HF Serial-Tone Waveform Design

M.B. Jorgenson and K.W. Moreland
Communications Research Centre
3701 Carling Ave.
Ottawa, Canada, K2H 8S2

Summary

The intent of this paper is to provide the reader with an understanding of the fundamental aspects of HF serial-tone waveform design. Factors which affect the design of a waveform are discussed. Some factors are regulatory, for example, the availability of bandwidth. Others are a consequence of nature, such as the characteristics of the HF channel. The technology used in the detection of serial-tone signals plays a key role in the design of the waveform. Dach of the component elements of a serial-tone modem is briefly described. Empirical waveform performance rules are presented and the design process is illustrated with an example based on the development of a 9600 bps HF waveform standard.

1.0 Introduction

With the increased demand for high data rate HF services, HF waveform design is enjoying a renaissance of sorts. New HF waveforms have been developed for NILE/Link-22 and STANAG 4444, and are being developed for STANAGs 4538, and 4539. A serial-tone standard for data rates up to 9600 bps was provided in STANAG 5066 Annex G as a non-mandatory component of the standard. The Annex G waveform was a commercial success in that it was implemented by vendors, even before adoption. When Annex G was removed from STANAG 5066, the same waveform, with minor modifications to the preamble, interleaver, and the addition of another intermediate data rate, found its way into draft Mil-Std 188-110B.

This paper will provide a tutorial guide to designing an HF serial-tone waveform and is intended to give the reader sufficient background to make judgements regarding the suitability of various proposed waveforms. We will begin by describing the components which make up a waveform and the key factors to consider in waveform design. For reference purposes, a brief description of the types of modulation commonly used at HF will be provided before focusing on the serial tone waveform. The essential elements of a modern serial-tone modem will be described, as will their impact on the design of a waveform. Empirical relationships for predicting the bit-error-ratio (BER) performance of serial-tone modems in a Doppler spreading environment are provided. Finally, the paper will conclude with a waveform design example, based on the authors experience in specifying Annex G of STANAG 5066.

2.0 What is a waveform?

A waveform specification is a description of the on-air signalling used to transmit a digital data signal over a radio channel. It includes a complete specification of the modulation to be used and prescribes the known symbols, commonly referred to as initial training (IT) or preamble, that are sent to establish synchronization as well as any other known symbols which may be inserted to aid in the demodulation process. Many standards, particularly more recent efforts, include forward error correction coding and data interleaving as an integral part of the definition.

As a rule, waveform standards only specify the details of the waveform to be transmitted. It is up to vendors to determine how best to demodulate the signal, although some standards, notably the US Mil-Stds, specify minimum performance figures which must be obtained in order for a vendor's equipment to be considered compliant with the various standards.

3.0 Factors affecting waveform design

A number of factors affect the design of a waveform. One of biggest constraints for HF waveform design, particularly in recent years as the emphasis has shifted to higher data rates, is the channel bandwidth. For military uses, HF spectrum is typically allocated with a 3 kHz channelization, although

Paper presented at the RTO IST Symposium on "Tactical Mobile Communications", held in Lillehammer, Norway, 14-16 June 1999, and published in RTO MP-26.

there are exceptions: some naval broadcasts operate in a 1.24 kHz bandwidth and there are some assignments of 6 kHz for use with independent sideband radios (typically for Link-11). 3 kHz is a respectable bandwidth for sending a 75 bps data signal, typified by the naval broadcasts run by many NATO nations, but it becomes a very narrow piece of spectrum for users who are attempting to transmit at data rates of up to 9600 bps.

The unique characteristics of the HF channel itself offers significant challenges. The ionospheric refraction which allows HF radio signals to propagate over long distances is not without its shortcomings. The received sky-wave signal may suffer distortion in the form of temporal dispersion (delay spread) as well as fluctuation in the signal's amplitude and phase (Doppler spreading). Recent high latitude DAMSON measurements have observed multipath signals of more than 10 ms duration and other signals have shown evidence of Doppler spreading greater than 50 Hz [1,2]. More typical mid-latitude sky-wave channels might show delay spreads of 1 - 4 ms with Doppler spreads of 1 Hz or less. In addition to the sky-wave channel, the HF surface wave channel offers interesting features and challenges. Over sea-water, the HF surface wave propagates far beyond line-of-sight, offering intriguing capabilities for Naval forces. As the surface wave begins to weaken at the periphery of the surface wave coverage region, a Rician channel is observed, with the non-fading component from the surface wave, and another, fading component, arising from a sky-wave path. The noise environment in the HF channel is also somewhat unique. CCIR Recommendation 322 provides a model for the HF noise environment. In general, it is much more impulsive than additive white Gaussian noise, with a much higher peak to mean ratio and tends to introduce burst error events.

The available bandwidth and the channel characteristics serve to limit the data rates which are achievable over HF. Until very recently most naval broadcasts were run at 75 bps, and were often unreliable even at these rates because of the lack of forward error correction (FEC) coding. With the adoption of the modern serial-tone modem, naval broadcasts are moving to a 300 bps data rate and other services are being provided at rates up to 2400 bps. The demand for increased data rates imposed by modern networking protocols has led to a determined effort to push achievable data rates upward and has resulted in new and developing standards for waveforms offering rates of 9600 bps and greater. The context in which these systems are being developed is also changing. At one time, EMCON considerations were considered paramount and most HF transmissions were one way, with no acknowledgment to give away information on the recipient's position. Increasingly, military forces are considering options which involve radiating detectable emissions for substantial portions of the time during operations. This evolving change in philosophy, coupled with the advent of modern automatic repeat request (ARQ) systems, has led to the transmission of data types which require error free reception (executable computer programs, for example).

4.0 Modulation Types

One way of characterizing modulations is to describe them as either bandwidth efficient or power efficient. Examples of bandwidth efficient modulations are phase-shift-keying (PSK), quadrature amplitude modulation (QAM) and multi-carrier modulations. Power efficient modulation types include direct sequence spread spectrum and M-ary frequency shift keying (MFSK). It is noteworthy that all of the above listed modulations have been used at HF. Older naval fleet broadcasts, some of which are still in service today, used an uncoded 75 bps 2-FSK modulation. Both STANAG 4415 and Mil-Std 188-110A specify a 75 bps in-band direct sequence spread spectrum modulation. The Mil-Std 188-141A Automatic Link Establishment (ALE) standard uses an 8-FSK signalling format. Multicarrier modulations have been used for both Link-11, where 16 tones are employed, and AN/DVT, where 39 tones are used. In the past decade and a half, PSK serial tone modulations have become the modulation of choice for high performance HF systems, with both Mil-Std 188-110A and STANAG 4285 employing PSK modulations for user data rates up to 2400 bps. QAM modulations have been specified in STANAG 5066 Annex G and in draft Mil-Std 188-110B for data rates beyond 2400 bps.

QAM constellations are chosen over PSK constellations for higher rates because of their better signal space distance properties. The QAM constellations fill the entire space, while the PSK constellations are confined to the periphery. The result is, for a given average transmitted power, the QAM modulation offers a much better signal space distance.The advantage to the QAM constellation becomes more pronounced as the number of bits per symbol is increased. Another factor which must be considered at HF is the impact of the peak power limited amplifier. This means that it is not sufficient to consider the performance of a system solely in terms of average power. Peak power performance must also be considered. As well, when considering the design of a constellation, improving the minimum distance is not the primary design criteria. Improving the overall BER is the aim. Simulation studies have shown that a constellation with good minimum distance properties and a good Gray code often outperform other constellations with superior minimum distance properties, but poor Gray coding structures for use with convolutional codes. This is a function of the FEC code employed and if, for example, a Reed-Solomon code were used, the signal space distance would again become paramount.

5.0 Synchronization and Frequency Offset Removal

Synchronization is the process at the receiver of identifying that a transmission is present and determining its timing with sufficient accuracy to permit demodulation. Again, HF offers some unique challenges in this area.

The extreme fading experienced over HF circuits means that it is possible that a fade could encompass the entire duration of the preamble, making detection difficult or impossible, even in channels where the average signal level is sufficient to permit fairly high rate communications. Designers have tried to mitigate this in two ways. STANAG 4285, for example, specifies an 80 symbol preamble which is reinserted every 256 symbols. This ensures that when the signal level rises to levels which will support communications, it can be detected and synchronized to. The disadvantage with this approach is that the ratio of data to known symbols is decreased, with the result that for the same data rate, the strength of the FEC code which can be used is decreased. The alternative to this is to use an initial long preamble to ensure synchronization, and then only include known symbols where they are directly required to assist in demodulation. Long preambles, with durations of up to 4.8 s have been used. It is very unlikely that a channel which would support communications at a rate of 75 bps or more would have a fade of that length with sufficient depth to preclude detection of the signal. This is the approach which was taken in Mil-Std 188-110A where the length of the preamble has been tied to the interleaver used; when short or no interleaving is selected, a 0.6 s preamble is sent while when long interleaving is specified, a 4.8 s preamble is used.

The degree of synchronization required depends upon the algorithm used to demodulate the data. Early techniques required synchronization which was accurate to the symbol. More modern algorithms operate effectively with synchronization which is accurate to within several symbols. This distinction can be critical at HF, where multi-path fading can result in a continually changing synchronization point.

The other use for the synchronization preamble is frequency offset removal. The known symbols in the preamble are used to estimate and remove any frequency offset in the received signal.

From a waveform design perspective, the trade-offs which must be considered are the delay and reduced data rate resulting from adding symbols dedicated to synchronization versus the probability of missing a signal which could have been successfully demodulated if an insufficient number of symbols is used for synchronization.

6.0 Equalization

In the absence of channel induced distortions, band-limited signals can be received with no intersymbol interference if the Nyquist criteria is met. The raised cosine response is a typical example of the kind of filtering that is used eliminate intersymbol interference by utilizing pulse shapes which have zero crossings at T spaced intervals.

Equalization provides compensation for channel induced distortions. At HF, the equalization must be adaptive in nature, changing as the channel itself changes with time. Equalization is required for serial tone modulations where the symbol duration, typically 0.4167 ms, is small relative to the expected time dispersion, which is often as severe as several milliseconds. Multi-carrier and M-FSK modulations, on the other hand, do not, as a rule, require equalization since their symbol spacing, typically between 8 and 13 ms, is sufficiently large as to mitigate the effect of multipath delay spread for most channels.

The simplest form of equalizer is based on a linear tapped-delay-line with the coefficients adapted directly based on some error criteria, usually a minimum mean squared error norm. An improvement on this is the decision feedback equalizer. This is composed of a feed-forward filter and a feed-back filter which operates on previous decisions. Again, the coefficients are adapted directly, based on the estimated error signal. However, in this case the characteristics of the data signal are considered. Further gains can be obtained by separating the equalization problem into two distinct tasks. The first is to identify the channel impulse response, the second is to compute the optimal weights for the decision feedback equalizer based on the estimated channel impulse response. Alternatively, when the channel impulse response is known or can be estimated a block equalization process can be used.

Most modern HF modems use equalization which requires estimation of the channel impulse response. As a consequence, the waveform designer must provide sufficient opportunity to make channel estimations and to maintain and update them as required. The main purpose of the known data segments in the framing structure of HF serial tone waveforms is to provide for isolated blocks of unknown data for block oriented detection and to facilitate the maintenance of an accurate channel estimate for both block and conventional equalizers. If the span of the known data segments is twice the expected delay spread, channel estimates can be made directly using only the known data. Otherwise, an LMS update procedure based on known data and past decisions made on unknown data is used to track and adjust the channel estimate.

7.0 Error Correction Coding

A critical feature of any modern HF data communication system is the forward error correction coding. There are a number of criteria which must be considered in selecting an FEC code. Performance, complexity, compatibility and proprietary rights issues are all significant factors in the choice of a code.

The relative performance of various coding schemes varies with code rate, modulation and the acceptable error thresholds.

7.1 Convolutional codes

Convolutional codes are soft decision, bit-error correcting codes which are usually decoded with a near maximum likelihood detection process known as Viterbi decoding. They are asymptotically optimal in an additive white Gaussian noise environment and, when combined with adequate interleaving, provide good performance in the fading channels found at HF. Convolutional codes perform poorly in burst error environments, which makes it critical to achieve sufficient interleaving to break up fades.

The rate 1/2, constraint length 7 convolutional code is commonly used for HF serial tone data transmission. Both Mil-Std 188-110A and STANAG 4285 Annex E call for this code. When rates greater than 1/2 are required, they can be achieved by puncturing the code, an exercise where selected output bits are not transmitted and a puncturing mask is applied in the decoding process. Rates lower than 1/2 are achieved by repeating the bits output by the encoder. Repeating output bits, to achieve a rate 1/4 code, for example, incurs only a small performance loss relative to what could have been obtained with a true rate 1/4 code. However, the advantage of the repetition strategy is that it is much simpler than developing alternate codecs for each data rate.

7.2 Reed-Solomon codes

Reed-Solomon codes are a class of symbol error correcting codes which provide good burst error performance, particularly when erasures are used. Reed Solomon codes with large symbol sizes, (typically 6 bits or more) usually eliminate the need for a cyclic redundancy check (CRC) for validating data fidelity. The code itself provides an indication of error when it is not possible to correctly decode the received data. This feature can be very valuable in packet data systems. Relative to other coding schemes, RS codes work best at high rates or when the acceptable BER thresholds are particularly stringent.

The major disadvantages associated with RS codes are their relatively poor performance in AWGN and in the difficulty in incorporating soft decision information into the decoding process in a form more sophisticated than simple erasures.

Reed-Solomon codes are used in STANAG 4444.

7.3 Concatenated codes

Concatenated codes attempt to use multiple encodings to overcome the shortcomings of some codes. Powerful concatenated codes have be formed by using convolutional inner codes with RS outer codes. These codes are a good match for one another when used in this way. Convolutional codes are sensitive to burst errors and, when they fail to decode properly, often produce extended error bursts. RS codes, on the other hand, work well with burst errors, so the bursty errors produced by the inner convolutional code can often be corrected by the outer RS codec.

The main disadvantage of concatenated codes is that they require two interleavers to be effective. This limits the amount of interleaving which can be applied to the inner code, with the result that for the error rates usually considered adequate at HF, i.e., in the 10^{-3} to 10^{-5} range, concatenated codes generally do not perform as well as convolutional codes by themselves. However, if a very stringent BER criteria is required, they will perform very well.

7.4 Trellis Coded Modulation (TCM)

This class of codes exploits the improved minimum distance properties which can be obtained by combining coding and modulation. TCM has proven very effective in wireline modems, where it is used extensively. The large gains seen in the wireline environment have not translated into comparable performance in fading channels, although research is ongoing to improve the performance of TCM in fading channel environments. HF in particular represents a difficult environment for TCM because of the interaction which takes place between the coding and the equalizer.

7.5 Iterative codes

Turbo codes are the best known example of iterative codes. These codes are composed of two or more component codes and iterate back and forth between the two codes to achieve performance which is very close to the Shannon capacity bound. These codes achieve results which were believed to be unobtainable with reasonable complexity until very recently. There are three drawbacks associated with these codes. To obtain the impressive performance that they offer, substantial interleaving is required. The computational complexity associated with these codes is substantially greater than convolutional codes, although significant strides have recently been made in reducing the computational complexity. Most, if not all, of these codes are protected by patents and, as such, are subject to proprietary rights which makes their adoption for use at HF very problematic.

7.6 Interleaving

Depending on the kind of code employed, the interleaver may interleave symbols or bits. In the case of Reed-Solomon codes, in order to preserve the burst error capabilities of the code, Reed-Solomon symbols are interleaved. With convolutional and other bit error correcting codes, it is the bits which are interleaved.

Both block and convolutional interleavers are used for HF data communications. The block interleaver has the advantage that if the data packets are sized to fit within an interleaver block, no flush is required. The drawback to the block interleaver is that it is only possible to synchronize at interleaver block boundaries. With a convolutional interleaver, on the hand, synchronization is possible once every cycle through the interleaver and, for the same end-to-end delay, better performance is achieved. The major disadvantage to the convolutional interleaver is that it requires a flush to clear out the interleaver at the end of the transmission.

8.0 Predicting Performance

The performance of serial tone modems as a function of delay spread is such that there is little variation or dependence of BER on delay spread up to a certain critical value. For delay spread values exceeding this value, failure is usually catastrophic. The span of this delay spread operating window is dictated by the span of the known training segments, (also called the probe or mini-probe). Within the delay spread region where the modem is functioning correctly, it is possible to develop empirical relationships between the uncoded BER as a function of SNR and Doppler spread. The example channels that are presented here are two path Rayleigh fading channels. We emphasize that the results are uncoded BERs and a further level of refinement would be necessary to make this applicable to the output of a modern modem which includes FEC.

In the high signal to noise ratio case, with delay spreads in the operating range, it has been found that the following simple relationship results in a very good fit to observed simulation results for the probability of uncoded bit error achieved with an advanced block equalizer.

$$P_b = k_m\left(\frac{\Phi_D}{R_{cs}}\right)^4 \tag{1}$$

where k_m is a modulation dependent constant, Φ_D is Doppler spread in Hz, and R_{cs} is the channel sampling rate in Hz. Thus, for high SNR, the B.E.R. is proportional to the ratio of the channel sampling rate divided by the Doppler spread, raised to the fourth power, where the channel sampling rate is simply the reciprocal of the time between successive probe sequences. Empirically, the proportionality constant k_m has been determined for BPSK, QPSK, 8PSK and 16 QAM. The values obtained are:

$$\begin{aligned} k_{bpsk} &\sim 0.2 \\ k_{qpsk} &\sim 0.6 \\ k_{8psk} &\sim 3 \\ k_{16qam} &\sim 10 \end{aligned} \tag{2}$$

An example of the fit obtained using these parameters, in a two path Rayleigh fading channel with 1 ms of delay spread, for a waveform using a repeated framing structure of 19 data symbols followed by 19 training symbols (denoted as 19-19), with BPSK, QPSK, 8PSK and 16QAM modulations is shown in Figure 1. Further analysis of the simulation results would allow refinement in the accuracy of these numbers.

It is possible to use this result to derive some design equations which would be useful in the development and specification of waveforms expected to use advanced block equalization techniques. The bit rate throughput of the waveform, R_b, can be expressed in terms of the symbol rate, R_s, the channel sampling rate, the number of training symbols per probe segment, T, and the number of bits per symbol for a given modulation, n_m.

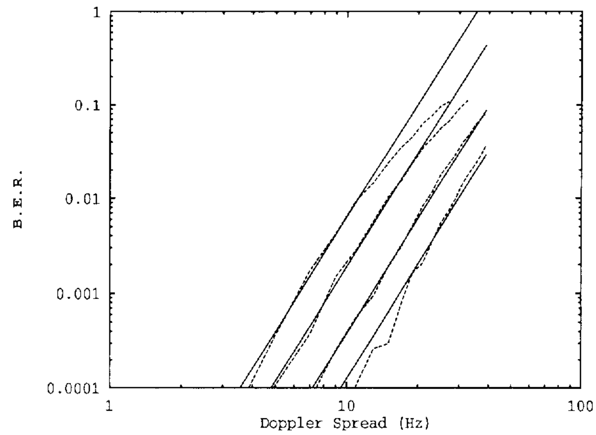

Figure 1 - Fourth power relationship of B.E.R. and Doppler spread for BPSK, QPSK, 8PSK and 16QAM modulations. Simulation (dashed), empirical relationship (solid).

$$R_b = (R_s - R_{cs}T)n_m \tag{3}$$

Substituting (1) into (3) results in

$$R_b = \left(R_s - \left(\frac{k_m}{P_b}\right)^{1/4}\Phi_D T\right)n_m \tag{4}$$

which provides the highest bit rate achievable for specified symbol rate, modulation, bit error rate, Doppler spread and probe sequence length. Note that because the probe sequence length determines the maximum delay spread handling capability, this equation provides a very good analytical tool for determining what sort of throughput goals are achievable for a specified environment. Simple algebraic manipulations of this equation lead to formulations in terms of bit error probability,

$$P_b = k_m\left(\frac{\Phi_D T}{R_s - \frac{R_b}{n_m}}\right)^4 \tag{5}$$

and Doppler spread

$$\Phi_D = \left(\frac{R_s - \frac{R_b}{n_m}}{T}\right)\left(\frac{P_b}{k_m}\right)^{1/4} \tag{6}$$

The previous relationships are valid only at high SNR. It would be desirable to establish similar relationships which are valid over a large range of SNR's. Such a relationship would take the form of

$$P_b = P_{b0}(N_0) + f(\Phi_D, N_0) + k_m\left(\frac{\Phi_D}{R_{cs}}\right)^4 \quad (7)$$

where, in addition to the fourth power term above, there is a constant term dependent upon SNR (and probably other factors such as length of data and training segments), as well as a term which might be dependent upon both Doppler spread and SNR. Here $P_{b0}(N_0)$ is the bit error rate for 0 Hz Doppler spread and is a function of the SNR. For convenience, this dependence will not be shown in the following, but is implicit. Empirical study has shown that a very good approximation to the bit error rate is provided by the following:

$$P_b = \left(\sqrt{P_{b0}} + \sqrt{k_m}\left(\frac{\Phi_D}{R_{cs}}\right)^2\right)^2 \quad (8)$$

A typical example of the accuracy of this fit is shown in Figure 2 below. The solid lines are the empirical relationship of (8), while the dashed lines represent simulation results. Each bit error rate point in this figure represents the simulation of 5000 modem frames, each containing 3 signalling blocks of 27 data and 27 training (27-27). For the 27 dB curve, this still results in a large degree of uncertainty in the B.E.R. for low Doppler spread values. Thus, an additional run of 50000 modem frames were run at 0 Hz Doppler spread and the P_{b0} values obtained from that run are used computing the curves shown in the figure.

The agreement between the empirically derived relationship and the simulation results in the above figure is quite remarkable. Figure 3 shows the fit of the relationship to the results obtained for the 27-27 waveform with the various modulations at Eb/No = 27dB. Again, agreement is quite good. Figure 4 shows the fit to a waveform with 75% efficiency (81 data, 27 training).

It should be noted that the agreement with the relationships of (5) and (8) is not quite as good over the ensemble of results for 8PSK and 16QAM as it is for BPSK and QPSK. As well, waveforms with very high efficiencies, particularly for the high delay

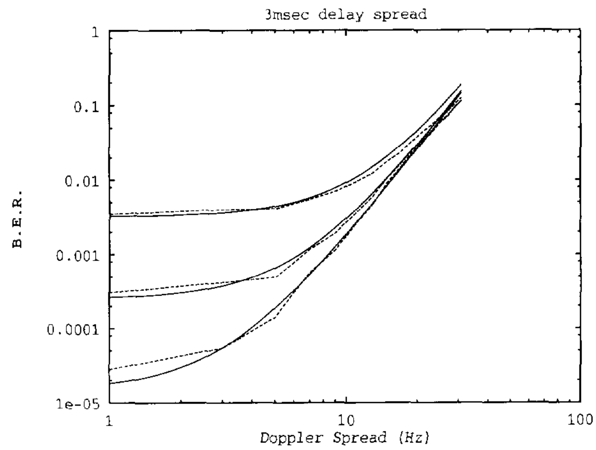

Figure 2 - Fit of empirical B.E.R. vs. Doppler spread relationship (solid) to simulated results (dashed) for qpsk signalling using a 50% efficient waveform (27-27) with Eb/No of 15, 21, and 27 dB.

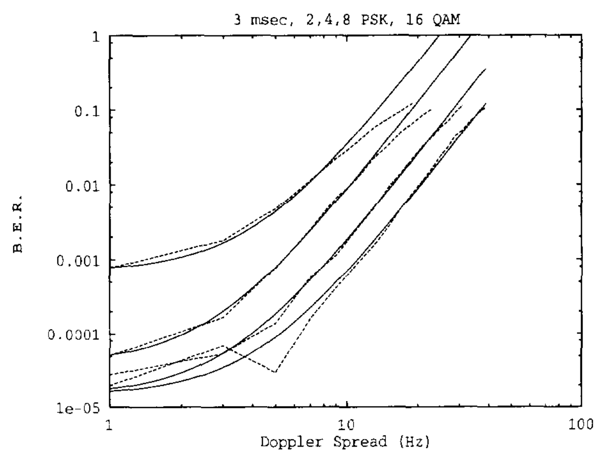

Figure 3 - Fit of empirical B.E.R. vs. Doppler spread relationship (solid) to simulated results (dashed) for various modulation techniques using a 50% efficient waveform (27-27) with Eb/No of 27 dB.

spread cases, show greater departures from (5) and (8) than do those with efficiencies near 50%. In any case, the agreement of simulation with (8) is sufficiently good that a set of design equations is proposed using (8) as a basis:

$$P_b = \left(\sqrt{P_{b0}} + \sqrt{k_m}\left(\frac{\Phi_D T}{R_s - \frac{R_b}{n_m}}\right)^2\right)^2 \quad (9)$$

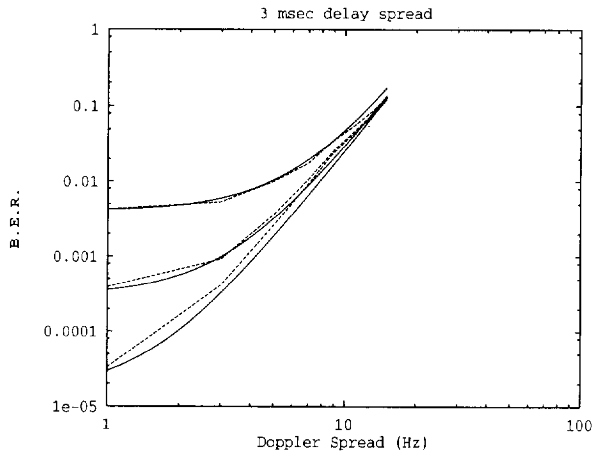

Figure 4 - Fit of empirical B.E.R. vs. Doppler spread relationship (solid) to simulated results (dashed) for qpsk signalling using a 75% efficient waveform (81-27) with Eb/No of 15, 21, and 27 dB.

$$R_b = n_m\left(R_s - \left(\sqrt{\frac{\sqrt{k_m}}{\sqrt{P_b} - \sqrt{P_{b0}}}}\right)\Phi_D T\right) \qquad (10)$$

$$\Phi_D = \left(\frac{R_s - \frac{R_b}{n_m}}{T}\right)\sqrt{\frac{\sqrt{P_b} - \sqrt{P_{b0}}}{\sqrt{k_m}}} \qquad (11)$$

These equations have proven to be a useful tool in the initial waveform design stage.

9.0 Waveform Design Example

STANAG 5066 is an HF subnet protocol developed by NATO in support of the BRASS program. During the development of the protocol, a requirement for waveforms providing data rates beyond 2400 bps was identified and the Communications Research Centre (CRC), who had published results of diversity combining trials with 9600 bps waveforms [3], were asked to submit a draft waveform to satisfy the requirement.

In this case, increased data rate was the clear driving factor in the design. This limited the choice of modulation to the bandwidth efficient modulations. The serial-tone approach was taken because of its better peak-to-average performance relative to the multi-carrier approaches.

The user data rate can be expressed as

$$r_d = r_b \eta_{wf} n_b r_c \qquad (12)$$

where r_d is the data rate, r_b is the baud rate or symbol rate, η_{wf} is the waveform efficiency or ratio of data symbols to total symbols sent, n_b is the number of bits per symbol, and r_c is the code rate. Increased data rates can be achieved with increased baud rate, improved waveform efficiency, increased modulation complexity or reduced code rate.

Some initial experimentation was carried out on the effects of increasing the baud rate above 2400 symbols per second. The result was that for typical commercial HF radios with 2.7-2.8 kHz passbands, the increased coding which was available improved BERs for the AWGN channel, but no improvement was seen for a CCIR poor channel where there was 2 ms of multipath delay spread. Performance degraded markedly as delay spread was increased when the signal was passed through real radios. This result has since been confirmed by Nieto, who provides strong evidence of the dangers of comparing modem implementations without adequately factoring in radio effects [4].

Improving the waveform efficiency is a relatively straightforward exercise. Neither STANAG 4285 or Mil-Std 188-110A, the key HF waveform standards, were designed with these high data rates in mind. As a consequence, STANAG 4285 has a waveform efficiency of 50% while the Mil-Std does slightly better with a waveform efficiency of 66.7% for the 2400 bps mode. Combinations of waveform efficiency, number of bits per symbol and standard coding rates which achieved user data rate of 9600 bps were considered.

A decision was made to maintain the waveform structure regardless of the selected user data rate and to achieve the various rates by changing the modulation. This also allowed for auto-baud to be implemented as a modulation recognition. As much as possible, compatibility with STANAG 4285 was to be maintained for ease of implementation.

The specification of the known data is the first major task. The preamble must be long enough so that it extends significantly beyond any radio induced transients and is designed to have good correlation properties to aid in signal detection and acquisition. At the end of the preamble, a sequence optimized for channel estimation was appended to ensure that a good channel estimate was available to the receiver at the beginning of the data.

The required delay spread capability is what drives the length of the training segment. With conventional implementations where a channel estimate is made on the preamble and only updated, not re-estimated, on subsequent known data segments, the length of the known data segments is just slightly more than the expected delay spread. The higher modulation complexities contemplated in the Annex G waveform require better channel estimates than previous implementations. For the Annex G waveform, where it was envisioned that the user would make a direct estimation of the channel estimate with each known data segment, the training segment duration needed to be more than twice the expected delay spread. The incentive for making the training segment short is that lengthening the training segment forces the data segment to be lengthened by a proportionate amount in order to maintain the waveform efficiency. The bigger the data segments get, the more susceptible the implementation is to Doppler spreading and frequency offset errors. On the other hand, longer known data segments allow for better channel estimates.

In the end, after simulation of several candidates, the frame structure shown in Figure 5, along with STANAG 4285 and Mil-Std 188-110A shown for comparison, was adopted. Another feature of the Annex G waveform design was the inclusion of a regularly reinserted preamble which could be used for late acquisition or reacquisition.

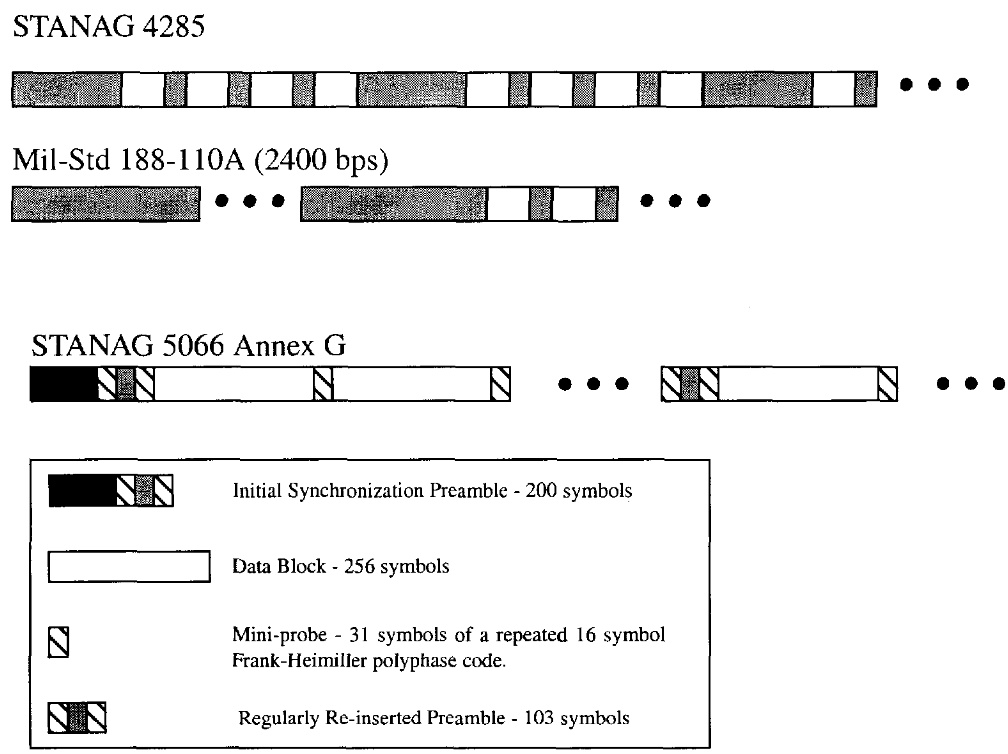

Figure 5 - Waveform structure.

A constellation optimization was undertaken for the QAM modulations used in the standard. The resultant constellations are shown in Figures 6 and 7. These clearly improve signal space distance properties relative to the conventional square constellations, but what is not so readily apparent is that they also preserve the excellent Gray-coding properties of the square constellations.

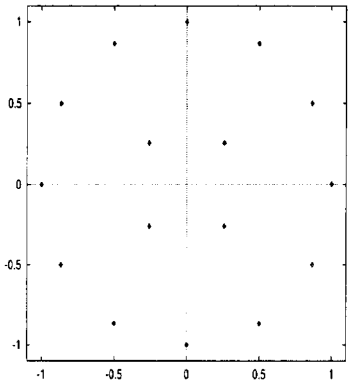

Figure 6 - 16QAM constellation used to achieve 6400 bps.

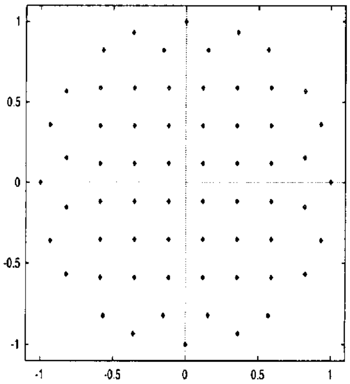

Figure 7 - 64QAM constellation used to achieve 9600 bps.

The FEC was based on the rate 1/2 constraint length 7 convolutional code specified in both STANAG 4285 and Mil-Std 188-110A, although in this case it is punctured to rate 3/4. A convolutional interleaver, much like the one in STANAG 4285, but using a nominal 48 row structure, was specified. This allows the interleaver to synchronize immediately after any preamble and every data block thereafter.

A summary of the Annex G waveform, with the STANAG 4285 2400 bps Annex E coding option shown for reference is provided in tabular form below.

Waveform (data-training)	Modulation	Code Rate	User Data Rate (bps)
256-31	64-QAM	3/4	9600
	16-QAM	3/4	6400
	8-PSK	3/4	4800
	QPSK	3/4	3200

For comparison, STANAG 4285 2400 bps mode:

Waveform (data-training)	Modulation	Code Rate	User Data Rate (bps)
32-16	8-PSK	2/3	2400

In comparing the two waveforms, it should be noted that they are designed for different purposes and services. The Annex G 4800 bps mode requires just slightly more SNR than the STANAG 4285 waveform in a benign channel. This is not surprising as both use the same modulation, 8PSK, and the code rates are not all that different, rate 3/4 vs. rate 2/3. On the other hand, the Annex G waveform will fall apart completely in the presence of large Doppler spreads, while the STANAG 4285 waveform can cope with substantial Doppler spreading.

10.0 Summary

This paper has provided a tutorial overview of the design of HF serial tone waveforms.

11.0 Acknowledgments

This work was partially funded by the Department of National Defence, Canada.

12.0 References

[1] T.J. Willink et. al., "Robust HF data communications at high latitudes", submitted to *IEE Proceedings - Microwaves, Antennas and Propagation*, 1999.

[2] M.J. Angling et. al., "Measurements of Doppler and multipath spread on oblique high latitude HF paths and their use in characterising data modem performance", *Radio Sci.*, Jan-Feb, 33(1), 1998.

[3] M.B. Jorgenson et. al., "Polarization diversity for HF data transmission", *7th International Conference on HF Radio Systems and Techniques, Nottingham*, UK, July 1997.

[4] John Nieto, "An implementation of STANAG 5066 Annex G Waveforms", *Nordic Shortwave Conference*, Faro, Sweden, August 1998.