Low SNR FreeDV Mode

David Rowe VK5DGR

October 17, 2023

1 Introduction

After 10 years development and on air experience with various FreeDV waveforms, we would like to develop a new waveform that outperforms and replaces a variety of existing modes such as 700C/D/E and 1600. Requirements include [6]:

1. Better performance than SSB at 0dB SNR on MPP and MPD channels.
2. A single mode that can handle MPP, MPD, GEO (e.g. QO-100), and replace several existing FreeDV modes, simplifying the end user experience.
3. For compliance with Export Control regulations, the minimum speech codec bit rate is 700 bit/s.
4. Use of legacy analog HF radio sets with a RF bandwidth of around 2000 Hz.
5. Compliant with the 300 baud per carrier limit set by the United States FCC, which implies a parallel tone or OFDM modem.

It is acceptable for performance to gradually decrease as the multipath channel quality declines, but we would like the decline to be gradual, e.g. a few dB more power for operation on MPD versus MPP.

This document explores ways we can improve the existing OFDM modem waveforms in order to meet these requirements.

1.1 Summary of Improvements

Compared to the baseline FreeDV 700D/E waveforms:

1. New equalisation algorithm that performs well on a range of AWGN and multipath channels, removing the need for modes, and providing several dB improvement for 700E users.
2. Diversity in frequency and possibly time providing around 3dB improvement over 700D on multipath channels.
3. Dynamic threshold and pilot symbol selection for ± 200 Hz acquisition (four times wider that 700D/E) at lower SNR, simplifying end user tuning.

1.2 Glossary

Acronym	Explanation
AWGN	Additive White Gaussian Noise - a communications channel with flat frequency response and additive noise
CP	Cyclic Prefix
FEC	Forward Error Correction
ISI	Inter Symbol Interference
LEO	Low earth orbit satellite channel, AWGN with large freq offset and Doppler shift (high rate of change of freq offset)
GEO	Geosynchronous satellite channel, AWGN but high phase noise and large freq offset
OTA	Over The Air
PTT	Push To Talk - voice communications where only one person is transmitting at any one time. Common in two way radio but not mobile telephones
MPG	Multipath Good channel, 0.1 Hz Doppler spread, 0.5ms delay spread
MPP	Multipath Poor channel, 1 Hz Doppler spread, 2ms delay spread, typical for US and Australian inter-state propagation
MPD	Multipath Disturbed channel, 2 Hz Doppler spread, 4ms delay spread, typical for UK Winter NVIS propagation

Table 1: Glossary of Acronyms

[1] Can be expressed as a linear ratio E_b/N_0 or $10log_{10}(E_b/N_0)$ dB

Symbol	Explanation	Units
B	Noise or signal bandwidth	Hz
B_d	Doppler spreading bandwidth for HF channel model	Hz
D	Algorithmic delay, how much speech we need to buffer before processing starts	seconds
E_b/N_0	Energy per bit on spectral noise density	dimensionless, dB[1]
N_c	Number of carriers	
N_d	Number of diversity Tx signals	
N_s	Number of symbols in a *modem frame*, pilot insertion rate	
R_b	Bit rate	Bits/second
R_s	Symbol rate	symbols/second
T_s	Symbol period	seconds
SNR	Signal to Noise Ratio	dB
S	Signal Power	Watts
N	Noise Power	Watts
N_c	Number of carriers in an OFDM waveform	
ω_c	Angular frequency of OFDM carrier c	rads/second

Table 2: Glossary of Symbols

2 Modem and Channel Models

In this section we will develop theoretical models to help us explore performance limits.

2.1 SNR and Bandwidth Limits

For practical PTT voice systems algorithmic delay is limited to a few 100ms, which limits the FEC codeword size and hence the performance of the code. For PSK channels a threshold $E_b/N_0 = 2\,\text{dB}$ and a code rate $R = 0.5$ is typical, where E_b/N_0 is the energy per payload data bit (coded E_b/N_0). The lowest (threshold) SNR for a viable voice link is given by:

$$\begin{aligned} \frac{S}{N} &= \frac{E_b R_b}{N_0 B} \\ SNR &= 10log_{10}\left(\frac{E_b}{N_0}\right) + 10log_{10}\left(\frac{R_b}{B}\right) \quad \text{[dB]} \end{aligned} \tag{1}$$

where R_b is the payload data bit rate, and B is the bandwidth in which we measure SNR. Given $Rb = 700$ and $B = 3000$ we have:

$$\begin{aligned} SNR &= 2 + 10log_{10}(700/3000) \\ &= -4.3\,\text{dB} \end{aligned} \tag{2}$$

This is ideal performance for an AWGN channel. In practice we must allocate some power to symbols used for synchronisation, such as pilot symbols used for frequency and phase estimation, or unique word bits used for frame synchronisation. Synchronisation algorithms often struggle at low SNRs, introducing additional "implementation" losses.

Performance on multipath channels is significantly worse, in our use cases typically 5 dB. On these channels, we may allocate some carrier power to deal with intersymbol interference (for example a cyclic prefix in OFDM modems).

A more complete model is:

$$SNR = 10log_{10}\left(\frac{E_b}{N_0}\right) + 10log_{10}\left(\frac{R_b}{B}\right) + L_p + L_{il} + L_{cp} \tag{3}$$

where L_p is the loss from power allocated to pilot symbols, L_{il} is the real world implementation loss, and L_{cp} is the loss in SNR due to the power allocated to the cyclic prefix.

To combat intersymbol interference and frequency selective fading on HF multipath channels, it is convenient to use parallel tone or OFDM modems. The signal is divided into N_c carriers, each carrying R_b/N_c bits/s. Due to the spectral efficiency of OFDM, the same bit rate, symbol rate, and occupied bandwidth is required for any N_c. For example:

1. A QPSK signal of $R_s = 1000$ symbols/s and $R_b = 2R_s = 2000$ bits/s occupies $B = 1000$ Hz (central lobe).

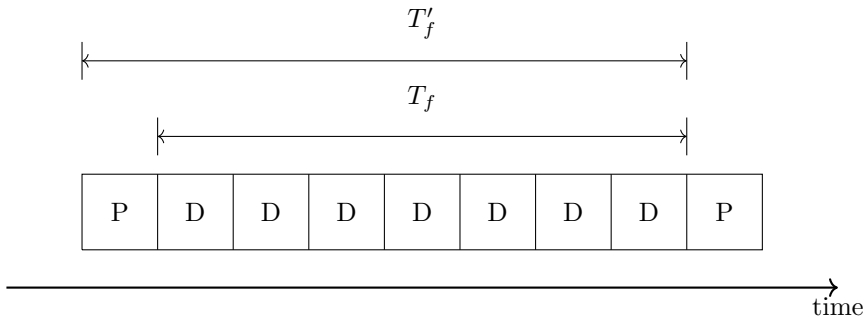
2. An OFDM signal of $N_c = 20$ carriers each at $R_s = 50$ symbol/s gives us a total $R_b = 2N_cR_s = 2000$ bits/s and occupies a RF bandwidth of $B = N_cR_s = 1000$ Hz.

This simple relationship between per-carrier and total R_s allows us to use them interchangeably for some calculations.

FreeDV signals are typically transmitted over the air using analog HF radios. We assume a RF bandwidth of 2000 Hz is available, which allows a maximum of 4000 bit/s using QPSK. In practice the payload data rate is much less, due to various overheads that we will discuss below.

2.2 Pilot Symbol Overhead

Figure 1: Modem Frame with $N_s = 8$, the pilot of the next modem frame is also shown.



In this section we explore the effect of inserting pilot symbols on the threshold SNR (1). Consider a sequence of $N_s - 1$ PSK data symbols that carry the modulated FEC codeword bits (e.g. data and parity bits) over the channel. We denote this sequence a modem *modem frame*. The frame of $N_s - 1$ symbols has a period of $T_f = (N_s - 1)T_s$ seconds, where T_s is the period of each symbol. We wish to insert a single pilot symbol after the data symbols, creating a new frame N_s symbols long, with period $T'_f = N_sT_s$. To maintain the same payload data rate:

$$\begin{aligned} T_f &= T'_f \\ (N_s - 1)T_s &= N_sT_s \\ R'_s &= R_s\frac{N_s}{N_s - 1} \end{aligned} \tag{4}$$

where the symbol rate $R_s = 1/T_s$. Expressing S/N (1) in terms of E_s and R_s:

$$\begin{aligned} \frac{S}{N} &= \frac{E_s R_s}{N_0 B} \\ \frac{S'}{N} &= \frac{E_s R'_s}{N_0 B} \\ &= \frac{E_s R_s N_s}{N_0 B(N_s - 1)} \\ \frac{S'/N}{S/N} &= \frac{N_s}{N_s - 1} \end{aligned} \tag{5}$$

Thus when we insert pilots, the threshold S/N increases by a factor of $N_s/(N_s - 1)$. Expressed in dB:

$$\begin{aligned} 10log_{10}\left(\frac{S'}{N}\right) &= 10log_{10}\left(\frac{S}{N}\right) + 10log_{10}\left(\frac{N_s}{N_s - 1}\right) \\ SNR' &= SNR + 10log_{10}\left(\frac{N_s}{N_s - 1}\right) \\ SNR' &= SNR + L_p \quad \text{[dB]} \end{aligned} \tag{6}$$

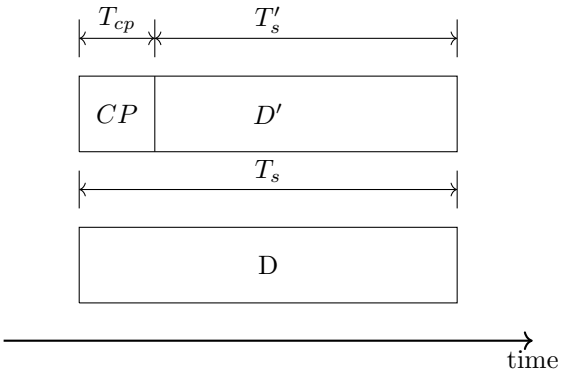
where L_p can be considered the pilot symbol *loss* - the SNR degradation from the ideal performance (1) due to the insertion of pilot symbols. For example FreeDV 700D uses a pilot insertion rate of $N_s = 8$ results in $L_p = 10log_{10}(8/7) = 0.58$ dB, thus we need 0.58 dB more SNR to achieve the threshold SNR for the voice link. For this example let Rs be the total symbol rate over all N_c carriers. To maintain $Rs = 700$ data symbols/second over the channel, we require $R'_s = (700)8/7 = 800$ symbols/second which introduces a 100 Hz bandwidth overhead.

2.3 Cyclic Prefix Overhead

Now we consider the SNR overhead for the Cycle Prefix (CP) used in OFDM modems to cope with delay spread on multipath channels. To achieve our payload data rate (e.g. 700 bits/s), we FEC encode and map the bits to PSK symbols D, and distribute the symbols across N_c parallel carriers. The transmitter power is spread equally across all carriers.

We send symbols D across the channel at a constant symbol rate R_s, or one symbol every $T = T_s$ seconds. To cope with delay spread, we construct a composite symbol by pre-pending a Cyclic Prefix (CP) T_{cp} seconds in duration to a new symbol D' of T'_s seconds in duration. D and D' contain the same PSK symbol, and convey the same information over the channel. The new composite symbol is now $T' = T_{cp} + T'_s$ seconds long. The CP contains no additional information, it is just a cyclic extension of the single symbol D'. Thus we still send one symbol of data over the channel every T' seconds. To maintain the payload data rate over the channel, we must send the new composite symbol at

Figure 2: Construction of composite symbol with a Cyclic Prefix CP pre-pended to a shortened data symbol D'.



the same rate as the original symbol:

$$\begin{aligned} T &= T' \\ T_s &= T_{cp} + T'_s \\ R'_s &= \frac{R_s}{1 - T_{cp}/T_s} \end{aligned} \tag{7}$$

It can be observed that $R'_s > R_s$, to account for the portion of the composite symbol allocated to the CP. For example with $R_s = 50$, $T_s = 0.02$, $T_{cp} = 0.002$, $R'_s = 50/(1 - 0.002/0.02) = 55.56$ symbols/second. Thus additional bandwidth is required to send the composite symbol including the cyclic prefix.

The increase in symbol rate does not directly affect BER performance if E_s/N_0 remains the same. For example if $Rs' = 2R_s$ we could send the symbol across the channel in $T_s/2$ seconds at power $2S$, followed by $T_s/2$ seconds of silence. The energy per symbol E_s and BER would remain the same.

For the composite symbol, the transmitter power S' is spread between the CP and D'. Given a constant power S', the energy for the symbol D' is given by:

$$\begin{aligned} E'_s &= \frac{S'}{R'_s} \\ &= \frac{S'}{R_s}(1 - T_{cp}/T_s) \end{aligned} \tag{8}$$

For the link BER to be maintained the energy per symbol must be unchanged,

i.e. $E'_s = E_s$:

$$\begin{aligned}
E_s &= \frac{S'}{R_s}(1 - T_{cp}/T_s) \\
\frac{S}{R_s} &= \frac{S'}{R_s}(1 - T_{cp}/T_s) \\
\frac{S}{N} &= \frac{S'}{N}(1 - T_{cp}/T_s) \\
\frac{S'}{N} &= \frac{S}{N}\left(\frac{1}{1 - T_{cp}/T_s}\right) \\
SNR' &= SNR - 10log_{10}(1 - T_{cp}/T_s) \\
&= SNR + L_{cp} \quad \text{[dB]} \\
L_{cp} &= -10log_{10}(1 - T_{cp}/T_s)
\end{aligned} \tag{9}$$

Thus to close the link with the composite symbol the S/N must be increased by a factor of $1/(1 - T_s/T_{cp})$ compared to our ideal modem, to account for the energy allocated to the CP. For example FreeDV 700E has $T_s = 0.02$, $T_{cp} = 0.006$, giving $L_{cp} = -10log10(1 - 0.006/0.02) = 1.55$dB.

2.4 Bandwidth

Using the expressions above, we can estimate the total RF bandwidth of candidate waveforms:

$$B = \frac{R_s N_s}{(N_s - 1)(1 - T_{cp}/T_s)} \tag{10}$$

where R_s is the total symbol rate over the channel. For example for 700D, $B = 700(8)/((8 - 1)(1 - 0.002/0.02)) = 888.89$ Hz. In practice 700D has some extra bits allocated for frame synchronisation and auxiliary text, resulting in $B = 944$ Hz.

2.5 Latency

Latency is delay experienced by end users when using a speech processing system such as FreeDV. Latency is an important trade off for real time, Push To Talk (PTT) speech communication. Speech compression and FEC tend to work better with latency, but subjective quality tends to decrease whenever there is delay in conversational speech.

The algorithmic delay is defined as the number of seconds D of signal that must be collected in order to start processing. A typical speech codec will collect 20-40ms of speech samples before starting analysis. However much longer FEC codewords sizes are desirable in order to ride over fades. For example FEC codeword sizes spanning several seconds are common for HF data modes. For maximum robustness, interleaving require all bits in the codewords to be received before FEC decoding starts. The modem may also introduce delay, for example FreeDV 700D has a one frame look ahead for pilot symbols.

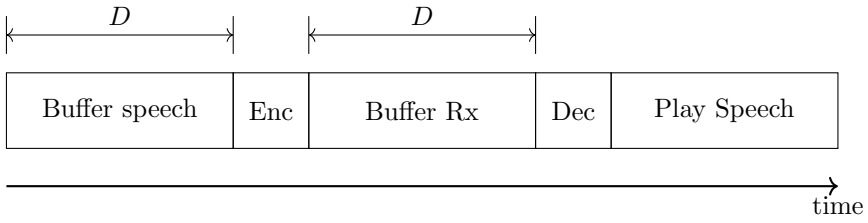
Consider a FreeDV mode with algorithmic delay D seconds (Figure 3)

1. First we must collect D seconds of speech samples.
2. The speech samples are encoded and a buffer of D seconds of modem samples generated.
3. The modem samples are sent over the radio channel, where they are buffered at the receiver.
4. When a D second buffer of modem samples is collected at the receiver, the modem samples can be decoded to a buffer of output speech samples.
5. The output speech samples are played to a listener.

Latency is the time from when a speech sample enters the input buffer to when it is played at the receiver. If we assume infinite CPU such that the encode and decode times are zero, and zero propagation delay across the channel, the minimum latency experienced by an end user is twice the algorithmic delay, or $2D$.

The real world impacts of latency can be even more severe. On poor channels, it may take a few frames to syncronise. If one frame is lost or cannot be decoded there will be silence or garbled speech for D seconds. On marginal channels, several attempts may be required to determine if an "over" was received - each attempt will experience long PTT turn around delays.

Figure 3: Latency of a waveform with algorithmic delay D seconds



2.6 Summary of FreeDV modes

Using the framework developed above and the information in [1], we can analyse the existing FreeDV waveforms (Table 3).

2.7 HF Channel Model

A common two path HF channel model [3] is given by:

$$y(t) = x(t)G_1(t) + x(t-d)G_2(t) \tag{11}$$

where G_1 and G_2 are two time varying, complex, Gaussian filtered random variables with *Doppler Spread* bandwidth B_d Hz, d is the path delay in seconds.

Mode	N_s	R_s	Ts	T_s'	T_{cp}	L_p (dB)	L_{cp} (dB)
700C	3	75	0.013	0.013	0.000	1.77	0.00
700D	8	50	0.020	0.018	0.002	0.58	0.46
700E	4	50	0.020	0.014	0.006	1.24	1.55
datac4	5	50	0.020	0.014	0.006	0.97	1.38

Table 3: Summary of FreeDV waveforms, note the differences in the overheads L_p and L_{cp}. R_s and T_s is per-carrier. 700C is a parallel tone modem with no cyclic prefix, the modem frame is comprised of two pilot symbols and 4 data symbols

Other models with varying delay, number of paths, and fixed frequency offset components are also possible, for example Appendix B of [5].

As $B_d << R_s$ we assume G_1 and G_2 are complex constants for the duration of a single symbol. Expressed in discrete time for the current symbol:

$$y(n) = x(n)G_1 + x(n - dF_s)G_2 \tag{12}$$

where F_s is the sample rate in Hz. Taking the z-transform:

$$\begin{aligned} Y(z) &= X(z)G_1 + X(z)z^{-dF_s}G_2 \\ \frac{Y(z)}{X(z)} &= G_1 + z^{-dF_s}G_2 \\ H(z) &= G_1 + z^{-dF_s}G_2 \\ H(e^{j\omega}) &= G_1 + e^{-j\omega dF_s}G_2 \end{aligned} \tag{13}$$

Lets examine the effects of intersymbol interference due to the delay spread. Consider the transmitter sample $x(0)$ where we transition from one PSK phase ϕ_1 to the next ϕ_2. Near the transition we can define $x(n)$ in terms of a step function $s(n)$:

$$\begin{aligned} x(n) &= \begin{cases} e^{j(\omega n+\phi_1)}, & n < 0 \\ e^{j(\omega n+\phi_2)}, & n \geq 0 \end{cases} \\ &= s(n)e^{j(\omega n+\phi_2)} + (1 - s(n))e^{j(\omega n+\phi_1)} \\ s(n) &= \begin{cases} 0, & n < 0 \\ 1, & n \geq 0 \end{cases} \end{aligned} \tag{14}$$

where $\omega = 2\pi cR_s/F_s$ is the frequency of OFDM carrier c. Substituting into (12):

$$\begin{aligned} y(n) = {} & G_1 s(n)e^{j(\omega n+\phi_2)} + G_1(1 - s(n))e^{j(\omega n+\phi_1)} \\ & + G_2 s(n - dF_s)e^{j(\omega(n-dF_s)+\phi_2)} + G_2(1 - s(n - dF_s))e^{j(\omega(n-dF_s)+\phi_1)} \end{aligned} \tag{15}$$

At $n = 0$ we have a mixture of both symbols:

$$y(0) = G_1 s(n) e^{j(\omega n + \phi_2)} + G_2 e^{j(\omega(n - dF_s) + \phi_1)} \tag{16}$$

However by $y(dF_s)$ the ISI from ϕ_1 has gone:

$$y(dF_s) = G_1 s(n) e^{j(\omega dF_s + \phi_2)} + G_2 e^{j\phi_2} \tag{17}$$

Or more generally for $n \geq dF_s$:

$$\begin{aligned} y(n) &= G_1 s(n) e^{j(\omega n + \phi_2)} + G_2 e^{j(\omega(n - dF_s) + \phi_2)} \\ &= e^{j(\omega n + \phi_2)} [G_1 + G_2 e^{-j\omega dF_s}] \\ &= e^{j(\omega n + \phi_2)} H(e^{j\omega}) \end{aligned} \tag{18}$$

Thus if we start detecting $y(n)$ after the longest delay term dF_s we can recover the transmitted PSK symbol without ISI, and equalise it by estimating a single complex coefficient $H(e^{j\omega})$. In practice detection means integration over a complete symbol T_s, therefore the symbol $e^{j\omega n + \phi_n}$ must be transmitted for $d + T_s' = T_{cp} + T_s'$ seconds, with the cyclic prefix length $T_{cp} \geq d$.

2.8 Waveform Improvements

In the following sections waveform improvements are presented.

The general strategy is to propose an innovation, and explore with analysis/maths and Octave simulation. To simulate performance with voice codec, we use a threshold of PER=0.1, BER=0.01 as the threshold for establishing a voice link.

3 Equalisation

Pilot symbols are used to the estimate the channel, and equalise (correct) the phase of the received symbols. We require an equalisation system that works with $B_d = 2$ Hz Doppler Spread that has a low implementation loss L_{il} on noisy channels, and has a reasonable latency (we can't use symbols far in the future).

If we detect symbols after ISI has settled, the channel for each OFDM carrier resolves to a single complex constant $H(e^{j\omega})$. This can be considered a complex random variable with bandwidth B_d.

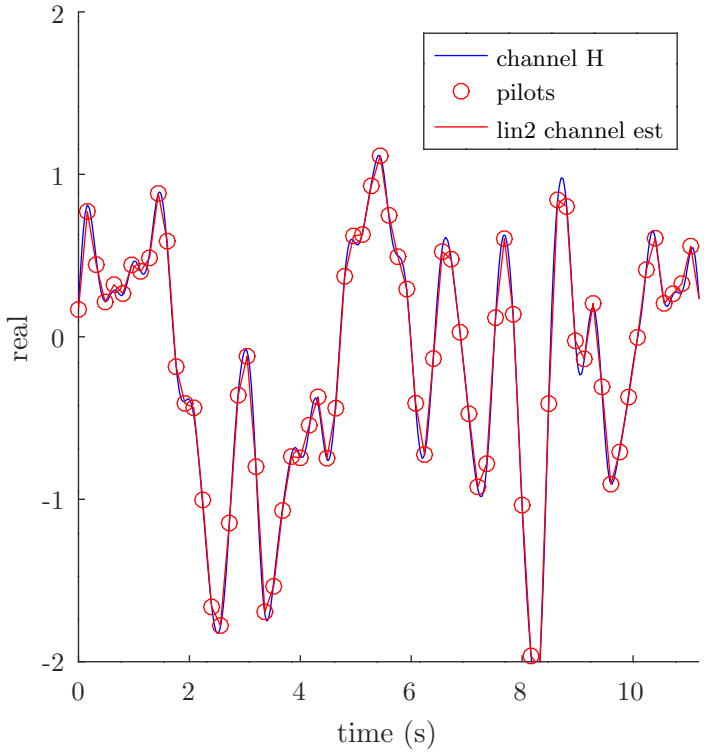
Consider the two path time domain channel model (12). Two additive terms of bandwidth B add linearly, so by linearity the result also has bandwidth B. The sum is a random modulation of bandwidth B_d about the symbol centre frequency. The Doppler bandwidth B_d therefore defines the bandwidth required for equalisation. One caveat - if B_d an appreciable fraction of R_s the DFT orthogonality may break down to some extent as energy falls into adjacent DFT bins.

FreeDV 700D samples pilots and hence $H(e^{j\omega})$ at $1/(N_s Ts) = 6.25$ Hz. As we are sampling a complex signal, this implies a Nyquist bandwidth of 6.25 Hz,

which should be adequate to recover $H(e^{j\omega})$ which has a maximum $B_d = 2$ Hz on the MPD channel. 700D currently uses a block average over a 2D array (4 pilots in time, across 3 carriers in frequency) of 12 pilots, this can be interpreted as a 2D FIR filter with all coefficients $c_i = 1/12$. It is effective on MPP channels ($B_d = 1$ Hz), but breaks down on MPD channels ($B_d = 2$ Hz). This suggests a resampler with a wider bandwidth might enable a waveform similar to 700D to be used on MPD channels.

To explore pilot resampling algorithms, a simulation was written to evaluate candidate algorithms by comparing uncoded BER versus E_b/N_0 curves. One interpretation of resampling 4 pilots in time is a 4 point FIR filter. This was explored, along with the existing *mean12* resampler (700D), and 2 point linear *lin2* resampler (as used in FreeDV 700E). Figure 4 shows the *lin2* resampler in action.

Figure 4: *lin2* resampler in action for a single OFDM carrier on a MPD channel. The blue continuous line is the simulated channel $H(e^{j\omega})$ at each symbol, the red dots are the pilot symbols, and the red line the channel estimates from the linearly interpolated pilots. Just the real part is plotted.



Attempts were made to design 4 sample FIR filters (e.g. with a *sinc*() impulse response) however these performed poorly. The *mean12* algorithm used for 700D works well on AWGN and MPP channels, but has a sluggish frequency response, and can't follow fading with $B_d > 1$ Hz. However on AWGN and slower fading channels it filters the channel noise well, resulting in good low SNR performance. This is consistent with on air reports of 700D. The *lin2* algorithm (700E) doesn't filter channel noise very well, but is hard to beat on fast fading channels, and even works with $B_d > 2$ Hz.

Attempts were therefore made to improve the performance of *lin2* on AWGN and slower fading channels. Simple averaging of pilots from adjacent carriers is used in FreeDV 700D to reduce estimation noise. Consider (13) when the carriers are spaced R_s Hz apart such that $\omega_c = 2\pi c R_s/F_s$:

$$\begin{aligned} H(e^{j\omega_c}) &= G_1 + G_2 e^{-j2\pi c R_s d} \\ H(e^{j\omega_{c-1}}) &= G_1 + G_2 e^{-j2\pi c R_s d + 2\pi c R_s d} \\ H(e^{j\omega_{c+1}}) &= G_1 + G_2 e^{-j2\pi c R_s d - 2\pi c R_s d} \end{aligned} \tag{19}$$

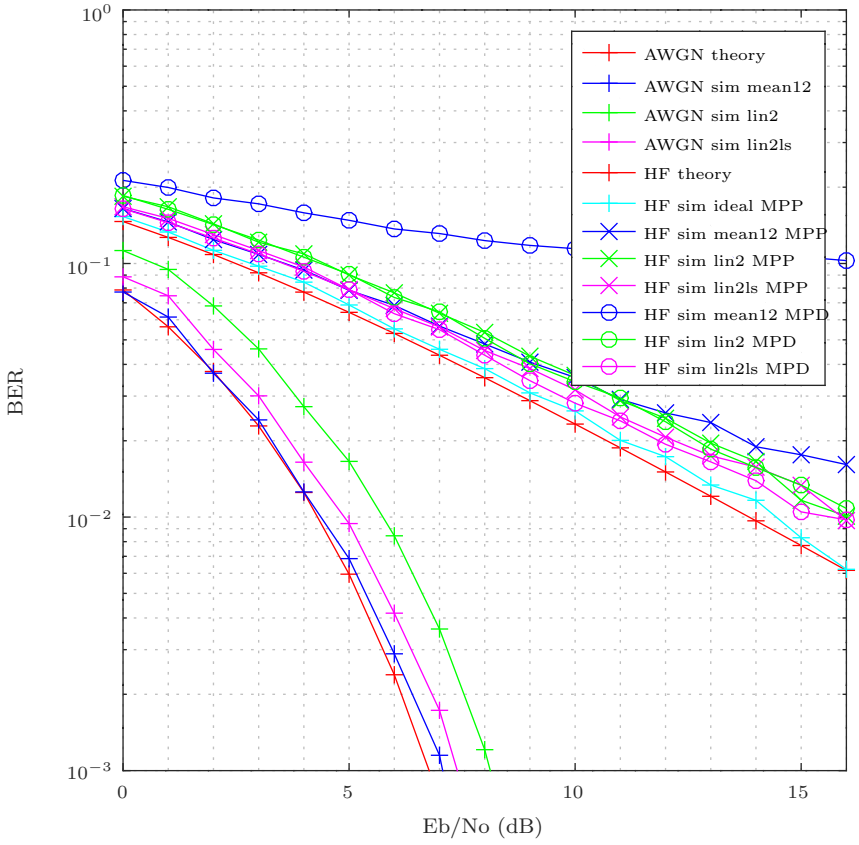
We can see some symmetry in the RH term in the phase between carrier $c-1$ and carrier $c+1$ which supports averaging over adjacent carriers to estimate $H(e^{j\omega_c})$, especially when the term $2\pi R_s d$ is small. A more robust approach is to perform a least squares fit [4] of three equations with G_1 and G_2 as the two unknowns, and treating the three pilot samples centred around the current carrier as samples of the channel at three frequencies $H(e^{j\omega_{c-1}}), H(e^{j\omega_{c-1}}), H(e^{j\omega_{c+1}})$:

$$\begin{aligned} G_1 + G_2 e^{-j\omega_{c-1} dF_s} &= H(e^{j\omega_{c-1}}) \\ G_1 + G_2 e^{-j\omega_c dF_s} &= H(e^{j\omega_c}) \\ G_1 + G_2 e^{-j\omega_{c+1} dF_s} &= H(e^{j\omega_{c+1}}) \\ \begin{bmatrix} 1 & e^{-j\omega_{c-1} dF_s} \\ 1 & e^{-j\omega_c dF_s} \\ 1 & e^{-j\omega_{c+1} dF_s} \end{bmatrix} \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} &= \begin{bmatrix} H(e^{j\omega_{c-1}}) \\ H(e^{j\omega_c}) \\ H(e^{j\omega_{c+1}}) \end{bmatrix} \\ Ag &= h \\ g &= (A^T A)^{-1} A^T h \\ \overline{H(e^{j\omega_c})} &= G_1 + G_2 e^{-j\omega_c dF_s} \end{aligned} \tag{20}$$

where $\overline{H(e^{j\omega_c})}$ is the smoothed estimate of the channel at ω_c. Note that the channel delay d is actually an unknown (and indeed the entire model is an approximation of the real channel). It was found that in simulation, choosing $d = 0.002$ or $d = 0.004$ produced reasonable results, possibly due to the symmetry of the RHS of (13), and the dominance of channel over resampler noise at low SNRs. If d is approximated with a known value, the $(A^T A)^{-1} A^T$ term can then be precomputed as all the parameters of A are known.

This algorithm is denoted *lin2ls*. It's BER performance is plotted in Figure 5 and three resamplers summarised in Table 4.

Figure 5: Uncoded BER performance curves for equaliser algorithms over various channels. *mean12* is the algorithm used for 700D, and *lin2* for 700E. Note *mean12* (blue) is very close to theory for AWGN but breaks down on MPP. The *lin2ls* algorithm performs quite well on MPP and MPD, and about 0.5dB poorer than *mean12* on AWGN.



The new *lin2ls* resampler has the following benefits:

1. Using a single waveform design, it can operate from AWGN to beyond MPP with low SNR performance similar to 700D.

2. On fast fading (Doppler Spread $B_d > 2$ Hz) channels it will outperform 700E at low SNRs.

Algorithm	Mode	Colour	AWGN	MPP	MPD
mean12	700D	blue	0.0	0.5	unusable
lin2	700E	green	1.5	2.0	2.0
lin2ls	700X	magenta	0.5	0.5	0.5

Table 4: Comparison of equaliser algorithms for $N_s = 8$. The last three numbers are implementation loss in dB, smaller is better. 700X has the advantage of low implementation loss across channels with a single waveform, but is 0.5dB poorer than 700D on the (rare) AWGN channel. The 700E results are an approximation, as that waveform uses $N_s = 5$.

3. The $N_s = 8$ pilot insertion rate means the pilot symbol loss $L_p = 0.58$dB is low, giving good low SNR performance.

4. Large N_s means the amount of bandwidth used per carrier is small, reducing on air RF bandwidth and giving us options for transmit diversity in the 2000 Hz SSB radio bandwidth.

5. The new estimator could be applied to the raw data modes to improve low SNR FreeDATA performance.

6. It uses 2 pilots in time centred around the current modem frame rather than 4 used in *mean12* so will have lower algorithmic delay (one modem frame less) than 700D.

The algorithm used to combine adjacent carriers (20) is new and makes an assumption around the estimate of d, so it should be carefully tested OTA on real world channels using objective, controlled tests.

4 Multipath

The multipath channel exhibits frequency selective fading. Spreading bits in a packet across multiple carriers with OFDM helps. Extending this is frequency diversity, where n copies of the entire OFDM signal signal are transmitted, then recombined at the receiver. Due to the spectral efficiency of OFDM, it is possible to transmit several copies of a narrow band digital voice waveform within a 2000 Hz bandwidth.

Figure 6 illustrates an example of $Nd = 2$ diversity over a multipath channel, using a $Nc = 8$ carrier waveform waveform similar to 700D. The original transmit signal has been copied to higher frequencies to produce $Nc = 16$ carrier signal of the same total power but with twice the bandwidth.

Frequency diversity helps, as there are times when one carrier is faded, but it's diversity copy at about 800 Hz offset is not faded. When recombined we can recover that carrier. However observing Figure 6 there are some periods where the entire channel at all frequencies is faded for up to 1 second. This is supported by Figure 7 which shows large burst of errors during fades. This

leads us to the idea of diversity in time, where the bits of packet are spread over a period longer than the duration of a fade, lowering the average bit error rate of that packet. Once the BER is reduced beneath a threshold, the FEC code can correct all errors in the packet, and we can recover intelligible speech. The trade off is latency, which makes fast turn around PTT speech difficult.

Figure 6: Spectrogram of 20 seconds of $Nd = 2$ diversity signal over a MPP channel. At some points in time the signal is faded at all frequencies, so frequency diversity cannot help.

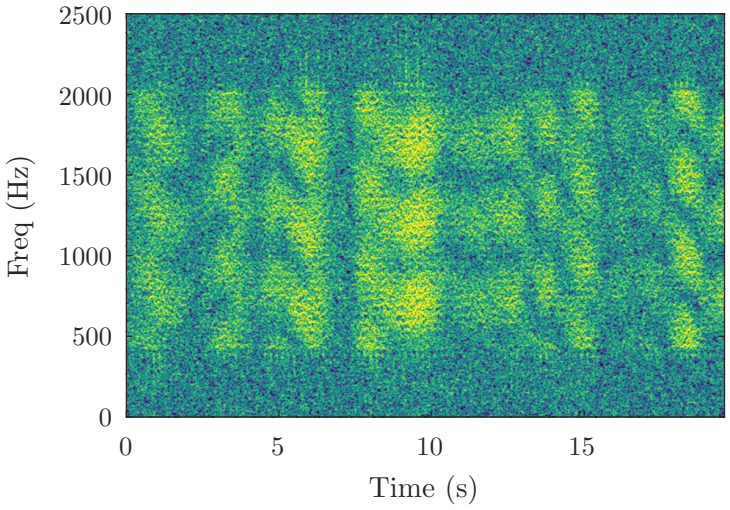
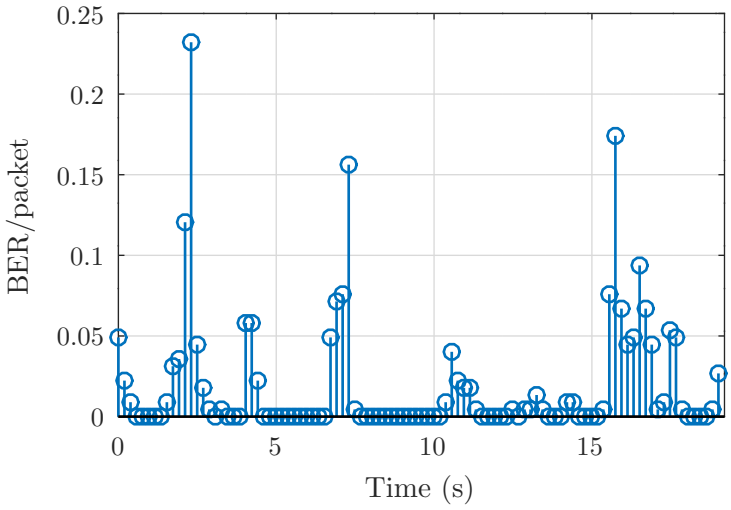


Figure 7: BER/packet for 0.16s packets, same channel as Figure 6. Larger packets (diversity in time) allow us to spread the errors over time, beneath a BER threshold where they can be corrected by FEC.



Figures 8 and 9 show the results of frequency and time domain ($D = 1.6$ and $D = 3.2$ second packet) diversity on BER and PER. The simulations were run for 300 seconds, which should give reasonable results down to a BER of 10^{-2} [3]. Care needs to be taken with HF simulation models such as [3] that have fixed delay, as it is possible for the resulting notches in the spectrum to be the same separation as a diversity carrier pair - a situation unlikely in a real world channel. The last few points plotted on the 1.6 and 3.2s curves had less than 10 packet errors so have a large margin of error.

Maximum Ratio Combining (MRC) is used for recombination, simulations indicate this is about 0.5dB better than Equal Gain Combining (EGC). As per Section 2.5 an algorithmic delay of D seconds implies a perceived latency of $2D$ seconds, so with $D = 3.2$, the user would experience a latency greater than 6.4 seconds, making PTT speech awkward. Such latencies may be useful for an emergency voice mode of last resort. Smaller D algorithmic delays can also be used, and will still have some benefit.

The waveform simulated is comparable to 700D, the *mean12* curve is the expected performance from FreeDV 700D. The combination of frequency and time domain interleaving gives impressive gains in PER performance, up to a total of 5dB with $D = 3.2$, and 3dB with $Nd = 2$ frequency diversity alone. With a more realistic $D < 1$ second and frequency diversity, a 4dB improvement over 700D should be possible.

Figure 10 shows the PER results of the C implementation of 700D, compared

to the Octave simulations of the proposed waveforms, expressed against SNR. It should be noted that the C implementation also includes timing and frequency offset estimation which are non ideal and will have small losses. The proposed waveforms have been adjusted to include the same L_p, L_{cp} and a 1dB loss for PAPR compression, to make the comparison with 700D realistic.

Table 5 summarises the performance improvements over 700D.

Waveform	Freq Div n	Alg Dly D (s)	Thresh SNR (dB)	Δ 700D (dB)
700D C impl	1	0.16	5.8	0.0
lin2s	1	0.16	4.8	1.0
div2 MRC	2	0.16	2.7	3.1
div2 MRC 1.8s	2	1.60	1.2	4.2
div2 MRC 3.6s	2	3.20	0.8	5.0

Table 5: Threshold SNR of proposed multipath waveforms on MPP Channel

Additional notes and further work:

1. The $N_d = 2$ diversity waveforms may suffer an additional PAPR loss due to the large number of carriers. PAPR compression with these waveforms should be simulated.

2. The longer D waveforms allow the use of longer FEC codes which will provide an additional boost to SNR performance.

3. An alternative to frequency domain diversity is a low rate FEC code applied across a large number of carriers spanning a wide bandwidth. This has not been simulated to date. An advantage of the diversity approach is we can selectively decode just one of the diversity copies - this can be used to decode speech in high SNR, but crowded band conditions where one of the diversity copies is impacted by strong co-channel interference.

4. Diversity works well when the symbols being combined are uncorrelated. Another approach is to send a low D waveform on one diversity channel, and a high D copy on another (or a low D waveform packet delayed by $D_2 > 1$ seconds). For high SNR channels the benefits of low D PTT speech could be enjoyed. For poor channels, the high latency copy can be combined for improved performance at the expense of longer latency. This idea exploits the fact that when $D_2 > D_{thresh}$, error bursts are uncorrelated in time, as shown in Figure 7. The performance would be 3dB less than 700D when just using the low diversity channel, due to the power split between the two channels.

5. These results simulated the use of a FEC code using a bit errors per packet metric. It would be useful to extend the simulation to run the actual LDPC codes.

6. To help ensure we are on the right track for a waveform usable on many channels it would be useful to test performance with other channel models such as those found in Appendix B of [5], and perform controlled Over the Air (OTA) experiments on real channels as early as possible.

Figure 8: $Nd = 2$ diversity simulation BER results. The 1.6 and 3.2 second curves simulate long delay D (long FEC codewords). Note there is no effect on BER with large D as these curves are the average BER over the entire simulation, they do not model the non-uniform distribution of errors over time on multipath channels as show in Figure 7.

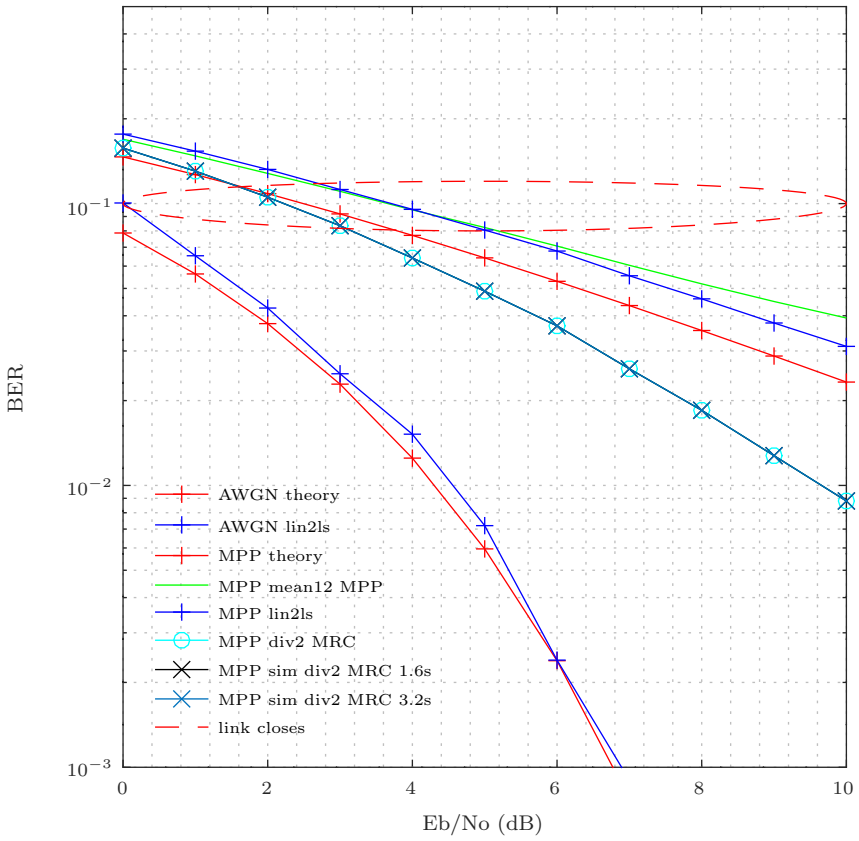


Figure 9: $Nd = 2$ diversity simulation PER results. A packet error is declared if the BER inside the packet exceeds 0.1 (approximating the performance of a rate 0.5 LDPC code). The packet length is 0.16s unless indicated in the legend. This metric better represents the effects of a non-uniform distribution of errors.

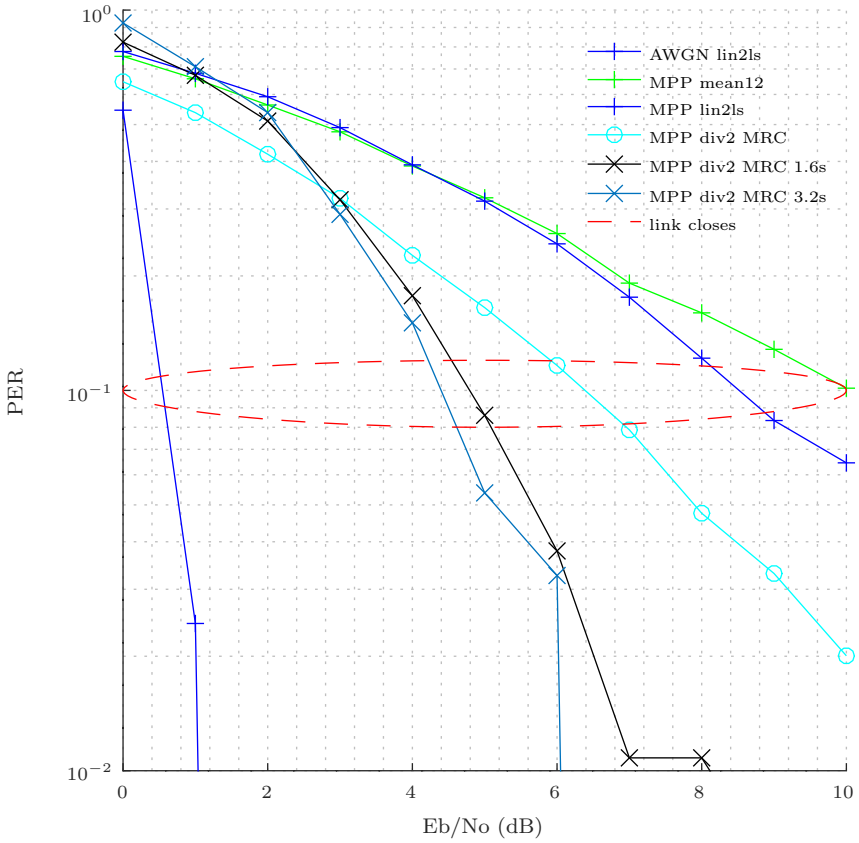
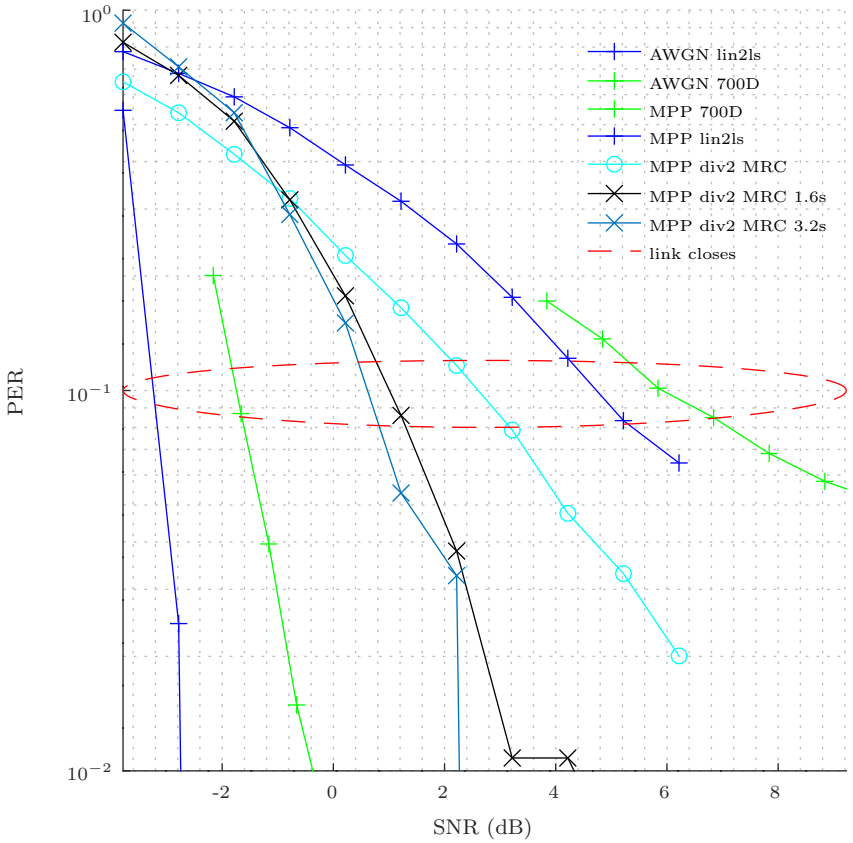


Figure 10: $Nd = 2$ diversity simulation PER results, mapped to a SNR axis and compared to 700D. 700E (not shown) is about 2dB worse than 700D on the MPP channel. The red oval is where a voice link becomes possible.



5 Delay Spread (ISI)

1. We need a CP long enough to handle MPD (4ms plus guard)

2. Try longer T_s which will mean less overhead. However this implies lower R_s which may be impacted by frequency spreading effect of Doppler. Caveat (as in equalisation section) is possible issues with Doppler spread and frequency offset tracking as R_s reduces.

3. Measure implementation loss or EVM against ISI, we might be able to get away with some ISI, it is acceptable to have performance drop off for MPD, but it needs to be gradual rather than breaking.

4. 700C had just 13ms symbols but dealt with ISI pretty well - this might be worth exploring.

6 Acquisition

PTT digital voice systems require fast acquisition - when a station starts transmitting we would like decoded speech to emerge from the receiver within a few 100ms, even on low SNR multipath channels. Some considerations:

1. A key part of acquisition is coarse timing and frequency offset estimation. After the signal is acquired, we need to track frequency and timing offsets as they evolve.

2. Acquisition can be specified as a probability, e.g. 90% probability of acquisition at the threshold link SNR, with an initial frequency offset of +/- 200 Hz.

3. False acquisition can occur, for example non-FreeDV signals such as noise, other HF signals, or sine wave "carriers". This can also be specified in terms of probability, e.g. no more than 1 false acquisition every 60 seconds on AWGN noise.

4. A digital voice signal operates in streaming rather than burst mode - a continuous sequence of modem frames is sent by the transmitter until the "over" is complete. A receiver should be able to acquire the signal at any time, and not be dependant on training sequences at the start of transmission.

5. OFDM carriers are narrow and hence quite sensitive to frequency offset errors, for example a 5% error of $0.05Rs = 2.5$ Hz is enough to impact performance. This sets a requirement for initial frequency offset estimation and tracking.

6. Acquisition schemes should be robust to slowly varying multipath channels which may have stationary notches in the spectrum that last for several seconds. This suggests spreading pilot symbols, carriers, and unique word bits over frequency to avoid frequency selective effects.

7. LEO channels have rapid rate of change of frequency due to Doppler shift, a challenge for frequency tracking.

6.1 Coarse Timing Estimation

The FreeDV OFDM waveforms such as 700D/E uses a correlation algorithm which compares the time domain received $r(n)$ signal to the transmitted pilot symbols $p(n)$. The correlation from two adjacent modem frames are combined. The correlation function is maximised over a 2D grid of time and frequency points. If the maximum exceeds a threshold the state machine enters a trial sync phase where acquisition is checked by demodulating a unique word present in every modem frame.

The time domain pilot signal is the IDFT of the frequency domain pilots for each carrier P_c:

$$p(n) = \sum_{c=1}^{c=N_c} P_c e^{j\omega_c n} \quad n = 0, ...M-1 \tag{21}$$

where $M = FsT_s'$ is the number of samples in one symbol (excluding cyclic prefix). For the simulation work above $P_c = 1$ for convenience, however other values are possible, for example FreeDV 700D uses a random sequence of BPSK symbols $P_c = \pm 1$.

The correlation algorithm can be derived from a least squared error formulation. Consider the total squared error between the time domain pilot signal $p(n)$ and the ideal received pilot sequence $p'(n)$:

$$\begin{aligned} E &= \sum_{n=0}^{M-1} |p(n) - p'(n)|^2 \\ E &= \sum_{n=0}^{M-1} |p(n) - gr(n+N_t)|^2 \end{aligned} \tag{22}$$

where $r(n)$ is the received off air samples at the current time offset N_t (in samples), and g is a complex scalar gain that models magnitude scaling and phase shift over the channel. E will be mimimised when the received signal is a close match to the pilot sequence. Expressing $p(n)$ and $r(n+N_t)$ as column vectors:

$$\begin{aligned} E &= (\boldsymbol{p} - g\boldsymbol{r})^H(\boldsymbol{p} - g\boldsymbol{r}) \\ &= \boldsymbol{p}^H\boldsymbol{p} - 2g\boldsymbol{r}^H\boldsymbol{p} + g^2\boldsymbol{r}^H\boldsymbol{r} \end{aligned} \tag{23}$$

where $\boldsymbol{a}^H = \overline{\boldsymbol{a}}^T$, the Hermitian (conjugate transpose) operator. Differentiating E and solving for g gives us:

$$g = \frac{\boldsymbol{r}^H\boldsymbol{p}}{\boldsymbol{r}^H\boldsymbol{r}} \tag{24}$$

Substituting back into (23):

$$\begin{aligned} E &= \boldsymbol{p}^H\boldsymbol{p} - \frac{(\boldsymbol{r}^H\boldsymbol{p})^2}{\boldsymbol{r}^H\boldsymbol{r}} \\ E &= \boldsymbol{p}^H\boldsymbol{p} - C_t^2 \\ C_t^2 &= \frac{(\boldsymbol{r}^H\boldsymbol{p})^2}{\boldsymbol{r}^H\boldsymbol{r}} \\ C_t &= \frac{\boldsymbol{r}^H\boldsymbol{p}}{\sqrt{\boldsymbol{r}^H\boldsymbol{r}}} \end{aligned} \tag{25}$$

where C_t is the correlation between the received signal at time N_t and the known transmitted pilot signal. Minimising E is equivalent to maximising the magnitude of the correlation $|C_t|$ or equivalently C_t^2. Typically C_t is sampled on a grid of N_t values. When C_t exceeds a threshold, we consider the current N_t as a candidate coarse time offset for further processing.

C_t is maximised with $\boldsymbol{r} = g\boldsymbol{p}$:

$$\begin{aligned} C_{tmax} &= \frac{g\boldsymbol{p}^H\boldsymbol{p}}{\sqrt{g\boldsymbol{p}^H g\boldsymbol{p}}} \\ C_{tmax} &= \frac{\boldsymbol{p}^H\boldsymbol{p}}{\sqrt{\boldsymbol{p}^H\boldsymbol{p}}} \end{aligned} \tag{26}$$

By scaling $\boldsymbol{p}$ we can set C_{tmax} to a convenient value (e.g. $C_{tmax} = 1$). Consider a prototype pilot vector $\hat{\boldsymbol{p}}$ such that:

$$\begin{aligned} \boldsymbol{p} &= s\hat{\boldsymbol{p}} \\ C_{tmax} &= s\frac{\hat{\boldsymbol{p}}^H\hat{\boldsymbol{p}}}{\sqrt{\hat{\boldsymbol{p}}^H\hat{\boldsymbol{p}}}} \\ s &= C_{tmax}\frac{\sqrt{\hat{\boldsymbol{p}}^H\hat{\boldsymbol{p}}}}{\hat{\boldsymbol{p}}^H\hat{\boldsymbol{p}}} \end{aligned} \tag{27}$$

where s is a scale factor chosen to achieve our desired C_{tmax}.

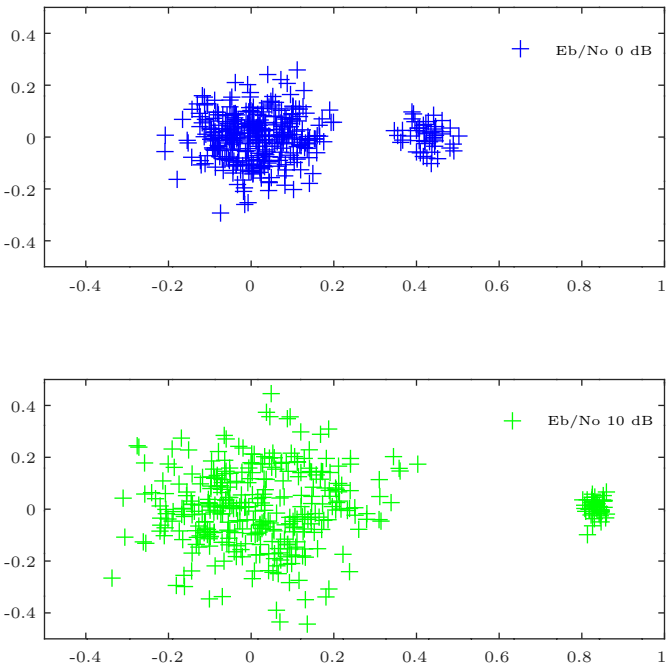
Figure 11 illustrates an issue with the use of a fixed threshold. Consider the denominator of (25) when we have a received signal $\boldsymbol{y}$ in presence of channel noise $\boldsymbol{n}$:

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{y} + \boldsymbol{n} \\ \boldsymbol{r}^H\boldsymbol{r} &= \boldsymbol{y}^H\boldsymbol{y} + 2\boldsymbol{y}^H\boldsymbol{n} + \boldsymbol{n}^H\boldsymbol{n} \end{aligned} \tag{28}$$

The first term is the signal power and the last term is the noise power. The second term is a small contributor as $\boldsymbol{y}$ and $\boldsymbol{n}$ are uncorrelated so the dot product terms cancel. With a high SNR the first term dominates, normalising C_t as desired. As the SNR drops the last term becomes significant, reducing $|C_{tmax}|$ as SNR reduces. From these observations we can see that an improvement would be a threshold that changes dynamically based on the noise power. Curiously,

the variance of C_t is still quite large even with small amounts of channel noise, due to random matches in the transmitted data symbols with $p(n)$.

Figure 11: Scatter plot of C_t sampled once a symbol over 50 modem frames, for low and high E_b/N_0. The $\sqrt{\boldsymbol{r}^H\boldsymbol{r}}$ normalising term on the denominator of (25) reduces $|C_{tmax}|$ in the presence of noise. Note there is no fixed threshold (e.g. 0.3 or 0.6) that would work in both cases. Thus a fixed threshold (as presently used in 700D/E) will be problematic.



Consider the numerator of $|C_t|$ (25) when only noise $\boldsymbol{n}$ with variance σ^2 is present:

$$D_t = \boldsymbol{r}^H\boldsymbol{p} = \boldsymbol{n}^H\boldsymbol{p} \tag{29}$$

For a typical sequence of pilots with equal amplitudes $P_c = \pm 1$ the variance of D_t will be the same as the variance as $\boldsymbol{n}$. The magnitude of a complex Gaussian noise signal has a Rayleigh distribution [2] with CDF:

$$CDF(x) = 1 - e^{-x^2/(2\sigma_r^2)} \tag{30}$$

where σ_r is the Rayleigh scale parameter. In our case we have spread the noise power σ^2 between two orthogonal channels such that $\sigma_r = \sigma/\sqrt{2}$. $CDF(x)$ is he probability that a sample is less than x. The probability of samples greater than x is therefore $1 - CDF(x)$. We define P_t as the probability of a sample greater than or equal to $|D_t|$:

$$\begin{aligned} P_t &= P(\geq |D_t|) \\ &= e^{-|D(t)|^2/\sigma^2} \end{aligned} \tag{31}$$

Note that large values of D_t are unlikely (low P_t) with noise alone. By rearranging (33) let us set a threshold:

$$D_{thresh} = \sigma\sqrt{-ln(P_{thresh})} \tag{32}$$

using a sufficiently low probability P_{thresh}. Now if $D_t > D_{thresh}$, we can conclude that $p(n)$ has been detected.

When a signal is present the distribution of $|D_t|$ will no longer be Rayleigh due to the correlation peaks with $p(n)$. Figure 12 is a histogram of $|D_t|$, the RHS hump in the histogram is due to the correlation peaks. This is even more obvious in the scatter plot of D_t (Figure 13). It was found that by removing the outliers in a set of D_t samples and recalculating the standard deviation, a good fit can be found with the Rayleigh distribution (green in Figure 12). The outlier removal algorithm is:

$$\begin{aligned} \sigma &= Std(D_t) \\ D'_t &= D_t < 2\sigma \\ \sigma' &= Std(D'_t) \end{aligned} \tag{33}$$

Any outlier greater than 2σ is removed from the set used to calculate the standard deviation, and σ' is used to calculate D_{thresh}. Note the variance of D_t with signal present is larger than the variance with noise alone, due to random correlations of the modem signal with $p(n)$. The circle in Figure 13 shows the threshold calculated for this example.

Figure 12: Histogram of D_t sampled once a symbol over 450 modem frames, for $E_b/N_0 = 0$ dB. After removing the outliers due to strong correlation with $p(n)$ on the RHS, the LHS of the histogram is a good fit to the Rayleigh PDF.

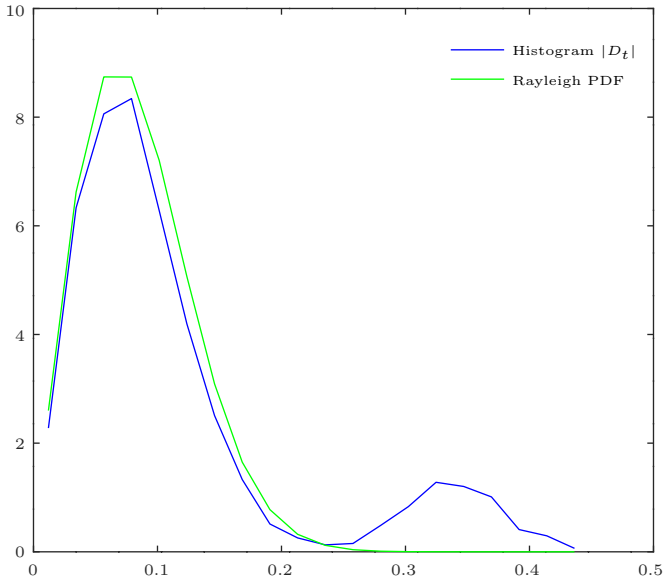
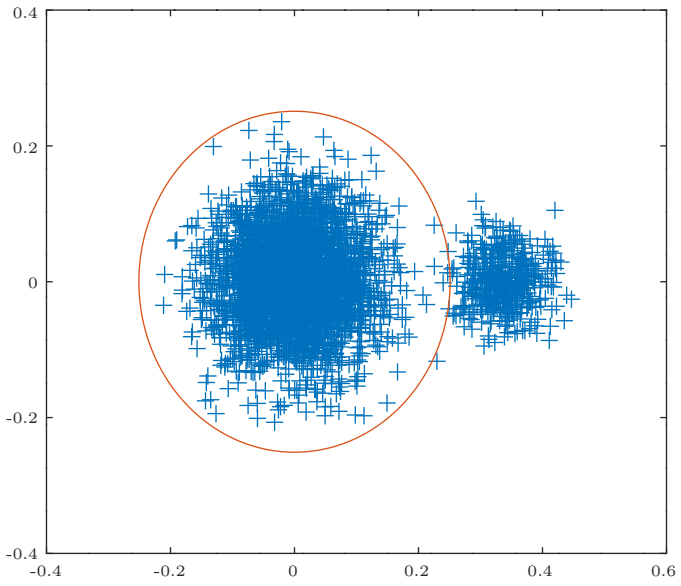


Figure 13: Scatter plot of D_t on the complex plane from the same simulation as Figure 12. The circle is D_{thresh} from (32) using σ' and $P_{thresh} = 10^{-3}$.



The new acquisition algorithm is therefore:

1. Using simulation results set a threshold P_{thresh} such that $D_t < D_{thresh}$ with noise alone but $D_t > D_{thresh}$ when signal is present for $E_b/N_0 > 0$ dB.
2. At run time, extract a T_f long window of samples (one modem frame).
3. Calculate a set of D_t values over the window.
4. Find the peak D_{tmax} over the set of D_t samples.
5. Estimate the standard deviation σ, then remove outliers to obtain σ'.
6. Calculate D_{thresh} using (32).
7. If $D_{tmax} > D_{thresh}$ we have found a candidate $p(n)$ sequence and the start of a modem frame.

Note that unlike the existing 700D/E acquisition algorithm D_{thresh} is dynamic, and based on the standard deviation of the received signal.

6.2 Coarse Frequency Estimation

Consider D_t as a function of frequency offset ω:

$$\begin{aligned} r(n) &= ge^{j\omega n}p(n) \\ D_t &= \boldsymbol{r}^H \boldsymbol{p} \\ D_t &= \sum_{n=0}^{M-1} r^*(n)p(n) \\ &= \sum_{n=0}^{M-1} ge^{-j\omega n}p^*(n)p(n) \end{aligned} \tag{34}$$

where the complex gain g accounts for the magnitude and phase shift of the channel and $p^*(n)$ is the conjugate of $p(n)$. Substituting the expression for $p(n)$ from (21):

$$\begin{aligned} D_t &= g\sum_{n=0}^{M-1} e^{-j\omega n}\sum_{c=1}^{N_c} P_c^* e^{-j\omega_c n}\sum_{d=1}^{N_c} P_d e^{j\omega_d n} \\ D_t &= g\sum_{c=1}^{N_c}\sum_{d=1}^{N_c} P_c^* P_d \sum_{n=0}^{M-1} e^{j(\omega_d-\omega_c-\omega)n} \end{aligned} \tag{35}$$

Noting $\omega_x = 2\pi x R_s/F_s = 2\pi x/M$:

$$\begin{aligned} D_t &= g\sum_{c=1}^{N_c}\sum_{d=1}^{N_c} P_c^* P_d \sum_{n=0}^{M-1} e^{j2\pi qn/M} \\ q &= d - c - \frac{M\omega}{2\pi} \end{aligned} \tag{36}$$

Consider the ideal case where $\omega = 0$, and $q = d - c$ is an integer. Using the identity for the sum of a geometric series:

$$\sum_{n=0}^{M-1} r^n = \begin{cases} \frac{1-r^M}{1-r} & r \neq 1 \\ M & r = 1 \end{cases} \tag{37}$$

$$D_t = g\sum_{c=1}^{N_c}\sum_{d=1}^{N_c} P_c^* P_d \frac{1-e^{j2\pi q}}{1-e^{j2\pi q/M}} \tag{38}$$

For all integer $q \neq 0$ the numerator $1 - e^{j2\pi q} = 0$ and the denominator is non-zero, thus all terms in the summation apart from $d = c$ are zero. When $q = 0$ we use the $r = 1$ (limit) clause in (37):

$$D_t = g\sum_{c=1}^{N_c} |P_c|^2 M \tag{39}$$

If we use BPSK pilots $P_c = \pm 1$:

$$D_t = gN_cM \tag{40}$$

For the case where $\omega \neq 0$, the inner term of 38 can be shown to be a phase shifted $sinc(x) = sin(x)/x$ function:

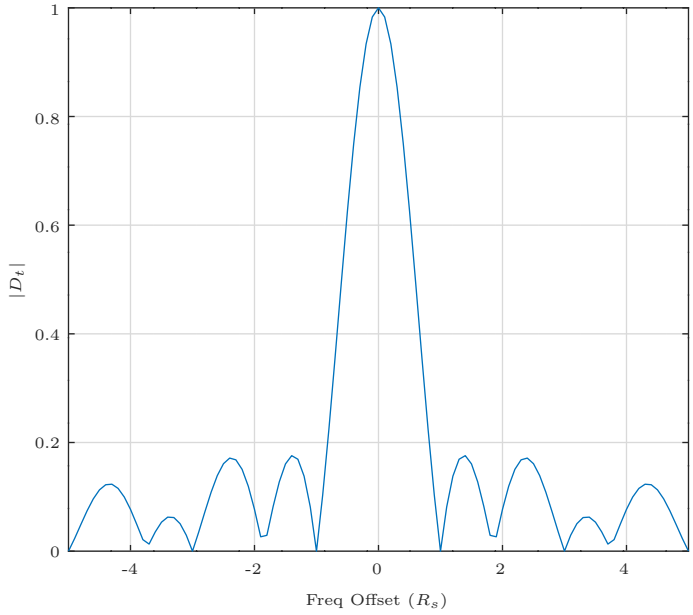
$$\begin{aligned} X(q) &= \frac{1 - e^{j2\pi q}}{1 - e^{j2\pi q/M}} \\ X(q)\frac{e^{-j\pi q}}{e^{-j\pi q/M}} &= \frac{1 - e^{j2\pi q}}{1 - e^{j2\pi q/M}} \frac{e^{-j\pi q}}{e^{-j\pi q/M}} \\ X(q)e^{-j(M+1)\pi q/M} &= \frac{e^{-j\pi q} - e^{j\pi q}}{e^{-j\pi q/M} - e^{j\pi q/M}} \\ X(q) &= e^{j(M+1)\pi q/M} \frac{sin(\pi q)}{sin(\pi q/M)} \end{aligned} \tag{41}$$

For the small values of $\pi q/M$ we are concerned with $sin(\pi q/M) \approx \pi q/M$:

$$\begin{aligned} X(q) &\approx Me^{j(M+1)\pi q/M} sinc(\pi q) \\ D_t &\approx gM \sum_{c=1}^{N_c} \sum_{d=1}^{N_c} P_c^* P_d e^{j(M+1)\pi q/M} sinc(\pi q) \end{aligned} \tag{42}$$

Unfortunately $sinc(x)$ is even so for $w \neq 0$ there no cancellation of the $d \neq c$ terms. The magnitude of D_t as ω (and hence q) varies is therefore highly dependant on the values chosen for P_c. Figure 14 is a an example when Pc is a set of Barker codes. This gives a sharp response when $\omega = 0$, which is a useful property for accurate frequency offset estimation in the presence of noise and channel distortions such as multipath. For 700D we simply chose random P_c values without consideration of their correlation properties, so employing a Barker sequence for the pilots P_c should improve frequency offset estimation performance.

Figure 14: $|Dt|$ as a function of frequency offset in multiples of R_s. There is ideal time alignment and no noise. Our requirement is ± 200 Hz or $\pm 4R_s$ (assuming $R_s \approx 50$ Hz). P_c is a 13 point Barker sequence, chosen for it's correlation properties.



The frequency offset can be estimated by adding a frequency correction term $\hat{\omega}$ to $r(n)$ from (34):

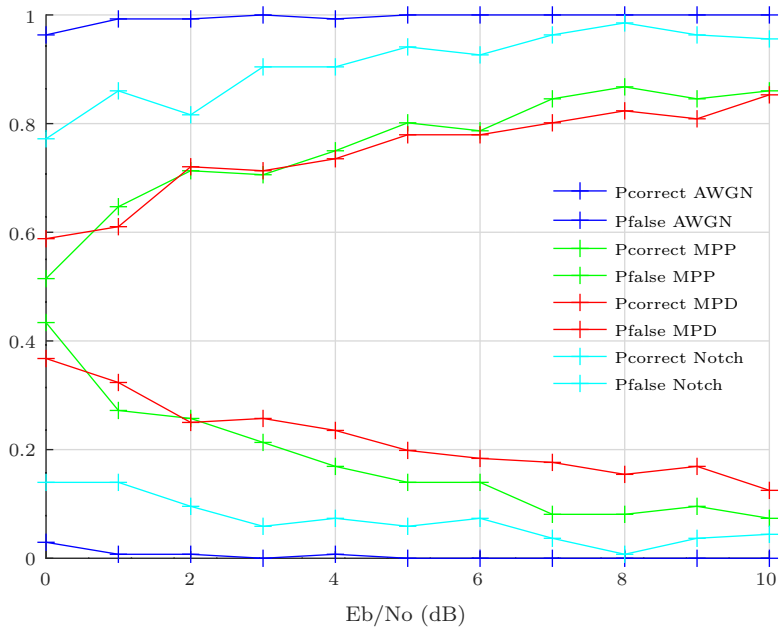
$$\begin{aligned} \hat{r}(n) &= e^{-j\hat{\omega}n} r(n) \\ \hat{D}_t &= \sum_{n=0}^{M-1} e^{j\hat{\omega}n} r^*(n) p(n) \\ &= \sum_{n=0}^{M-1} g e^{-j(\omega-\hat{\omega})n} p^*(n) p(n) \end{aligned} \tag{43}$$

When $\hat{\omega} = \omega$, the two frequency shift terms will cancel out and we will be left with the ideal case, maximising $\hat{D}_t$. The frequency offset can therefore be found by sampling (43) on a sufficiently fine grid of $\hat{\omega}$. Figure 14 suggests sampling $\hat{D}_t$ with $0.5R_s$ Hz steps would be sufficient to bracket the maxima.

Multipath will introduce frequency selective phase shifts, perturbing the phase of the received pilots P_c and lowering $|D_{tmax}|$. However the greatest change in phase will be near the notches in the spectra where the pilot am-

plitudes are smallest, and we tend to operate at higher SNRs in multipath channels.

Figure 15: Probability of acquisition for a range of channels and SNRs for 300 seconds of simulated $N_c = 16, N_d = 1,1400$ bits/s signal. Correct acquisition is defined as a timing error of $< \pm R_s$ and frequency error of $< \pm 0.5 R_s$. False acquisition occurs when we exceed the threshold for detection but the estimates are outside the success tolerances.



TODO: CPU estimates, especially with +/- 200 Hz frequency offset estimation.

6.3 Frequency Offset Refinement and Tracking

TODO

6.4 Pilot Carriers

A common acquisition strategy is to add one or more pilot carriers to aid frequency offset estimation. These carriers consume power and carry no payload data so an increase in the threshold SNR is required to maintain the voice link.

Consider the effect on E_b when adding an additional carrier:

$$\begin{aligned} E_b &= \frac{ST_s}{N_c} \\ E'_b &= \frac{S'T_s}{N_c + 1} \end{aligned} \tag{44}$$

Where E_b is the energy per bit with N_c data carriers and E'_b is the energy per bit after an additional pilot carrier has been added. To maintain the link at the same BER:

$$\begin{aligned} E'_b &= E_b \\ \frac{ST_s}{N_c} &= \frac{S'T_s}{N_c + 1} \\ \frac{S'}{S} &= \frac{S(N_c + 1)}{N} \\ SNR' &= SNR + 10log_{10}\left(\frac{N_c + 1}{N_c}\right) \\ SNR' &= SNR + L_{pc} \quad \text{[dB]} \\ L_{pc} &= 10log_{10}\left(\frac{N_c + 1}{N_c}\right) \end{aligned} \tag{45}$$

where L_{pc} is the loss in dB for adding a single pilot carrier. For example given a 16 carrier waveform, $L_{pc} = 10log_{10}(17/16) = 0.26$dB, which is quite modest for the potential gains in acquisition performance. A pilot carrier will also be visible on a real time waterfall and provide a useful manual tuning aid.

At this stage it does not appear necessary to add pilot carriers to meet our performance goals.

6.5 Acquisition TODO

1. Determine dimensions of time/freq sampling grid, for example is D_t is maximised despite +/-200 Hz offset, ie do we need to sample it at every time instant and every freq instant, or can we just scan time first.

2. Derive correlation approach as MMSE. Show issue with combining C_p from adjacent carriers. Can we establish reason it breaks down at +/- 25 Hz? Are there any cycles at T_f? Review blog posts.

3. Include frequency offset in E and C_p, explain how that can be included as function that we sample.

4. Can we come up with an expression for the C_t threshold?

5. Can we come up for an expression for probability of error?

6. Simply optimise what we have, make it work at a wider range.

7. Even if cost function with freq offset is non-linear, can we differentiate it and solve iteratively?

8. Can phase shift between successive pilots be used for fine freq estimation/refinement?

9. Simulation framework to test current performance, measure probability of sync and probability of failure. Why do we need UW? Lets drill into that.

10. Other methods (multiple by time shifted version) once we have a framework.

11. With a wider $Nd = 2$ signal, we might get better results. We could also consider different pilots for each copy, as they aren't used for combining.

12. There may be a good way to select an appropriate sequence for pilots to maximise the probability of detection, and mimimise PAPR.

13. Consider an "end of over frame", to signal a clean end of over and squelch.

14. Algorithm idea. Pilots can be considered a set of QPSK symbols, e.g. if $N_c = 16$ we have 32 bits. We could equalise then measure BER and use this as metric for sync. It also gives a us a tractable probability of error, and an easy way to combine with adjacent pilots.

15. test with impulse noise and sinusoids

16. plot Rayleigh on top of histogram

17. req: we want acquisition to work a few dB lower than FEC.

6.6 Timing estimation

TODO: make sure the current algorithm is doing sensible things on real world HF channels with complex impulse responses and large delay spreads.

6.7 PAPR reduction

TODO: expressions for PAPR reduction, model for improved SNR at Rx.

7 Further Work

This section presents topics useful to explore in future.

- Can we include PAPR into model? What can we do about improving PAPR, e.g. successive clipper/filter of ECSSB.

- Expression for Fading channels, block error rate, why 2020 is a lemon.

- Test against simulations other channel models, for example Appendix B of [5].
- Test with impulse noise. With a good code this should have a small effect. Does it cause problems with estimators?
- Table of FreeDV waveforms and values, plugged into formula, effect of increasing pilot symbol rate.
- Where we can gain, diversity, PAPR reduction, reduced overheads for fast fading and ISI (discuss)
- Wades MAP techniques (ref). This has a lot of promise, need an effective way to simulate and establish benefit with a modest amount of work. Can we combine MAP with extra bits for FEC? Index optimisation also a simple approach.
- Equalisation (20). Literature search, I'm sure this nothing new. Determine if (20) is a reasonable approximation for real HF channels - we have derived it from a simulation model. From VK5DSP review Sep 2023: Compare variance of $\overline{H(e^{j\omega_c})}$ to $H(e^{j\omega_c})$. Consider extending (20) to more carriers, it is essentially estimating the entire channel. Consider different estimators to MMSE, e.g. BLE. Consider modulating pilots as a side data channel or to aid frequency offset estimation.
- Equalisation: Consider combining decision directed approaches with pilots to improve estimates (McKillim et al).

8 References

[1] Codec 2 repo, modem codec frame design spreadsheet. `https://github.com/drowe67/codec2/blob/main/doc/modem_codec_frame_design.ods`.

[2] Rayleigh Distribution. `https://en.wikipedia.org/wiki/Rayleigh_distribution`.

[3] ITU-R F.1487: Testing of HF modems with bandwidths of up to about 12 kHz using ionospheric channel simulators, 2000.

[4] P Bergadà, RM Alsina-Pagès, JL Pijoan, M Salvador, JR Regué, D Badia, and S Graells. Digital transmission techniques for a long haul HF link: DSSS versus OFDM. *Radio Science*, 49(7):518–530, 2014.

[5] ES ETSI. 201 980 v4. 2.1 (2021-01) digital radio mondiale (drm). *System Specification. European Broadcast Union.*

[6] David Rowe. FreeDV-020 WP4000 Low SNR Mode.