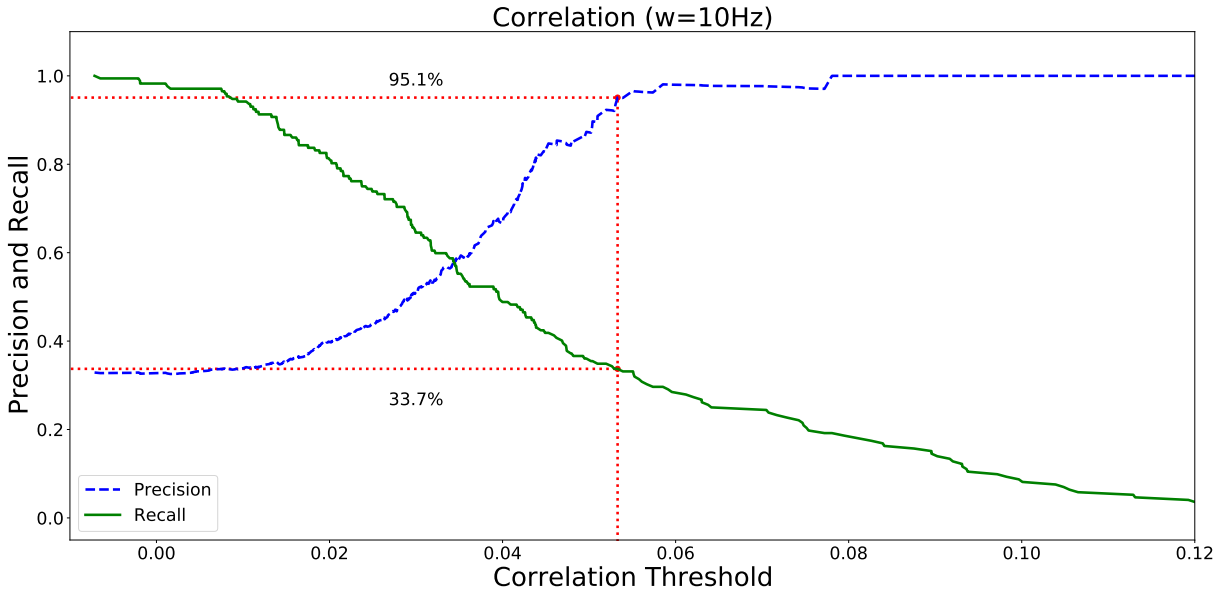

Figure 5.6: Precision and recall curves for a baseline 2D correlation model, which does not rely on ML techniques and is used solely to serve as a benchmark to evaluate the performance improvement of our ML application (Section 5.5.2). With the chosen threshold, the baseline model detects approximately a third of the RFI in the data with 95% precision.

The baseline model also helped define a distribution of frequency shifts that we used in building the training and train–dev sets (Section 5.3.4.3). We selected a subset of 5,750 signals from the 2017 observations [Pinchuk et al., 2019] that passed the UCLA DoO filter and had correlation values that exceeded the threshold of 0.0551. We used randomly selected values from the distribu-

tion of shifts of this subset to shift the signals in the second images of our training and train–dev sets (Section 5.3.4.3).

5.4.2 Model Selection

In order to select the best suitable model for the DoO filter, we carried out a scaled-down performance comparison of over 20 model architectures. For this comparison, we selected four ResNet variants [ResNet34, ResNet50, ResNet101, and ResNet152; He et al., 2016], two VGG variants [VGG16, VGG19; Simonyan and Zisserman, 2014], and the Xception architecture [Chollet, 2017]. In addition, we trained two Siamese model variants (see Appendix 5.B) for each of these 7 models. We did not perform any hyperparameter tuning at this stage, and we trained each model on only 10% of the training set, using 10% of the train–dev set and the full validation set to evaluate the results, because the goal was to quickly compare as many models as possible.

We found that all models outperformed their respective Siamese versions when comparing the loss and other relevant metrics. We also found that none of the seven standard model architectures significantly outperformed the others in terms of these metrics. However, we did notice that the Xception architecture did not exhibit significant overfitting during training, whereas all other models did. Although there are multiple model regularization techniques designed to overcome model overfitting, we decided to select the Xception model as our base architecture in order to reduce the amount of model tuning required later during training.

5.4.3 Hyperparameter Tuning

After selecting the best model architecture for the DoO filter, we were left with a significant number of hyperparameters to tune. Géron [2019] defines a hyperparameter as "a parameter of a learning algorithm (not of the model). As such, it is not affected by the learning algorithm itself; it must be set prior to training and remains constant during training." All of the hyperparameters that were considered during this process, as well as several suitable values for each, are listed in Table 5.2.

The "Optimizer", "Learning rate", "Batch size", and "Activation function" hyperparameters simply refer to the network hyperparameter that was tuned during this process. The hyperparameter "Fully connected layers on top" refers to the addition of one or more fully connected layers of neurons inserted immediately after the global average pooling layer but before the final prediction node. The number of layers and the number of neurons per layer were also tuned as part of this process. When the hyperparameter "Dropout rate" was set to None, no changes were made to the network architecture. Otherwise, a dropout layer [Srivastava et al., 2014] was added at the end of the network with the corresponding dropout rate. The hyperparameter "Include Squeeze-and-Excitation blocks" refers to the addition of Squeeze-and-Excitation (SE) blocks [Hu et al., 2018] at the end of every separable convolution[2] module of the Xception architecture. SE blocks are network units that are designed to adaptively recalibrate channel-wise feature responses by explicitly modeling interdependencies between the channels. Hu et al. [2018] demonstrated that SE blocks bring significant improvements in performance for state-of-the-art CNNs with only a slight addition to the computational cost. The "input batch normalization" hyperparameter controlled the normalization of the input data. Specifically, if this parameter was set to False, the input data would be normalized to zero mean and unit standard deviation, and no further modifications were made to the base network structure. When this parameter was set to True, the input data were not scaled, but an extra batch normalization layer [Ioffe and Szegedy, 2015] was added immediately after the input layer of the network.

While a comprehensive grid search for the best hyperparameter combination would yield the optimal model configuration, we found that hardware limitations made this approach impractical. A single training session with only 20% of the training data and 10 epochs, where each epoch represents a full pass of the training data through the neural network, took ∼10 hours on a single ML-enabled graphical processing unit (GeForce RTX 2060 SUPER 8 GB GPU), which would make a grid search prohibitively large considering the need to examine ∼4,000 combinations. With

[2]A convolution layer is the central building block of a CNN. It applies a convolution kernel to each pixel of an input image and produces a feature map. When the input image contains multiple channels (e.g., red, green, blue), the convolution kernel has a third dimension equal to the number of channels.

the current specifications, a grid search for the best hyperparameter combination from the set of values described in Table 5.2 would take ∼4.5 years to complete. Instead, we chose the best hyperparameter combination from the results of ∼30 different training sessions of 10–15 epochs each using judiciously chosen combinations of hyperparameters. Our selection approach was "semi-greedy" because we allowed the results of previous training sessions to have some influence over the hyperparameter choice for the next session. Although this approach does not guarantee a globally optimal model configuration, we found that the the hyperparameter combination obtained via this method yields satisfactory model performance (see Section 5.5.2).

The final combination of hyperparameters was determined by comparing the model performance over all ∼30 training sessions. The best values for each parameters are listed in the final column of Table 5.2.

5.4.4 Final Model

Our final model architecture is shown in Figure 5.7. The most important layer of the Xception architecture is the separable convolution layer, which consists of a spatial convolution performed independently over each channel of an input, where a channel refers to a slice along the depth dimension of the input matrix, followed by a 1×1 convolution projecting the outputs of the first convolution onto a new space. Chollet [2017] argues that the separable convolution layer is almost identical to an "extreme" version of the inception module, which is the backbone of the GoogLeNet architecture [Szegedy et al., 2015]. The Xception architecture prescribes the number of convolution kernels and output channels in each layer as well as the connections between layers. Some key differences between our model and the standard Xception architecture include a batch normalization layer in front of the network, an extra SE layer after every residual block in the middle portion of the architecture, and the addition of a dropout layer at the end of the network, which we included in place of the L2 weight regularization used in the original Xception model Chollet [2017].

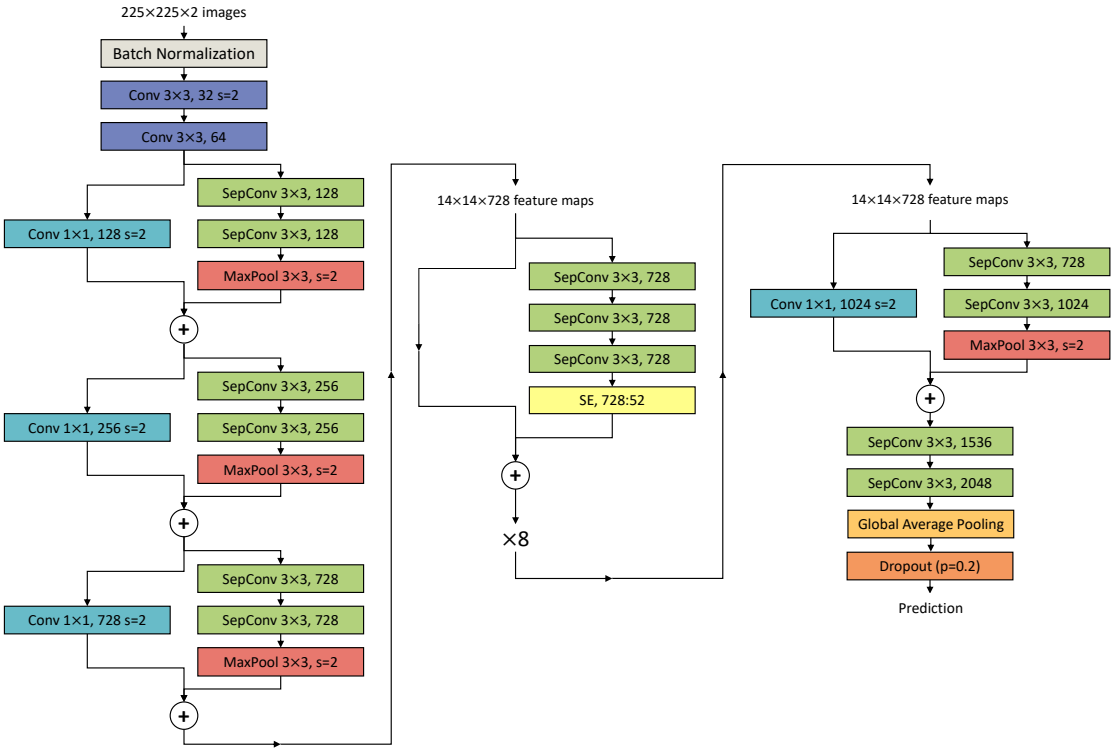

Figure 5.7: Architecture of the final model presented in this work. This figure is adapted from Figure 5 of Chollet [2017]. Batch normalization and activation (ReLu) layers that follow each convolution and separable convolution layer are omitted from the diagram. Data flow follows the arrows. The middle portion of the network is repeated to create 8 identical sections. For each layer, we list the name, the kernel size, and the number of output channels. Layers with a stride length of 2 instead of 1 are distinguished by "s=2." The reduction ratio (14) of the SE layer is presented as "728:52", which denotes the number of input and output channels as a ratio of the number of hidden layer channels [Hu et al., 2018].

We trained our final model for 25 epochs, where each epoch was a full pass of all 1,000,000 samples in the training data through the neural network. Model loss was calculated using binary cross-entropy, which is a standard loss function for binary classification problems that measures how well the predicted class probabilities match the target class. The model performance was monitored by

calculating performance metrics (Section 5.5.2) using the full train–dev and validation sets at the end of every epoch. The training was carried out on a single ML-enabled graphical processing unit (GeForce RTX 2060 SUPER 8 GB GPU) and took approximately 112 hours (~ 4.5 days) days to complete.

All of our models were implemented using TensorFlow [Abadi et al., 2015], and the source code to reproduce our final model is available online.

5.5 Results

5.5.1 Model Evaluation

Although we allowed some tolerance on the calculation of the extrapolated detection frequencies in the training set (Section 5.3.4.3), we found that a portion of the RFI signals in the validation set were still misclassified as valid technosignature candidates because the signal in the bottom image was not properly centered. Similarly, we found that a subset of signals were misclassified because the S/N difference between the top and bottom images was too large. These discrepancies are not surprising because our training data was generated by splitting a single scan into two parts, while our test data contains signals from two completely different scans. As a result, errors on the extrapolated detection frequencies as well as the S/N variability are not as pronounced in the training data as they are in the test data.

In order to address these issues, we applied several additional steps when evaluating the model on the validation, test, and production data. First, we evaluated the model multiple times for each image pair in the validation and test sets, applying a pixel shift in the range -4 to 4 to the bottom image each time. The largest of the resulting 9 values was chosen as the score for that data point. The range of pixel shifts used for this step was chosen by running this test with a larger set of pixel shift values and choosing a symmetric range that yielded the largest scores for $\sim$95% of the validation data. We found that this step increased the validation recall from 0.859 to 0.942 for a

total decrease of less than 0.1% in the validation set precision. Then, if the score after this step was still below the decision threshold value of 0.5, we rescaled both images so that the new pixel values ranged from zero to the average of the maximum pixel values of both images prior to scaling. We found that this step further increased the validation recall from 0.942 to 0.992, while retaining a validation precision of $> 99\%$.

5.5.2 Model Performance

We used several metrics to evaluate the performance of our model. First, we evaluated the precision and recall scores, which are defined in Section 5.4.1. We also calculated the F_1 score, which is defined as

$$\mathrm{F}_1 = 2 \times \frac{P \times R}{P + R}, \tag{5.2}$$

where P and R are the precision and recall, respectively. Another important metric often used for model performance evaluation is the *area under the curve* (AUC) score. In this context, the "curve" is the *receiver operating characteristic* (ROC) curve, which plots the true positive rate against the false positive rate (Figure 5.8). We also calculated the area under the precision-recall curve (AUPRC) as well as the average precision (AP) of the model, which is the precision averaged over all recall values. Together, these metrics offer a thorough picture of the performance of all models considered in this work. Table 5.3 lists the values of these metrics for the baseline model, the existing DoO filter, and our trained CNN.

We find that our CNN significantly outperforms the baseline model, with a 99.15% precision at a recall of 97.81%, compared to a baseline model precision of 95.08% at a recall of only 33.72%. The CNN also performs favorably with respect to the existing DoO filter. Although the DoO filter did not admit any false positives (100% precision) over the set of 1,238 hand-labeled signals (Section 5.3.2), its recall scored at only 80.68%. This translates to a significant portion of signals left over for manual inspection after application of the filter. The balance between precision and recall can be summarized with the F_1, AUC, or AUPRC score, all of which favor the CNN over the

baseline model or direction of origin filter. A visualization of the performance differential between the baseline model, the DoO filter, and the CNN is exemplified in Figure 5.8, which plots the ROC curve for each model.

We note that $\sim$36% of the misclassifications on the test set are attributable to signals detected in scans of TRAPPIST-1 and LHS 1140. This percentage is appreciably larger than the total fraction of signals from these two sources in the full test set ($\sim$23%). This disproportionate distribution may be due in part to the much larger sky separation between this source pairing (26.3°) compared to the typical angular separations between almost all other source pairings (1–3°), including the other source pairings in this observation run. Specifically, the time between the TRAPPIST-1 and LHS 1140 data acquisitions is approximately twice as large as the mean time between data acquisitions of the other source pairs, a unique circumstance that was driven by the desire to observe noteworthy exoplanets that had been recently discovered at the time. The increased time between data acquisitions of these two sources accentuates any errors in the frequency extrapolation that we used to center the signals in the second image. These errors account for a major failure mode discussed in Section 5.6.1.2. The increased error rate of the large-separation pairing suggests that such pairings should be avoided when practical. This information can help guide the design of future observing plans.

5.5.3 Application to Observational Data

We applied the trained model to a subset of the data presented by Margot et al. [2018], Pinchuk et al. [2019], and Margot et al. [2021]. Specifically, we evaluated the model on signals that passed the drift rate, S/N, and bandwidth data selection filters described in Section 5.3.3 as well as the existing DoO filter. Table 5.4 shows the total candidate signal counts (from the first scan of each source only) before and after application of the direction-or-origin filter and the CNN-based filter described in this work. Although the existing DoO performs remarkably well already, we found that the CNN can further reduce the number of signals left over to examine by a factor of 6-16 in nominal situations. In the atypical data set with unusually large angular separations between

sources, the reduction factor decreased to $\sim$3.

We did not evaluate the CNN on the 3% of signals that did not pass our data filters (Section 5.3.3) because we anticipate poor classification performance on these signals as they are outside of the domain of applicability for the network.

The evaluations took approximately 5, 18, and 6 hours on the Margot et al. [2018], Pinchuk et al. [2019], Margot et al. [2021] data sets, respectively, on a single ML-enabled GeForce RTX 2060 SUPER 8 GB graphical processing unit, i.e., several times slower than data acquisition. This performance is promising with respect to near-real-time data processing: the evaluation could keep up with data acquisition with the addition of one or more high-end graphical processing units.

5.6 Discussion

5.6.1 Failure Modes

In this section, we examine some of the CNN failure modes that we identified by examining the test set samples that the CNN misclassified.

5.6.1.1 Model-related Failure Modes

One set of failure modes fall under the category of model-related failures, which stem from the model's inability to learn an adequate representation of the data and therefore correctly classify a subset of signal types.

One such failure mode occurred when the S/N of the signal in a data sample with S/N $<\sim 30$ was lower in the top image compared to the bottom image. In these cases, the network would assign the pair of images a label of "0" when there is clearly a signal present in both. Figure 5.9 shows an example of such an image pair. Note that we did not find any evidence for the reverse failure mode – when S/N of the signal in a data sample is larger in the top image compared to the bottom image. Although we did attempt to introduce S/N variations between the two images

in each sample of the training set (Section 5.3.4.2), this failure mode suggests that we needed to allow even larger variations, specifically including cases where the S/N lower than the S/N of the signal in the bottom image.

5.6.1.2 Failure Modes Related to Simplifying Assumptions

A different set of failure modes stemmed from some of the simplifying assumptions that we made about the data. For example, one of the failure modes is related to the frequency extrapolation that we performed in order to center the signal in the second image. We assumed that the frequency drift rate would be linear, and we assumed that our estimate of the frequency drift rate would be accurate enough to ensure centering of the signal in the second image within a tolerance of $\sim$15 Hz. Although we included this tolerance directly into the model, both during training (Section 5.3.4.3) and evaluation (Section 5.5.1), we still found cases where the signal was clearly present in the second image but was not properly centered. Figure 5.10 (left) shows an example of such a signal from our test set. In this case, the model gave the sample a score of 0.0193. This score yields a label of "0" (i.e., no signal in the second scan) because it is below the decision threshold of 0.5. However, if we shift the bottom image 5 pixels ($\sim$15 Hz) to the left, the score jumps to 0.7545, which yields the correct label of "1." Shifts of 6-10 pixels to the left all yield scores > 0.99.

Unfortunately, this problem cannot be fixed by simply increasing the range of shifts allowed for the bottom image. In fact, it is likely that the same problem persists for any choice of the tolerance on frequency to accept/reject a match. More importantly, increasing the range of allowed shifts for the bottom image would also increase the risk of removing a technosignature candidate by pairing it with RFI detected in its vicinity. Instead, this problem can be better addressed by obtaining more accurate representations of the detected signal properties, which can then be used to more accurately localize and thus center the signal in a subsequent scan.

Another failure mode is related to the simplifying assumption that the signal power in the first half of a scan is comparable to the signal power in the second half of a scan. The input to the CNN

is limited to approximately half of each scan, assuming the data taking parameters of Pinchuk et al. [2019], by virtue of the training set parameters. Specifically, the input images are limited to a size of 225×225 pixels, which corresponds to ∼75 seconds of observation, whereas each scan typically lasts for a total of ∼150 seconds. This limitation significantly hinders the network's ability to identify RFI, because the CNN only examines the first temporal half of each scan, but there are instances where a signal is only present in the latter half of one or both scans. An example of this case is shown in Figure 5.10 (middle).

We performed three preliminary attempts at mitigating this issue. First, we tested the possibility of downsampling the scan in the time dimension by a factor of two, effectively allowing us to fit partial information from 450 rows (∼150 seconds) into 225 pixels along the time dimension. Second, we tested an image rescaling approach consisting of a linear interpolation [Virtanen et al., 2020] of the entire scan duration that was sampled at 225 equally-spaced time intervals. Neither approach reduced the number of signals left to examine after application of the filter, indicating that the issue persisted. In a third attempt, we applied the filter a total of four times to each set of scans in the test data set. Each filter evaluation paired a different set of temporal scan halves ($\text{scan}_{1,\text{top}}$ with $\text{scan}_{2,\text{top}}$, $\text{scan}_{1,\text{top}}$ with $\text{scan}_{2,\text{bottom}}$, etc.). We combined the results of these evaluations by taking the maximum score across the four trials. We found that this method did increase the recall score for the test set from 97.81% up to 98.91%. However, the precision score was heavily penalized, decreasing from 99.15% down to 97.92%. This trade-off increases the likelihood of finding additional pairings and therefore false positives. Taking the median score across the four trials yielded similar results. This four-execution mitigation attempt also increased the computational cost of the CNN filter by a factor of four. For these reasons, we did not apply this method when evaluating the CNN on observational data. Further investigation beyond the scope of this work is required to minimize the impact of this failure mode.

5.6.1.3 Other Failure Modes

The final failure mode that we observed is related to instances of human error in labeling the validation and test sets. Because the labels were supplied by a single classifier (PP), the margin for error on the validation and test labels is nonzero. Figure 5.10 (right) shows an example of a test signal that the CNN "misclassified". Upon further investigation, it is clear that the label provided with this data sample is incorrect. Although the network technically classified this signal correctly, it counted as a misclassification when computing model performance (Section 5.5.2). This problem could be substantially mitigated if multiple people examined and labeled the validation and test data.

5.6.2 Future Improvements

Though we have attempted to thoroughly search the parameter space for the best model to perform our classification task, there are still a number of options to consider for future improvements. For example, for all CNN models considered in this work, the input was comprised of $225 \times 225 \times 2$ images, where the last dimension distinguished the top half of the first scan from the top half of the second scan. An alternative approach would pass the data as a single 450×225 image, where the top halves of each scan are concatenated in the time dimension. It is worth investigating whether or not this variant on the input data improves network classification performance. Along the same lines, it may be beneficial to train a denoising auto–encoder [e.g., Xiang and Pang, 2018, and references therein] and apply it to the images prior to sending them through the CNN. If the denoising auto–encoder functions properly (i.e., reduces the noise around the signals in the image), it is likely that the CNN would receive a boost in classification performance.

During our model selection step (Section 5.4.2), we found that standard network architectures always outperformed their Siamese variants. However, those tests were performed without any hyperparameter tuning, so it may be worthwhile to investigate whether a tuned Siamese model still underperforms when compared to the base architecture model. On top of that, new state-of-the art

CNN architectures are still being rapidly developed and may offer significant improvements over the Xception architecture used as the final model in this work. For example, the novel EfficientNet architecture [Tan and Le, 2019], which was published after our model selection efforts, is almost an order of magnitude smaller and faster than other CNN architectures, yet has been shown to exhibit state-of-the-art performance on the ImageNet data.

Finally, there are some improvements that can be made to the overall training and evaluation process to mitigate the various failure modes discovered after evaluating the CNN on the test data. These improvements are included with the corresponding description of each failure mode in Section 5.6.1.

5.7 Conclusions

In this work, we designed a DoO filter using modern computer vision techniques to assist in the mitigation of RFI in the search for radio technosignatures. We began by randomly selecting 1,100,000 signals from a carefully selected set of over 8 million detections in order to obtain the cleanest training and train-dev data set possible. Both of these data sets consist of pairs of images that were obtained by splitting a single scan containing a signal into two parts. This approach allowed us to label a large amount of signals in a short period of time.

Using these data sets, we trained and evaluated a CNN designed to determine whether or not the signal in the first image is also present in the second image. This network can therefore be applied to determine if a detected signal is persistent in one and only one direction on the sky. This approach is similar to the one employed by traditional DoO filters, except that the CNN analyzes the dynamic spectra directly instead of relying on inferred signal properties, such as frequency and frequency drift rate.

We found that the CNN trained in this work outperformed both the baseline 2D correlation model and the existing DoO filters, with a precision value of 99.15% at a recall of 97.81%. We find that the CNN can reduce the number of signals left to analyze after applying the existing DoO

filter by a factor of 6-16 in nominal situations. In the atypical data set with unusually large angular separations between sources, the reduction factor decreased to $\sim$3.

We identified several failure modes of the trained network, labeling failures, and failures related to simplifying assumptions. Each failure mode can be addressed with future CNN versions to increase the classification performance. Integrating this ML-based DoO filter into existing radio technosignature search pipelines has the potential of providing accurate RFI identification in near-real-time.

5.A Sample training signals

We validated our choice for the threshold of a top-half S/N < 6 (Section 5.3.4.2) by examining a sample of signals below this cutoff value. Figure 5.11 depicts these signals, most of which are faint and difficult to detect visually. Similarly, Figure 5.12 depicts a sample of signals above this threshold, which are clearly visible in the dynamic spectra.

Figure 5.13 shows a sample of signals from the negative category with a $P_{\mathrm{bottom}}/P_{\mathrm{top}}$ ratio of 0.2 or lower and prominence value below 3 standard deviations of the noise (Section 5.3.4.2).

5.B Siamese models

The concept of a "Siamese" neural network was first introduced in 1993 by Bromley et al. [1993] for the purpose of signature verification. Siamese networks are defined as two identical sub–networks that are joined at the output, typically by subtracting the neuron values of the final layer of one model from the neuron values of the final layer of the other model. The input to these networks always consists of two data points, each of which are passed to one of the two sub–networks. The output of the Siamese networks is typically given as a similarity score between the two data points.

During the model selection portion of this work (see Section 5.4.2), we set the two identical

sub–networks of each Siamese network to be one of the architectures under consideration (Figure 5.14). Each sub-network received one scan as input. For each architecture, we tested two methods of joining the outputs of the final layers of the Siamese sub–networks. Specifically, we considered the standard method of subtracting the values of one output from the other, as well as a generalized version of this procedure. For the latter, we concatenated the output weights of both sub–networks and added another fully connected layer with N nodes immediately after the concatenated layer, where N is the number of neurons in the output layer of the sub–network. This method is a generalized version of the subtraction procedure because it can be recovered by setting the weights w_{ij} between the two layers to be

$$w_{ij} = \begin{cases} 1, & \text{if } i = j \\ -1, & \text{if } i = j + N \\ 0, & \text{otherwise} \end{cases} \tag{5.3}$$

where i and j represent the indices of the neurons of the concatenated and the fully connected layer, respectively.

Although Siamese networks seem like a promising solution to the problem of pairing signals from two different scans, we found that the standard network architectures always outperformed their Siamese variants (see Section 5.4.2).

Table 5.2: Hyperparameters that were considered in this work, as well as the set of possible values for each and the final value used for training the model. For a definition of these concepts, see Géron [2019]

Hyperparameter	Possible values	Final Value
Optimizer	{Stochastic Gradient Descent, RMSProp, Adam, Nadam, AdaMax}	Nadam
Learning rate	$\{1e^{-3}, 1e^{-4}, 1e^{-5}, 1e^{-6}\}$	$1e^{-3}$
Batch size	{16, 32, 64, 128}	16
Activation function	{ReLu, Swish}	ReLu
Fully connected layers on top	{True, False}	False
Dropout rate	{None, 0.2, 0.5}	0.2
Include Squeeze-and-Excitation blocks	{True, False}	True
Include input batch normalization	{True, False}	True

Table 5.3: Scoring metrics for the baseline model, DoO filter, and CNN validation and test sets.

Metric	Baseline Model	Direction-of-origin	CNN Validation	CNN Test
Precision	0.9508	1.0000	0.9973	0.9915
Recall	0.3372	0.8068	0.9919	0.9781
F_1	0.4979	0.8931	0.9946	0.9848
AUC	0.7324	0.9034	0.9951	0.9811
AUPRC	0.6914	0.9975	0.9998	0.9982
AP	0.6921	0.9950	0.9998	0.9991

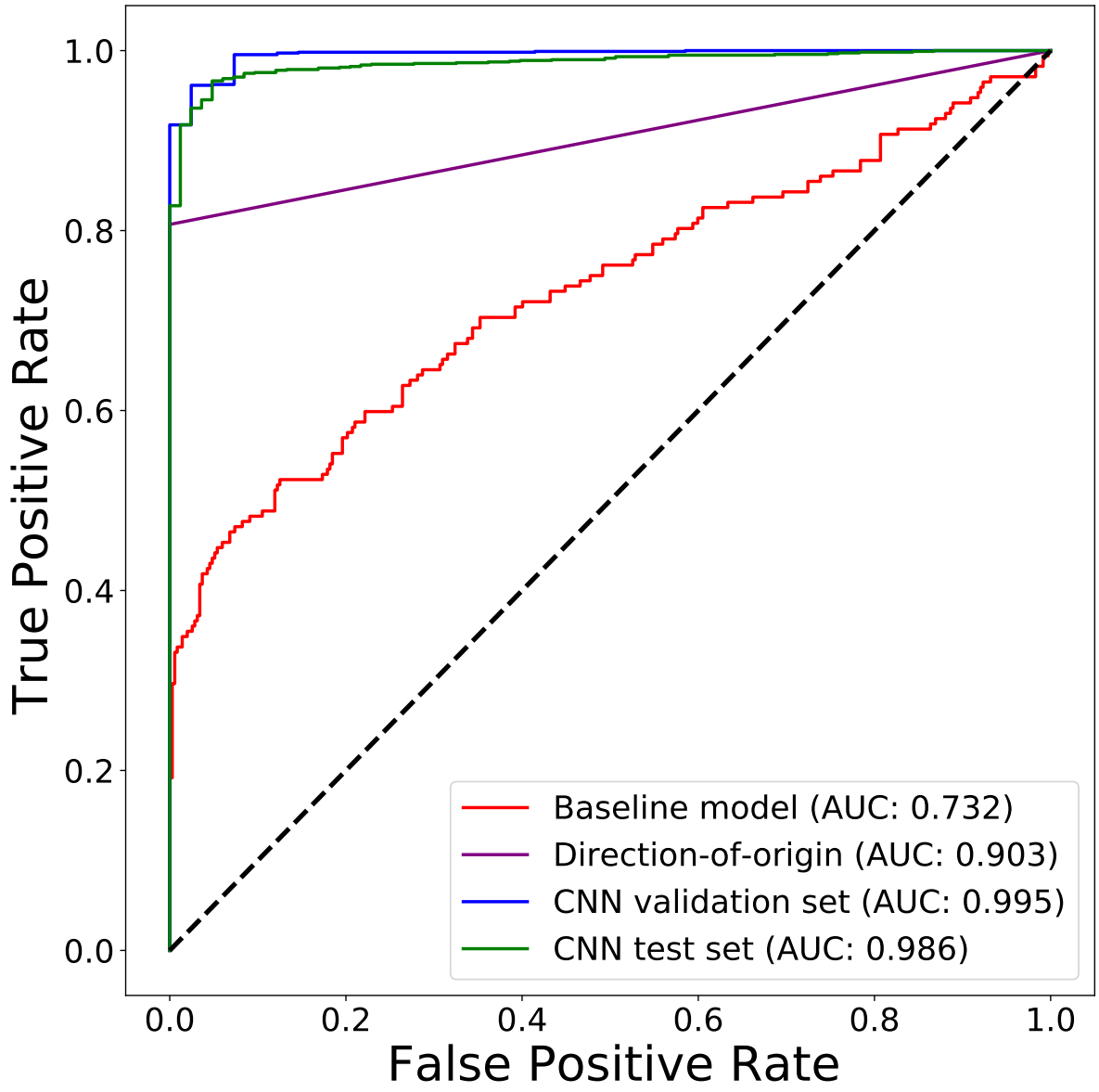

Figure 5.8: ROC curve and AUC scores for the baseline 2D correlation model, the existing DoO filter, as well as the CNN evaluated on both the validation and the test set. The DoO curve is linear because the filter only outputs binary scores of 0 or 1, unlike the other models, which output a score in the range from 0 to 1 for each sample. The dashed line shows the ROC curve for a purely random classifier with an AUC score of 0.5.

Table 5.4: Comparison of filter performance across different data sets. The "Data Set" column includes the journal and online data references for the data sets. The "Total Signals" column lists the total number of signals from the first scan of each source. The "Applicable Signals" column lists the total number of signals from the first scan of each source with an S/N between 10 and 600, a drift rate in the range ± 2 Hzs^{-1}, and a bandwidth with FWHM ≤ 100 Hz. The "DoO" column lists the number of candidate signals remaining after application of the existing DoO filter to the subset of signals from the "Applicable Signals" column. The final column lists the candidate technosignature counts after application of the CNN-based filter described in this work to the subset of signals that passed the existing DoO filter.

Data Set	Total Signals	Applicable Signals	DoO	DoO + CNN
UCLA search 2016 [Margot et al., 2018, Margot et al., 2020a]	2,230,659	2,142,964	16,168	2,772
UCLA search 2017 [Pinchuk et al., 2019, Margot et al., 2020b]	2,973,499	2,888,766	62,301	20,560
UCLA search 2018–9 [Margot et al., 2021, 2020c]	12,779,984	12,438,375	21,978	1,357

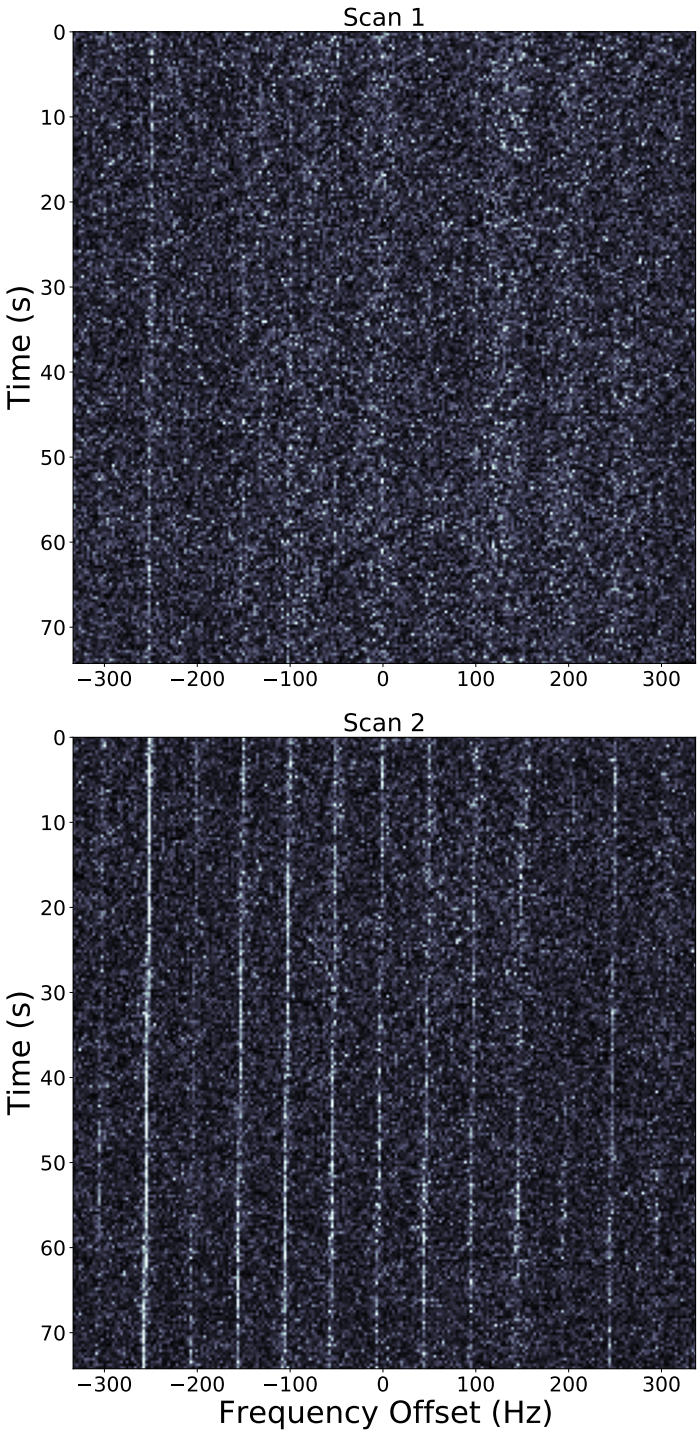

Figure 5.9: Sample signal with lower S/N in the top image compared to the bottom image. The CNN score for this image pair is 0.2706, which corresponds to a label of "0".

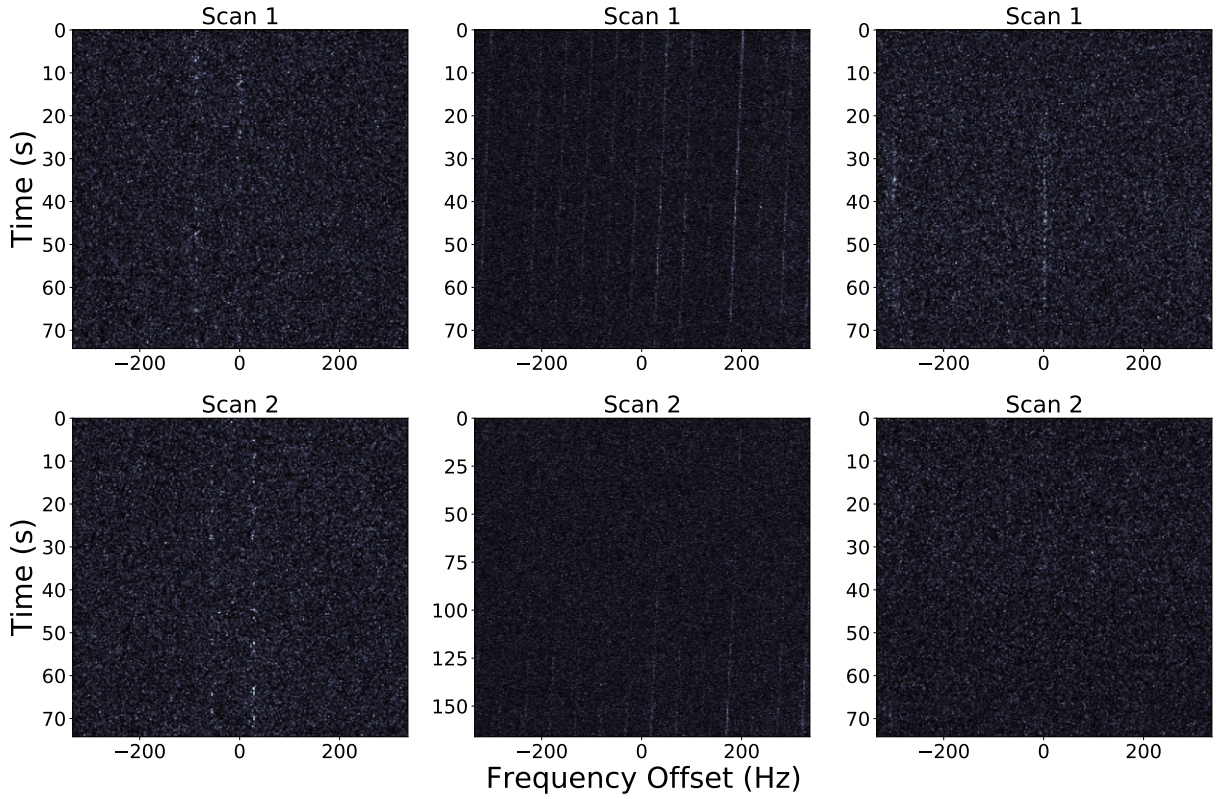

Figure 5.10: (Left) Example image pair where the signal in the second image is not centered. This occurs when the properties of the signal in the top image are inaccurately determined. (Middle) Example of a signal that does not appear until the second half of the scan in the second image. The standard application of the network mislabels this signal because the CNN looks at the top half of each scan only. (Right) Example signal that was incorrectly hand-labeled as "1", seemingly indicating that it contains a signal in the second image.

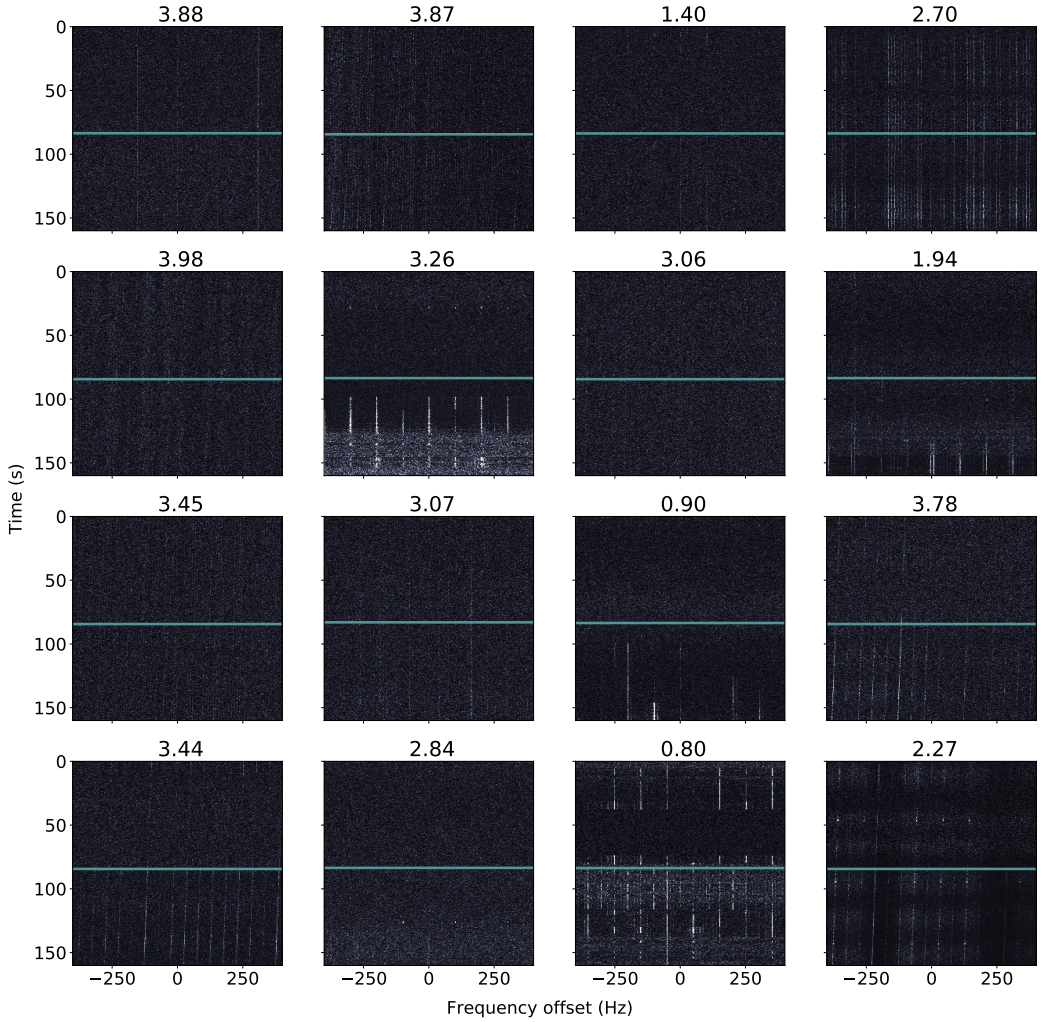

Figure 5.11: Example dynamic spectra of signals with a top-half S/N < 6 (exact values shown above each plot). The horizontal blue line delimits the top and bottom halves. Note that these signals (located in the center of the image starting at 0 Hz offset at time $t = 0$) are faint in the top half of the image and difficult to detect visually.

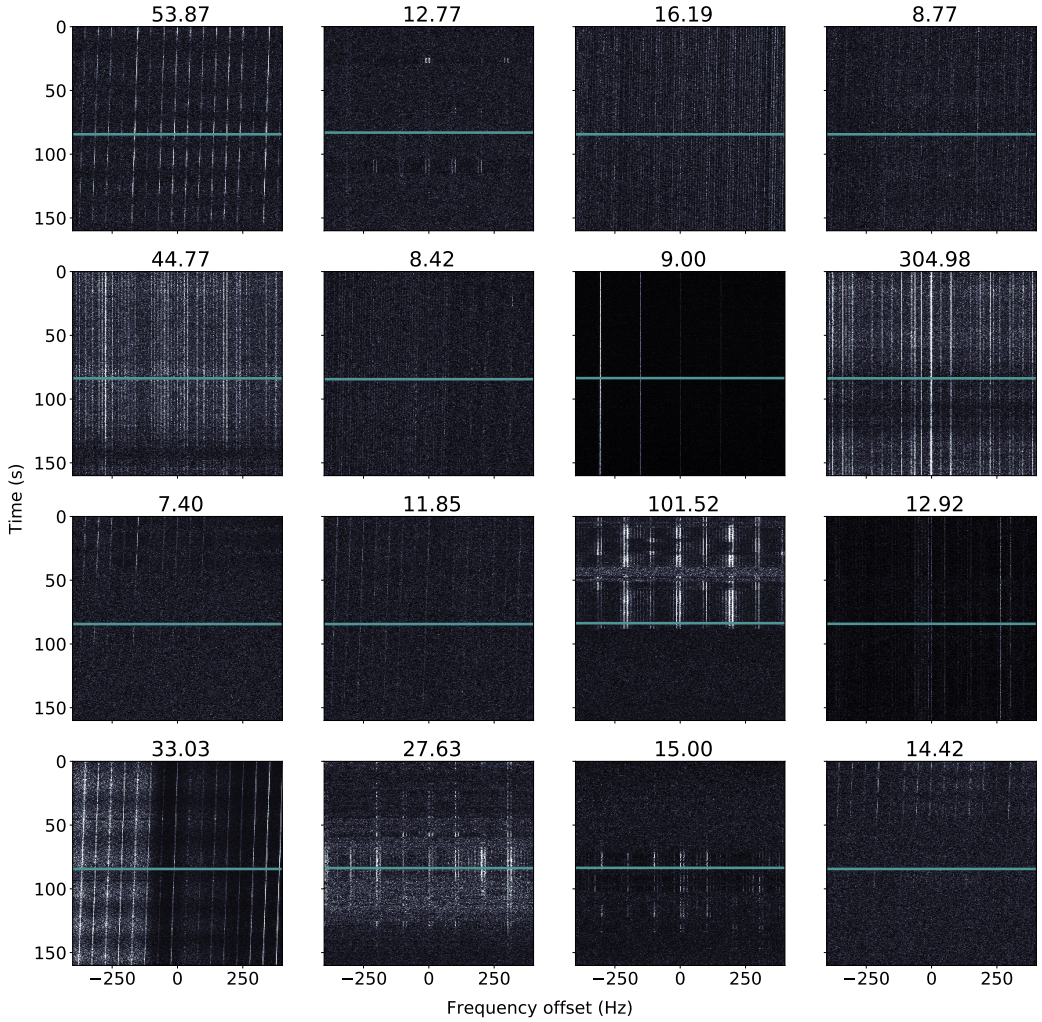

Figure 5.12: Example dynamic spectra of signals with a top-half S/N $\geq$ 6 (exact values shown above each plot). The horizontal blue line delimits the top and bottom halves. Note that all of these signals (starting at 0 Hz offset at time $t = 0$) are visually detectable in the top half of each sample.

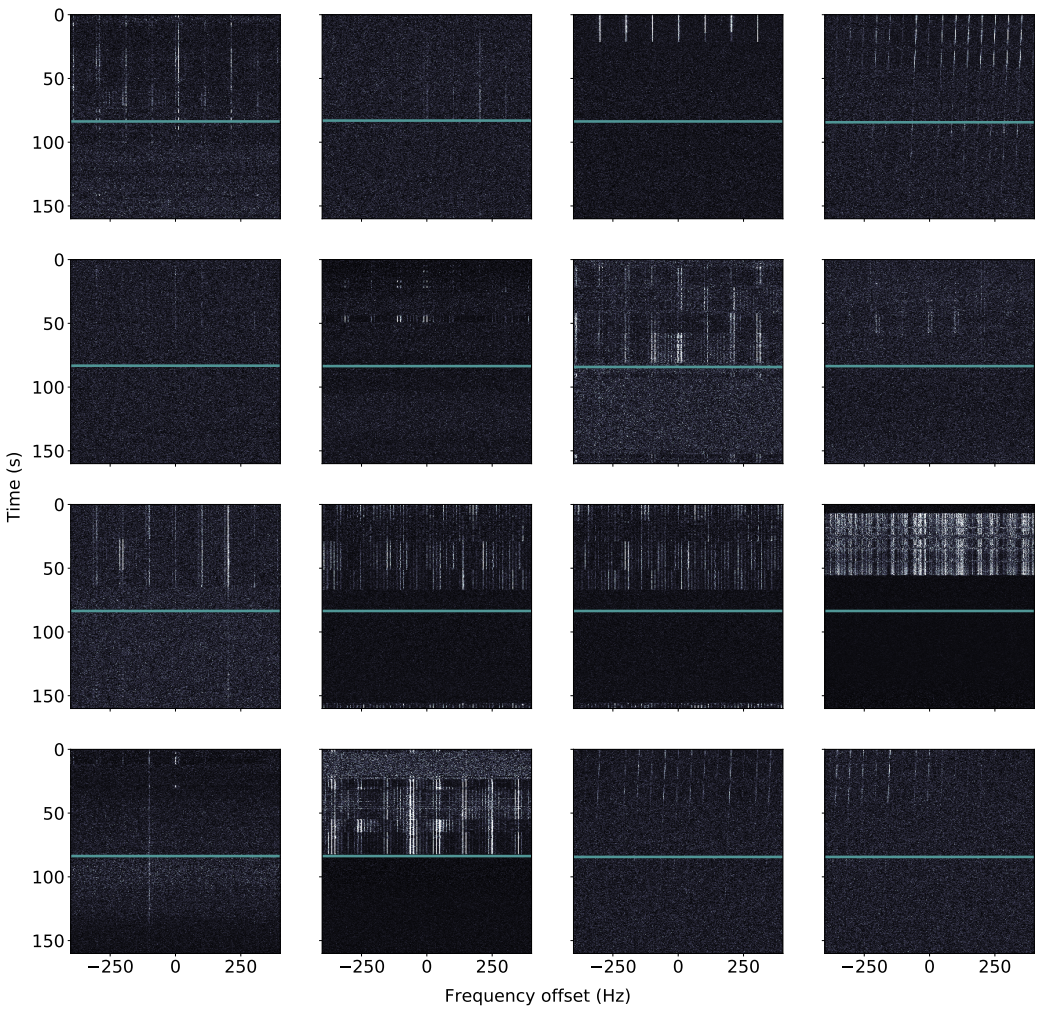

Figure 5.13: Example time-frequency diagrams of signals with a $P_{\mathrm{bottom}}/P_{\mathrm{top}}$ ratio of 0.2 or lower. The horizontal blue line delimits the top and bottom halves. No signals with a prominence value greater than 3σ are present in the bottom halves. All of these signals represent valid negative samples.

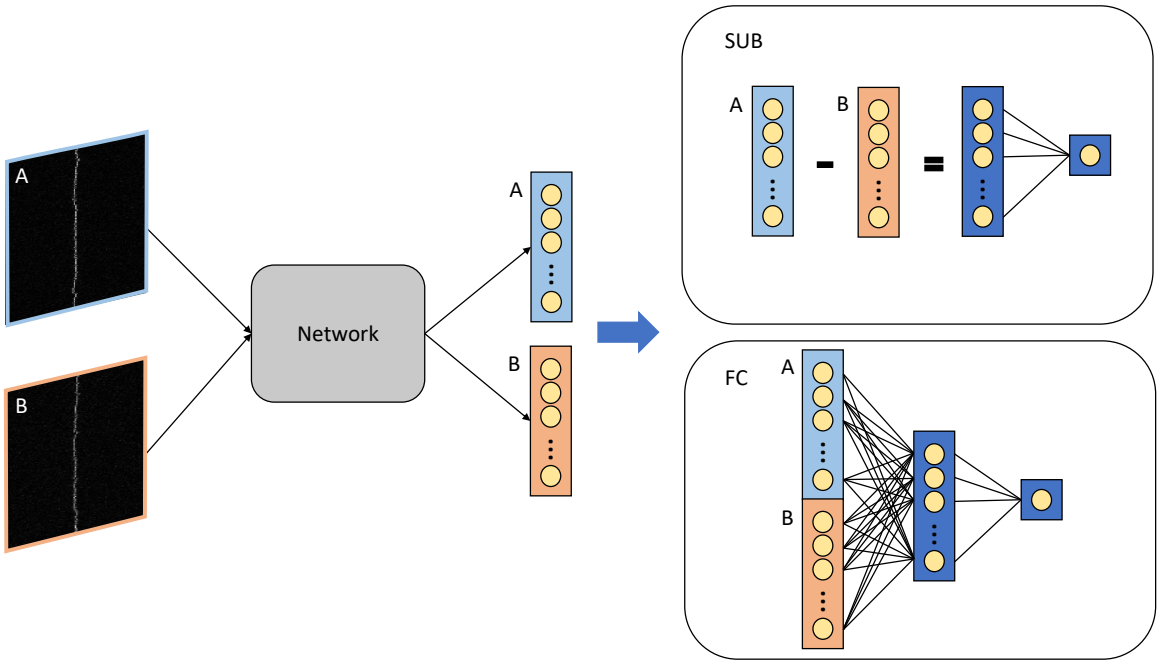

Figure 5.14: Example of a Siamese network tested in this work. The labels "A" and "B" represent input from two different scans. The "Network" in the middle was replaced with one of the architectures that was tested in this work (see Section 5.4.2). The output layers were joined in two different ways: (Top Right) In standard Siamese networks, the output layers are subtracted. (Bottom Right) In our generalized version, the output layers are concatenated and connected to another layer with N neurons. Equation 5.3 gives the set of weights for this configuration that reproduce the standard layer subtraction procedure.

CHAPTER 6

Conclusions

In this work, I described significant advancements to the signal detection and Radio Frequency Interference (RFI) identification capabilities of modern radio technosignature detection algorithms. These improvements were presented alongside the results of the analysis of four annual UCLA radio technosignatures searches spanning 2016–2019. These searches compare favorably to other recent searches in terms of end-to-end sensitivity, frequency drift rate coverage, and signal detection count per unit bandwidth per unit integration time. The searches focused on detecting signals that are narrow in the frequency domain, which would provide compelling evidence of the existence of another civilization. Although no technosignatures have been discovered to date, the tools and algorithms used to analyze the radio baseband data have undergone multiple levels of improvements.

I began by describing the UCLA SETI Group's initial versions of the signal detection and RFI identification algorithms, which identified approximately 850 000 candidates, corresponding to a hit rate density of 1.691×10^{-3} hits per hour per hertz (see Section 4.5 for a definition of hit rate density), and automatically classified most (99%) of these signals as human-generated RFI. A large fraction (>99%) of the remaining candidate signals were also flagged as anthropogenic RFI because they have frequencies that overlap those used by global navigation satellite systems, satellite downlinks, or other interferers detected in heavily polluted regions of the spectrum.

Then, I described an improved candidate signal detection algorithm that detected approximately 6 million signals in a 2017 search for technosignatures with identical observational parameters. The improved data processing pipeline yielded a hit rate density of 1.07×10^{-2} hits per

hour per hertz, a significant improvement over the inital version of the pipeline. The improved candidate signal detection algorithm helps mitigate problems associated with the common practice of ignoring frequency space around candidate detections in radio technosignature detection pipelines, which include inaccurate estimates of figures of merit and unreliable upper limits on the prevalence of civilizations.

Afterwards, I presented further improvements to these detection algorithms along whit their application to a search for technosignatures in regions surrounding 31 Sun-like stars near the plane of the Galaxy. The improved candidate signal detection procedure nearly doubles the signal detection count of some previously analyzed data sets. With a hit rate density of 2.21×10^{-2} hits per hour per hertz, these algorithms significantly outperformed the algorithms used previously by the UCLA SETI Group as well as those used by the Breakthrough Listen (BL) team. I also described an improvement to the direction-of-origin filters that promotes unique links between signals observed in separate scans.

Finally, I discussed the implementation of a novel machine-learning-based RFI mitigation algorithm, which helped address a major remaining challenge in the search for radio technosignatures. The newly-designed Convolutional Neural Network (CNN) can determine whether or not a signal detected in one scan is also present in another scan. This CNN-based DoO filter outperforms both a baseline 2D correlation model as well as existing DoO filters over a range of metrics. Importantly, the CNN reduced the number of signals requiring visual inspection after the application of traditional DoO filters by a factor of 6-16 in nominal situations. Integrating this ML-based DoO filter into existing radio technosignature search pipelines has the potential of providing accurate RFI identification in near-real-time.

Bibliography

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL https://www.tensorflow.org/. Software available from tensorflow.org.

Dalya Baron. Machine learning in astronomy: A practical overview. arXiv preprint arXiv:1904.07248, 2019.

N. M. Batalha. Exploring exoplanet populations with NASA's Kepler Mission. Proceedings of the National Academy of Science, 111:12647–12654, 2014. doi: 10.1073/pnas.1304196111.

W. J. Borucki. KEPLER Mission: development and overview. Reports on Progress in Physics, 79 (3):036901, 2016. doi: 10.1088/0034-4885/79/3/036901.

William J Borucki. Kepler mission: development and overview. Reports on Progress in Physics, 79(3):036901, 2016.

Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. Signature verification using a" siamese" time delay neural network. Advances in neural information processing systems, 6:737–744, 1993.

Bryan Brzycki, Andrew P. V. Siemion, Steve Croft, Daniel Czech, David DeBoer, Julia DeMarines, Jamie Drew, Vishal Gajjar, Howard Isaacson, Brian Lacki, and et al. Narrow-band signal localization for seti on noisy synthetic spectrogram data. Publications of the Astronomical Society

of the Pacific, 132(1017):114501, Sep 2020. ISSN 1538-3873. doi: 10.1088/1538-3873/abaaf7. URL http://dx.doi.org/10.1088/1538-3873/abaaf7.

François Chollet. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1251–1258, 2017.

Jessie L. Christiansen, Bruce D. Clarke, Christopher J. Burke, Jon M. Jenkins, Thomas S. Barclay, Eric B. Ford, Michael R. Haas, Anima Sabale, Shawn Seader, Jeffrey Claiborne Smith, Peter Tenenbaum, Joseph D. Twicken, Akm Kamal Uddin, and Susan E. Thompson. Measuring Transit Signal Recovery in the Kepler Pipeline. I. Individual Events. 207(2):35, 2013. doi: 10.1088/0067-0049/207/2/35.

R. J. Cohen, G. Downs, R. Emerson, M. Grimm, S. Gulkis, G. Stevens, and J. Tarter. Narrow polarized components in the OH 1612-MHz maser emission from supergiant OH-IR sources. Monthly Notices of the RAS, 225(3):491–498, 1987. doi: 10.1093/mnras/225.3.491.

G. A. Cox, S. Egly, G. R. Harp, J. Richards, S. Vinodababu, and J. Voien. Classification of simulated radio signals using wide residual networks for use in the search for extra-terrestrial intelligence, 2018.

Jason A Dittmann, Jonathan M Irwin, David Charbonneau, Xavier Bonfils, Nicola Astudillo-Defru, Raphaëlle D Haywood, Zachory K Berta-Thompson, Elisabeth R Newton, Joseph E Rodriguez, Jennifer G Winters, et al. A temperate rocky super-earth transiting a nearby cool star. Nature, 544(7650):333–336, 2017.

F. Drake. SETI Science Working Group Report, chapter Estimates of the Relative Probability of Success of the SETI Search Program, pages 67–69. Number TP-2244 in NASA Technical Paper. 1984.

F. Drake. The search for extra-terrestrial intelligence. Philosophical Transactions of the Royal Society of London Series A, 369(1936):633–643, February 2011. doi: 10.1098/rsta.2010.0282.

F. D. Drake. The Radio Search for Intelligent Extraterrestrial Life, pages 323–345. 1965.

R. DuPlain, S. Ransom, P. Demorest, P. Brandt, J. Ford, and A. L. Shelton. Launching GUPPI: the Green Bank Ultimate Pulsar Processing Instrument. In Advanced Software and Control for Astronomy II, volume 7019 of Proceedings of the SPIE, page 70191D, August 2008. doi: 10.1117/12.790003.

J. E. Enriquez, A. Siemion, G. Foster, V. Gajjar, G. Hellbourg, J. Hickish, H. Isaacson, D. C. Price, S. Croft, D. DeBoer, M. Lebofsky, D. H. E. MacMahon, and D. Werthimer. The Breakthrough Listen Search for Intelligent Life: 1.1-1.9 GHz Observations of 692 Nearby Stars. Astrophysical Journal, 849:104, 2017. doi: 10.3847/1538-4357/aa8d1b.

Julia Fang and Jean-Luc Margot. ARCHITECTURE OF PLANETARY SYSTEMS BASED ONKEPLERDATA: NUMBER OF PLANETS AND COPLANARITY. The Astrophysical Journal, 761(2):92, nov 2012. doi: 10.1088/0004-637x/761/2/92. URL https://doi.org/10.1088/0004-637x/761/2/92.

Matteo Frigo and Steven G. Johnson. The design and implementation of FFTW3. Proceedings of the IEEE, 93(2):216–231, 2005. Special issue on "Program Generation, Optimization, and Platform Adaptation".

Gaia Collaboration. The Gaia mission. 595:A1, 2016. doi: 10.1051/0004-6361/201629272.

Gaia Collaboration. Gaia Data Release 2. Summary of the contents and survey properties. 616: A1, 2018. doi: 10.1051/0004-6361/201833051.

Vishal Gajjar, Karen I Perez, Andrew PV Siemion, Griffin Foster, Bryan Brzycki, Shami Chatterjee, Yuhong Chen, James M Cordes, Steve Croft, Daniel Czech, et al. The breakthrough listen search for intelligent life near the galactic center i. arXiv preprint arXiv:2104.14148, 2021.

D. Gale and L. S. Shapley. College admissions and the stability of marriage. 69(1):9–15, 1962. ISSN 00029890, 19300972.

Aurélien Géron. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, 2019.

Michaël Gillon, Amaury HMJ Triaud, Brice-Olivier Demory, Emmanuël Jehin, Eric Agol, Katherine M Deck, Susan M Lederer, Julien De Wit, Artem Burdanov, James G Ingalls, et al. Seven temperate terrestrial planets around the nearby ultracool dwarf star trappist-1. Nature, 542 (7642):456–460, 2017.

R. H. Gray and K. Mooley. A VLA Search for Radio Signals from M31 and M33. Astronomical Journal, 153:110, 2017. doi: 10.3847/1538-3881/153/3/110.

G. R. Harp, J. Richards, J. C. Tarter, J. Dreher, J. Jordan, S. Shostak, K. Smolek, T. Kilsdonk, B. R. Wilcox, M. K. R. Wimberly, J. Ross, W. C. Barott, R. F. Ackermann, and S. Blair. SETI Observations of Exoplanets with the Allen Telescope Array. Astronomical Journal, 152:181, 2016. doi: 10.3847/0004-6256/152/6/181.

G. R. Harp, Jon Richards, Seth Shostak Jill C. Tarter, Graham Mackintosh, Jeffrey D. Scargle, Chris Henze, Bron Nelson, G. A. Cox, S. Egly, S. Vinodababu, and J. Voien. Machine Vision and Deep Learning for Classification of Radio SETI Signals. art. arXiv:1902.02426, 2019.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

P. Horowitz and C. Sagan. Five years of Project META - an all-sky narrow-band radio search for extraterrestrial signals. Astrophysical Journal, 415:218–235, 1993. doi: 10.1086/173157.

Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7132–7141, 2018.

Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training

by reducing internal covariate shift. In International conference on machine learning, pages 448–456. PMLR, 2015.

P. R. Jewell and R. M. Prestage. The Green Bank Telescope. In J. M. Oschmann, Jr., editor, Ground-based Telescopes, volume 5489 of Proceedings of the SPIE, pages 312–323, October 2004. doi: 10.1117/12.550631.

S. R. Kane, M. L. Hill, J. F. Kasting, R. K. Kopparapu, E. V. Quintana, T. Barclay, N. M. Batalha, W. J. Borucki, D. R. Ciardi, N. Haghighipour, N. R. Hinkel, L. Kaltenegger, F. Selsis, and G. Torres. A Catalog of Kepler Habitable Zone Exoplanet Candidates. Astrophysical Journal, 830:1, 2016. doi: 10.3847/0004-637X/830/1/1.

Leonid Kogan. Corrections functions for digital correlators with two and four quantization levels. Radio Science, 33(5):1289–1296, 1998. ISSN 1944-799X. doi: 10.1029/98RS02202.

Ravi Kumar Kopparapu, Ramses Ramirez, James F Kasting, Vincent Eymet, Tyler D Robinson, Suvrath Mahadevan, Ryan C Terrien, Shawn Domagal-Goldman, Victoria Meadows, and Rohit Deshpande. Habitable zones around main-sequence stars: new estimates. The Astrophysical Journal, 765(2):131, 2013.

Eric J. Korpela. SETI@home, BOINC, and Volunteer Distributed Computing. 40:69–87, 2012. doi: 10.1146/annurev-earth-040809-152348.

Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25:1097–1105, 2012.

Matthew Lebofsky, Steve Croft, Andrew P. V. Siemion, Danny C. Price, J. Emilio Enriquez, Howard Isaacson, David H. E. MacMahon, David Anderson, Bryan Brzycki, Jeff Cobb, Daniel

Czech, David DeBoer, Julia DeMarines, Jamie Drew, Griffin Foster, Vishal Gajjar, Nectaria Gizani, Greg Hellbourg, Eric J. Korpela, Brian Lacki, Sofia Sheikh, Dan Werthimer, Pete Worden, Alex Yu, and Yunfan Gerry Zhang. The Breakthrough Listen Search for Intelligent Life: Public Data, Formats, Reduction, and Archiving. 131(1006):124505, December 2019. doi: 10.1088/1538-3873/ab3e82.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998.

R. N. Manchester, G. B. Hobbs, A. Teoh, and M. Hobbs. The Australia Telescope National Facility Pulsar Catalogue. Astronomical Journal, 129:1993–2006, 2005. doi: 10.1086/428488.

J. L. Margot, D. B. Campbell, R. F. Jurgens, and M. A. Slade. Digital elevation models of the Moon from Earth-based radar interferometry. IEEE Trans. Geoscience and Remote Sensing, 38 (2):1122–1133, 2000.

J.-L. Margot, A. H. Greenberg, P. Pinchuk, A. Shinde, Y. Alladi, S. Prasad MN, M. O. Bowman, C. Fisher, S. Gyalay, W. McKibbin, B. Miles, D. Nguyen, C. Power, N. Ramani, R. Raviprasad, J. Santana, and R. S. Lynch. A Search for Technosignatures from 14 Planetary Systems in the Kepler Field with the Green Bank Telescope at 1.15–1.73 GHz. Astronomical Journal, 155:209, 2018. doi: 10.3847/1538-3881/aabb03.

J. L. Margot, A. H. Greenberg, P. Pinchuk, et al. Data from: A search for technosignatures from 14 planetary systems in the kepler field with the green bank telescope at 1.15–1.73 ghz, v4. Dataset, 2020a.

J. L. Margot, P. Pinchuk, A. H. Greenberg, et al. Data from: A search for technosignatures from trappist-1, lhs 1140, and 10 planetary systems in the kepler field with the green bank telescope at 1.15–1.73 ghz, v4. Dataset, 2020b.

J. L. Margot, P. Pinchuk, et al. Data from: A search for technosignatures around 31 sun-like stars with the Green Bank Telescope at 1.15–1.73 GHz. Dataset, 2020c.

Jean-Luc Margot, Pavlo Pinchuk, Robert Geil, Stephen Alexander, Sparsh Arora, Swagata Biswas, Jose Cebreros, Sanjana Prabhu Desai, Benjamin Duclos, Riley Dunne, Kristy Kwan Lin Fu, Shashwat Goel, Julia Gonzales, Alexander Gonzalez, Rishabh Jain, Adrian Lam, Briley Lewis, Rebecca Lewis, Grace Li, Mason MacDougall, Christopher Makarem, Ivan Manan, Eden Molina, Caroline Nagib, Kyle Neville, Connor O'Toole, Valerie Rockwell, Yoichiro Rokushima, Griffin Romanek, Carlyn Schmidgall, Samar Seth, Rehan Shah, Yuri Shimane, Myank Singhal, Armen Tokadjian, Lizvette Villafana, Zhixian Wang, In Yun, Lujia Zhu, and Ryan S. Lynch. A search for technosignatures around 31 sun-like stars with the green bank telescope at 1.15–1.73 GHz. The Astronomical Journal, 161(2):55, jan 2021. doi: 10.3847/1538-3881/abcc77. URL https://doi.org/10.3847/1538-3881/abcc77.

S. P. Naidu, L. A. M. Benner, J. L. Margot, M. W. Busch, and P. A. Taylor. Capabilities of Earth-based Radar Facilities for Near-Earth Asteroid Observations. Astronomical Journal, 152:99, 2016. doi: 10.3847/0004-6256/152/4/99.

S. J. Ostro. Planetary radar astronomy. Rev. Modern Phys., 65:1235–1279, 1993.

E. A. Petigura, A. W. Howard, and G. W. Marcy. Prevalence of Earth-size planets orbiting Sun-like stars. Proceedings of the National Academy of Science, 110:19273–19278, 2013. doi: 10.1073/pnas.1319909110.

Pavlo Pinchuk and Jean-Luc Margot. A Machine-Learning-Based Direction-of-Origin Filter for the Identification of Radio Frequency Interference in the Search for Technosignatures. submitted, 2021.

Pavlo Pinchuk, Jean-Luc Margot, Adam H. Greenberg, Thomas Ayalde, Chad Bloxham, Arjun Boddu, Luis Gerardo Chinchilla-Garcia, Micah Cliffe, Sara Gallagher, Kira Hart, Brayden Hesford, Inbal Mizrahi, Ruth Pike, Dominic Rodger, Bade Sayki, Una Schneck, Aysen Tan, Yinxue "Yolanda" Xiao, and Ryan S. Lynch. A Search for Technosignatures from TRAPPIST-1, LHS 1140, and 10 Planetary Systems in the Kepler Field with the Green Bank Telescope at 1.15-1.73 GHz. The Astronomical Journal, 157(3):122, 2019. doi: 10.3847/1538-3881/ab0105.

Danny C. Price, J. Emilio Enriquez, Bryan Brzycki, Steve Croft, Daniel Czech, David DeBoer, Julia DeMarines, Griffin Foster, Vishal Gajjar, Nectaria Gizani, Greg Hellbourg, Howard Isaacson, Brian Lacki, Matt Lebofsky, David H. E. MacMahon, Imke de Pater, Andrew P. V. Siemion, Dan Werthimer, James A. Green, Jane F. Kaczmarek, Ronald J. Maddalena, Stacy Mader, Jamie Drew, and S. Pete Worden. The Breakthrough Listen Search for Intelligent Life: Observations of 1327 Nearby Stars over 1.10-3.45 GHz. 159(3):86, 2020. doi: 10.3847/1538-3881/ab65f1.

Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), 115(3):211–252, 2015. doi: 10.1007/s11263-015-0816-y.

Abraham Savitzky and Marcel JE Golay. Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry, 36(8):1627–1639, 1964.

Kevin Schawinski, Ce Zhang, Hantian Zhang, Lucas Fowler, and Gokula Krishnan Santhanam. Generative adversarial networks recover features in astrophysical images of galaxies beyond the deconvolution limit. Monthly Notices of the Royal Astronomical Society: Letters, 467(1): L110–L114, 01 2017. ISSN 1745-3925. doi: 10.1093/mnrasl/slx008. URL https://doi.org/10.1093/mnrasl/slx008.

Christopher J Shallue and Andrew Vanderburg. Identifying exoplanets with deep learning: A five-planet resonant chain around kepler-80 and an eighth planet around kepler-90. The Astronomical Journal, 155(2):94, 2018.

Sofia Z. Sheikh, Jason T. Wright, Andrew Siemion, and J. Emilio Enriquez. Choosing a Maximum Drift Rate in a SETI Search: Astrophysical Considerations. 884(1):14, 2019. doi: 10.3847/1538-4357/ab3fa8.

A. P. V. Siemion, P. Demorest, E. Korpela, R. J. Maddalena, D. Werthimer, J. Cobb, A. W. Howard, G. Langston, M. Lebofsky, G. W. Marcy, and J. Tarter. A 1.1-1.9 GHz SETI Survey of the Kepler

Field. I. A Search for Narrow-band Emission from Select Targets. Astrophysical Journal, 767: 94, 2013. doi: 10.1088/0004-637X/767/1/94.

Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1):1929–1958, 2014.

Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1–9, 2015.

Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning, pages 6105–6114. PMLR, 2019.

J. C. Tarter, A. Agrawal, R. Ackermann, P. Backus, S. K. Blair, M. T. Bradford, G. R. Harp, J. Jordan, T. Kilsdonk, K. E. Smolek, J. Richards, J. Ross, G. S. Shostak, and D. Vakoch. SETI turns 50: five decades of progress in the search for extraterrestrial intelligence. In Instruments, Methods, and Missions for Astrobiology XIII, volume 7819 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, page 781902, 2010. doi: 10.1117/12.863128.

Jill Tarter. The search for extraterrestrial intelligence (seti). Annual Review of Astronomy and Astrophysics, 39(1):511–548, 2001.

J. H. Taylor. A Sensitive Method for Detecting Dispersed Radio Emission. Astronomy and Astrophysics, Supplement, 15:367, 1974.

The staff at the National Astronomy and Ionosphere Center. The Arecibo message of November, 1974. Icarus, 26:462–466, 1975. doi: 10.1016/0019-1035(75)90116-5.

Raffy Traas, Steve Croft, Vishal Gajjar, Howard Isaacson, Matt Lebofsky, David HE MacMahon, Karen Perez, Danny C Price, Sofia Sheikh, Andrew PV Siemion, et al. The breakthrough listen search for intelligent life: Searching for technosignatures in observations of tess targets of interest. The Astronomical Journal, 161(6):286, 2021.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, CJ Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake Vand erPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1. 0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2. URL https://doi.org/10.1038/s41592-019-0686-2.

M. Wenger, F. Ochsenbein, D. Egret, P. Dubois, F. Bonnarel, S. Borde, F. Genova, G. Jasniewicz, S. Laloë, S. Lesteven, and R. Monier. The SIMBAD astronomical database. The CDS reference database for astronomical objects. 143:9–22, 2000. doi: 10.1051/aas:2000332.

B. S. Wlodarczyk-Sroka, M. A. Garrett, and A. P. V. Siemion. Extending the Breakthrough Listen nearby star survey to other stellar objects in the field. art. arXiv:2006.09756, 2020.

Jason T. Wright, Shubham Kanodia, and Emily Lubar. How Much SETI Has Been Done? Finding Needles in the n-dimensional Cosmic Haystack. Astronomical Journal, 156:260, 2018. doi: 10.3847/1538-3881/aae099.

Qian Xiang and Xuliang Pang. Improved denoising auto-encoders for image denoising. In 2018

11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), pages 1–9, 2018. doi: 10.1109/CISP-BMEI.2018.8633143.

Y. G. Zhang, K. Hyun Won, S. W. Son, A. Siemion, and S. Croft. Self-supervised anomaly detection for narrowband seti. In 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pages 1114–1118, 2018. doi: 10.1109/GlobalSIP.2018.8646437.

Yunfan Gerry Zhang, Vishal Gajjar, Griffin Foster, Andrew Siemion, James Cordes, Casey Law, and Yu Wang. Fast radio burst 121102 pulse detection and periodicity: a machine learning approach. The Astrophysical Journal, 866(2):149, 2018.