

Demodulation of Wi-Fi 802.11g Packets

Shamman Noor

Email: sxn170028@utdallas.edu

Abstract. In this project, IEEE 802.11g Wi-Fi packets, captured using NI USRP and LabVIEW, are demodulated in MATLAB to extract information of the signal.

Background. IEEE 802.11g Wi-Fi uses OFDM (Orthogonal Frequency Division Multiplexing) and operates in frequency band 2.400-2.4835 GHz. The entire band is divided into 14 channels, spaced 5 MHz apart from each other except for a 12 MHz space before channel 14. Each channel is 16.25 MHz wide and is divided into 64 subcarriers with 12 null (one DC), 48 data and 4 pilot subcarriers. The pilot subcarriers are located at subcarriers -21, -7, 7 and 21. The sampling rate is 20MHz and only channel 1 is used in this project due to Wi-Fi signal availability in that channel.

802.11g Wi-Fi packets start with a preamble field of 16 μ s, consisting of short and long preamble fields of 8 μ s each. There are 10 repeated fields, each of 0.8 μ s, in the short preamble field. This repetitive pattern is used for frame synchronization and carrier frequency offset compensation (CFO). The long preamble field uses OFDM and consists of two repeated fields of 3.2 μ s each and a guard interval of 1.6 μ s, which is a cyclic prefix of the long preamble field. The long preamble is used for channel estimation. The pilot subcarriers are used for residual frequency (RF) compensation.

The SIGNAL field is 4 μ s, with a guard interval (cyclic prefix) of 0.8 μ s in the beginning. It is BPSK modulated and contains the data rate, length information of the Wi-Fi packet. From these information, other information for example: modulation scheme, convolutional coding rate, coded bits and data bits per sub carrier etc. can be extracted from the 802.11 IEEE standard [1]. Rate depended variables' table is presented below:

Rate	Data Rate (Mbps)	Modulation	Coding Rate (R)	Coded Bits per Subcarrier (NBPSC)	Coded Bits per OFDM Symbol (NCBPS)	Data Bits per OFDM Symbol (NDBPS)
1101	6	BPSK	1/2	1	48	24
1111	9	BPSK	3/4	1	48	36
0101	12	QPSK	1/2	2	96	48
0111	18	QPSK	3/4	2	96	72
1001	24	16-QAM	1/2	4	192	96
1011	36	16-QAM	3/4	4	192	144
0001	48	64-QAM	2/3	6	288	192
0011	54	64-QAM	3/4	6	288	216

Table 1. Rate depended variables for IEEE 802.11g Wi-Fi packet

Description of ERP-OFDM packet receiver implementation. The demodulation algorithm can be divided into the following 18 steps:

1. **Signal detection.** The signal received is run through a power estimator (1) to detect if a signal is present. The threshold to detect a signal is set to 0.003, based on the available data.

$$p(n) = (1 - \rho) |r(n)|^2 + \rho p(n-1) \quad (1)$$

2. **Signal truncating.** After detecting the presence of a signal, the entire received signal is divided into pieces of valid signals and silent frames are discarded. Using the power estimator, the beginning of a valid signal is taken as 10 samples before the detection time due to delay in signal detection.
3. **Autocorrelation.** For each truncated signal, the autocorrelation algorithm searches for the 10 repeated patterns, by correlating two segments of 7.2 μ s each, with a time lag of 0.8 μ s between them and taking the maximum value.
4. **Frame synchronization.** The Wi-Fi packet starting index or time is taken where the correlator finds its value maximum within the truncated valid signal.
5. **Carrier Frequency Offset (CFO) compensation.** After detecting the starting of the short preamble field, carrier frequency offset is determined from the phase information at the starting index. Using the frequency offset, CFO is compensated only for the Long preamble field and the signal field.
6. **Time domain to frequency domain.** 64 length FFT is taken of each long preamble fields of 64 bits to convert them from time domain signals to frequency domain signals. The length is reduced to 52 after removing the null subcarriers (DC and 28-38).
7. **Channel estimation and channel effect compensation.** After CFO compensation, channel coefficients of length 52 bits are estimated using the known 52 length OFDM LT signal and detected long preamble field. Both LT signals are used here and the channel coefficients are average of the two estimations from two LT fields. After channel estimation, channel effect is compensated for the SIGNAL block.
8. **Residual frequency compensation.** The residual frequency is compensated using the four pilot bits located in the SIGNAL field. The bit signs are given in the standard [1]. The length of SIGNAL field is reduced to 48 after removing the pilot subcarriers.
9. **Demodulation - SIGNAL field.** The SIGNAL field is BPSK modulated. Only the real parts of the symbols are used to demodulate symbols to bits.
10. **Deinterleaving - SIGNAL field.** The deinterleaver rearranges the bits in the SIGNAL field according to the formulas in [1].

11. **Convolutional Decoder - SIGNAL field.** The SIGNAL field is coded using convolutional coding rate of $\frac{1}{2}$. MATLAB built-in functions 'poly2trellis' and 'vitdec' are used for decoding the bits. After $\frac{1}{2}$ rate decoding, the length of SIGNAL field becomes 24 bits.
12. **SIGNAL field information.** SIGNAL field information is following:
 - a. Rate (1-4): First four bits of decoded signal indicate the rate. The four bits are matched with the standard table in [1] to extract other information of the Wi-Fi packet.
 - b. Reserved (5)
 - c. LENGTH (6-17): The length (in octets) of OFDM DATA is given in bits 6 to 17, where bit 6 is the LSB and bit 17 is the MSB.
 - d. Parity bit (18): The 18 bit acts as even parity bit for the entire SIGNAL field.
 - e. Tail bits (19-24): The last 6 bits are all zeros so the convolutional coder registers are flushed properly.
13. **OFDM DATA extraction.** Using the length information from SIGNAL field, OFDM data is truncated from the signal.
14. **CFO, RF and Phase offset compensation for DATA field.** After extracting DATA field, the CFO, RF and Phase offset is compensated for the entire length of the DATA field.
15. **Demodulation of DATA field.** The DATA symbols are demodulated using the modulation scheme information found from the standard table [1] using the rate information in SIGNAL field.
16. **Deinterleaving - DATA field.** Each OFDM block is 80 OFDM symbols long, with 64 subcarriers and cyclic prefix of length 16. First, the data bits are reshaped into a matrix of [80] x [number of OFDM blocks]. The deinterleaver used previously for the SIGNAL field is used here for each of the columns of that matrix.
17. **Convolutional Decoder - DATA field.** From the standard table, using the rate information, the convolutional coding rate information was obtained. If the coding rate is not $\frac{1}{2}$, MATLAB built-in function 'vitdec' is used with an additional parameter 'puncpat', which specifies the puncturing pattern of the decoder.
18. **Descrambler.** The first 7 bits of the decoded signal are taken as the initial state of the descrambler. A descrambling sequence is generated using this initial state and bitwise XOR operation is performed between the scrambling sequence and the decoded signal.
19. **CRC-32 check.** Cyclic Redundancy Check is performed on the descrambled data. For the CRC-32, a polynomial of length 33 is used. Discarding the MSB, the polynomial is 0x04C11DB7. Before calculating CRC on the data, it is performed on the character '123456789' or hex value 0x313233343536373839. The result of this CRC (0xCB43926) is checked to make sure the CRC is functioning properly. If the CRC is matched, the algorithm moves on to the next stage. If the CRC is not matched, the current truncated signal is discarded and the algorithm moves on to the next detected signal.
20. **Information extraction.** If the data passes CRC-32 check, the hex dump is created by taking groups of 4 bits and converting them from binary to hex (LSB first operation). The following information are extracted from the hex dump:
 - a. **Frame type:** Management, Control or Data frame
 - b. **Frame subtype:** frame subtype information - Beacon, QoS Data, ACK, RTS etc.
 - c. **toDS and fromDS flags:** whether the packet is sent from access point to destination or otherwise
 - d. **Duration:** time allocated for the transmission of this packet.
 - e. **BSSID address:** 6 octet address, position depends on the values of toDS and fromDS flags.
 - f. **DS address:** 6 octet address, position depends on the values of toDS and fromDS flags.
 - g. **SA address:** 6 octet address, position depends on the values of toDS and fromDS flags.
 - h. **SSID:** variable length. First octet gives the Element ID (0) and the second octet gives the length.
 - i. **Basic and supported rates:** variable length. First octet gives the Element ID (1) and the second octet gives the length.

Issues encountered during implementation.

1. **Frame synchronization using long preamble:** At first, frame synchronization was tried using the repetitive patten in the long preamble field. But it gave different frame index values than what was found using the short preamble. The later one synchronized the starting of the packet very well than long preamble.
2. **Multiple Wi-Fi packets and incomplete Wi-Fi packets:** The algorithm at first computed the autocorrelation on the entire received signal and took the index where the autocorrelation was maximum to be the packet starting index. That way, two problems were faced:
 - a. **Multiple packets:** This algorithm couldn't detect if there were multiple packets present in the received signal.
 - b. **Incomplete packets:** If there are incomplete Wi-Fi 802.11g packets with high energy (as was in the case of a given dataset), where only the fields until the data fields are present, the auto-correlator would find strong peaks for the short preambles. But they would be wrong peaks as the packet is not complete and the data field is not present.

The algorithm was modified to detect all the Wi-Fi packets present in the received signal. It can also discard incomplete packets as for those wrong detections, the CRC-32 doesn't match.
3. **CRC-32 algorithm:** At first, it wasn't obvious whether the CRC-32 check is giving the right output. As the validity of a Wi-Fi 802.11g packet is based upon the CRC-32 check in this algorithm, the CRC-32 algorithm is first checked with a character '123456789' and the answer is checked to be sure that the written CRC-32 code is reliable.
4. **Speed of program:** As the algorithm designed is extracting all possible 802.11g Wi-Fi packets in the received signal and using CRC-32 for validity of the packets, initially the code was taking long time to show results (for example, for the last dataset of sample number 1 million!). So, instead of normal 'for' loop, variable slicing and 'parfor' were used to speed up the execution time.

Constellation diagrams of signal and data before and after the necessary compensation methods.

Below, results of the algorithm are shown for the uploaded dataset: I_OFDM1.txt, Q_OFDM1.txt

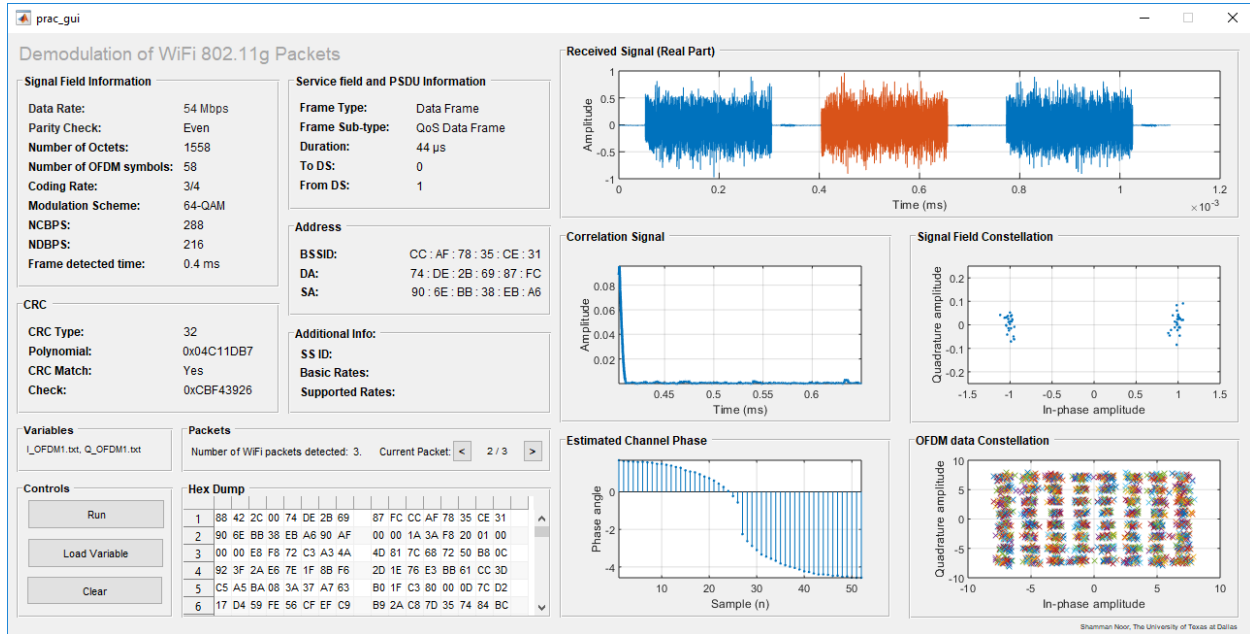


Fig 1. GUI result for dataset: I_OFDM1.txt, Q_OFDM1.txt

Additional Plots and data:

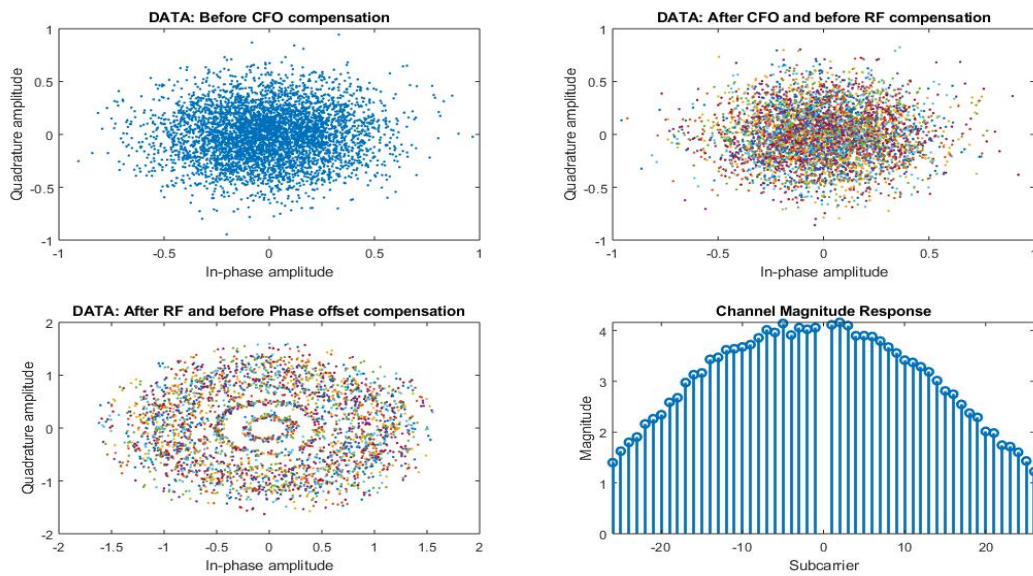


Fig 2. (a) Channel magnitude response, Constellation diagram of DATA field (b) before CFO compensation, (c) after CFO and before RF compensation and, (d) after RF and before phase compensation

Parameter	Value
Frame Synchronization Index (second Wi-Fi packet)	8082
Frequency Offset (kHz)	19.849
Phase Offset (degrees)	-5.2889
Error Vector Magnitude (dB)	-11.5173
Error Vector Magnitude (%)	26.5542

MATLAB Interational GUI Structure.

Following is the MATLAB Interational GUI used to display the information of the Wi-Fi packet information in this project.

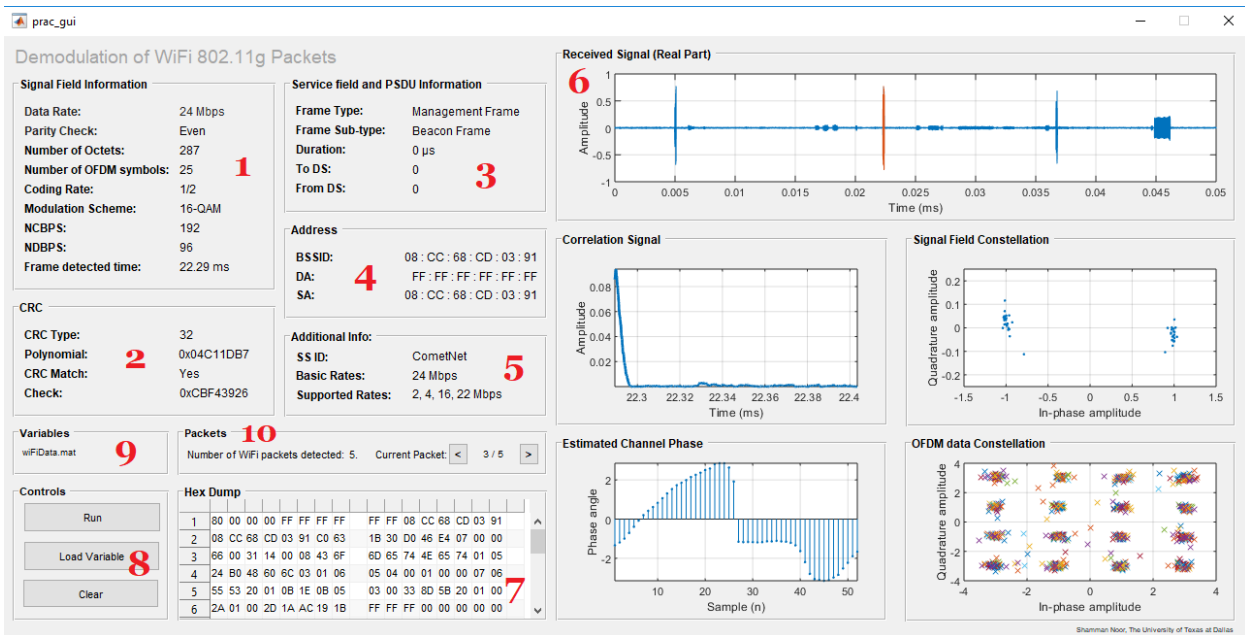


Fig 2. MATLAB GUI Structure

1. **Signal field Information:** In this field, those data are displayed which are directly obtained from the SIGNAL field and from the standard 802.11g table in [1] using SIGNAL field rate information.
 - Data Rate: obtained from table 1 using SIGNAL Field rate information.
 - Parity Check: calculated using SIGNAL field decoded data bits.
 - Number of Octets: extracted directly from SIGNAL field.
 - Number of OFDM symbols: calculated using formula:

$$\text{Number of OFDM Symbols} = \text{ceil}\left(\frac{16 + 8 * \text{Number of Octets} + 6}{\text{NDBPS}}\right)$$
 - Coding Rate: obtained from table 1 using SIGNAL Field rate information.
 - Modulation Scheme: obtained from table 1 using SIGNAL Field rate information.
 - NCBPS (Number of Data Bits per Symbol): obtained from table 1 using SIGNAL Field rate information.
 - NDBPS (Number of Data Bits per Symbol): obtained from table 1 using SIGNAL Field rate information.
 - Frame detected time: calculated from the frame index and sampling rate of 20 MHz.
2. **CRC:** The CRC field contains the following information:
 - CRC Type: The type of CRC performed. According to the standard of IEEE 802.11g [1], CRC-32 is used.
 - Polynomial: According to the standard [1], the polynomial used for CRC-32 is 0x04C11DB7, excluding the MSB.
 - CRC Match: If the CRC (last 32 bits or 4 octets) of the descrambled data matches with the calculated CRC of the rest of the data (excluding service bits), then this string shows “Yes”, otherwise “No”.
 - Check: At the beginning of the algorithm, the CRC-32 algorithm is applied on the character “123456789” or hex 0x313233343536373839. From the standard, the CRC-32 of this character should be 0xCB43926. If this value doesn’t match with the standard, then the CRC-32 isn’t working properly and the validity decision taken by the algorithm cannot be trusted.
3. **Service Field and PSDU Information:** This field contains the following:
 - Frame Type: Displays whether the frame is Management, Control or Data type.
 - Frame Subtype: Shows to which subtype the frame type belongs.
 - Duration: Shows how much time was allocated for transmission of this packet.
 - ToDS flag: If this flag is 1, then transmission occurred from access point to a client and vice versa.
 - fromDS flag: If this flag is 1, then transmission occurred from a client to an access point and vice versa.
4. **Address:** This field contains 3 addresses:
 - BSSID: The BSSID is the MAC address of the access point.
 - DA: MAC address of the destination station
 - SA: MAC address of the transmitting station

The position of these three addresses depend on the two flags in the frame control field, toDS and fromDS.

toDS flag	fromDS flag	Address 1	Address 2	Address 3
1	0	BSSID	SA	DA
0	1	DA	BSSID	SA

5. **Additional Info:** This field contains the following information
- I. **SSID:** The service set ID (SSID) identifies the wireless LAN over which a frame is transmitted. The SSID field in the hex dump is of variable length. The first octet represents the element ID, for SSID element which is 0. The second octet gives the length of the SSID. Starting from the third octet, the SSID name begins up to the specified length.
 - II. **Basic and Supported rates:** Similar to SSID, the Rates field has an element ID in the first octet, which is 1. The second octet gives us the length. The rates can be divided into two groups:
 - i. If the MSB of the rate is 1, the rate is a basic rate.
 - ii. If the MSB of the rate is 0, the rate is a supported rate.The value of the rate is the number excluding the MSB times 500 kbps or .5 Mbps.
6. **Plots:** There are 5 plots displayed in the MATLAB Interactional GUI:
- I. **Received signal:** The real part of the received signal is displayed in this axes with the current Wi-Fi packet signal highlighted in orange.
 - II. **Correlation Signal:** The correlation is performed only inside each truncated signal that are detected using the power estimation of threshold 0.003. Thus, the length of the correlation signal is reduced and the peak is found only at the beginning, as the short preamble is present in the beginning of a Wi-Fi packet.
 - III. **Signal Field constellation:** Signal field constellation is displayed in this axes.
 - IV. **Estimated Channel Phase:** Estimated channel phase (length 52) is unwrapped and displayed in this axes.
 - V. **DATA Constellation:** The final constellation of DATA fields is displayed in this axes.
7. **Hex Dump:** In this field, the hex dump of the DATA is displayed. This field came in very handy when cross checking for different fields, for example: checking the addresses, SSID element ID and length, rates element ID and length etc.
8. **Controls:** The control panel has three buttons:
- I. **Load:** This button loads the variable that is decoded by the algorithm. This GUI can load two types of variables:
 - i. **Single variable:** one single variable can be loaded in the GUI using 'Load' button. The variable name in this single variable should be named 'Y'.
 - ii. **Double variable:** Two variables can also be loaded in the GUI in case the real and complex values of the received signal are in separate files, which was the case for this project. The received signal's real and imaginary parts were saved in two different text files. For user convenience, this feature is added to the GUI.
 - II. **Run:** This button executes the algorithm on the loaded variable.
 - III. **Clear:** This button clears the fields, axes and the command prompt. It doesn't clear the variables.
9. **Variables:** This field shows the names of the variables that are loaded in the GUI. Otherwise, confusion is created whether a variable is loaded or not and also on which variable the user is working.
10. **Packets:** This field of the GUI lets the user interact with it. It has two parts:
- I. **Total number of packets detected:** After finishing executing the algorithm, a final count on detected and decoded valid OFDM Wi-Fi 802.11g packets is computed and that count is displayed here.
 - II. **Current Packet and Navigation Keys:** This is the user interaction section of the GUI. This field has two buttons to move forward and backwards and the number of the current packet out of total packets. The forward button takes the user to the next Wi-Fi packet in the received signal and highlights that portion of the received signal in orange in the first plot. The back button does the similar but takes the user to the previous packet. If the detected number of Wi-Fi 802.11g packets is 0, this buttons won't do anything. In MATLAB, all decoded Wi-Fi packets' information are stored in a structure and when the user wants to display a particular packet information using the navigation keys, that particular packet information is displayed.

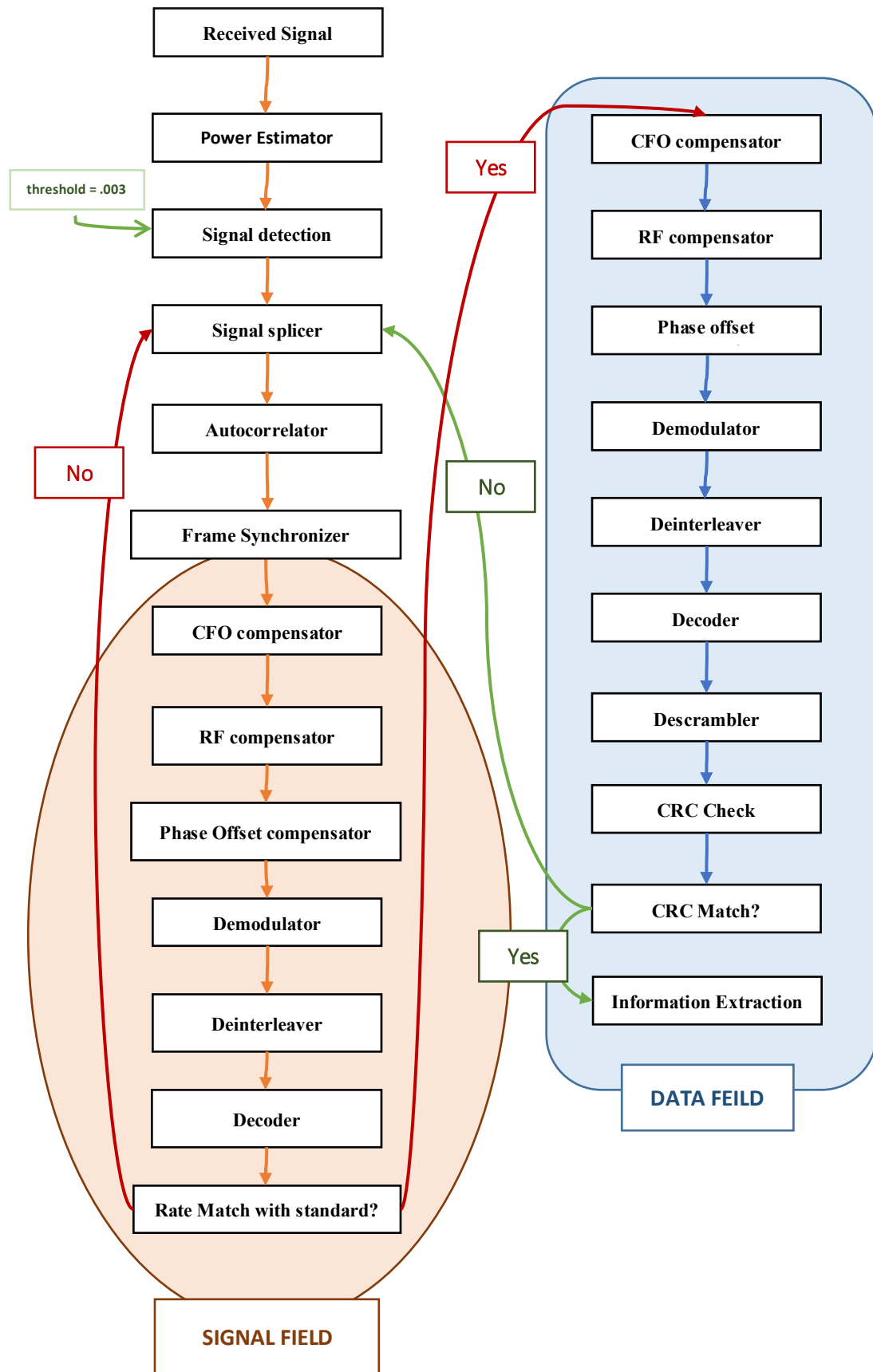


Fig. Flow diagram of developed algorithm

[Results on Datasets.](#) The total number dataset that are tested using this algorithm is 11: 5 collected and 6 given. Three types of SSIDs were found in these datasets: UTDGuest, CometNet and eduroam. All these packets' sub types were beacon frames and used 16-QAM modulation. All of the results are not shown in this report. Some significant results are shown below.

1. SSID: eduroam
Subtype: Beacon (Management)
Mod: 16-QAM
Variable Name: 'labviewWiFiDataNew.mat' – given
Packet Number/ Total Packets: 1/2 (2.6 ms) and 2/2 (4.48 ms)

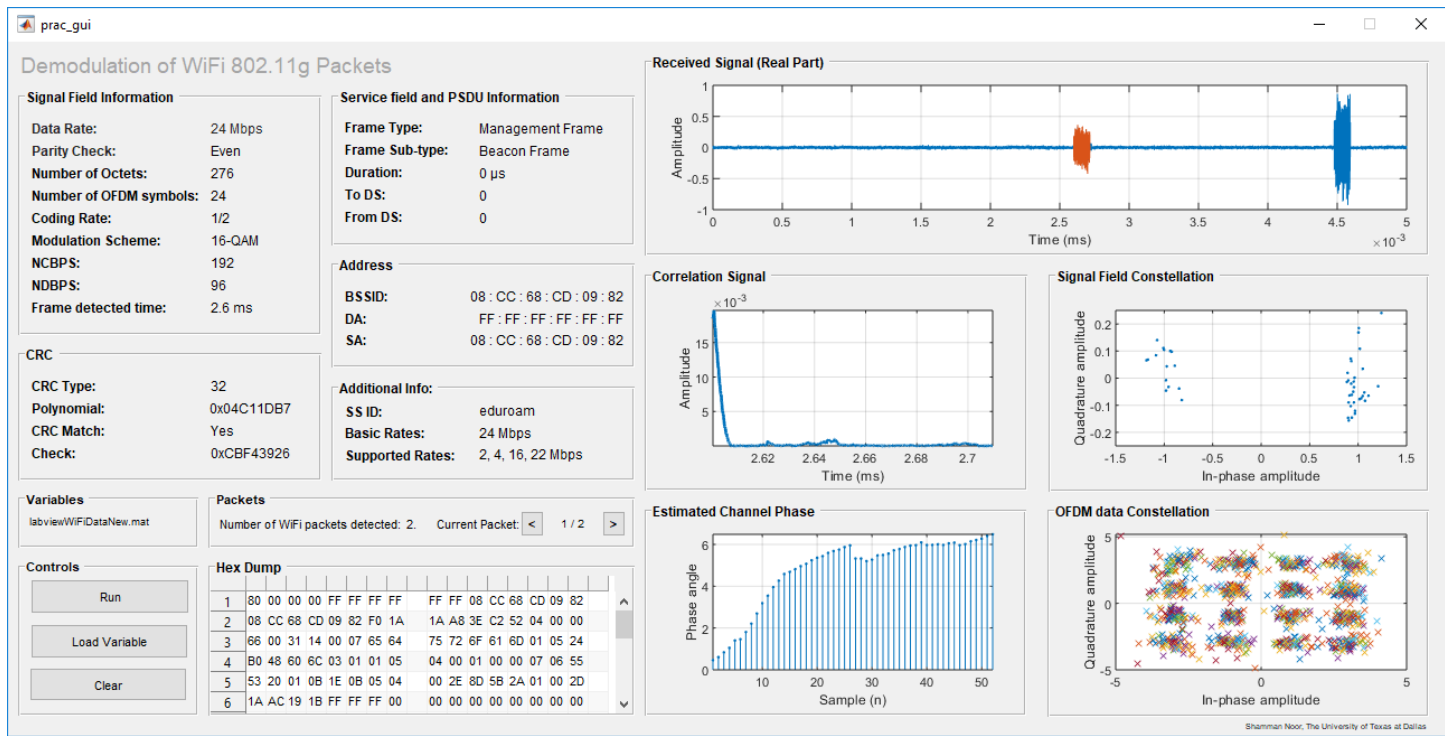


Fig. 3: SSID – ‘eduroam’, Subtype – Beacon, Variable: ‘labviewWiFiDataNew.mat’ – given, Packet number/ Total packets – ½ (2.6 ms)

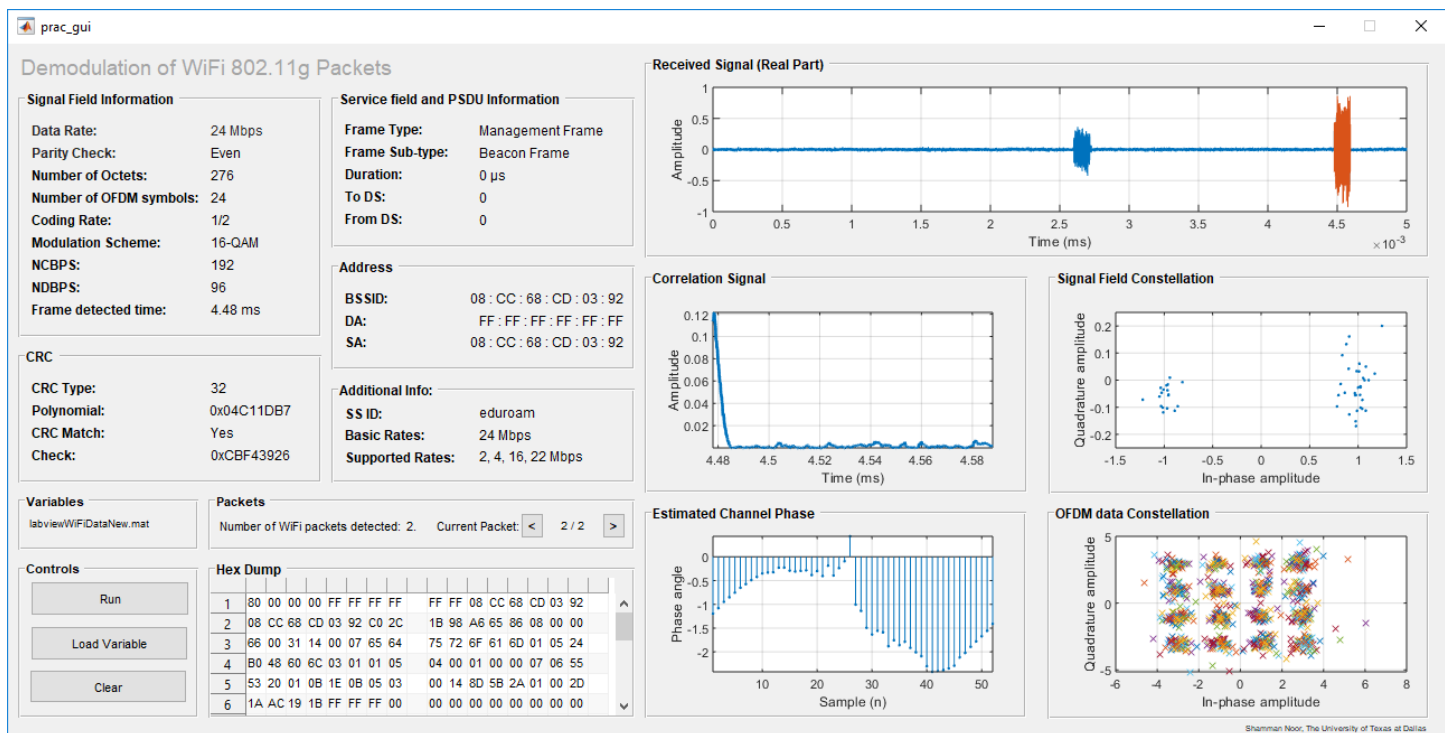


Fig. 4: SSID – ‘eduroam’, Subtype – Beacon, Mod – 16-QAM, Variable: ‘labviewWiFiDataNew.mat’ – given, Packet number/ Total packets – 2/2 (4.48 ms)

2. **SSID:** CometNet
Subtype: Beacon (Management)
Mod: 16-QAM
Variable Name: 'wifiData.mat' (given)
Packet Number/ Total Packets: 3/5 (22.29 ms)

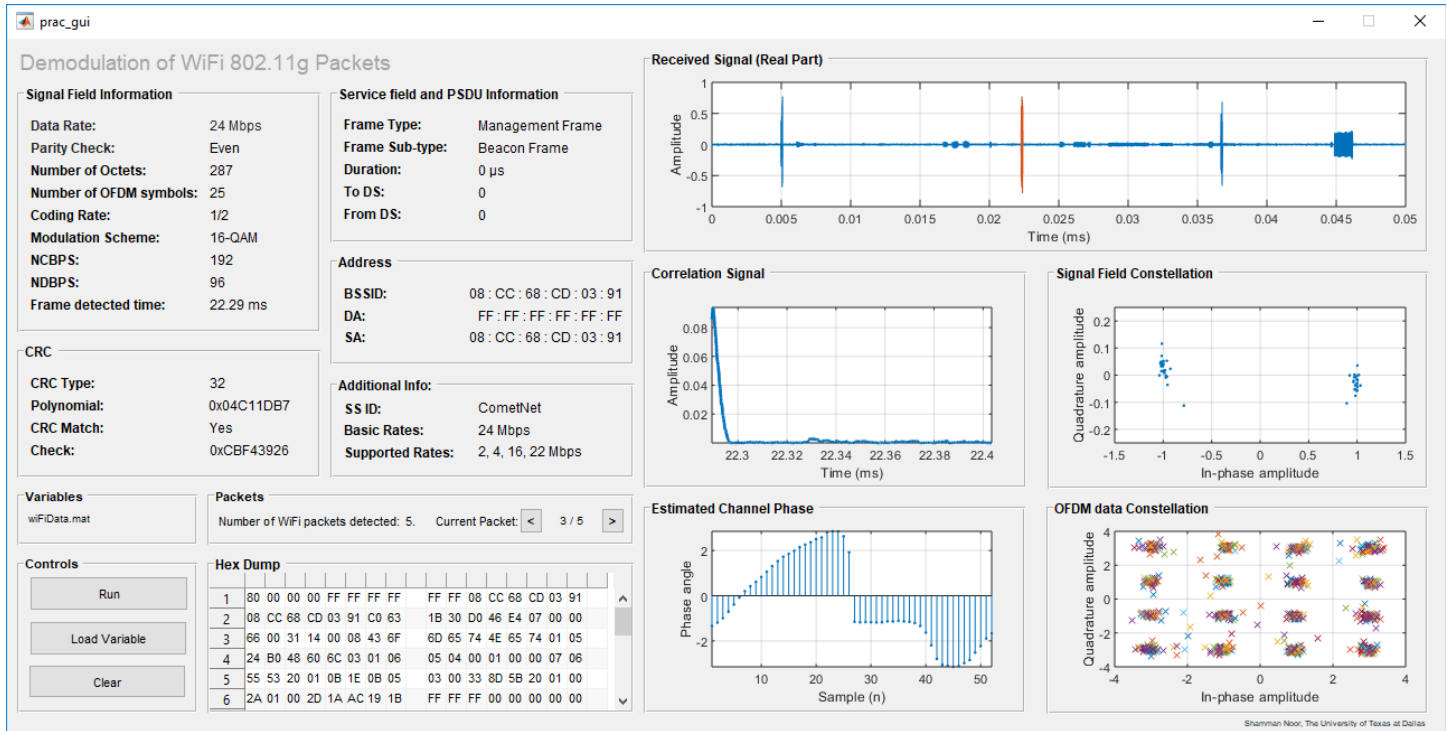


Fig. 5: SSID – 'CometNet, Subtype – Beacon, Mod – 16-QAM, Variable: wifiData.mat' (given), Packet number/ Total packets – 3/5 (22.29 ms)

3. **SSID:** UTDGuest
Subtype: Beacon (Management)
Mod: 16-QAM
Variable Name: Collected Data - 1
Packet Number/ Total Packets: 1/1 (3.76 ms)

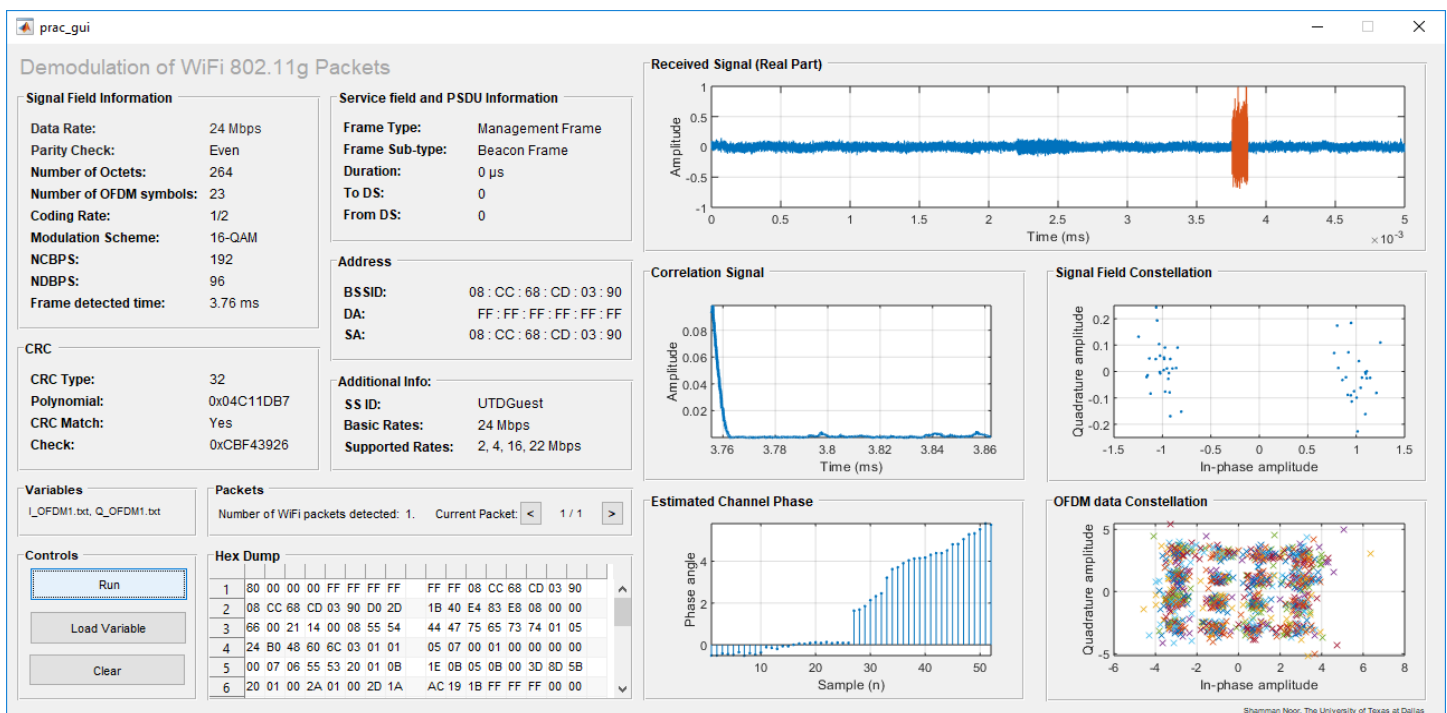


Fig. 6: SSID – 'UTDGuest, Subtype – Beacon, Mod – 16-QAM, Variable: Collected 1, Packet number/ Total packets – 1/1 (3.76 ms)

4. **Type:** Data
Subtype: QoS Data
Mod: 64-QAM
Variable Name: wifiData_index_915909_code_3_4_54qam (given)
Packet Number/ Total Packets: 17/184 (4.11 ms)

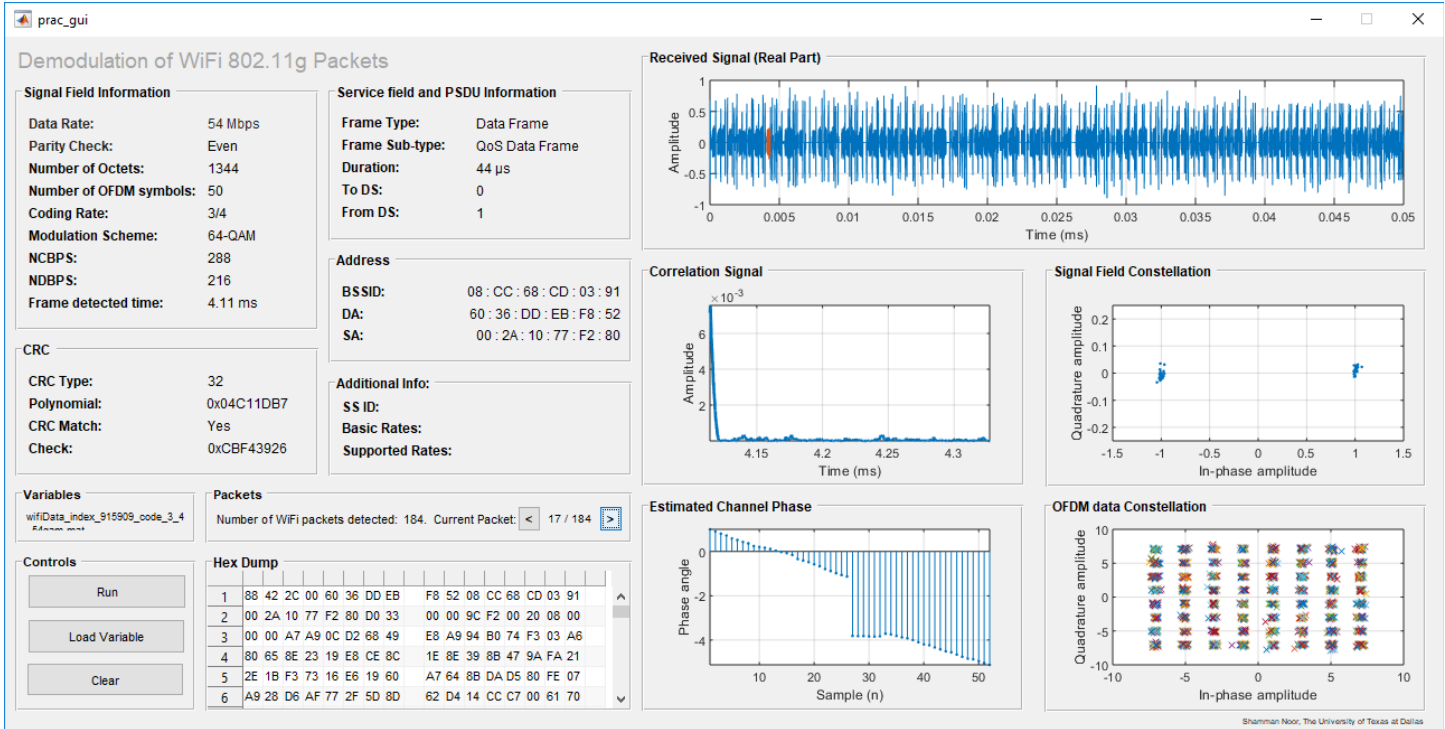


Fig. 7: Type - Data, Subtype – QoS Data, Mod – 64-QAM, Variable: wifiData_index_915909_code_3_4_54qam (given), Packet number/ Total packets – 17/184 (4.11 ms)

5. **Type:** Data
Subtype: QoS Data
Mod: QPSK
Variable Name: wifiData_index_751455_code_3_4_54qam (given)
Packet Number/ Total Packets: 20/128 (6.87 ms)

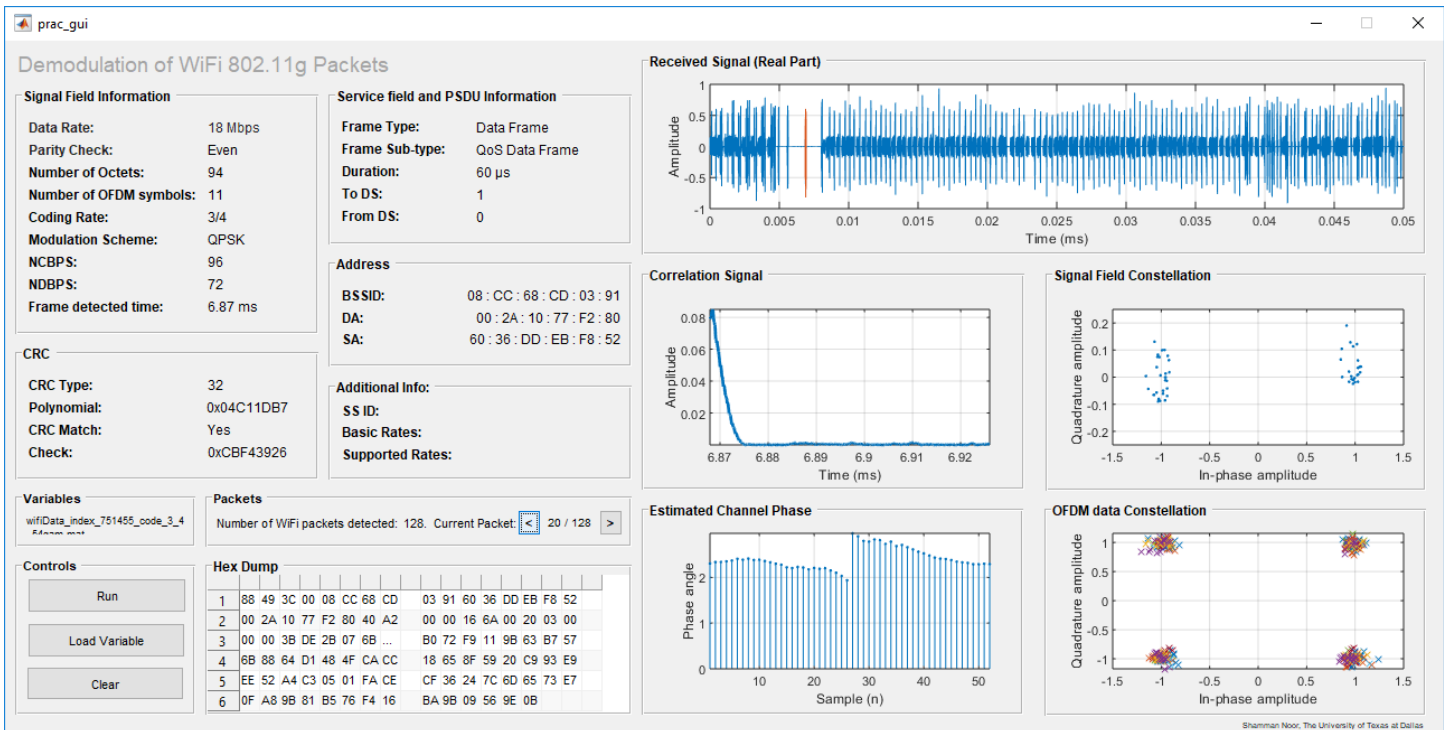


Fig. 8: Type - Data, Subtype – QoS Data, Mod – QPSK, Variable: wifiData_index_751455_code_3_4_54qam (given), Packet number/ Total packets – 20/128 (6.87 ms)

6. **Type:** Control
Subtype: RTS Frame
Mod: 16-QAM
Variable Name: Collected 2
Packet Number/ Total Packets: 1/1 (3.03 ms)

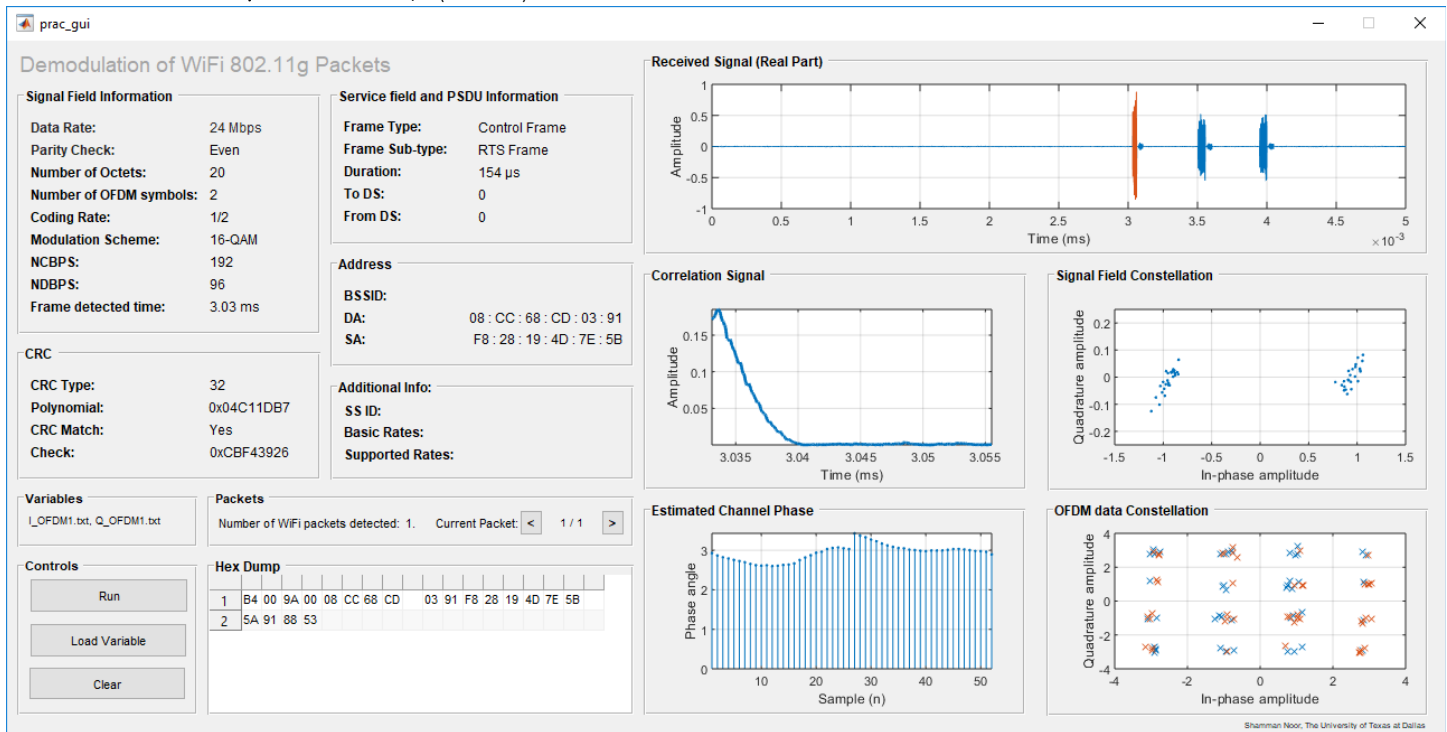


Fig. 9: Type - Control, Subtype – RTS frame, Mod –16-QAM, Variable: collected 2, Packet number/ Total packets – 1/1 (3.03 ms)

7. **Type:** Control
Subtype: ACK Frame
Mod: BPSK
Variable Name: wiFiData.mat (given)
Packet Number/ Total Packets: 5/5 (36.75 ms)

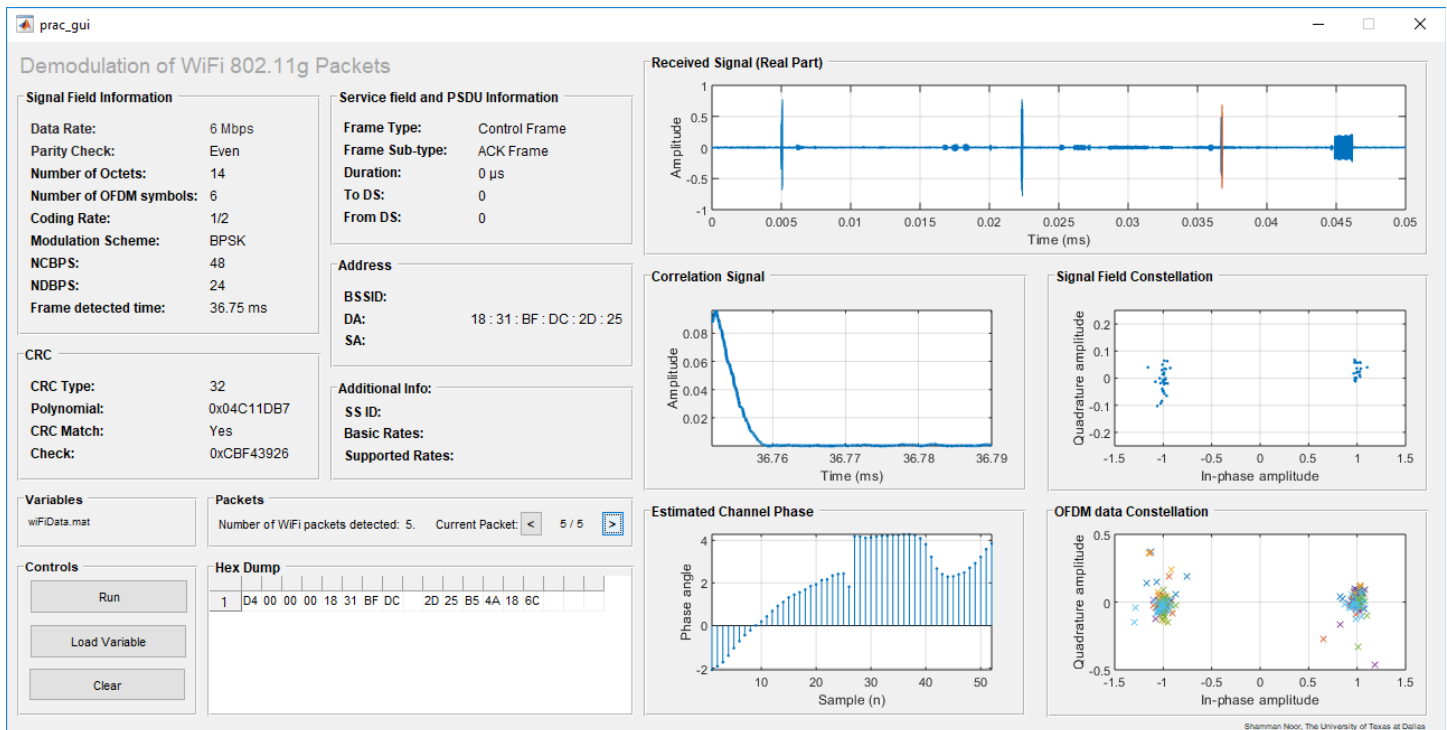


Fig. 10: Type - Control, Subtype – ACK frame, Mod –BPSK, Variable: wiFiData.mat (given), Packet number/ Total packets – 5/5 (36.75 ms)

Besides these results, there are hundreds of results from the given data.

References.

- [1] 802.11-2016 - IEEE Standard for Information technology--Telecommunications and information exchange between systems Local and metropolitan area networks--Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.