



Applicazione della trasformata di Karhunen-Loève per la ricerca di segnali telemetrici extraterrestri

Corso di Laurea Triennale in Fisica

Candidato: Giovanni Nicola D'Aloisio (N85002412)

Relatori: Prof. Giovanni Covone, Dott. Matteo Trudu, Dott. Andrea Melis

16 maggio 2025



UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II



Indice dei contenuti

1 SETI: Search for Extra-Terrestrial Intelligence

- ▶ SETI: Search for Extra-Terrestrial Intelligence
- ▶ Teoria della trasformata di Karhunen-Loève
- ▶ Gli strumenti di un cacciatore di alieni
- ▶ Implementazione della KLT con turboSETI
- ▶ Conclusioni



Il SETI, Breakthrough Listen e il Sardinia Radio Telescope

1 SETI: Search for Extra-Terrestrial Intelligence

**BREAKTHROUGH
LISTEN**



**BERKELEY SETI
RESEARCH CENTER**

$$N = R_* f_p n_e f_l f_i f_c L$$

(Equazione di Drake, 1961)

OAC

**Osservatorio
Astronomico
di Cagliari**





Tipologia di dati e numero di target analizzati

1 SETI: Search for Extra-Terrestrial Intelligence

Tipo di target	Quantità	Bande	Fonte	Ore di osservazione
Fast Radio Burst	4	P, L	SRT ¹ , VLA ² , Parkes ³	40
Veicoli spaziali artificiali	1	X	GBT ⁴	> 2
Esopianeti (TESS)	6	C, K	SRT	> 3
Esopianeti (nucleo Via Lattea)	6	C, K	SRT	> 3

¹Sardinia Radio Telescope

²Karl Guthe Jansky Very Large Array

³Parkes Observatory

⁴Robert C. Byrd Green Bank Telescope



Tecniche per la ricerca di vita extraterrestre intelligente

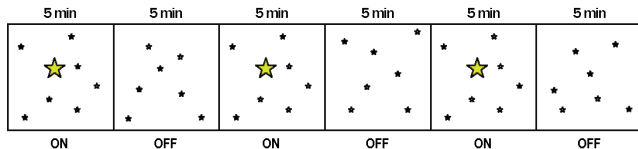
1 SETI: Search for Extra-Terrestrial Intelligence

Spostamento Doppler – Un veicolo spaziale con $v(t) = v_0 + at$ emette segnali del tipo $S(t) = A \sin(2\pi\nu t)$.

A Terra si misura una frequenza ν' :

$$\begin{aligned}\nu' &= \nu \frac{c}{c \pm v(t)} \simeq \left(1 \mp \frac{v(t)}{c}\right) \nu \\ &\simeq \left(1 \mp \frac{v_0}{c}\right) \nu \mp \underbrace{\frac{a \nu}{c}}_{\text{drift rate}} t\end{aligned}$$

ON & OFF – Osservazione alternata tra bersaglio e zona casuale per distinguere potenziali segnali legati a tecnofirme artificiali dalle interferenze radio (RFI).





I Fast Radio Burst: fenomeni astrofisici ad alta energia

1 SETI: Search for Extra-Terrestrial Intelligence

A bright millisecond radio burst of extragalactic origin

D. R. Lorimer,^{1,2*} M. Bailes,³ M. A. McLaughlin,^{1,2}
D. J. Narkevic,¹ F. Crawford⁴

¹Department of Physics, West Virginia University, P.O. Box 6315, WV 26506 USA

²National Radio Astronomy Observatory, P.O. Box 2, Green Bank, WV 24944

³Centre for Astrophysics and Supercomputing, Swinburne University of Technology,
P.O. Box 218, Hawthorn, Vic, 3122, Australia

⁴Department of Physics and Astronomy, Franklin and Marshall College, Lancaster, PA 17604 USA

*To whom correspondence should be addressed; E-mail: Duncan.Lorimer@mail.wvu.edu.

Accepted for publication in the journal Science

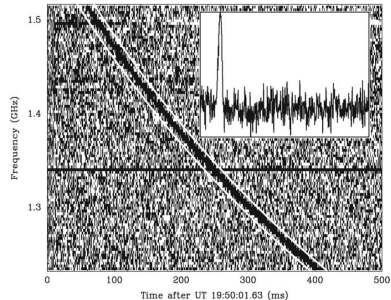


Figura: A destra: evoluzione in frequenza e convoluzione del lampo di Lorimer, il primo FRB mai osservato.

A sinistra: il titolo con cui fu annunciato.

[astro-ph] 27 Sep 2007



Indice dei contenuti

2 Teoria della trasformata di Karhunen-Loève

- ▶ SETI: Search for Extra-Terrestrial Intelligence
- ▶ Teoria della trasformata di Karhunen-Loève
- ▶ Gli strumenti di un cacciatore di alieni
- ▶ Implementazione della KLT con turboSETI
- ▶ Conclusioni



Definizione matematica della *Karhunen-Loève Transform*

2 Teoria della trasformata di Karhunen-Loève

Dato un processo stocastico $X(t)$, la KLT è definita come:

$$X(t) = \sum_{m=0}^{+\infty} \zeta_m \phi_m(t) + \mu(t)$$

dove:

- $\mu(t) = \mathbb{E}[X(t)]$ - Valore atteso;
- ζ_m - Variabili aleatorie indipendenti;
- $\phi_m(t)$ - Autofunzioni dell'operatore di *autocovarianza*:

$$R(t, s) = \mathbb{E}[(X(t) - \mu(t))(X(s) - \mu(s))^*]$$

Le $\phi_m(t)$ soddisfano $\int_0^T R(t, s) \phi_m(s) ds = \lambda_m \phi_m(t)$, dove $\lambda_m \in \mathbb{R}^+$ sono gli autovalori. La varianza del processo è $\sigma^2 = \sum_{m=0}^{+\infty} \lambda_m$.



La trasformata di Karhunen-Loève su dataset discreti

2 Teoria della trasformata di Karhunen-Loève

Per un segnale discreto $X = x_{\alpha i}$, la matrice di covarianza campionaria è:

$$R_{ij} = \frac{1}{M-1} \sum_{\alpha=0}^{M-1} (x_{\alpha i} - \mu_i)(x_{\alpha j} - \mu_j)^*$$

Risolvendo il seguente problema agli autovalori:

$$\sum_{j=0}^{N-1} R_{ij} \phi_{jm} = \lambda_m \phi_{im}$$

si ottengono:

- gli autovalori λ_m – Varianza per ciascuna componente;
- gli autovettori ϕ_{im} – Base ortogonale del segnale.

La KLT consente una rappresentazione compatta e filtraggio del segnale.



Indice dei contenuti

3 Gli strumenti di un cacciatore di alieni

- ▶ SETI: Search for Extra-Terrestrial Intelligence
- ▶ Teoria della trasformata di Karhunen-Loève
- ▶ **Gli strumenti di un cacciatore di alieni**
- ▶ Implementazione della KLT con turboSETI
- ▶ Conclusioni



turboSETI e un caso di studio: il Voyager 1

3 Gli strumenti di un cacciatore di alieni

turboSETI – Strumento sviluppato per il SETI per la ricerca di segnali *poco dispersi* in frequenza e dotati di *deriva Doppler*. Può identificare potenziali tecnofirme e calcolarne il *drift rate*.

Voyager 1 – Sonda spaziale lanciata nel 1977 dalla NASA che, dopo aver esplorato Giove e Saturno, è diventata il primo oggetto interstellare costruito dall'uomo. Ancora operativa al 2025.

Idea – Utilizzare le osservazioni ON-OFF di una fonte artificiale nota per collaudare turboSETI.





Risultati di turboSETI sul Voyager 1 e waterfall plot

3 Gli strumenti di un cacciatore di alieni

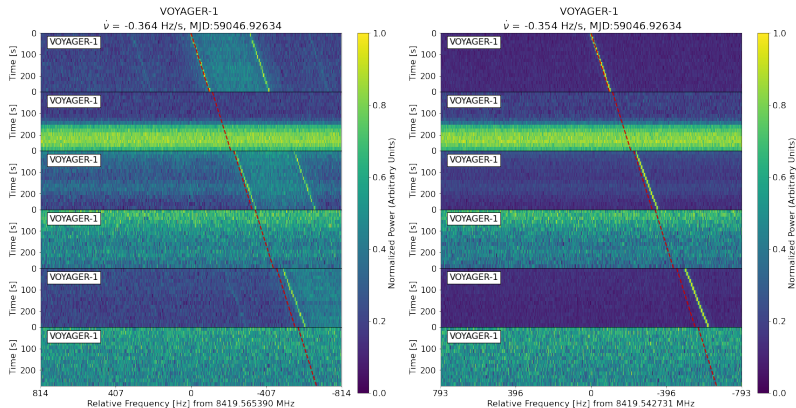


Figura: Waterfall plot delle osservazioni del Voyager 1 con il *drift rate* evidenziato in rosso, per le due frequenze rilevate in banda X (7 – 12.5 GHz). I colori indicano la potenza del segnale, l'asse orizzontale la frequenza, quello verticale il tempo.



Indice dei contenuti

4 Implementazione della KLT con turboSETI

- ▶ SETI: Search for Extra-Terrestrial Intelligence
- ▶ Teoria della trasformata di Karhunen-Loève
- ▶ Gli strumenti di un cacciatore di alieni
- ▶ **Implementazione della KLT con turboSETI**
- ▶ Conclusioni



Verifica preliminare dei *waterfall plot* di alcuni FRB ⁵

4 Implementazione della KLT con turboSETI

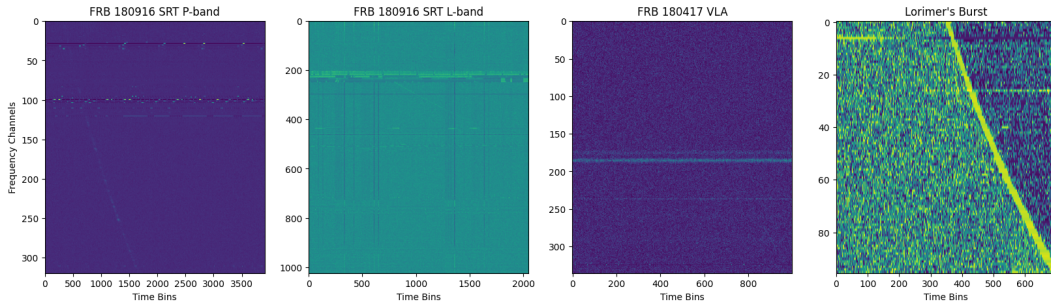


Figura: *Waterfall plot* delle osservazioni in banda P e banda L di FRB 180916 (visto da SRT), FRB 180417 (visto dal Karl Guthe Jansky Very Large Array, VLA) e lampo di Lorimer (visto da Parkes il 24 agosto 2001 in banda L).

⁵Fast Radio Burst



Produzione dei template delle RFI con la KLT⁶

4 Implementazione della KLT con turboSETI

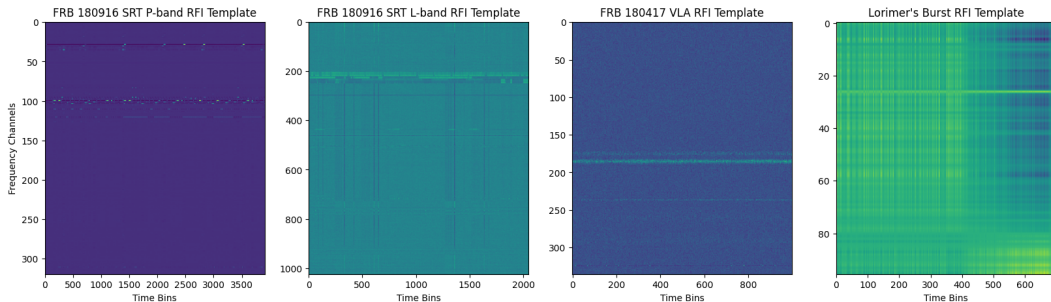


Figura: Template delle Radio-Frequency Interferences (RFI) contenute in FRB 180916 in banda P ed L e in di FRB 180417 e nel lampo di Lorimer ottenute tramite KLT. Si noti che in FRB 180916 la reattività è la stessa in entrambe le bande.

⁶Karhunen-Loève Transform



Ricostruzione dei segnali senza le interferenze radio

4 Implementazione della KLT con turboSETI

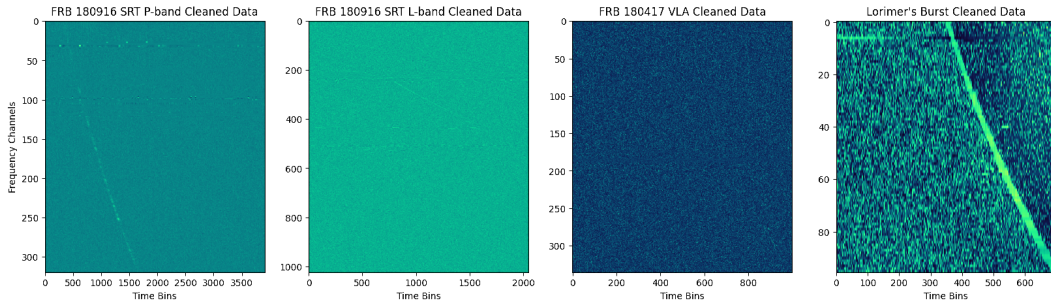
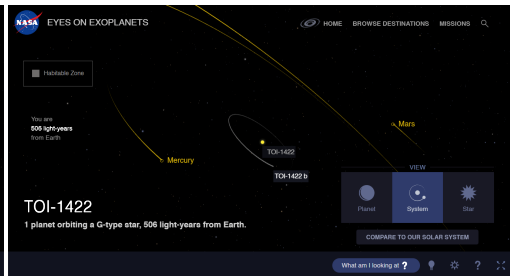


Figura: *Waterfall plot* dei dataset ricostruiti sottraendo i template delle RFI dai dati delle osservazioni originali di FRB 180916 e FRB 180417 e del lampo di Lorimer.



TOI-1422 b: un esopianeta nettuniano caldo in orbita attorno a una stella di tipo solare

4 Implementazione della KLT con turboSETI



Raggio	Tipo	Scoperta	Massa	Raggio orbitale	Periodo orbitale	Eccentricità
$0.35 R_{\text{Giove}}$	Nettuno-like	Per transito, 2022	$9 M_{\oplus}$	0.108 UA	13 giorni	0.04



Osservazioni ON e OFF di TOI-1422 b per SETI

4 Implementazione della KLT con turboSETI

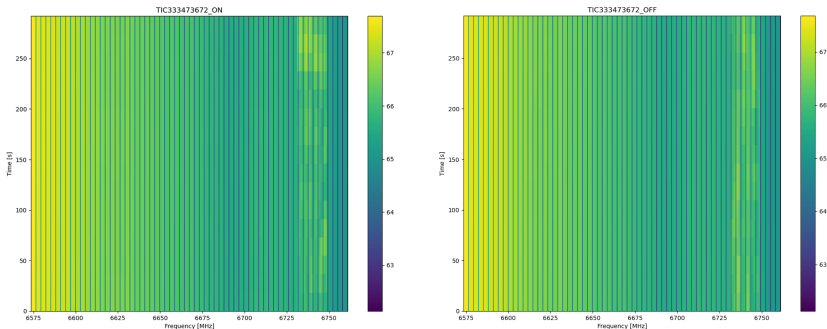


Figura: Da sinistra, una coppia di dati originali in banda C, presi da SRT il 13 giugno 2021 puntando verso TOI-1422 b e verso una direzione casuale in un ciclo "ON-OFF". Si notino le RFI, prive di *drift rate* e dunque verticali.



Applicazione di tecnofirme artificiali e KLT su TOI-1422 b

4 Implementazione della KLT con turboSETI

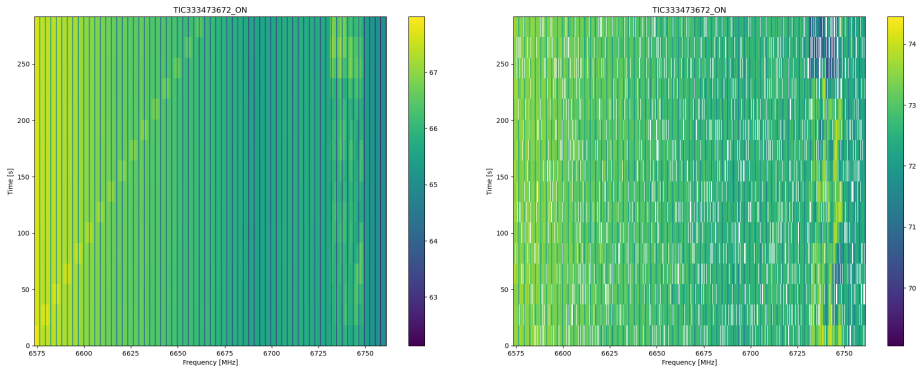


Figura: Da sinistra, l'osservazione del 13/06/2021 già riportata con l'aggiunta di una tecnofirma di $\text{SNR} = 2$ iniettata con setigen (dello stesso tipo del Voyager 1), e la maschera di RFI ottenuta con la KLT.



Indice dei contenuti

5 Conclusioni

- ▶ SETI: Search for Extra-Terrestrial Intelligence
- ▶ Teoria della trasformata di Karhunen-Loève
- ▶ Gli strumenti di un cacciatore di alieni
- ▶ Implementazione della KLT con turboSETI
- ▶ Conclusioni



Conclusioni

5 Conclusioni

In questo lavoro:

- **Ho partecipato attivamente** alle attività dell'Osservatorio Astronomico di Cagliari, fornendo il mio contributo nell'**analisi dei dati**;
- Ho contribuito a mettere a punto una **nuova procedura di analisi**, sulla base del lavoro teorico del dott. Claudio **Maccone**.

Tra i risultati principali:

- Si è riscontrata una **riduzione significativa delle interferenze radio** e un aumento dei rilevamenti utili nel caso dei dati con tecnofirme artificiali, specie con segnali deboli;
- Si è dimostrato che la trasformata di Karhunen-Loève è in grado di **evidenziare segnali simili a quelli dei Voyager** e distinguerli dalle interferenze radio.



Grazie per l'attenzione!

Backup slides

5 Conclusioni



UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II



La carta stellare del SETI: il TESS Input Catalog

5 Conclusioni

Il TESS Input Catalog (TIC) è un catalogo pubblico di parametri stellari che comprende ogni oggetto otticamente persistente e stazionario nel cielo. Racchiude circa 1.5 miliardi di oggetti e un suo sottoinsieme è la Candidate Target List (CTL), che comprende alcuni target per i quali è più probabile che si osservi il transito di un pianeta.

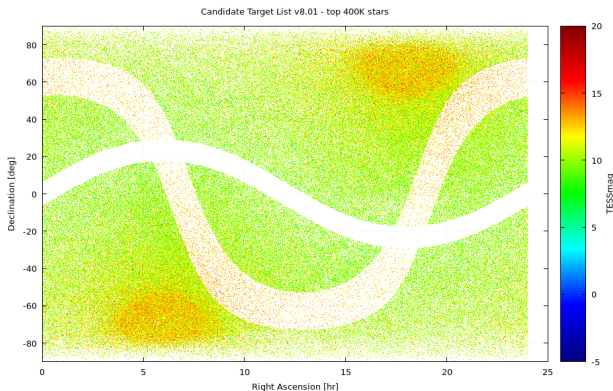


Figura: Una mappa del CTL v8.01.



Un'introduzione ai file filterbank e HDF (.fil / .h5)

5 Conclusioni

- Un file filterbank (.fil o .h5) contiene:
 - **Spettrogrammi**: rappresentano la potenza ricevuta da un radiotelescopio per varie frequenze nel tempo.
 - **Intestazione**: informazioni sui parametri dell'osservazione, come data e nome del bersaglio.
- Formato binario non leggibile direttamente in editor di testo.
- Analizzabile con il pacchetto Python `blimpy`, che fornisce:
 - La classe `waterfall` per rappresentare i dati in forma di array `numpy`.
 - Funzioni per visualizzare e analizzare i dati.
- **Dati in tre dimensioni**:
 - **Frequenza, potenza, tempo**.
 - Visualizzati in un *grafico a cascata*, con tempo sull'asse delle ascisse.
 - "Taglio orizzontale": spettro di potenza per una specifica integrazione.
- Esempi: Dati Voyager 1 dal telescopio di Greenbank (16 luglio 2020).



Dati per SETI vs dati di FRB

5 Conclusioni

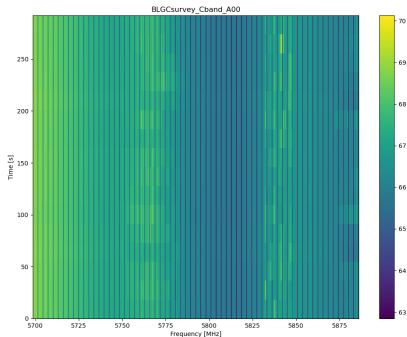


Figura: Osservazione di un punto della Via Lattea in banda C, caratterizzata da 67.1 milioni di canali, larghezza di banda (per canale) 2.79 Hz e tempo di campionamento 18.25 s.

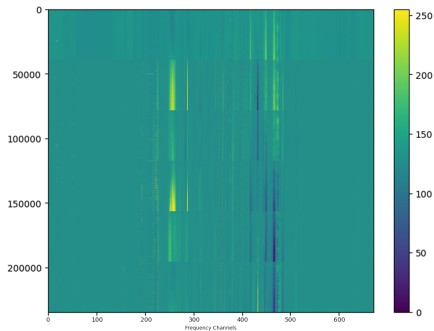


Figura: Osservazione di FRB 121102 in banda L, caratterizzata da un tempo di campionamento di 0.256 ms, 672 canali e una larghezza di banda di 0.5 MHz per canale.



Operatori, autovalori e autovettori

5 Conclusioni

Sia V su $\mathbb{K}$ uno spazio vettoriale e $f : V \rightarrow V$ un operatore lineare. Uno scalare $\lambda \in \mathbb{K}$ si dice autovalore di f se esiste un $v \in V$, $v \neq 0$, tale che $f(v) = \lambda v$.

Il sottospazio $V(\lambda) = \{v \in V : f(v) = \lambda v\} = \ker(f - \lambda \text{id})$ è l'autospazio relativo a λ .

Per calcolare gli autovalori di una matrice A , si risolve:

$$\det(A - \lambda I) = 0 \tag{1}$$



Definizione matematica della KLT

5 Conclusioni

Dato un processo stocastico $X(t)$, la trasformata di Karhunen-Loève è definita come:

$$X(t) = \sum_{m=0}^{+\infty} \zeta_m \phi_m(t) + \mu(t) \quad (2)$$

dove:

- $\mu(t) = \mathbb{E}[X(t)]$ è il valore atteso;
- ζ_m sono variabili casuali indipendenti;
- $\phi_m(t)$ sono autofunzioni dell'operatore di autocovarianza $R(t, s)$.



Operatore di autocovarianza, varianza e spettro

5 Conclusioni

L'operatore di autocovarianza è definito come:

$$R(t, s) = \mathbb{E}[(X(t) - \mu(t))(X(s) - \mu(s))^*] \quad (3)$$

Le autofunzioni $\phi_m(t)$ soddisfano:

$$\int_0^T R(t, s) \phi_m(s) ds = \lambda_m \phi_m(t) \quad (4)$$

dove λ_m sono autovalori reali e positivi. La varianza σ^2 del processo è data da:

$$\sigma^2 = \int_0^T \sigma^2(t) dt = \sum_{m=0}^{+\infty} \lambda_m \quad (5)$$

Gli autovalori λ_m rappresentano lo spettro dell'operatore $R(t, s)$.



Applicazione a dataset discreti

5 Conclusioni

Per un processo (o segnale) discreto $X = x_{\alpha i}$, la matrice di covarianza campionaria è:

$$R_{ij} = \frac{1}{M-1} \sum_{\alpha=0}^{M-1} (x_{\alpha i} - \mu_i)(x_{\alpha j} - \mu_j)^* \quad (6)$$

Risolvendo:

$$\sum_{j=0}^{N-1} R_{ij} \phi_{jm} = \lambda_m \phi_{im} \quad (7)$$

si ottengono gli autovalori e gli autovettori per ricostruire il processo, ovvero il segnale.



turboSETI – Metodo FindDoppler

5 Conclusioni

Parametri principali:

- `max_drift`: tasso massimo di deriva Doppler analizzato;
- `snr`: soglia minima del rapporto segnale-rumore;
- `out_dir`: directory in cui salvare i registri `.dat` dei segnali utili rilevati.

Altri attributi utili:

- `gpu_backend`: accelerazione hardware;
- `precision`: `float32` o `float64`;
- `blank_dc`: rimozione di picchi di offset.

```
1 # Create a .lst file of the .h5 files
2 h5_list = sorted(glob.glob(os.path.join
   (data_directory, '*.h5')))
3
4 # Write the .h5 files into a .lst
5 h5_list_path = os.path.join(
   data_directory, 'h5_files.lst')
6 with open(h5_list_path, 'w') as f:
7     for h5_path in h5_list:
8         f.write(h5_path + '\n')
9
10 for file in h5_list:
11     doppler = FindDoppler(file,
12         max_drift = 4,
13         snr = 10,
14         out_dir = data_directory)
15     doppler.search()
```



Struttura di un file .dat in uscita da turboSETI

5 Conclusioni

```
1 # ----- o -----
2 # File ID: single_coarse_guppi_59046_80672_DIAG_VOYAGER-1_0013.rawspec.0000.h5
3 # ----- o -----
4 # Source:VOYAGER-1
5 # MJD: 59046.933703703704    RA: 17h12m40.482s    DEC: 12d24m13.656s
6 # DELTAT: 18.253611    DELTAF(Hz): -2.793968    max_drift_rate: 4.000000
   obs_length: 292.057776
7 # -----
```

#	Drift Rate (Hz/s)	SNR	Frequency (Hz)	Index	Start Freq (Hz)	End Freq (Hz)
1	-0.408170	10.997	8419.565152	651964	8419.567343	8419.562965
2	-0.387762	82.803	8419.542493	660074	8419.544684	8419.540306
3	-0.377557	10.477	8419.519834	668184	8419.522025	8419.517647



find_event_pipeline.py – Ricerca di segnali di interesse

5 Conclusioni

Se il segnale rilevato è presente in tutte le osservazioni "ON" di un ciclo e assente in tutte quelle "OFF", allora è una potenziale tecnofirma.

```
1 csvf_path = os.path.join(  
    data_directory,  
2 'found_event_table.csv')  
3 find_event_pipeline(  
    dat_list_path,  
    filter_threshold = 3,  
    number_in_cadence =  
4 len(dat_list), csv_name =  
    csvf_path, saving = True)
```

	0	1
TopHitNum	1	2
DriftRate	-0.367353	-0.357149
SNR	22.329107	192.895111
Freq	8419.565390	8419.542731
ChanIndx	651879	659989
FreqStart	8419.567580	8419.544921
FreqEnd	8419.563202	8419.540543
...	...	...
DELTAf	-2.793968	-2.793968
max_drift_rate	4.000000	4.000000
Hit_ID	VOYAGER-1_0	VOYAGER-1_1

Tabella: Contenuto del file found_event_table.csv



plot_spectrum.py – Power plot delle osservazioni

5 Conclusioni

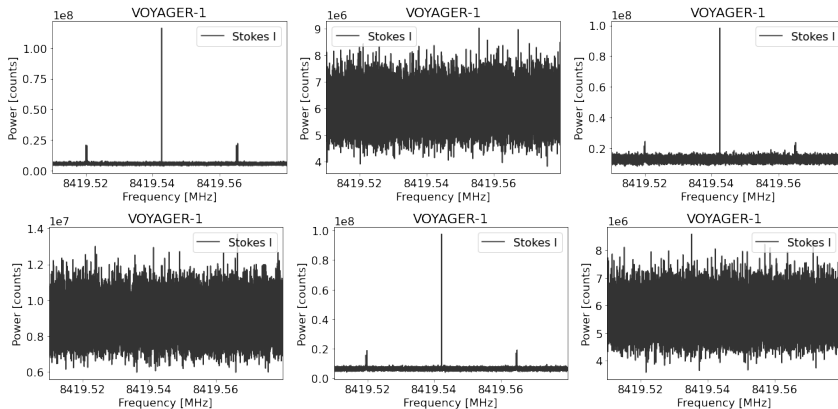


Figura: Power plot delle osservazioni “ON/OFF” (in senso orario) del Voyager 1. I picchi principali rappresentano la frequenza portante; gli altri, invece, le bande secondarie.



Calcolo della velocità del Voyager 1 con turboSETI

5 Conclusioni

- Dal drift rate rilevato da turboSETI si trova una velocità relativa $v \approx 30.97$ km/s, molto diversa dal valore riportato sulle effemeridi dal Jet Propulsion Laboratory (≈ 17 km/s).
- L'accuratezza di v è influenzata dalla durata della finestra osservativa (per Green Bank, 30 minuti), durante la quale la Terra ruota su sé stessa.
- Invece, le effemeridi del JPL sono basate su misure annuali per maggiore precisione.



Identificazione delle componenti principali

5 Conclusioni

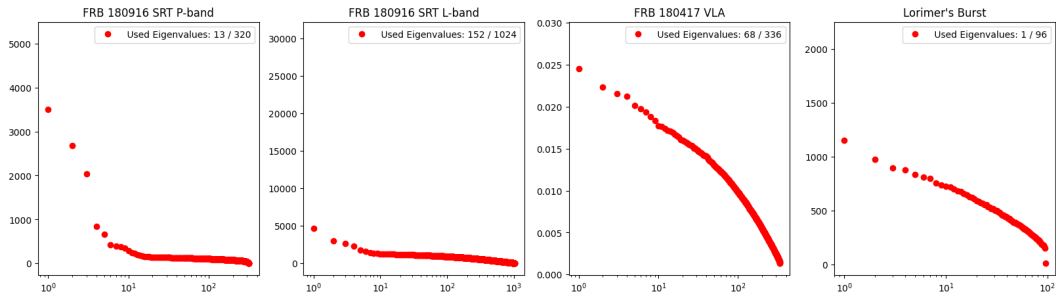


Figura: Autovalori delle matrici di FRB18O916 e del lampo di Lorimer, in ordine decrescente.



Algoritmo di KLT tramite matrici di Toeplitz per FRB

5 Conclusioni

```
1 def tklt(signal,neig):
2     acf = correlate(signal-np.mean(signal),signal-np.mean(signal), mode='full'
3         )/len(signal)
4     acf = acf/acf[np.argmax(np.abs(acf)))]
5     T = toeplitz(acf[nmax:nmax+len(signal)])
6     eigsp,eigvec = eigenbasis(T)
7     coeff = np.matmul((signal[:]-np.mean(signal)),np.conjugate((eigvec[:,:])))
8     rec = np.matmul(coeff[0:int(neig)],np.transpose(eigvec[:,0:int(neig)])) +
9         np.mean(signal)
10    return eigsp, eigvec, rec
11
12 eval, evec, rec = tklt(bandpass, 9) # Set the no. of eigvalues
13 badchans = [i for i, value in enumerate(rec) if abs(value - statistics.median(
14     rec)) > value * 0.05] # Set the threshold
15 for channel in badchans:
16     for i in range(header.nspectra):
17         input_matrix[i][channel] = sum(bandpass) / len(bandpass)
18 output_data.append_spectra(input_matrix, "PostKLT_FRB180417.fil")
```



Algoritmo di KLT by threshold per SETI

5 Conclusioni

```
1 def klt(signals, threshold):
2     logging.info("KLT function called with matrix element, threshold: "
3                 + str(signals[10,10]) + " | " + str(threshold))
4     R = np.cov((signals - np.mean(signals, axis=0)), rowvar=False)
5
6     eigenspectrum, eigenvectors = eigenbasis(R)
7     neig = count_elements_for_threshold(eigenspectrum, threshold)
8
9     coeff = np.matmul((signals[:,:] - np.mean(signals, axis=0)), np.conjugate((
10         eigenvectors[:,:]))) # KLT coefficients
11     recsignals = np.matmul(coeff[:,0:int(neig)], np.transpose(eigenvectors[:,0:
12         int(neig)])) + np.mean(signals, axis=0)
13
14     return neig, eigenspectrum, eigenvectors, recsignals
```



Algoritmo di iniezione tecnofirme con setigen (1/2)

5 Conclusioni

```
1 def setigen_injection(frame, snr, df, header, f_start, t_start, t_length):
2
3     if(df == 'nudot'):
4         nchans = header['nchans']
5         tsamp = header['tsamp']
6         fch0 = header['fch1']
7         foff = header['foff']
8         nsamp = frame.shape[0]
9         bw = nchans * foff
10        fc = fch0 + bw / 2
11        obslen = nsamp * tsamp
12        df = ((fc - fch0) / obslen) * u.MHz / u.s
13    else:
14        df = float(df) * u.Hz / u.s
15
16    # [omissis]
```



Algoritmo di iniezione tecnofirme con setigen (2/2)

5 Conclusioni

```
1  # [omissis]
2
3  signal = frame.add_signal(path = stg.constant_path(f_start=frame.
4      get_frequency(index = int(f_start)),
5      drift_rate = df),
6      t_profile = stg.constant_t_profile(level = frame.get_intensity(snr =
7          float(snr))),
8      f_profile = stg.gaussian_f_profile(width = 3000000 * u.Hz),
9      bp_profile = None
10 )
11 logging.debug('03a_setigen injection succesfully finished')
12 return frame.data
```

```
1 python 05a_signal_injector.py -v -w blimpy -f [INPUT_FILE.fil]
   -o [OUTPUT_FOLDER] -stg -snr 2 -df nudot
```



Applicazione di setigen e KLT su TIC333473672

5 Conclusioni

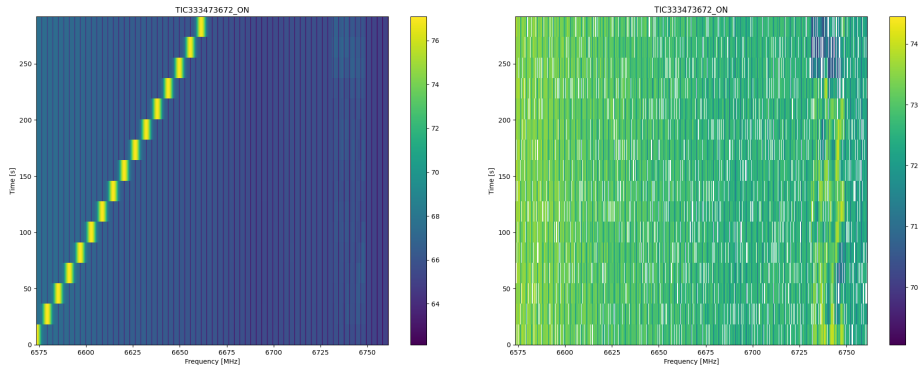


Figura: Da sinistra, l'osservazione del 13/06/2021 già riportata con l'aggiunta di una tecnofirma di SNR = 200 iniettata con setigen, e la stessa trattata con KLT.



Risultati di turboSETI sui dati processati con la KLT

5 Conclusioni

	RAW	SETIGEN		SETIGEN + KLT	
Tipo	C-band	SNR = 2	SNR = 200	SNR = 2	SNR = 200
ON	297	298	280	9	9
OFF	295	295	295	295	295
ON	262	262	249	9	9
OFF	223	223	223	223	223
ON	264	264	245	9	9
OFF	233	233	233	233	233

Tabella: Numero di *hit* (segnali dotati di deriva Doppler) ottenuti con turboSETI dalle osservazioni di TOI-1422 b del 13 e 15 giugno 2021, in banda C (5.7 – 7.7 GHz).

