Proceedings of the Institute of Acoustics

MODULATION TRANSFER FUNCTION MEASUREMENT WITH MAXIMUM-LENGTH SEQUENCES

Douglas D Rife

DRA Laboratories, Sterling, Virginia

INTRODUCTION

The complex modulation transfer function (CMTF) has important applications in many fields including audio. for example, the magnitude of the CMTF or MTF is the basis for the speech transmission index (STI). A well known theorem shows how the CMTF of a noiseless linear time-invariant system can be derived from its impulse response. When maximum-length sequence (MLS) methods are employed, this theorem can be extended to include noise-contaminated as well as weakly non-linear systems. Furthermore, with MLS methods, there is no requirement that the interfering noise be stationary.

0. BACKGROUND

Given a noiseless linear time-invariant system having an impulse response, h(t), Schroeder [1] has shown that its complex modulation transfer function (CMTF) denoted, $m_n(f)$, can be expressed as,

$$m_n(f) = \frac{\int_{-\infty}^{+\infty} h^2(t) e^{-j2\pi ft}\, dt}{\int_{-\infty}^{+\infty} h^2(t)\, dt} \qquad (1)$$

The numerator of (1) is the Fourier transform of the squared impulse response while the denominator is the total energy of the impulse response. The denominator can be interpreted as the DC value of the numerator. When system noise is present, however, (1) is no longer valid but it can be corrected to yield the true CMTF, m(f). This corrected form, derived by Houtgast and Steeneken [2] is,

$$m(f) = \frac{\int_{-\infty}^{+\infty} h^2(t) e^{-j2\pi ft}\, dt}{\int_{-\infty}^{+\infty} h^2(t)\, dt} \cdot \frac{\overline{p^2_s}}{\overline{p^2_s} + \overline{p^2_n}} \qquad (2)$$

where p^2_s denotes the mean-square signal level and p^2_n the mean-square system noise. Relation (2) can therefore, in principle, be used to determine the CMTF of a noise-contaminated linear system based on two independent measurements of a) the noiseless system impulse response and b) the system signal-to-noise ratio. While sometimes practicable, (2) is nonetheless always inconvenient because of the need to make two separate system measurements.

Furthermore, (2) cannot easily be applied in those cases where the system noise level is dependent upon the input signal level. This is often the case with digital systems (e.g. CODEC's) in which quantization noise appears when a signal is present but disappears entirely when the signal is absent. A noise measurement of such a system cannot be performed simply by measuring the residual output noise with the input terminal grounded. Another limitation of the direct application of (2) is the need to measure an essentially noiseless impulse response. This requirement might require very long averaging times when measuring particularly noisy systems.

These limitations of (2) could be overcome if somehow (1) could be used directly simply by substituting a noise-contaminated impulse response for the noiseless version h(t). Unfortunately this approach, in general, will not work. Although the Fourier transform generally exists for finite energy functions such as the noiseless impulse response, h(t), this is not the case for stationary (persistent) functions that generally characterize noise. That is, noise processes generally possess infinite energy when integrated over all time as in the denominator of (1) causing it to "blow up". The lack of convergence of (1) under this infinite energy condition therefore generally forces us to employ (2) instead. However, if the total noise energy contaminating h(t) could somehow be guaranteed to be finite there is at least a possibility that (1) would converge to (2). This last conjecture is, in essence, the subject of this paper.

A previous paper [3] presented a comprehensive analysis of the maximum-length sequence (MLS) methods which can measure the ordinary transfer function as well as the related coherence function of linear time-invariant systems. Maximum-length sequences are binary (two-level) periodic sequences of period L which is always on less than a power-of-two thus,

(3) $L = 2^N - 1$, where N is an integer

The important property of any MLS is that its circular autocorrelation sequence is essentially an impulse. This property is exploited by an MLS analyzer to measure the periodic impulse response (PIR) of a linear time-invariant system according to,

(4)
$$h_m(n) = \frac{1}{L+1} \sum_{k=0}^{L-1} s(k)y(n+k),$$

where the index (n + k) is evaluated modulo L.

Here $h_m(n)$ denotes the measured PIR sequence, s(n) the driving MLS and y(n) the raw system output sequence. It can be seen that (4) is essentially a circular cross-correlation operation. Furthermore, because the measured PIR sequence is periodic, it is naturally suited to analysis by means of the discrete Fourier transform (DFT).

1. THEOREM

In an MLS measurement a maximum-length sequence is applied to a system whose output, y(n), which may contain additive noise, is cross correlated related with original sequence according to (4) to yield,

(5) $h_m(n) = h(n) + n(n)$

where $h_m(n)$ is the measured PIR sequence of period L
h(n) is the noiseless system PIR sequence
n(n) is an additive noise component

Thus $h_m(n)$ is, in general, the MLS-derived PIR sequence of a noise-contaminated linear time-invariant system.

The sequel will show that simply applying the discrete-time version of (1) to $h_m(n)$ yields, as L increases without limit, the true CMTF of (2) for all modulation frequencies except zero. for large but finite L, the result will also be correct except for a small random error.

2. ENERGY CONSERVATION PROPERTY

The MLS circular cross-correlation operation that recovers the PIR from the system output can be interpreted as a linear filter having an impulse response that is the time-reverse of the original MLS [3]. Such a linear filter is often called a matched filter. The time-reverse of any MLS is also an MLS denoted here also as s(n) for convenience. Now, if y(n) is an arbitrary input signal to an MLS analyzer, then the measured PIR sequence, $h_m(n)$, can be express as,

(6) $$h_m(n) = \frac{1}{L+1} y(n) * s(n)$$

where * denotes circular convolution [3].

By the convolution theorem of the DFT,

(7) $$H_m(f) = \frac{1}{L+1} Y(f)S(f).$$

Taking the magnitude and squaring both sides gives,

(8) $$|H_m(f)|^2 = \frac{1}{(L+1)^2}|Y(f)|^2|S(f)|^2.$$

The squared magnitude of the S(f) denoted $|S(f)|^2$ can be shown to equal L+1 at all frequencies except DC yielding,

(9) $$|H_m(f)|^2 = \frac{1}{L+1} |Y(f)|^2, \quad Y(0) = 0.$$

summing both side over all frequencies gives,

(10) $$\sum_{f=0}^{L-1} |H_m(f)|^2 = \frac{1}{L+1} \sum_{f=0}^{L-1} |Y(f)|^2, \quad Y(0) = 0.$$

Equation (10) shows the relationship between the energy in the respective DFTs of the system output signal and the measured PIR. We can express (10) entirely in time-domain terms by use of Parseval's energy relation [4] given as,

(11) $$\sum_{k=0}^{L-1} x^2(k) = \frac{1}{L} \sum_{f=0}^{L-1} |X(f)|^2$$

which holds for any sequence x(n). this yields,

(12) $$\sum_{k=0}^{L-1} h_m{}^2(k) = \frac{1}{L+1} \sum_{k=0}^{L-1} y^2(k), \quad Y(0) = 0$$

$$\approx \frac{1}{L} \sum_{k=0}^{L-1} y^2(k), \quad L >> 1.$$

Thus the MLS circular cross-correlation operation can be interpreted as a linear filter having an absolutely flat frequency response, except at DC. Therefore, assuming the input signal contains no DC component, the energy in the MLS-derived PIR, $h_m(n)$, nearly equals the mean-square value of y(n) over the measurement period for large L. It's important to realize that (12) holds regardless of whether the system output, y(n), is correlated (signal) or uncorrelated (noise) with the MLS stimulus. That is, the signal-to noise ratio (SNR) in, $h_m(n)$, will exactly mirror the SNR of the actual measurement situation over the measurement interval T = Lt where t is the sampling interval.

3. PHASE RANDOMIZATION PROPERTY

While the MLS method preserves the system SNR in the measured PIR, it also randomizes the phase spectrum of any system output component that is not correlated with the input MLS. Let $n_a(n)$ be the actual system output noise defined as the sum of all system output components that are not correlated with the driving MLS. As defined by (5), n(n), is simply a filtered version of $n_a(n)$, the linear filter here being the MLS cross-correlation operation itself. The DFT of the noise appearing in the measured PIR is therefore,

(13) $$N(f) = \frac{1}{L+1} N_a(f) S(f),$$

or in polar form,

(14) $$N(f) = \frac{1}{L+1} |Na(f)||S(f)|e^{j\theta\{N_a(f)\}}e^{j\theta\{S(f)\}}$$

$$= \frac{1}{L+1} |Na(f)||S(f)|e^{j[\theta\{N_a(f)\} + \theta\{S(f)\}]}$$

so that,

(15) $$\theta\{N(f)\} = \theta\{N_a(f)\} + \theta\{S(f)\}.$$

MODULATION TRANSFER FUNCTION MEASUREMENT WITH MLS

By definition, the actual noise $n_a(n)$ is uncorrelated with the driving MLS. Therefore, their respective phase spectra $q\{N_a(f)\}$ and $q\{S(f)\}$ are statistically independent. Furthermore, and of greater significance, is that fact that an MLS exhibits a highly erratic and therefore essentially random phase spectrum having a uniform or constant probability density function (PDF) over its domain of -T p *to* + *T p* radians [see 3, figure 2]. The number of unique Fourier coefficients in the MLS spectrum is shown by Golom to equal the number of cyclotomic cosets of the sequence [5]. The number of unique Fourier coefficients (K) for various MLS orders (N) and periods (L) is reproduced in the table below.

Table I
Number of unique MLS Fourier coefficients (K) vs. MLS order (N) and MLS period (L) from Golomb.

N	L	K
12	4095	351
13	8191	631
14	16383	1181
15	32767	2191
16	65535	4115
17	131,071	7711
18	262,143	14,601
19	524,287	27,595
20	1,408,575	52,487

But since the magnitude of the MLS spectrum is a constant at all frequencies except DC, then table I can be taken to represent the number of unique phase angles in the MLS spectrum. As can be seen from this table, for the longer MLS periods there are a sufficient number of phase angles that the PDF of the MLS phase spectrum can be considered continuous for practical purposes. In the limit, this last condition is exactly true as the MLS period L approaches infinity.

The PDF of the sum of two independent random variables, as in (15), is equal to the convolution of the two PDFs [6]. In this case, however, the convolution must be performed on a circular basis because phase is circular over its domain. This means that the PDF of the sum of these two phase spectra must also be uniform because the circular convolution of a constant function with *any other function* is also constant. Thus the PIR noise, n(n), except for a scale factor, will have the same magnitude spectrum as the actual noise, $n_a(n)$, but its phase spectrum will be *completely randomized* by the MLS cross-correlation operation. Thus the additive noise component, n(n), will generally be stationary even if the actual system noise is non-stationary (eg. transient noise).

These two desirable properties namely, energy conservation and phase randomization are, individually, not unique to MLS methods. All methods which employ a truly random stimulus (eg. dual-channel FFT methods) exhibit phase randomization but they fail to conserve energy, this being due to the random nature of the stimulus which shows up in the measurement as variance or excess noise over and above the actual system background noise. On the other side, some other deterministic methods, for instance cross-correlation employing Legendre sequences [7], exhibit energy conservation but these fail to randomize the phase of uncorrelated system noise. the phase spectra of Legendre sequences, for instance, are two valued, not uniformly distributed. MLS methods are currently the only ones known which posses both of these desirable properties. In retrospect, this result should not be surprising for it turns out that maximum-length sequences satisfy all three of Golomb's randomness postulates and hence are, in his works, "the true pseudo-noise sequences" [5].

MODULATION TRANSFER FUNCTION MEASUREMENT WITH MLS

Golomb's randomness postulates derive from the observed properties of true randomness such as the flipping of a fair coin. In that case we observe that a) the number of heads nearly equals the number of tails (due to equal probability of heads and tails), b) there are twice as many runs of heads (or tails) of length r as there are runs of length r + 1 (due to the independence of successive coin flips) and c) by setting heads = -1 and tails = +1, the auto correlation sequence is nearly an impulse (due to an overall lack of a pattern in a sequence of coin flips). Maximum-length sequences satisfy all three of Golomb's randomness postulates while Legendre sequences violate postulate b. Thus although maximum-length sequences are not truly random they can, with confidence, be regarded as such in many applications including the present one.

4. COMPLEX MODULATION TRANSFER FUNCTION

Using relation (1) in discrete time, the MLS-derived CMTF or MLS-CMTF denoted, $m_m(f)$, is given as,

(16)
$$m_m(f) = \frac{DFT[h_m^2(n)]}{\sum_{k=0}^{L-1} h^2_m(k)} .$$

Substituting (5) in (16)

(17)
$$m_m(f) = \frac{DFT[h^2(n)] + DFT[2h(n)n(n)] + DFT[n^2(n)]}{\sum_{k=0}^{L-1} h^2(k) + \sum_{k=0}^{L-1} 2h(k)n(k) + \sum_{k=0}^{L-1} n^2(k)} ,$$

and applying the product theorem of the DFT gives

(18)
$$m_m(f) = \frac{H(f)*H(f) + 2H(f)*N(f) + N(f)*N(f)}{\sum_{k=0}^{L-1} h^2(k) + \sum_{k=0}^{L-1} 2h(k)n(k) + \sum_{k=0}^{L-1} n^2(k)}$$

Because n(n) has been phase-randomized by the MLS cross-correlation operation, the expected value of the second term of the numerator and the denominator must both be zero so that,

(19)
$$E\{m_m(f)\} = \frac{H(f)*H(f) + N(f)*N(f)}{\sum_{k=0}^{L-1} h^2(k) + \sum_{k=0}^{L-1} n^2(k)} ,$$

Where $E\{m_m(f)\}$ denotes the expected value of $m_m(f)$. The expected value of $N(f)^*N(f)$ is a special case. Because n(n) is a real sequence, its DFT, N(f), must be conjugate symmetric around zero frequency. Therefore, for a zero frequency shift, the indicated convolution denoted $\Omega_{NN}(0)$ is

(20) $$\Omega_{NN}(0) = \frac{1}{L}\sum_{k=0}^{L-1} N(k)N^*(k),$$

$$= \frac{1}{L}\sum_{k=0}^{L-1} |N(k)|^2,$$

where * denotes complex conjugate.

Then by Parseval's relation,

(21) $$\Omega_{NN}(0) = \sum_{k=0}^{L-1} n^2(k).$$

for all other frequency shifts, however, the expected value of $\Omega_{NN}(f) = N(f)^*N(f)$ will be zero given that N(f) exhibits uniformly distributed phase. [Since the complex sequence N(f) exhibits uniformly distributed phase then so must the convolutional lag product N(k)N(f-k) for any non zero f. The summation of this lag product over all k must therefore approach zero.] Therefore, in general, the expected value of $N(f)^*N(f)$ can be written as

(22) $$E\{\Omega_{NN}(f)\} = \delta(f)\sum_{k=0}^{L-1} n^2(k),$$

where,
d(f) = 1, f = 0
= 0, otherwise.

Substitution in (19) yields,

(23) $$E\{m_m(f)\} = \frac{H(f)*H(f) + \delta(f)\cdot\sum_{k=0}^{L-1} n^2(k)}{\sum_{k=0}^{L-1} h^2(k) + \sum_{k=0}^{L-1} n^2(k)},$$

which is the final result. In words, relation (23) simply states that, except at zero modulation frequency the expected value of the MLS-CMTF, $E\{m_m(f)\}$, will equal the true CMTF, m(f) of (2). To see this more clearly, (23) can be express in factored form as,

(24) $$E\{m_m(f)\} = \frac{H(f)*H(f)}{\sum_{k=0}^{L-1} h^2(k)} \cdot \frac{\sum_{k=0}^{L-1} h^2(k)}{\sum_{k=0}^{L-1} h^2(k) + \sum_{k=0}^{L-1} n^2(k)}, \quad f > 0$$

which is just a discrete-time version of (2).

But at zero modulation frequency $E\{m_m(f)\}$ must be unity because analogous to (21)

(25) $$n_{HH}(0) = \sum_{k=0}^{L-1} h^2(k).$$

Therefore, there is an error in the MLS-CMTF but only at zero modulation frequency. At all other modulation frequencies the result will be correct except for a random error which will decrease as the MLS period increases.

In the limit, as L goes to infinity, the MLS-CMTF and the true CMTF will agree exactly except at zero modulation frequency. We can be sure that this limit exists because the additive noise component, n(n), in the measured PIR is of finite energy even for infinite L. This can be appreciated by careful examination of (12) which shows that the total energy of n(n) will approach the mean-square value of the actual noise $n_a(n)$. And while it is true that $n_a(n)$ generally possesses infinite energy, its mean-square value is clearly finite and so therefore is the energy of n(n).

Another way to look at the convergence of $m_m(f)$ is to recognize that the mean-square amplitude of n(n) will go to zero as L increases without limit. Thus in one sense, the measured PIR will not contain any noise at all for infinite L. Yet paradoxically, n(n), will still possess a definite amount of energy. But there is no real contradiction here because, in the limit, the finite energy of n(n) is simply dispersed or diffused over all time. This apparent paradox explains why MLS methods are at once immune to background noise and yet still able to perfectly account for it in MLS-CMTF measurements!

5. VARIANCE ESTIMATE

In any practical measurement, of course, L must be finite leading to a certain amount of random error or variance in the MLS-CMTF. Obviously, the variance of $M_m(f)$ denoted $Var\{m_m(f)\}$ will be zero if no system noise is present. Conversely, it will be at a maximum when *only* noise is present. This last worst case condition will be the working assumption in the analysis to follow. The noise, n(n), is also assumed to be stationary, zero-mean Gaussian, white and ergodic (i.e. time averages equal ensemble averages). While these last two conditions are not always true in practice, they are necessary to make the analysis tractable. The noise-only MLS-CMTF denoted $m_{mn}(f)$ is

(26) $$m_{mn}(f) = \frac{DFT[n^2(n)]}{\sum_{k=0}^{L-1} n^2(k)}.$$

It is more transparent to express (26) entirely in terms of N(f) yielding,

MODULATION TRANSFER FUNCTION MEASUREMENT WITH MLS

(27)

$$m_{mn}(f) = \frac{\sum_{k=0}^{L-1} N(k)N(f-k)}{\sum_{k=0}^{L-1} |N(k)|^2} .$$

In practice it is often useful to pre-filter the measured PIR in order to determine the CMTF over a particular band of frequencies. If we denote B to be the filter bandwidth in Hz then,

(28)

$$m_{mn}(f) = \frac{\sum_{k=k_l}^{k_h} N(k)N(f-k)}{\sum_{k=k_l}^{k_h} |N(k)|^2} ,$$

where k_l is the lower cut-off frequency index and k_h the upper cut-off frequency index. The DFT frequency spacing will be 1/T Hertz where T = Lτ seconds is the PIR duration so that,

(29) $B = (1/T)(k_h - k_l + 1)$ or,
$BT = (k_h - k_l + 1)$

Note that because the DFT is a linear transformation, N(f) must also be a zero-mean Gaussian sequence. Assuming n(n) is also white and ergodic, the variance of N(f) is,

(30)

$$\mathrm{Var}\{N(f)\} = \frac{1}{L}\sum_{k=0}^{L-1} |N(k)|^2 = \sigma^2 ,$$

and so (28) can be expressed as,

(31)

$$m_{mn}(f) = \frac{1}{BT\sigma^2} \sum_{k=kl}^{kh} N(k)N(f-k) .$$

The variance of the lag product N(k)N(f-k) will equal the product of the variance of its two factors or σ^4 (assuming N(f) is zero-mean Gaussian). Further, if we add together BT points of this lag product the variance of the resulting sum will be $BT\sigma^4$, therefore,

(32)

$$\mathrm{Var}\{m_{mn}(f)\} = \frac{1}{(BT\sigma^2)^2} \cdot BT\sigma^4$$

$$= \frac{1}{BT}, \quad f > 0.$$

Thus the variance or random error of $m_{mn}(f)$ will be inversely proportional to the time-bandwidth product BT. This is the case simply because more points of the lag product, N(k)N(f-k), are averaged together as BT increases. When a signal is also present, equation (32) represents an upper bound on the variance so that in general,

(33)
$$\mathrm{Var}\{m_m(f)\} \leq \frac{1}{BT}$$

6. EXPERIMENTAL VERIFICATION

To illustrate the soundness of the above theory, the MLSSA acoustical measurement system was used to measure a noise-contaminated system whose output contained exactly 50% uncorrelated white noise. The test system's noiseless impulse response was that of MLSSA's 8-pole Chebyshev antialiasing filter -- essentially a delta function -- in order to focus exclusively on the noise effects. A 32767-point MLS was used to measure an equally long PIR. MLSSA's antialiasing filter bandwidth was set to 10 kHz with a sampling rate of 30.1 kHz. Thus T = 1.09 seconds and B = 10 kHz so that BT $\approx$ 10,000 and the maximum variance expected in the MLS-derived CMTF should be 10.0×10^{-5}.

Figure I shows the actual measured CMTF magnitude for these conditions. The measured variance below 100 Hz is shown to be 1.2×10^{-5} which is well below the upper bound of 10.0×10^{-5}. The mean value over the same range is shown to be 0.498 which is in close agreement with the expected value of 0.5. Note that this curve rolls-off significantly above a modulation frequency of about 100 Hz. This is to be expected for as the modulation frequency, f, increases, the lag product $H_m(k)H_m(f\text{-}k)$ exhibits less self-overlap assuming that $h_m(n)$ is band limited as it must always be in practice. Thus, the CMTF of any band limited system, even an otherwise perfect one, must always exhibit high frequency roll-off. In other words, any such roll-off is not an artefact or measurement error but is the logical result of band limiting the transmission channel.

Figure II shows the MLS-CMTF computed from a digitally filtered version of the same PIR. MLSSA's 6-pole Butterworth infinite impulse response (IIR) digital filter was set to a 1 kHz centre frequency and a 1 octave bandwidth. For these conditions, B = 707 Hz, T = 1.09 seconds, BT = 771 and the maximum expected variance is 1.3×10^{-3}. The actual variance below 10 Hz is shown to be 1.9×10^{-4}. Clearly, the amount of random error in an MLS-CMTF measurement does increase as the time-bandwidth product decreases.

The foregoing results were obtained using stationary white interfering noise. We now consider non-stationary, colored interfering noise. In place of white noise, impulsive noise was injected into the wideband signal path. this noise, shown in Figure III, consisted of 3.25 millisecond pulses repeated every 101 milliseconds, somewhat reminiscent of machine gun fire. The wideband PIR of the noise-contaminated system is shown in figure IV. Note the smooth stationary-like noise in the tail of the measured PIR bears no evidence of the impulsive nature of the actual noise of Figure III. Such is the expected result of MLS phase randomization. Figure V shows the computed wideband CMTF magnitude obtained from the PIR of Figure IV. Note the increased variance due mainly to the presence of spikes that fall on harmonics of the pulse repetition rate, that is, the spikes fall at modulation frequencies of about 10 Hz, 20 Hz and 30 Hz etc. The spikes also tend to alternate in direction thus tending to cancel each other out in the calculation of the mean. Clearly, impulsive and colored interfering noise does increase the MLS-CMTF variance over white noise but the results are still quite good. Recall that the previous analysis of MLS-CMTF variance assumed *white* interfering noise which is not the case for Figure V. the residual variance can be reduced somewhat by smoothing the CMTF magnitude. Figure VI shows the same data as figure V which has now been smoothed to 0.33 octave. Such smoothing is accomplished as follows. Every smoothed CMTF point is computed as the RMS value of the raw CMTF curve taken over a 1/3 octave interval centred on that point. Note that the spikes have disappeared in the smoothed version.

Figure VII shows the CMTF with impulsive noise for a 1 kHz, 1 octave band by passing the wideband PIR through the same 6-pole bandpass filter. Note that the interfering noise level was readjusted here to give a mean CMTF magnitude of about 0.5 for easier comparison to figure II. Note that the variance for the 1 kHz octave case with stationary interfering noise shown in Figure II. Thus MLS-CMTF measurements seem to be quite tolerant of non-stationary noise as predicted by theory.

7. REQUIREMENTS FOR STI MEASUREMENT

One of the most important applications of the CMTF is the determination of the Speech Transmission Index (STI). This important measure of speech intelligibility is well documented elsewhere [2, 8, 9] but little attention has been given in the past to instrumentation requirements when applying equations (1) or (2). STI is computed from the magnitude of the CMTF, often named just the modulation transfer function (MTF), measured at 14 discrete 1/3 octave modulation frequencies ranging from 0.63 Hz to 12.5 Hz in each of 7 octave bands ranging from 125 Hz to 8,000 Hz..

An oft-neglected consideration is that, due to the time-frequency uncertainty principle, the 0.63 Hz modulation frequency demands a minimum measured impulse response duration of 1/0.63 or about 1.6 seconds for good accuracy. While it is true this requirement can be relaxed if the actual system impulse response decays to a negligible value in somewhat less time, in practice, STI is often used to evaluate auditoria many of which have long reverberation times. Indeed, the most reverberant halls are the ones most likely to suffer from speech intelligibility problems and hence are most likely to be evaluated by the STI method.

The upper cut-off frequency of the 8,000 Hz octave passband is 1.414.8,000 = 11,310 Hz so that a total measurement bandwidth of about 12 kHz is also required. Thus a minimum overall time-bandwidth product of 12000.1.6 = 19,200 is needed to assure accurate STI measurements using any method. These requirements can be fully met by employing a 65535-point MLS to measure an equally long PIR.

To test the effects of noise on the MLS-STI, MLSSA measured a similar noise-contaminated test system whose output contained exactly 50% MLS signal and 50% uncorrelated stationary white noise. This time, however, the instrument bandwidth was set to 12 kHz and 65535-point MLS was used to measure an equally long PIR which was then analyzed by MLSSA's built-in STI function. The STI calculation required just 3 seconds to complete running on a 33 MHz 486 computer. The PIR duration was 1.8 seconds which is well above the 1.6 second minimum requirement for the 0.63 Hz modulation frequency. The results are presented in Figure VIII.

As can be seen, the individual MTF values as well as the final STI value all lie near 0.5 as would be predicted. The MTF values of the lower bands show more variance than those of the upper bands also as predicted. All of these random MTF errors are reduced further because the STI procedure effectively averages together the 14 individual MTF values to determine each octave's transmission index (TI) shown on the bottom row of the MTF matrix of Figure VIII. The largest TI error is in the 125 Hz band as expected. Finally, the STI procedure combines these 7 individual TI values into a weighted average to form the final STI value. This further reduces the random error. In this example the total error in the final STI value is -0.003 which is within 0.6% of the expected value.

As previously demonstrated, impulsive interfering noise tends to generate spikes in the MLS-CMTF at harmonics of the pulse repetition rate but that these spikes tend to occur in alternate directions. Thus such spikes will tend to cancel each other out in MLS-STI measurements involving impulsive noise. As a practical point, therefore, one should not put too much weight on the individual MTF values in the STI matrix especially when non-stationary noise is present. What counts is rather their mean as reflected by the octave TI values and the final overall STI value.

Note that when the actual STI value is very close to zero a positive bias error will appear in the MLS-STI. That is, although the real and imaginary parts of $m_m(f)$ are unbiased estimates of the real and imaginary parts of $m(f)$, the magnitude, $|m_m(f)|$, will show a positive bias error when the actual CMTF is near zero, this being due to squaring and summing the random errors. Restated in statistical terms, if the real and imaginary parts of $m_m(f)$ each exhibit a zero-mean Gaussian PDF then $|m_m(f)|$ will exhibit a Rayleigh PDF which always shows a positive mean [6]. Nevertheless, this bias error only becomes significant for very low SNRs in which case exact knowledge of the true STI is of little value anyway.

Note also, in order to make a real MLS-STI measurement that properly accounts for the background noise, the white MLS stimulus must first be passed through a speech-weighting filter prior to applying it to the system under test. A speech-weighting filter is one whose frequency response corresponds to the long-term average spectrum of normal speech.

As discussed in [3], MLS methods can give rise to time aliasing, this being due to the periodic nature of maximum-length sequences. Time aliasing occurs when the MLS period is short relative to the reverberation time of the system being measured. In such cases the tail of the true impulse response will wrap around and add to its initial portion to form the measured PIR [see 3, Figure 1]. Because of this possibility, a theoretical analysis is required to estimate the effects of time aliasing upon MLS-STI measurements in reverberant environments.

This is most easily accomplished by regarding the wrapped or aliased trailing portion of the impulse response as excess noise which would not be present in the PIR if time aliasing did not occur. According to the definition of STI, however, octave noise levels 15 dB or more below the signal level have negligible effect upon and octave's TI value. Therefore, a room's particular octave RT_{60} can be up to 4 times longer than the MLS period when making MLS-STI measurements. This result is obtained as follows. Assuming exponential reverberant decay, if we divide the true impulse response at a point in time measured from the first arrival equal to $RT_{60}/4$ seconds, then the energy in the trailing (time-aliased) portion will be $60/4 = 15$ dB below the total PIR energy. Thus a PIR duration of 1.8 seconds can accurately measure the STI of rooms having an octave RT_{60} of up to $4.1.8 = 7.2$ seconds with negligible error. Normally, the 125 Hz octave will have the longest RT_{60} and this octave is given little weight in computing the final STI value. In conclusion, when employing a 65535-point MLS for MLS-STI measurements, time aliasing can generally be disregarded as a source of error except in the most pathologically reverberant environments.

8. SOME EXAMPLES OF MLS-STI MEASUREMENTS

In response to the request of one reviewer, this section was added to give some more concrete example of MLS-STI measurements. The first example is a real STI measurement of a large church whose impulse response is shown in Figure IX. The MLS-STI calculation for this impulse response is shown in Figure X. Note that the rather poor intelligibility in this case is due mostly to reverberations and reflections with the background noise playing only a minor role.

The next example demonstrates just the opposite situation. A telecommunications CODEC typically exhibits no reverberation or reflection effects yet can easily degrade speech intelligibility due to its internally generated quantization noise. MLS-STI measurements of such noise-dominated systems must always include a speech-weighting filter for valid results. Figure XI shows the MLS-STI of a CODEC having a 3.4 kHz bandwidth. The CODEC transmits digital voice but the data are encoded with only 2 bits per sample of resolution. Figure XII shows the MLS-STI of the same CODEC but now with 3 bits per sample of resolution. The improvement in intelligibility is notable. Both of these CODEC measurements employed a speech-weighting filter to properly account for the quantization noise.

9. DISCUSSION

It's important to realize that any CMTF measurement based on equations (1) or (2), whether MLS-derived or otherwise, applies to linear time-invariant systems only. The CMTF of strongly non-linear or time-variant systems must be measured by the direct method [1, 8] which involves passing bandlimited amplitude-modulated white noise through the system and measuring the reduction in modulation at the output. Examples of systems that must be measured by the direct method include vocoders, which are both strongly non-linear and time-variant and, certain time-variant artificial reverberators. Furthermore, in such cases, direct CMTF measurements must be made one octave at a time leading to long measurement times. Note also that even if direct CMTF measurements of such systems are possible, the STI derived from them might be invalidated due to other effects such as pitch translation in vocoders.

Moreover, even when used to measure linear time-invariant systems, the direct method can suffer from severe errors if the interfering noise is non-stationary. This is the case when, as is typical with direct methods, ordinary non-synchronous envelope detection is employed to recover the output modulation. Anyone who has observed the lack of noise immunity of ordinary AM radio can attest to this problem which can only be alleviated by resorting to long averaging times. The MLS approach, in contrast, due to synchronous detection and phase randomization of interfering noise, is generally much faster and more tolerant of non-stationary interfering noise than the direct method.

Incidentally, the energy conservation and phase randomization properties of MLS methods are also the fundamental reason why they can measure the coherence function of weakly non-linear systems [3]. That is, weak non linearity creates harmonic and intermodulation (distortion) components that are largely (though not entirely) uncorrelated with the driving MLS due to the fact that such components appear mostly at frequencies that differ from the fundamental MLS frequencies that gave rise to them. (See [3] for a more detailed discussion on this point.) These uncorrelated harmonic and intermodulation components, like the background noise, also get phase-randomized by the MLS cross-correlation operation and simply add to the stationary noise, $n(n)$, of the measured PIR. Therefore, weak non-linear distortion will, in general, have an effect upon the measured MLS-CMTF that is identical to the effect of background noise. This conclusion must be qualified, however, when such MLS-CMTF measurements are used to compute STI.

MODULATION TRANSFER FUNCTION MEASUREMENT WITH MLS

As pointed out in [8], to precisely account for non linearity in an STI measurement, the MTF values for each octave must be measured separately, one octave at a time, with running speech present simultaneously in the other octaves. The reason for this is not only because the non-linear distortion spills over into adjacent STI analysis bands but that its level in the current analysis band depends strongly upon the statistical properties of the signal present in the adjacent bands. As discussed in [8], the amount of distortion generated by a non-linear system when fed an artificial test signal will usually be lower than the amount of distortion generated by the same system when fed running speech. Due to its random envelope and speech-like spectrum, a speech-filtered MLS signal is at least a rough approximation at running speech but, like most noise-like artificial test signals, its envelope crest factor is generally lower than that of actual running speech. Thus parallel MLS-STI measurements will, in general, only *partially* account for the detrimental effects of distortion upon speech intelligibility while *fully* accounting for the detrimental effects of the background noise, reflections and reverberation.

Note, however, that with proper modifications, MLS-STI methods could be extended to fully account for weak non-linearity by following the example of [8], that is, by measuring the MLS-CMTF in each STI octave in turn (serially) with running speech present simultaneously in the adjacent octaves. Of course, such a modified MLS-STI method would be more complicated and would require longer measurement times thanthe parallel method which measures all 7 STI octaves simultaneously in one fell swoop. It is doubtful whether the modestly increased accuracy of a serial MLS method would outweigh the speed and simplicity of the parallel MLS method.

Compared to time-delay spectrometry (TDS) or dual-channel FFT methods (FFT), only MLS methods can satisfy the enormous time-bandwidth product requirement of STI ($\approx$ 20,000) in reasonably short measurement times. Furthermore, because the present theorem does not apply to them, both TDS and FFT methods will still require a separate measurement of the system background noise, where applicable, in order to correct the noiseless CMTF of (1) according to (2). Moreover, when using these other methods to make such noiseless CMTF measurements of particularly noisy systems one may need to employ averaging techniques which use can lead to even longer measurement times. Not only is averaging not required with MLS-STI methods but such averaging can actually invalidate the measurement by artificially reducing the background noise in the measured PIR thus leading to an overly optimistic (too high) STI value.

In conclusion, MLS-STI measurements are applicable to a wide range of intelligibility verification applications including auditoria, sound systems, telecommunications systems, as well as digital and analog tape recorders. Unlike TDS or FFT methods, they do not require any special correction for background noise effects provided only that the MLS stimulus is first passed throng a speech-weighting filter. Nevertheless, MLS methods are not applicable to strongly non-linear systems nor to time-variant systems. Examples of common time-variant systems include certain types of artificial reverberators and certain sound-effects (fuzz) boxes that work by modulating the frequency or phase of the audio signal. In such cases only direct MTF measurements will suffice in determining the STI, where applicable, as neither MLS, TDS nor FFT methods are valid under conditions of strong non linearity and/or time-variance.

10. ACKNOWLEDGEMENTS

The author would like to thank Dick Campbell of Bang-Campbell Associates for making and providing the church impulse response measurement. He would also like to thank Mark Dismukes of Loral Fairchild corporation for making and providing the CODEC measurements. Finally, the author would like to thank the reviewers for their kind and helpful comments.

REFERENCES

[1] M. R. Schroeder, "Modulation Transfer Functions: Definition and Measurement.", Acustica, vol. 49, pp179-182 (1981)

[2] T. Houtgast and H. J. M. Steeneken, "Predicting Speech intelligibility in rooms from the Modulation Transfer Function. I. General Room Acoustics", Acustica, vol. 46, pp60-72 (1980)

[3] D. Rife and J. Vanderkooy, "Transfer-Function Measurement with Maximum-Length Sequences", Journal of the Audio Engineering Society (JAES), vol. 37, No.6, pp419-443, 1989 June[4] A. V. Oppenheim and R. W. Schafer, Digital Signal Processing, Prentice-Hall

[5] S. W. Golomb, Shift Register Sequences, Revised Edition, Aegean Park Press, Laguna Hills, California, 1982.

[6] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill Inc., NY, 1982

[7] M. R. Schroeder, "Integrated-impulse Method Measuring Sound decay without using Impulses", Journal of the Acoustical Society of America (JASA), vol. 66, pp497-500, August 1979

[8] H. J. M. Steeneken and T Houtgast, "A physical method for measuring speech-transmission quality", JASA, vol. 67, No.1, pp318-326 (1980)

[9] K. D. Jacob, "Correlation of Speech Intelligibility tests in Reverberant Rooms with Tree Predictive Algorithms", JAES, vol. 37, No. 12, pp 1020-1030, December 1989

FIGURE CAPTIONS

I. MLS-MTF measurement of 10 kHz lowpass system contaminated with stationary white noise with 0 dB SNR. Note that the mean value is near 0.5 as expected.

II. MLS-MTF measurement of a 1 kHz, 1 octave bandpass system contaminated with stationary white noise also with 0 dB SNR. Note the increased variance as compared to Figure I.

III. Artificially produced non-stationary impulsive noise source consisted of 3.25 millisecond pulses repeated every 101 milliseconds, somewhat reminiscent of machine gun fire.

IV. MLS periodic impulse response (PIR) measurement of a 10 kHz lowpass system contaminated with the impulsive colored noise of Figure III. Note that, due to phase randomization, there is no evidence of the impulsive nature of the interfering noise in the measured PIR.

V. Wideband (10 kHz) MLS-MTF magnitude calculated from the impulse response of Figure IV shows that the impulsive colored interfering noise increases the variance due to the appearance of spikes that occur at harmonics of the pulse repetition rate, namely 10 Hz, 20 Hz and 30 Hz etc.

VI. Same MLS-MTF as figure V but smoothed to 0.33 octave to reduce the variance.

VII. MLS-MTF measurement of a 1 kHz, 1 octave bandpass system with colored impulsive interfering noise. Note that the variance here is actually less than for stationary white interfering noise shown in Figure II.

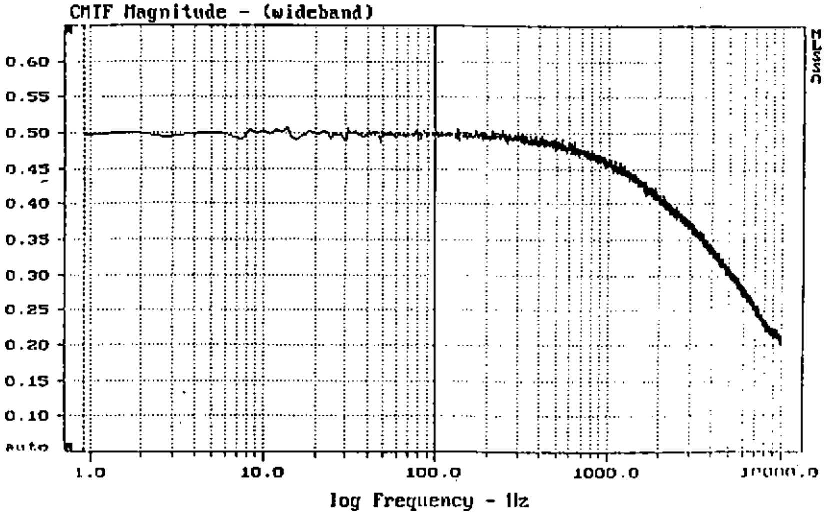
VIII. MLS-STI analysis of a 65535-point PIR measurement of a 12 kHz lowpass system with a 0dB SNR and stationary white interfering noise. The top row labelled "m-correction" is the auditory masking correction of the STI method as described by Steeneken and Houtgast in [8]. The modified STI value shown in parentheses was calculated from the same octave TI values as the normal STI value but with different octave weighting factors recommended by French and Steinberg.

IX. MLS-PIR measurement of a large but quiet church. Note the extended reverberation and the obvious late reflection at 200 milliseconds.

X. MLS-STI computed from the impulse response of Figure IX. The poor intelligibility of this space is due mostly to reflections (or echoes) and reverberation. The background noise often plays a minor role in such cases.

XI. MLS-STI of a CODEC having a 3400 Hz signal bandwidth but, a resolution of only 2 bits per sample . The less than perfect intelligibility of this CODEC is due entirely to quantization noise since there are no significant reflections or reverberation to hamper intelligibility.

XII. MLS-STI of a CODEC having a 3400 Hz signal bandwidth but higher resolution of 3 bits per sample. Note the improved intelligibility as compared to Figure XI.

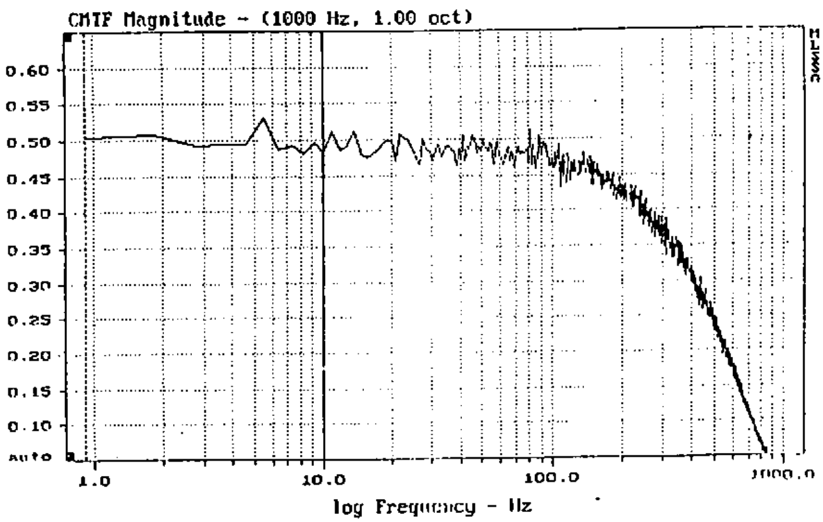


variance: 1.23945e-005, mean: 0.49777

6-27-92 1:01 PM

I.

MLSSA: Frequency Domain

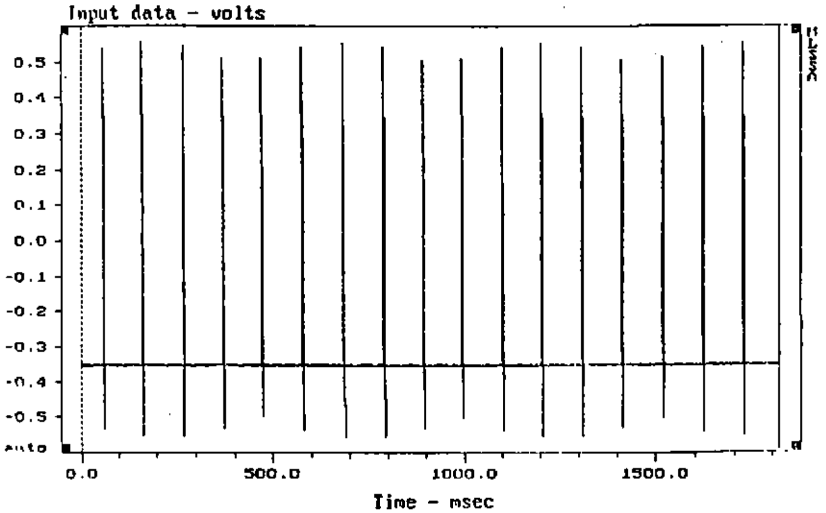


variance: 0.000188167, mean: 0.496412

6-27-92 1:04 PM

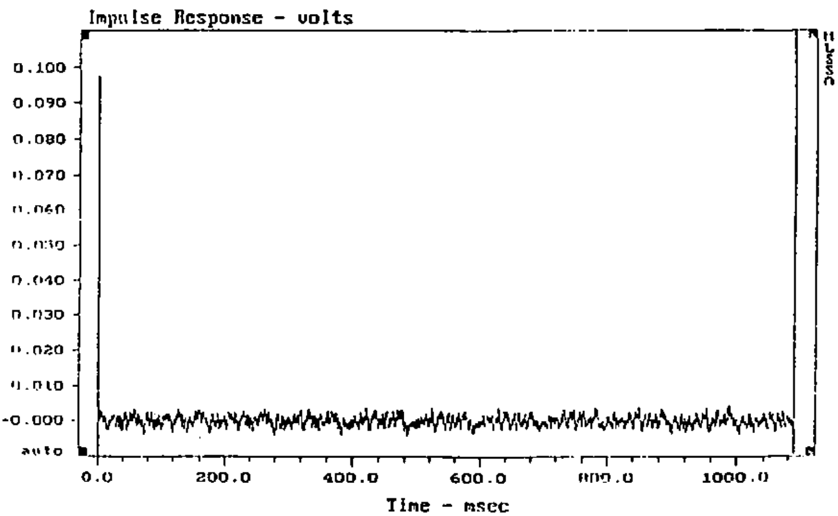
II.

MLSSA: Frequency Domain



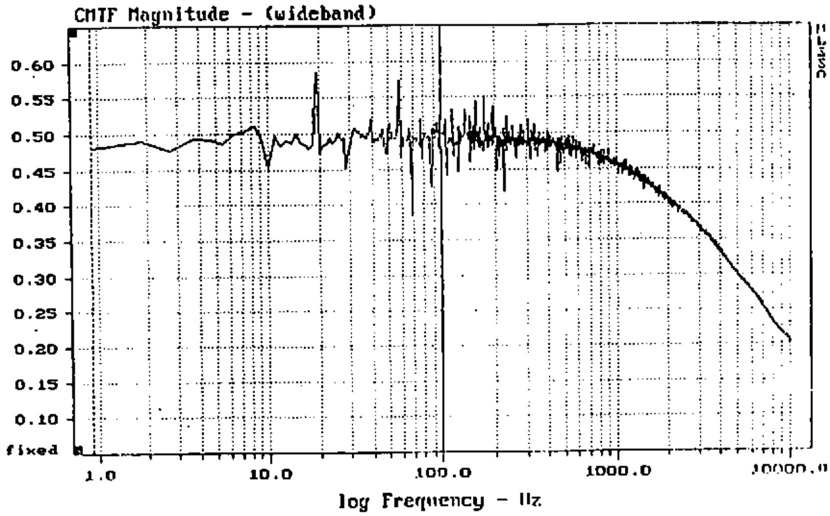
CURSOR: y = -0.35346 x = 1818.5962 (65535)

·27-92 11:28 PM **III.** MLSSA: Time Domain



CURSOR: y = -0.000122554 x = 1089.4695 (32766)

·28-92 2:08 PM **IV.** MLSSA: Time Domain

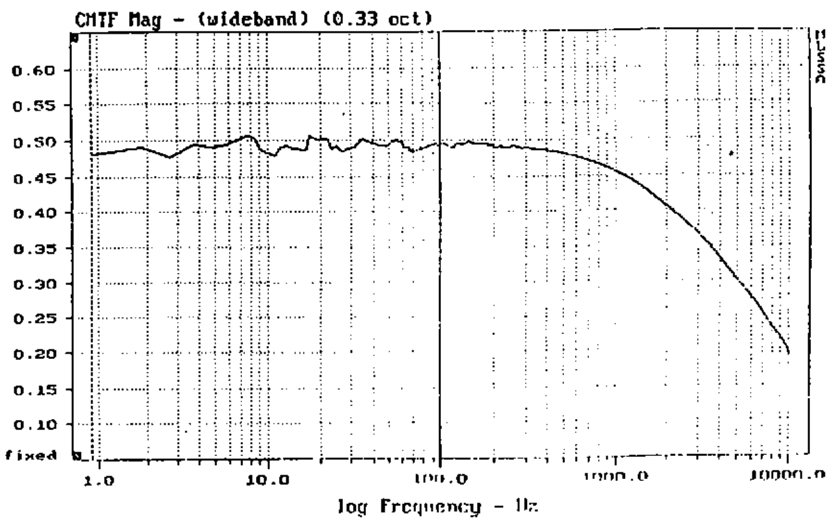


variance: 0.000145696, mean: 0.492003

6-28-92 2:02 PM

V.

MLSSA: Frequency Domain

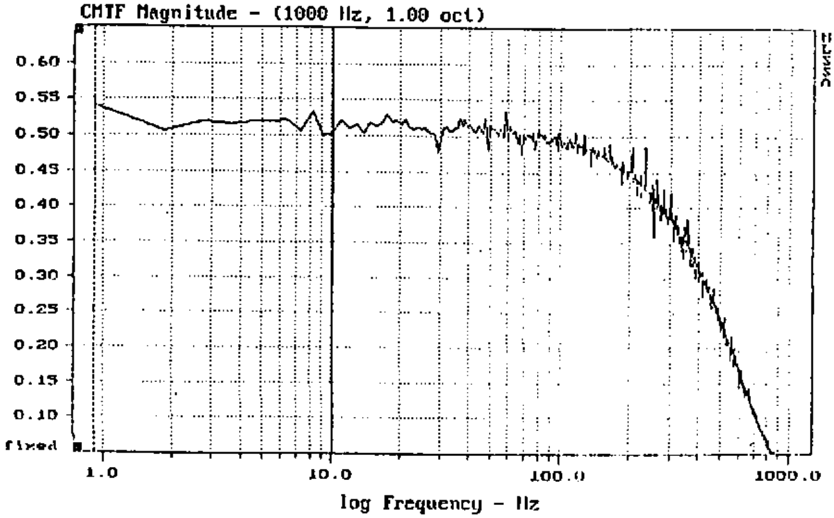


variance: 3.12786e-005, mean: 0.492306

6-28-92 2:04 PM

VI.

MLSSA: Frequency Domain



variance: 0.000142506, mean: 0.516027

6-28-92 2:07 PM MLSSA: Frequency Domain

VII.

Modulation Transfer Function Matrix

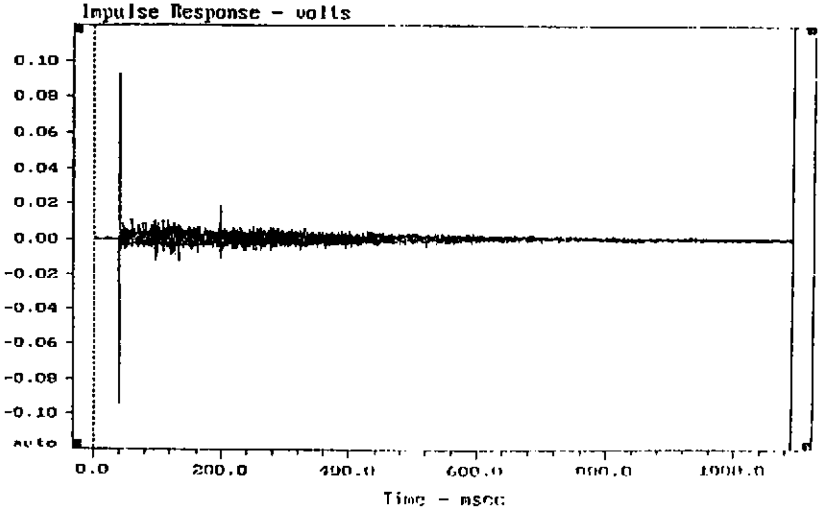
Frequency Hertz	1 125	2 250	3 500	4 1000	5 2000	6 4000	7 8000
m-correction	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.63	0.465	0.522	0.491	0.519	0.505	0.501	0.504
0.80	0.356	0.513	0.509	0.526	0.497	0.507	0.502
1.00	0.356	0.513	0.509	0.526	0.497	0.507	0.502
1.25	0.463	0.520	0.489	0.516	0.504	0.501	0.502
1.60	0.463	0.520	0.489	0.516	0.504	0.501	0.502
2.00	0.465	0.492	0.478	0.515	0.507	0.504	0.504
2.50	0.457	0.520	0.488	0.517	0.503	0.501	0.501
3.15	0.452	0.513	0.486	0.515	0.503	0.500	0.500
4.00	0.446	0.513	0.484	0.515	0.501	0.500	0.502
5.00	0.443	0.512	0.482	0.514	0.503	0.500	0.500
6.30	0.505	0.511	0.499	0.496	0.508	0.512	0.503
8.00	0.423	0.503	0.478	0.512	0.501	0.500	0.501
10.00	0.350	0.520	0.465	0.511	0.504	0.493	0.506
12.50	0.399	0.488	0.468	0.508	0.499	0.499	0.503
octave TI	0.460	0.507	0.492	0.509	0.502	0.501	0.501

STI value= 0.497 (0.502 modified) ALcons= 11.6% Rating= FAIR

File: C:\VIEW\CMTF64K.TIM 3-2-92 2:57 PM MLSSA: STI

VIII.

File: C:\VIEW\CHURCH.TIM 9-11-90 8:59 PM
Impulse Response - volts



Comment: ASA Meeting 9/11/90

6-27-92 11:33 PM

IX.

MLSSA: Time Domain

Modulation Transfer Function Matrix

Frequency Hertz	1 125	2 250	3 500	4 1000	5 2000	6 4000	7 8000
m-correction	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.63	0.628	0.520	0.581	0.575	0.604	0.689	0.833
0.80	0.628	0.520	0.581	0.575	0.604	0.689	0.833
1.00	0.628	0.520	0.581	0.575	0.604	0.689	0.833
1.25	0.628	0.520	0.581	0.575	0.604	0.689	0.833
1.60	0.414	0.259	0.350	0.305	0.383	0.453	0.632
2.00	0.414	0.259	0.350	0.305	0.383	0.453	0.632
2.50	0.262	0.148	0.295	0.211	0.300	0.366	0.519
3.15	0.262	0.148	0.295	0.211	0.300	0.366	0.519
4.00	0.183	0.095	0.239	0.168	0.237	0.359	0.492
5.00	0.170	0.187	0.184	0.150	0.209	0.338	0.487
6.30	0.107	0.113	0.070	0.035	0.191	0.304	0.468
8.00	0.082	0.144	0.142	0.112	0.206	0.319	0.426
10.00	0.152	0.205	0.251	0.092	0.252	0.333	0.471
12.50	0.112	0.101	0.188	0.157	0.305	0.367	0.523
octave TI	0.375	0.342	0.384	0.337	0.415	0.475	0.574

STI value= 0.419 (0.404 modified)   Alcons= 17.6%   Rating= POOR

File: C:\VIEW\CHURCH.TIM 9-11-90 8:59 PM   MLSSA: STI

X.

Modulation Transfer Function Matrix

Frequency Hertz	1 125	2 250	3 500	4 1000	5 2000	6 4000	7 8000
m-correction	1.000	1.000	1.000	1.000	1.000	0.999	0.605
0.63	0.988	0.987	0.977	0.936	0.791	0.552	0.071
0.80	0.982	0.984	0.975	0.933	0.794	0.542	0.067
1.00	0.982	0.984	0.975	0.933	0.794	0.542	0.067
1.25	0.978	0.982	0.974	0.935	0.793	0.553	0.068
1.60	0.978	0.982	0.974	0.935	0.793	0.553	0.068
2.00	0.973	0.979	0.973	0.936	0.798	0.548	0.068
2.50	0.967	0.976	0.972	0.933	0.793	0.545	0.052
3.15	0.958	0.971	0.968	0.930	0.791	0.558	0.058
4.00	0.946	0.967	0.966	0.929	0.787	0.535	0.047
5.00	0.935	0.961	0.962	0.929	0.793	0.536	0.060
6.30	0.920	0.952	0.958	0.925	0.787	0.544	0.072
8.00	0.905	0.946	0.953	0.923	0.786	0.555	0.074
10.00	0.880	0.933	0.947	0.920	0.786	0.543	0.059
12.50	0.854	0.920	0.940	0.918	0.786	0.551	0.057
octave TI	0.926	0.967	0.975	0.874	0.693	0.528	0.109

STI value= 0.704 (0.751 modified) ALcons= 3.8% Rating= GOOD

File: C:\VIEW\BPS2BW34.TIM 5-13-92 8:16 AM HLSSA: STI

XI.

Modulation Transfer Function Matrix

Frequency Hertz	1 125	2 250	3 500	4 1000	5 2000	6 4000	7 8000
m-correction	1.000	1.000	1.000	1.000	[illegible]	0.992	0.739
0.63	0.993	0.995	0.992	0.980	0.931	0.825	0.080
0.80	0.988	0.992	0.991	0.979	0.933	0.819	0.079
1.00	0.988	0.992	0.991	0.979	0.933	0.819	0.079
1.25	0.983	0.990	0.989	0.979	0.931	0.824	0.078
1.60	0.983	0.990	0.989	0.979	0.931	0.824	0.078
2.00	0.978	0.987	0.988	0.978	0.933	0.822	0.073
2.50	0.973	0.985	0.987	0.977	0.933	0.820	0.072
3.15	0.963	0.979	0.984	0.976	0.931	0.821	0.065
4.00	0.952	0.974	0.981	0.974	0.930	0.817	0.059
5.00	0.942	0.969	0.978	0.972	0.928	0.815	0.069
6.30	0.926	0.961	0.974	0.971	0.920	0.814	0.082
8.00	0.911	0.953	0.969	0.968	0.927	0.822	0.060
10.00	0.886	0.940	0.962	0.964	0.925	0.811	0.064
12.50	0.861	0.926	0.955	0.960	0.925	0.804	0.067
octave TI	0.934	0.975	0.994	0.995	[illegible]	0.718	0.133

STI value= 0.794 (0.886 modified) ALcons= 2.3% Rating= EXCELLENT

File: C:\VIEW\BPS3BW34.TIM 5-13-92 8:25 AM HLSSA: STI

XII.