UI MARSIS-7-0

Signal Processing Requirements for Ionospheric Sounding Measurements with MARSIS*

by

Donald A. Gurnett

April 8, 1999

Department of Physics and Astronomy
The University of Iowa
Iowa City, IA 52242

Document 97-60007

*Prepared in response to an action item at the MARSIS Team Meeting, Jet Propulsion Laboratory, March 22 - 26, 1999.

Revision List

1.	Date

Objectives of this Document

The scientific objectives of the MARSIS ionospheric sounding measurements are to study spatial and temporal variations of the electron density in the Martian ionosphere using ionospheric soundings. The purpose of this document is to describe the signal processing requirements required to carry out ionospheric sounding with the MARSIS instrument and to provide an initial description of the in-flight modes of operation.

The Ionospheric Sounding Technique

The basic ionospheric measurement technique consists of transmitting a radar pulse at a frequency f, and then measuring the intensity of the reflected radar echo as a function of the time delay t_d. Two types of radar reflections occur in a planetary ionosphere: vertical and oblique. For a radar signal incident on a horizontally stratified ionosphere a strong vertical reflection occurs from the level where the wave frequency is equal to the electron plasma frequency, which is given by

$$f_p = 8{,}890\sqrt{N_e}\ \text{Hz}, \qquad (1)$$

where N_e is the electron density in cm^{-3}. This reflection process is illustrated in Figure 1 which shows a typical vertical plasma frequency profile through the ionosphere at Mars. The frequency f_p(Max) is the maximum plasma frequency in the ionosphere. For vertical incidence ionospheric reflections only occur at frequencies less than f_p(Max). Typical values for f_p(Max) range from a few hundred Hz on the nightside of Mars to about 4 MHz on the dayside of Mars. The round trip time delay for the vertically reflected signal is controlled by the group velocity, $v_g = c[1 - (f/f_p)^2]^{1/2}$, and is given by

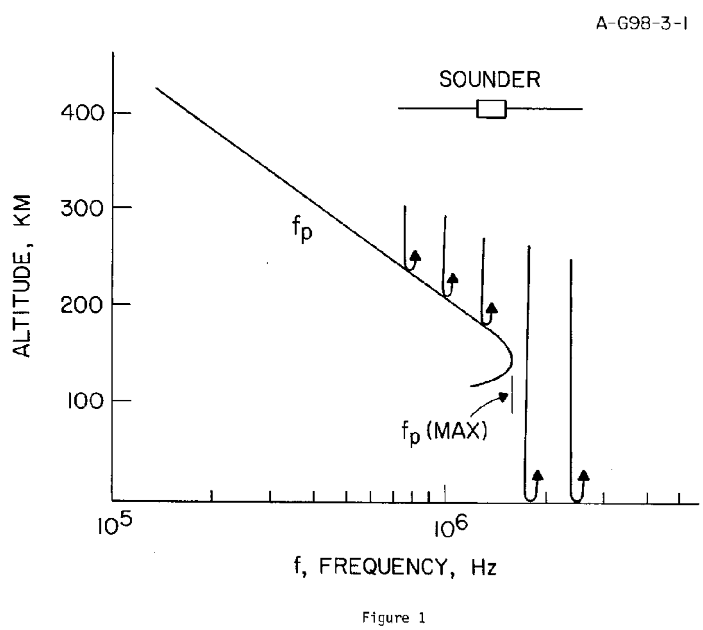

Figure 1

$$t_d(f) = \frac{2}{c}\int_0^{s*} \frac{ds}{\sqrt{1 - (f_p(s)/f)^2}}\ , \qquad (2)$$

where s is the vertical distance downward from the spacecraft, and s* is the reflection point. By measuring the time delay as a function of frequency, it is possible to invert the above equation to give the plasma frequency, hence the electron density, as a function of height. Note that it is only possible to obtain information on the topside of the ionosphere, at altitudes above f_p(Max), since no reflections can be obtained from the bottom side of the ionosphere. To measure the time delay as a function of frequency the transmitter frequency must be sequentially stepped after each radar pulse over a frequency range from well below the local electron plasma frequency at the spacecraft, f_p(Local), to well above the maximum plasma frequency in the ionosphere, f_p(Max). After each transmitter pulse the received signal strength must be recorded as a function of time delay after transmission of the radar pulse, $t_d(f)$. The resulting signal strengths are usually displayed in the form of an ionogram, which consists of a grey scale plot of the signal strength as a function of frequency and time delay. A representative ionogram from the Aloutte 2 spacecraft, which is in orbit around the Earth is shown in Figure 2, which has been adapted from [Benson, 1982], . Because of the relatively strong magnetic field in the Earth's ionosphere, terrestrial ionograms are quite complicated and show a variety of resonance and propagation effects. Resonances appear as vertical spikes in the ionogram and can be seen at the electron plasma frequency f_p, at the electron cyclotron frequency, $f_{c,}$ which is given by f_c = 28B Hz, where B is the magnetic field in nT, at harmonics of the electron cyclotron frequency $2f_c$, $3f_c$, etc., and various frequencies that involve combinations of f_p and f_c (i.e., f_{UH}, $f_{L=0}$, and $f_{R=0}$). Since Mars has a very weak magnetic field, the effects associated with the electron cyclotron frequency are expected to be completely absent. Therefore, ionograms at Mars should be much simpler than at Earth.

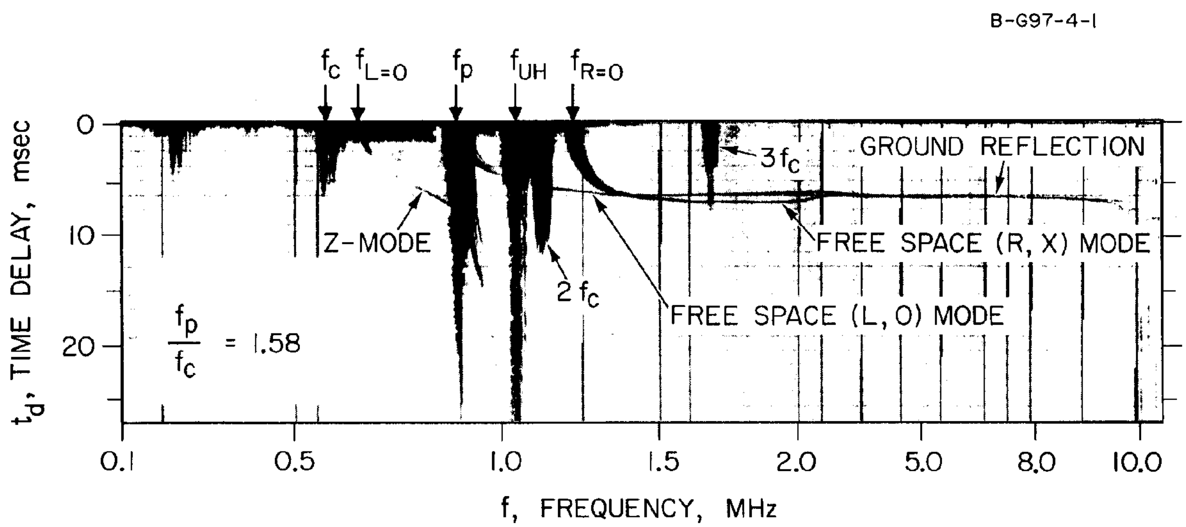

Figure 2

A representative sketch of an ionogram at Mars is shown in Figure 3. Dark lines indicate strong echos. Starting at the lowest frequency, the first response is expected to be a strong “spike” at the local plasma frequency, f_p(Local). This spike is caused by the excitation of electrostatic oscillations (Langmuir waves) at the local electron plasma frequency. Since these oscillations have very low damping they decay very slowly, thereby causing a long duration echo (i.e., a vertical spike) in the ionogram. The frequency of this spike can be used to calculate the local electron density at the spacecraft. Using Equation 1 it is easy to show that the relevant equation is

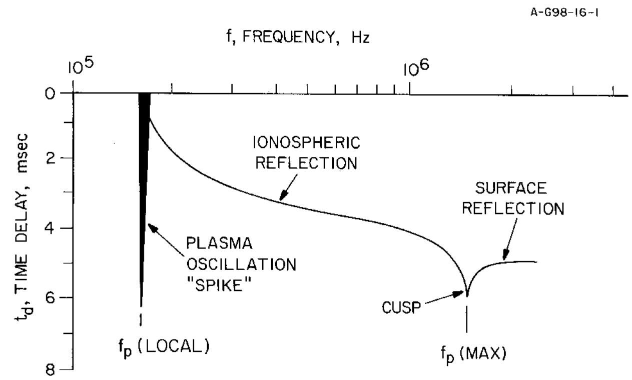

Figure 3

$$N_e = (f_p/8980)^2 \text{ cm}^{-3}, \quad (3)$$

where f_p is the frequency of the plasma oscillation spike in Hz. Immediately above the plasma oscillation spike a second echo is expected to appear with a time delay that increases with increasing frequency. This echo corresponds to the vertical reflection of the radar signal from the ionosphere, and terminates in a cusp at f_p(Max). At frequencies above f_p(Max) the radar signal penetrates through the ionosphere and reflects from the surface of the planet. The cusp in the time delay occurs because of the low group velocity and long path length in the vicinity of f_p(Max).

In addition to the vertical echos, oblique echos can also occur. Oblique echos can be caused by a variety of effect, for example, normal incidence reflections in regions with strong horizontal gradients, such as due to waves or discontinuities in the ionosphere, and scattering from small scale plasma density irregularities in disturbed region of the ionosphere. In the Earth’s ionosphere oblique echos are frequently observed from small scale structures in the auroral zones. Although Mars does not have auroral zones comparable to Earth, oblique echos are likely to occur in regions of enhance low energy electron precipitation, such are known to occur on the night side of Mars. Since the bow shock and ionopause are located just above the ionosphere at Mars, it is also possible that reflections could be observed from small scale density structures associated with these discontinuities. Since the conditions for oblique echos depends sensitively on the range and geometry of the reflecting or scattering region, both of which vary rapidly as the spacecraft moves along its orbital path, the shape of these features in the ionogram are much more complicated and difficult to predict than for echos from the horizontally stratified regions of the ionosphere.

Signal Processing and Implementation

In order to minimize the complexity and cost, wherever possible the frequencies, bandwidths, and sample rates for the ionospheric mode of operation have been selected to correspond as close as possible with the subsurface sounder mode of operation. In the original proposal [Picardi et al., 1998], it was proposed that the ionospheric sounding pulse would be a linear frequency modulated chirp with a duration of 530 μsec and a bandwidth of 20 kHz. Because of doubts about the processing gain that could be achieved by using a chirp the increased data handling required to process the chirp signal, it was decided at the March 22 -26, 1999, meeting that ionospheric sounding would be performed using an unmodulated (i.e., fixed frequency) pulse with a duration of about 100 μsec, which would give a range resolution of about 15 km.. The loss in signal to noise ratio due to the absence of pulse compression will be made up by increasing the transmitter power, which due to the shorter pulse duration can be increased without increasing the average transmitter power. Our present design assumes a peak pulse power into the final amplifier

To discuss the preliminary implementation arrived at during the March 22 -26 meeting it is useful to first start with a discussion of the frequency sweep that must be used for ionospheric sounding. As a baseline mode we propose to implement a linear frequency sweep that starts at a frequency f_{Min} and sweeps linearly up to a frequency f_{Max} over a time period ΔT. This basic frequency sweep is illustrated in Figure 4. For the minimum and maximum frequencies we adopt the same values used in the proposal, namely f_{Min} = 0.1 MHz and f_{Max} = 5.4 MHz. The time required for a complete frequency scan is determined by the dwell time Δt_1 between successive pulses, and the frequency step size Δf. The dwell time between successive pulses is determined by the maximum sounding range, which in the proposal was listed as 1200 km. Adopting this same value gives Δt_1 = 8 msec. This dwell time corresponds to a pulse repetition rate of 125 pulses/sec. For a contiguous frequency scan the frequency step size is determined by the pulse duration Δt_0, which from the uncertainty principle, is give by $\Delta f = 1/\Delta t_0$. To achieve maximum commonality with the subsurface sounder processing and a pulse duration of about 100 μsec the pulse duration has been chosen to 256 times the basic sampling time of the fast analog-to-digital converter used in the data processing system, which has a sampling rate chosen to be 2.8 megasamples/sec. For this sample rate of the pulse duration works out to be Δt_0 = 91.43 μsec, which is close to our target value of 100 μsec. The frequency step size is then Δf = 10.937 kHz. The total number of

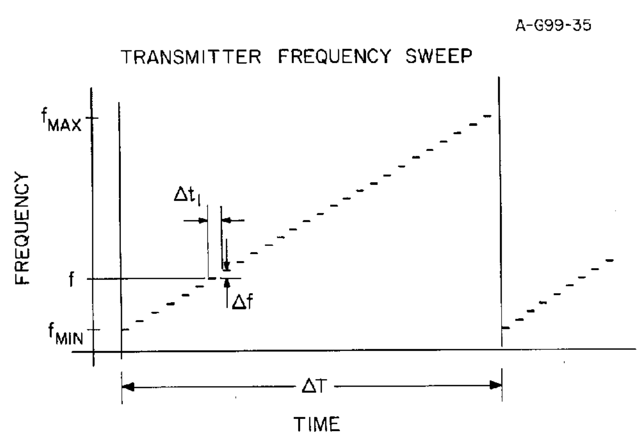

Figure 4

frequencies (from f_{Min} = 0.1 MHz to f_{Max} = 5.4 MHz) is then n = 485, and the time to scan all of these frequencies is $\triangle T$ = 485 x 8 msec = 3.88 sec.

Next we consider the receiver that is used to process the returning ionospheric echo. The subsurface sounder has four frequency bands, all with bandwidths of 1 MHz: Band 1 from 1.3 to 2.3 MHz, Band 2 from 2.5 to 3.5 MHz, Band 3 from 3.5 to 4.5 MHz, and Band 4 from 4.5 to 5.5 MHz. These bands are indicated by the shaded regions in Figure 5. Note that for the subsurface sounding there is no frequency coverage at frequencies below 1.3 MHz, and no coverage in the frequency range from 2.3 to 2.5 MHz. The gap from 2.3 to 2.5 MHz is to provide frequency separation in the impedance matching network between Band 1 and Bands 2, 3 and 4. Since continuous frequency coverage is required for ionospheric sounding, certain provisions must be make in the receiver design to assure that continuous frequency coverage is achieved. To provide continuous frequency coverage from 0.1 to 1.3 MHz a switch is to be included in the receiver to bypass the mixer so that signals from the antenna can be sampled directly, with no frequency conversion. This direct sampling channel is labeled Band 0 in Figure 5. To avoid aliasing, conventional practice is to sample the waveform at least 2.4 times the highest frequency component, which work out to be 3.12 megasamples/sec. Since the sample rate of the fast analog-to-digital converter is only 2.8 megasamples/sec, special care must be taken in the design of the filter for Band 0 to minimize aliasing effects in this band.

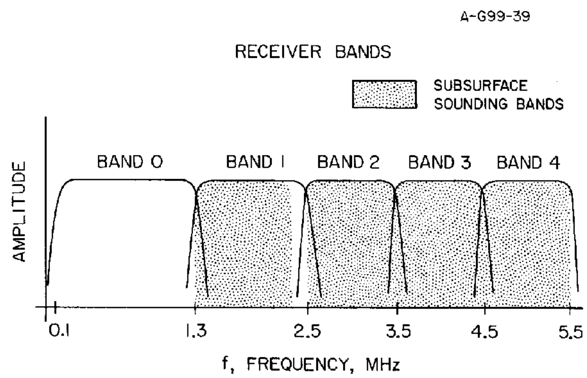

Figure 5

Since the noise level in the ≃ 1 MHz bandwidth of the receiver would be too high for efficient detection of the returning radar signal, some method must be used to effectively construct a narrow band filter centered on the frequency of the transmitted radar pulse. We propose to construct this filter by Fourier transforming the received waveform using exactly the same scheme that is used for processing the subsurface radar signals. This processing consists of Fourier transforming 128 in-phase and quadurature-phase (I and Q) 8-bit samples of the received waveform to give 128 complex Fourier amplitudes. By selecting the transmitter frequencies such that they exactly match the frequencies of the Fourier transforms, the amplitude of returning radar signal can be obtained by simply computing the amplitude of the appropriate Fourier component. This process is illustrated in Figure 6. The effective filter passband can be shaped by using an appropriate weighting function in the Fourier transform. We recommend that a simple Hamming weighting function be used to reduce the sidelobe response. If the waveform sample rate is 2.8 megasamples/sec,the frequency resolution of the Fourier transform would be $\triangle f$ = 10,937 kHz. Assuming that the conversion frequencies are appropriated multiples of this basic frequencies, the transmitter frequencies can be simple integer multiples of this frequency.

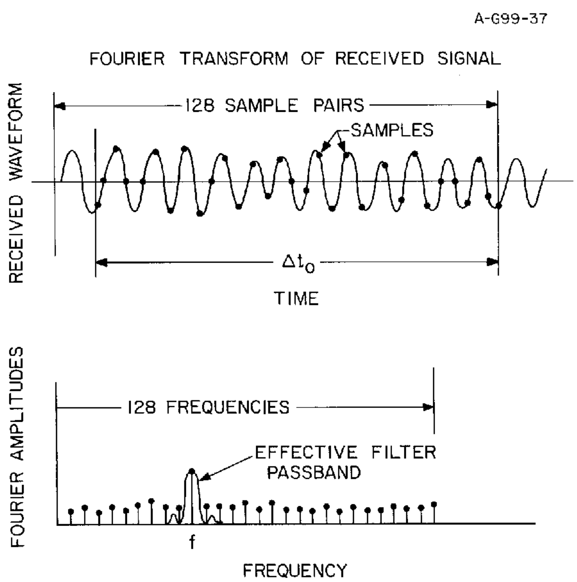

Figure 6

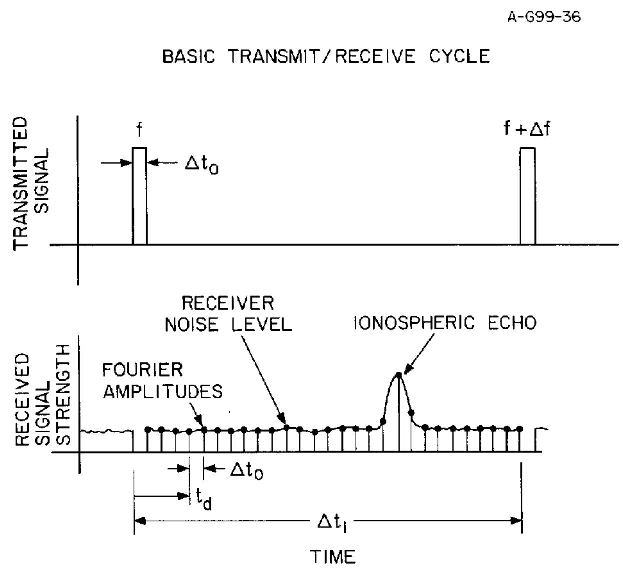

Figure 7

Since the time at which the returning radar signal is not known, a contiguous sequence of Fourier amplitudes (each at over an interval of Δt_0 = 91.43 µsec and at the frequency f of transmitted radar pulse) must be processed and stored in a memory over the entire Δt_1 = 8 msec dwell time interval until the next transmission of the next radar pulse. This sampling process is illustrated in Figure 7, and requires storage of 87 Fourier amplitudes, each of which consists of a 16-bit word. Sixteen bit words arise because the multiply operation in the Fourier transform converts the 8-bit waveform samples to 16-bit Fourier components. Since a total of n = 485 pulses are required to perform a complete frequency scan the total number of 16-bit Fourier amplitudes that must be stored for one ionogram (as in Figure 3) is 87 x 485 = 42,195 words, or 675,120 kbits. Since it takes 3.88 sec for one complete frequency scan, the bit rate that would be required to continuously transmit this data to the ground is174.0 kbits/sec. Since this bit rate clearly exceeds the capability of the spacecraft, we must next consider various forms of data compression and/or decimation.

Data Compression

A number of options exist for reducing the data rate without seriously compromising the scientific objectives of the ionospheric sounding. First, it is obvious that it is not necessary or even sensible to transmit the Fourier amplitudes with 16-bit accuracy, since these amplitudes were derived from only 8-bit samples of the waveform. We propose to reduce these Fourier amplitudes to 8-bit words. To maintain a constant fractional accuracy at all amplitudes, it would be desirable to convert the amplitudes to a decibel scale by computing 10 $\log_{10}$ of the Fourier amplitude and then transmitting this number with 8-bit accuracy. Since we are not certain that floating point arithmetic can be done sufficiently fast, an alternate scheme would be to transmit a mantissa and a characteristic by performing a bit shift so as to keep only the 4 most significant bits and then use the remaining 4 bits to transmit the number of bit shifts. Other somewhat similar quasi-logarithmic schemes are also well known.

To perform the primary data compression task we propose to use a well-known imaging compression scheme that is based on a 2-dimensional integer cosine transform (ICT). The basic approach is to regard the ionogram (i.e., Figure 3) as an image, and then to perform a 2-dimensional Fourier transform on 8 x 8 blocks of this "image"(see Figure 8). The ICT does not by itself reduce the data volume. However, by properly processing the ICT coefficients it is possible to use a compression scheme to significantly reduce the transmitted data volume. To select the proper ordering of the ICT coefficients, a "training" process is required on a sample of either real or simulated data. We have used this technique to perform on-board processing of frequency-time spectrograms (which are very similar to ionograms) from the plasma wave instrument on the Galileo spacecraft. The process work especially well on "images" where the main feature of interest is quasi-linear features or smooth curved lines, such as the spike or ionospheric reflection trace in an ionogram. Essentially the algorithm emphasizes the high wave number features in the image, thereby giving high resolution to small scale quasi-linear features. With Galileo we have consistently obtained high quality spectrograms with compression factors of 5 to 10. Assuming that the 16-bit Fourier transform amplitude are reduced to 8-bits, this would reduce the data rate for ionospheric sounding from 188.2 kbits/sec to 9.4 to 18.8 kbits/sec, which we hope is an acceptable rate.

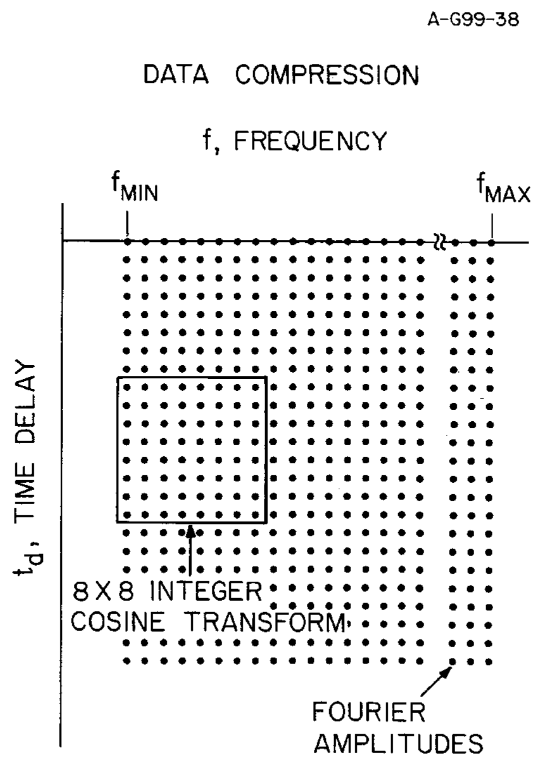

Figure 8

To give a rough idea of the processing requirements required to implement the ICT , we have consulted with Kar-Ming Cheung of JPL (telephone 818-393-0662) who was responsible for implementing the ICT used on Galileo. According to him it takes 512 integer addition, 64 integer division and192 shift operations for each 8 x 8 block in the image (in our case the 453 x 94 point ionogram) Since this was well within the capabilities of the more than 20-year technology used in the Galileo processor, we hope that this data compression scheme can be implemented in the MARSIS processor. Kar-Ming Cheung has volunteered to provide the source code for the 8 x8 ICT, provided appropriate arrangement are made within JPL. We presume that this will possible. One problem with the ICT compression, as with all lossy compression schemes, is that the compressed data rate is not entirely predictable. If the image has a high degree of redundancy (i.e., smooth and flat, with little structure) a large compression factor is possible. However, if the image becomes complicate, as might occur in the ionograms if many oblique echos are received from small scale structures, then the compression factor may get quite small. Just how this variability is managed so as to stay within the MARSIS data volume allocation is a topic that cannot be adequately addressed in this preliminary document. However, the ICT does have a control parameter called the Q-factor (quantization factor) that controls the compression factor. Increasing Q has the effect of increasing the compression factor and decreasing the quality of the image. On Galileo we introduced a "throttle" in the software that monitored the long term averaged data rate and adjusted the Q-factor so that the data volume stayed within acceptable limits. Such a scheme could easily be implemented in the MARSIS software.

If the above reductions do not reduce the data rate to an acceptable level, then there are a number of other reductions that could be made. However, none lead to the dramatic reductions that can be achieved with the ICT. These include, (1) reducing the upper frequency limit, $f_{Max,}$ of the frequency scan, particularly on the night side, where the maximum plasma frequency is quite low, (2) increasing the step size between adjacent frequency pulses, particularly in the upper part of the frequency scan, so as to maintain a more nearly constant fractional frequency resolution, $\Delta f/f$, (3) introduce a time gap between adjacent frequency scans (this would have the effect of reducing the horizontal spatial resolution), and (4) restrict the sampling in time delay to a small window around the time of the expected ionospheric reflection. Of these, the largest reduction could probably be achieved by implementing number (4). However, we do not recommend this option, since it would be operationally very complex and presupposes that we could come up with an ionospheric model that can predict the range to the ionospheric reflection with sufficient accuracy to assure that the vertical ionospheric return falls within a known small time window. Given the uncertainties of the Martian ionosphere, we doubt that this could be done with sufficient accuracy to justify the operational complexities that would be involved. Also, restricting the sampling to a small time window would preclude the possibility of studying oblique echos from small scale structures, which will most likely involve diffuse echos with completely unpredictable arrival times.