Return Path Linear Distortion and Its Effect on Data Transmissions

Tom Williams
Holtzman Inc.

Abstract

Much has been written on the operational challenges for cable systems caused by upstream noise, especially burst noise, radio frequency ingress, and common path distortion. There is also an upstream problem with linear distortions caused by discrete echoes, amplifier tilt, group delay and micro-reflections. Diplex and lightning filters in amplifiers, missing terminators, damaged components, design problems, and construction mistakes create these linear distortions. The effect of uncorrected linear distortions is to increase the bit error rate in the presence of random noise or other additive impairments. Uncorrected linear distortion also limits an operator's ability to go to higher-order modulation schemes. In the return band a micro-reflection problem is aggravated by the low loss of the coaxial cable. This paper presents a burst reference signal measurement technique for measuring linear distortions on working plant as well as test data from active usptream cable plant. Simulation plots and curves show increases in intersymbol interference from linear distortions.

BACKGROUND

Linear distortions can severely impair the performance of digitally modulated carriers or render them useless. The problem that linear distortion creates is an inter-symbol interference (ISI) between the symbols comprising the data stream. This impairment is also referred to as dispersion [1]. The bit error rate will be increased in the presence of random noise if ISI is small. If the ISI is severe, the data will be useless.

Fortunately, a solution to the linear distortion problem exists: adaptive equalizers that cancel linear distortion. Adaptive equalizers are typically digitally implemented filters that are programmed with the inverse of the channel's frequency response, causing them to "undo" linear distortion. For bursty upstream cable traffic, using adaptive equalizers is complicated by the reality that each subscriber's modem has its own unique frequency response determined by its upstream path characteristics, including the inside wiring of the individual house. Thus the adaptive equalizer's programming solution is unique for every active upstream modem. Multiple solutions are currently being discussed. A first solution is placing the adaptive equalizer in the modem and pre-distorting the signal so it arrives in the headend corrected. This solution is the one that is currently being planned for DOCSIS 1.1. Programming coefficients for the adaptive equalizer must be transmitted over the downstream portion of the cable plant for each active modem in the system. A second solution is to store the programming for the adaptive equalizer in the headend for each active modem. The second method is complicated when the headend controller does not know in advance which modem's transmission will be received next, such as when modems are operating in a contention mode. A third solution is to accompany an upstream burst transmission with a reference signal. The reference signal may be used to remove linear impairments from the upstream transmissions.

Each of these approaches has their own drawbacks and work-arounds. Pre-distorting the signals works well, but as the number of modems attached to a headend controller increases, additional traffic will be generated if there is a sudden change in the echo environment of each modem's signal path. For example, a changing echo on a main line will cause any modems using that same line to re-adapt when their common signal path changes. Likewise, if the adaptive equalizer predistortion coefficients in the modems are not stored in non-volatile memory, and there is widespread power outage, the recovery sign-on storm will create additional traffic and delay. This is especially true if each modem's coefficients are not stored in the headend controller.

Placing the adaptive equalizer at the headend also can work, but this places a computational load on the headend controller. The problem with contention mode can be alleviated by using a robust modulation for contention (e.g. QPSK) and a higher rate modulation such as 16-QAM for reservation packets.

Accompanying the modem's burst transmission with a reference signal is another possible solution. Using a reference signal for just the contention mode packets is another possible variant.

At this time, however, adaptive equalizers are not required by the DOCSIS 1.0 specification and are not widely deployed. The net effect is that the cable operators typically place digitally modulated carriers in the middle portion of the return band where the linearity is best, avoid higher order modulations (such as 16-QAM) in all but the best of circumstances, and attempt to maintain their plant with a flat frequency response.

If a conventional sweep test is performed on an upstream cable plant, the results will give an operator an indication of the flatness of the 5-42 MHz band, but this is only half of the answer. The other half is the phase response of the plant. That is, a frequency response is a set of complex numbers with both magnitude and phase components. The derivative of the phase component with respect to frequency is known as group delay. If the channel has group delay variation, ISI is created. Upstream group delay typically has a "bathtub" shaped response in the 5-42 MHz band with the group delay variation being worse between 5 and 10 MHz and between 35 and 42 MHz. Most of the group delay variation above 35 MHz is caused by the low pass portion of the diplex filter, and most of the group delay variation below 10 MHz is caused by the "lightning filter" which is a 5 MHz high pass filter. Unfortunately, while the amplifier designers have generally done an excellent job of maintaining a flat magnitude frequency response over the entire return band, the group delay response on commercially available amplifiers can cause serious problems on relatively short cascades even using a modulation as robust as QPSK.

Group Delay

Both amplitude non-flatness and non-linear phase can cause ISI. If the signal path can be classified as a minimum phase network, any deviation in the phase linearity will be accompanied by amplitude non-flatness. Echoes (multipath distortion) are examples of responses that are minimum phase. Group delay is a measure of the slope of the phase vs. frequency curve. Group delay variation occurs when the

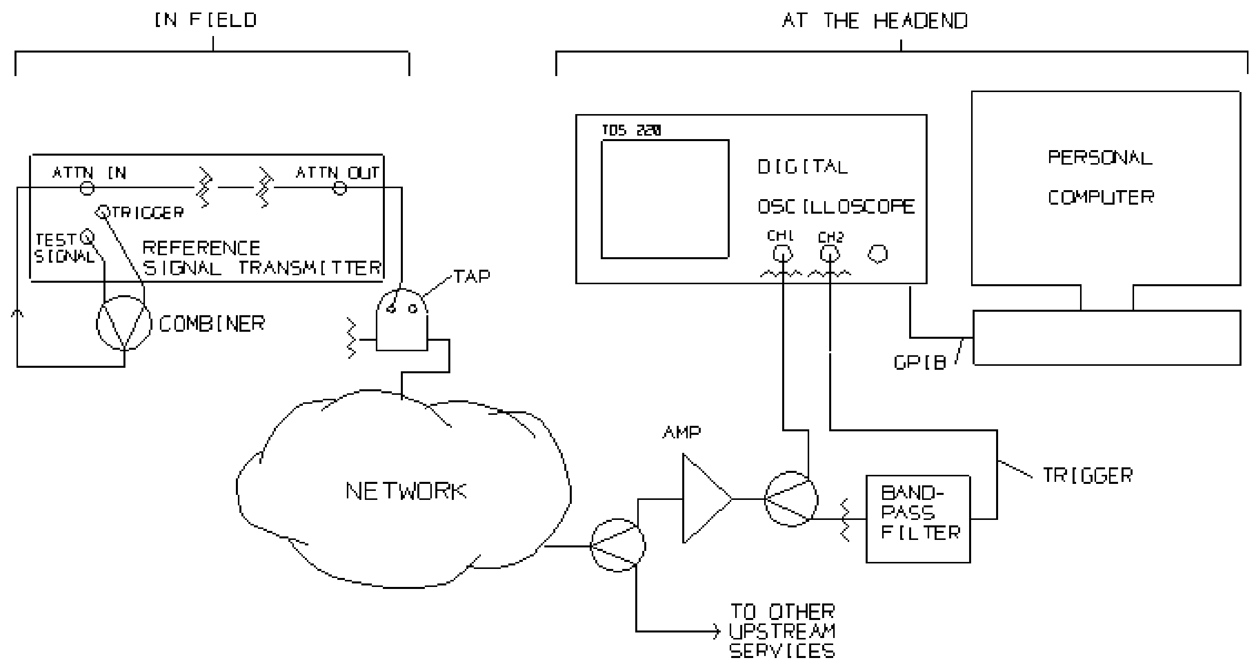

Figure 1 Wiring Diagram for Measuring Upstream Frequency Response

group delay changes with respect to frequency. Thus, if ω is frequency, ϕ is phase, gd is group delay, gdv is group delay variation, ω_1 is the bottom of the channel and ω_2 is the top of the channel:

$$gd(\omega) = -\frac{d\phi}{d\omega} \qquad (1)$$

$$gdv = gd(\omega_2) - gd(\omega_1) \qquad (2)$$

While amplitude non-flatness is an easy-to-understand impairment, group delay variation causes confusion to some cable operators. A tilt in the magnitude portion of the response means that the signals at one frequency are attenuated relative to signals at another frequency. Group delay variation means that signals at one frequency can make it through a network faster than signals at another frequency. For example, the effect of severe group delay variation on a Channel 2 downstream analog TV picture would be a misregistration between the luminance and the chrominance signals because the chrominance subcarrier, which is 3.58 MHz higher in frequency than the luminance signal, arrives at the subscribers's TV sets a fraction of a second before or after the luminance signal. Misregistration means that an actress's red lip color would be shifted to the left or right of the outline of her mouth. Reference [2] provides a definition of group delay.

In fiber optic cable, group delay variation has been renamed chromatic dispersion. The problem with fiber optic cable is that different colors (frequencies or wavelengths) can traverse through the glass fiber at different rates, distorting fast pulses, which contain energy over a wide range of frequencies [3].

Measurement of Complex Frequency Response

There is equipment that can measure complex frequency response. For example, a conventional network analyzer can make this measurement, but the input and output of the network must be in the same location. This works for test beds and temperature chambers, but is not applicable for fieldwork where the ends of the signal path are separated by many kilometers. One way to measure complex frequency response with the cable plant in the field is to use equipment that was originally designed for microwave links. These microwave link analyzers typically employ a sweep carrier that is frequency modulated (FM) as it is swept across the band. Unfortunately, this equipment requires the plant to be taken out of service to make the measurement.

MEASUREMENTS USING BURST REFERENCE SIGNALS

A technique that can be used on plant that is in service is to transmit a 5-42 MHz burst reference signal that is so brief that it does not cause interference. That is, the dwell time of the test signal in a channel is so short that any symbols that are damaged can be easily corrected by the FEC (forward error correction) circuits in the receiver. Equipment that uses this test method is commercialized under the name Cable Scope™. The brief broadband test signal is received on a digital acquisition unit (an A-D converter with memory) and is processed with an unimpaired copy of the signal to give the complex frequency response. If a subscriber service is using an upstream channel the instant that the reference signal is transmitted, the complex frequency response results will be contaminated over only the occupied band of the subscriber service.

The process works as follows. First a technician connects a reference signal transmitter directly into the receiver at the headend (or hub site) and captures an unimpaired reference signal. This unimpaired reference signal is stored directly onto a computer's hard drive. Next, the technician takes the reference signal transmitter to the field and connects its output to a tap port. Figure 1 is a block diagram that illustrates this step. At the tap location the technician injects a burst reference signal into the network on command. The reference signal is preceded by a 5 microsecond sine-wave burst that triggers the capture of the reference signal. The signal propagates back through a coaxial and a fiber portion of the network to the hub site or headend where the trigger burst passes through a bandpass filter and triggers a digital storage oscilloscope (DSO). The DSO captures the impaired reference signal and downloads it to a PC over a GPIB cable. The PC processes the impaired reference signal with a stored version of the unimpaired reference signal and displays the results to the operator in the hub site or headend.

Figure 2 (top) is a time domain trace of the burst trigger signal followed by an unimpaired reference signal that lasts about 18 microseconds. (Figures 2-6 are located at the end of the paper.) The grayed-out parts of the signal are displayed but not used by the process. Figure 2 (bottom) is a spectral (frequency domain) amplitude plot of the same trace that was obtained by performing a FFT on the time domain series. Figure 3 (top) is a time domain trace of a received impaired reference signal after it has passed through a number of amplifiers. Figure 3 (bottom) is a spectral amplitude plot of the received signal. At this point an operator in the hub site or headend instructs the technician in the field to increase or decrease the transmit level and send another burst signal. If the signal was transmitted too strong the network may be briefly clipped, distorting the test results. If the launched signal was too weak, any background noise on the network may degrade the quality of the reference signal, producing a poor quality frequency response result.

The PC processes both signals in the frequency domain to produce the complex frequency response. A complex frequency response plot is illustrated in Figure 4 with a magnitude component, a phase component, and a group delay component. The grayed-out areas below 2 MHz and above 42 MHz are the regions where the reference signal has too-little energy to give reliable results. The magnitude vs. frequency plot is displayed with a linear vertical scale. The phase vs. frequency plot actually “wraps-around” where the top of the plot (+180 deg.) and the bottom of the plot (-180 deg.) are the same phase. The group delay vs. frequency is derived as the slope of the phase vs. frequency plot. There is an adjustable smoothing function on the plots. Smoothing uses a running average of 16, 8 or 1 frequency points. Smoothing reduces the effects of noise at the expense of detail. At the receiver in the hub site an analog-to-digital converter running at 100 megasamples per second captures 2048 points in 20.48 microseconds. Frequency domain samples occur every 48.8 kHz.

The complex frequency response is also processed with an IFFT (inverse fast Fourier transform) to produce an impulse response, which is the top plot in Figure 4. An impulse response is nothing more than a voltage vs. time plot of how the upstream signal path would respond to an impulse (narrow voltage spike) injected at the tap.

While a single strong echo shows up as a ripple in magnitude and group delay plots, it shows up on the impulse responses as smaller pulse, which is delayed relative to the main pulse.

Reference Signals

There are many possible waveforms that can be used as reference signals. The characteristics that make a good reference (or training) signal are flat energy spectrum, a low crest factor (crest factor is the peak to average energy), ease of generation, and a good autocorrelation function. When a signal is convolved with itself, the autocorrelation function is found. A good autocorrelation has a single large impulse at t=0 and low energy at all other time samples.

Examples of well-known reference signals are:
1. A sin(x)/x waveform has a flat energy spectrum, but the crest factor of the waveform is high.
2. Pseudonoise (PN) sequences are a classic favorite. They have a low crest factor and are cheap and easy to generate. Unfortunately, the PN sequences do not produce uniform energy in the frequency domain; the spectral plots normally have frequencies with low energy.
3. The Koo training signal, which is used on line 19 in the vertical interval of some US analog television broadcasts, makes a good reference signal since it has uniform energy in the frequency domain, as well as high energy.
4. Chirps also work well. A chirp is a well-known waveform that has a frequency that increases (or decreases) linearly with time.

Stepped Frequency Reference Signal

Holtzman used a stepped-frequency reference signal to test two of the three systems. Figure 2 illustrates this signal. This stepped-frequency reference signal has characteristics that are similar to a chirp, but its instantaneous frequency increments in steps. This signal holds a frequency for a short period of time before stepping to a new frequency in a continuous-phase manner. Its frequency vs. time plot looks like a staircase. The advantage of this signal is that it can be easily generated by rapidly reprogramming a numerically controlled oscillator (NCO) integrated circuit. NCOs are also known as direct digital synthesis (DDS) chips.

The stepped frequency reference signal generates less interference with traffic relative to a PN sequence reference signal because the time that the stepped frequency signal spends in any frequency band is much less than the 18 microseconds needed to traverse the entire 5-42 MHz return band.

Perhaps the biggest advantage of the stepped-frequency reference signal is that a DDS circuit function is already inside most cable modems. Thus cable modems can, for very little extra cost, act as test signal generators at the end points of the network. This diagnostic capability should be invaluable to cable operators when homeowners start self-installing modems.

One of the three systems tested was characterized with a PN sequence. The stepped frequency signal produces a smoother frequency response plot relative to the PN sequence.

TEST RESULTS

As mentioned previously, three systems were tested. All systems were working two-way systems carrying modem and other traffic. Systems A and B are relatively recent rebuilds with 5-40 MHz return bands. Systems A and B used hybrid amplifiers. System C is an older system with a mix of upstream amplifiers. Some of system C's reverse amplifiers were capable of 30 MHz and some were capable of 40 MHz. The amplifiers in System C were the discrete transistor type. Systems A and B used upstream FP (Fabry Perot) lasers, while system C used DFB (distributed feedback) lasers.

A typical plot from System B is presented in Figure 4 and a typical plot from system A is presented in Figure 5. Figure 6 is a typical plot from System C. Some general observations can be made on the data. The first is that the group delay has a "bathtub" shaped response. This can also be observed as a "bending" or slope change of the phase response. Group delay variation tends to be very bad near the low end of the band and the high end of the band. Another observation is that the magnitude and group delay plots have a rough appearance. This roughness is caused, in the author's opinion, by micro-reflections. Many small echoes created by signals reflecting off taps, amplifiers, power inserters, cable kinks, and drop wiring anomalies can cause micro-reflections. If the equipment is within specifications, no single reflection is large, but all reflections summed together can produce large excursions in magnitude and phase. Micro-reflections were first predicted for the downstream plant, but were not strongly observed outside the home wiring. The probable reason for this is the reflections were damped by the high loss of the cable in the forward frequency bands. In

Node Number	Number of Actives	∆Group Delay 7.5-10 MHz	∆ Group Delay 35-37.5 MHz	∆ Group Delay 37.5-40 MHz
A1	na	-261 ns.	100 ns.	278 ns.
A2	na	-175	92	394
A3	na	-160	119	408
A4	3	-100	145	301
A5	4	-124	67	278
A6	4	-134	85	305

Table 1 Group Delay Variation Summary for System A

Node Number	Number of Actives	∆ Group Delay 7.5-10 MHz	∆ Group Delay 35-37.5 MHz	∆ Group Delay 37.5-40 MHz	∆ Group Delay 40-42 MHz
B1	6	-201 ns.	94 ns.	158 ns.	160 ns.
B2	5	-174	79	115	201
B3	4	-106	48	79	174
B4	3	-71	35	71	124
B5	3	-125	31	84	130
B6	3	-87	64	79	149
B7	3	-150	44	94	180
B8	4	-93	56	84	153
B9	4	-130	66	125	228

Table 2 Group Delay Variation Summary for System B

Node Number	Number of Actives	∆ Group Delay 7.5-10 MHz	∆ Group Delay 25-27.5 MHz	∆ Group Delay 27.5-30 MHz
C1	6	-114 ns.	26 ns.	76 ns.
C2	6	-154	65	175
C3	6	-160	13.5	77
C4	6	-172	41	48
C5	5	-106	23	64
C6	4	-127	33	135
C7	6	-132	52	96
C8	5	-130	22	25

Table 3 Group Delay Variation Summary for System C

the return band cable loss is exceedingly small. The periodic ripple on the magnitude plot of Figure 5 show a distinct echo.

Test Signals and System Clip

The levels at which networks clip can sometimes be determined by transmitting the reference signal at a high level and observing distortion products or compression in the received signal. This technique did not work reliably in system B because the distribution equipment (presumably the fiber optic transmitter) shuts the signal path down quickly when it senses a high level signal. This feature was probably installed to protect the laser diode.

SIMULATION RESULTS

non-flatness similar to the measured values, and demodulate the distorted signal. After the demodulation takes place, the impairment is characterized by MER (modulation error ratio). MER is a signal-to-noise related measure of the desired signal relative to the undesired noise and distortion components. MER is given by:

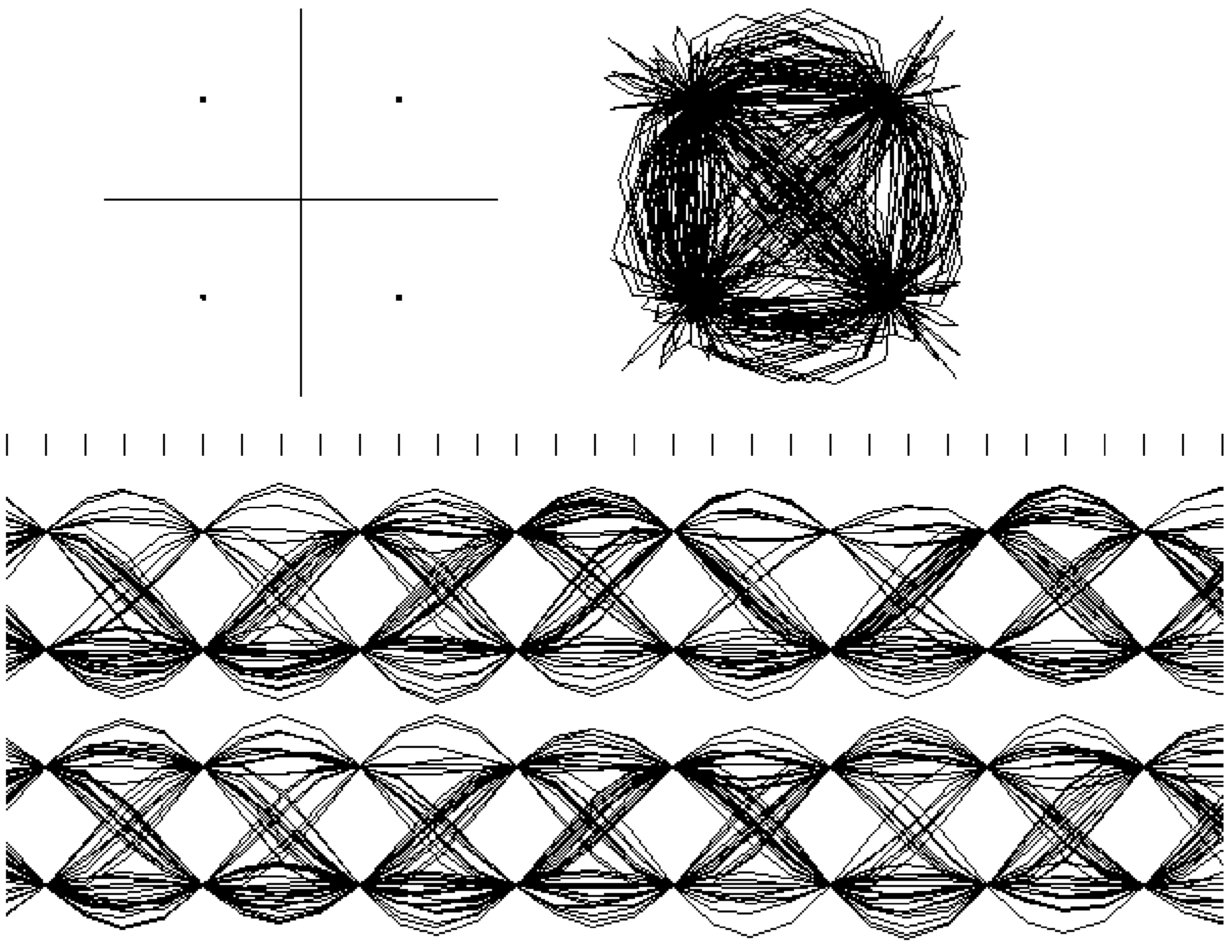

Figure 7 A QPSK Constellation Plot and Eye Diagram of an Unimpaired QPSK Signal

The next step is to take the measured results of amplitude non-flatness and group delay variation and evaluate their effect on signal integrity. To accomplish this task a simulation program was written in the C programming language to create a hypothetical test QPSK packet, distort it with group delay variation and amplitude

$$MER = 10 \cdot \log_{10}\left(\frac{\sum_{j=1}^{N}(I_j^2 + Q_j^2)}{\sum_{j=1}^{N}(\delta I_j^2 + \delta Q_j^2)}\right) dB \qquad (3)$$

where I and Q are ideal coordinates and δI and δQ are the errors in the received data

Modulation Type	QPSK
Packet Duration	179.2 microseconds
Symbol Rate	2.5 M Symbols/sec.
Number of Symbols/packet	448
Baseband Nyquist Frequency	1.25 MHz
Roll-Off Factor	0.25
Occupied Bandwidth	3.125 MHz
FFT Size	2048 points
Number of radians/sec per freq. sample	30.68E3 radians/sec.
Number of Time Samples/Symbol	4

Table 4 Characteristics of the Test Packet

points created by impairments. j is the symbol number.

Since no random noise was added in the simulation, the distortion energy comprises the undesired terms. Table 4 lists the characteristics of the test packet. This packet is similar to the modulation used for DOCSIS at its highest upstream rate.

The program works by creating a QPSK burst in the time domain comprising rectangular pulses and converting the burst into the frequency domain. In the frequency domain the burst is filtered to produce an inverse sin(x)/(x) pulse shape correction followed by raised cosine rolloff filtering using a rolloff factor of 0.25. Keep in mind that a multiplication in the frequency domain is equivalent to a convolution in the time domain.

Figure 7 (bottom) is a resulting set of I and Q eye diagrams showing an unimpaired channel with no group delay variation and a flat magnitude response. Note that the eye diagram is fully open. Figure 7 (upper left) is a constellation plot of four dots showing no constellation point spread due to linear distortion. Figure 7 (upper right) is a vector plot showing the trajectory of the instantaneous magnitude and phase over the duration of the packet.

In the frequency domain the phase is assumed to be distorted by the formula:

$$phase(n) = gdf \cdot n^2 + dly \cdot n + phase_O \tag{4}$$

where phase(n) is the phase angle as a function of frequency index, n. In this example n is equal to 30,680 radians/sec. “gdf” is a group delay factor, which produces a quadratic phase curve vs. frequency, dly is a time delay offset, and $phase_O$ is a static phase offset. The center of the channel is the reference point. The assumption that the group delay variation distortion creates a quadratic phase vs. frequency term will be less valid as the occupied band of the test signal increases.

The variation of group delay is the derivative of (4) with respect to frequency, evaluated over the channel. If n_2 is the top of the channel, and n_1 is the bottom of the channel, and gdf holds constant over the channel:

$$\begin{aligned} gd &= (2 \cdot gdf \cdot n_2 + dly) - (2 \cdot gdf \cdot n_1 + dly) \\ &= 2 \cdot gdf\,(n_2 - n_1) \end{aligned} \tag{5}$$

Figure 8 (bottom) is a resulting set of I and Q eye diagrams showing an impaired channel with a group delay variation of 333 ns. and a flat magnitude response. Note that the eye is partially closed. Figure 8 (upper left) is a constellation plot showing constellation point spread due to group delay variation. Figure 7 (upper right) is a vector plot showing the trajectory of the instantaneous magnitude and phase over the duration of the packet. If this level of distortion is placed on a 16-QAM modulated signal, the eye diagrams will be completely closed.

Another approach to modeling would be to use the measured magnitude and phase data from the plots to predict ISI. Interpolation would be needed to fill-in the vacant frequencies between sample points. This approach was not taken because any noise on the frequency response data would increase the ISI of the resulting demodulated packet.

Linear distortion caused by tilt on the magnitude vs. frequency plots is modeled by:

$$A(n) = 1 + m \cdot n \qquad (5)$$

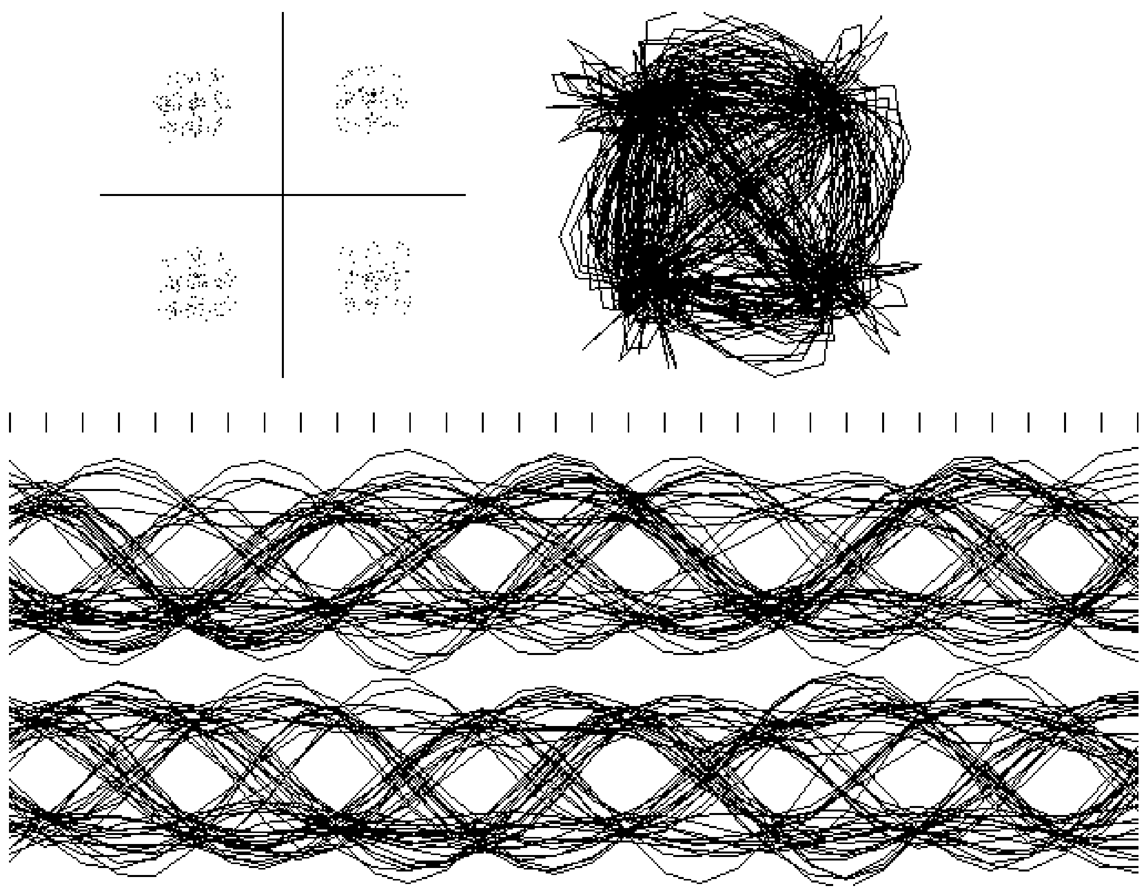

Figure 8 A QPSK Constellation Plot and Eye Diagram of a Signals Distorted with 333 ns. of Group Delay Variation over 2.5 MHz. , MER = 13.82 dB

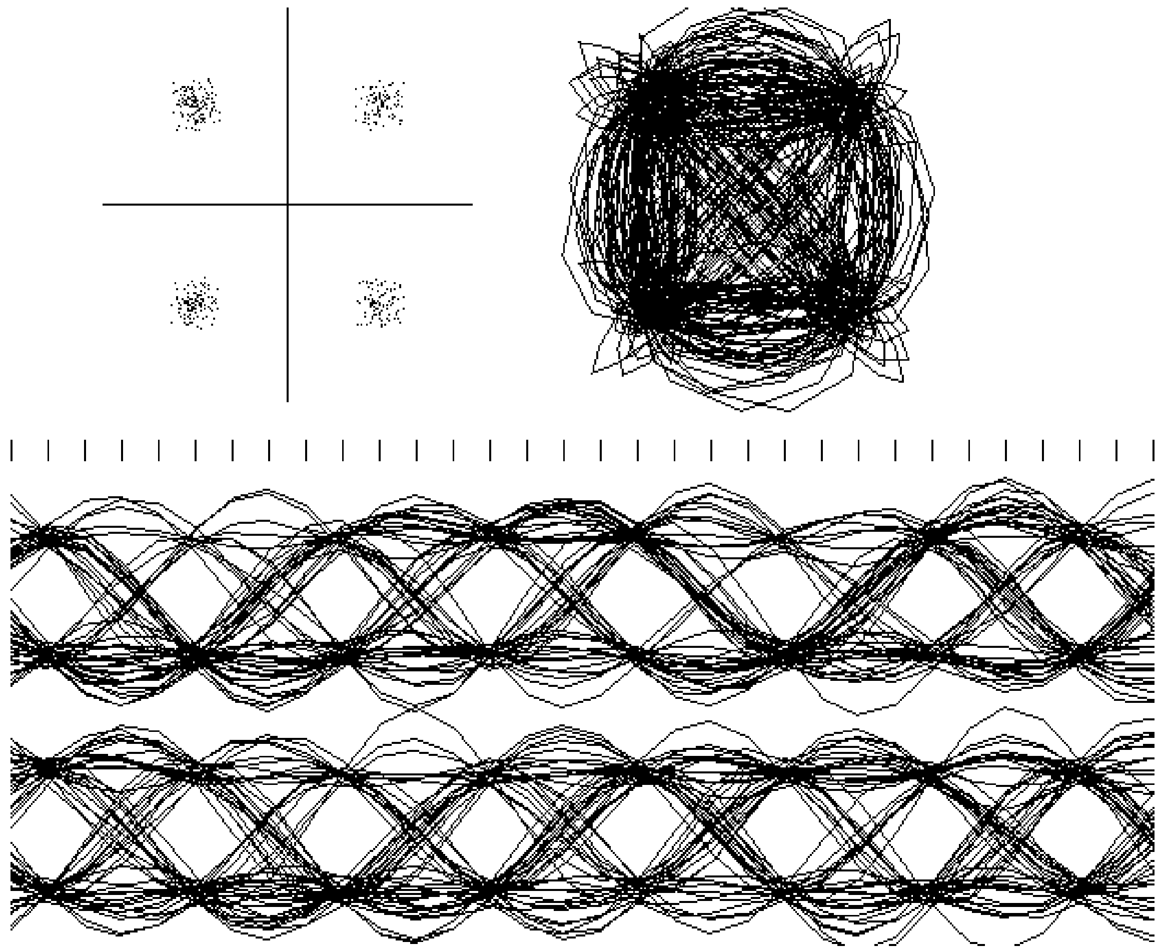

Figure 9 A QPSK Constellation Plot and Eye Diagram of a Signals Distorted with 4.4 dB of Amplitude Tilt over 2.5 MHz. , MER = 18.5 dB

where A(n) is the magnitude of the channel response vs. the frequency index, n. m is the slope factor. It is assumed that the center of the channel is the reference frequency. In decibels the channel tilt is expressed as:

$$tilt = 20 \cdot \log\left(\frac{A(n_2)}{A(n_1)}\right) \quad (6)$$

Figure 9 is a QPSK constellation plot showing the results of a 4.4 dB tilt over 2.5 MHz. with no group delay variation.

Figure 10 is a plot of MER vs. group delay variation showing the increasing corruption of a signal with increasing group delay variation. Figure 11 is a plot of MER vs. magnitude tilt showing that a channel with large amplitude tilt also has large intersymbol interference.

Patents are pending on this technology.

Observations

The data show that un-equalized 16-QAM is not going to work well over some parts of the 5-42 MHz band on some systems because of excessive group delay distortion. In particular, group delay variation below 10 MHz and above 35 MHz is too great in some cases, even for the relatively short cascades of new equipment tested. System A's upstream links have higher group delay variation than system B's links in the 37.5-40 MHz band. Furthermore, QPSK carriers will have degraded performance in the presence of random noise because of linear distortion. In evaluating the effects of amplitude tilt or group delay variation, the occupied bandwidth of the signal should be considered. If you reduce the occupied bandwidth of a signal by one half the level of amplitude tilt or group delay variation may be roughly reduced by one half.

Adaptive equalizers on the upstream path should fix problems with linear distortion

Any proposed adaptive equalization system should be thoroughly field tested prior to deployment to ensure that the system can work with real-world impairments, such as intermittent echoes.

CONCLUSION

This paper has presented field measurements on linear distortion in upstream cable plants. A simulation tool was created to quantify the effects of data measured in the field. The data show that there is a microreflection phenomenon visible on the upstream plant. The simulation shows that in many amplifier cascades the group delay variation at the high and low ends of the return band will stop 16-QAM without adaptive equalization and will degrade the performance of QPSK in the presence of noise.

Acknowledgments

The author wishes to thank the system operators, the technicians, the engineers, and technical reviewers for their support in gathering this data and reviewing this report.

References

[1] Digital Communication by E. Lee and D. Messerschmitt, p. 126, 2nd , Kulwer Academic Publishers

[2] Discrete Time Signal Processing by A. Oppenheim and R. Schafer, pp. 202-205, Prentice Hall

[3] Digital Communication by E. Lee and D. Messerschmitt, pp. 128-135 2nd ed, Kulwer Academic Publishers

Contact Information

Tom Williams
President, Holtzman Inc.
6423 Fairways Dr.
Longmont, CO 80503
303-444-6140
e-mail:tom@holtzmaninc.com

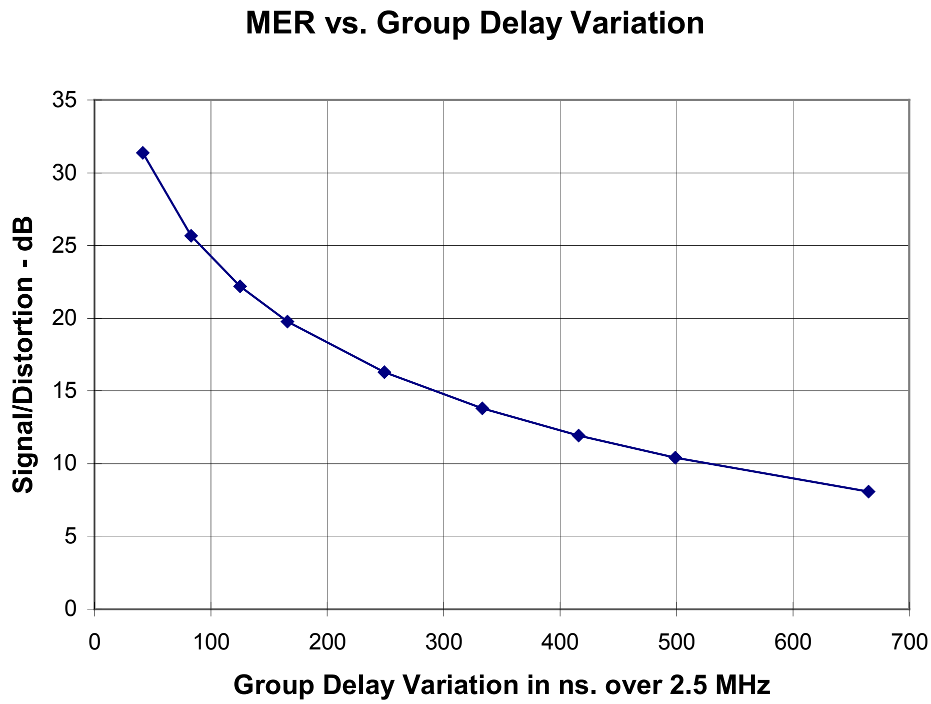

Figure 10 Eye Closure vs. Group Delay Variation over 2.5 MHz.

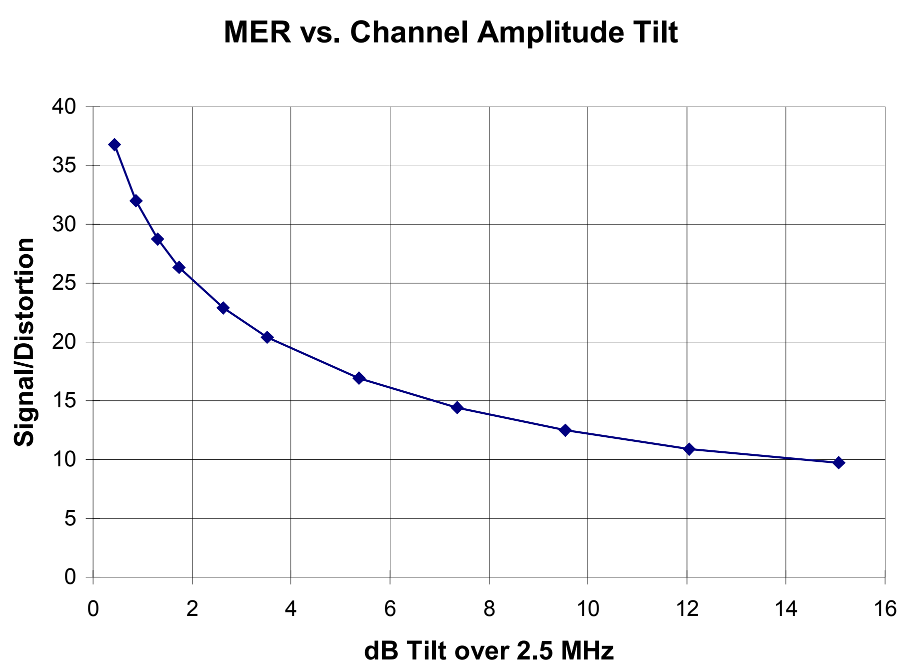

Figure 11 Eye Closure vs. Amplitude Tilt

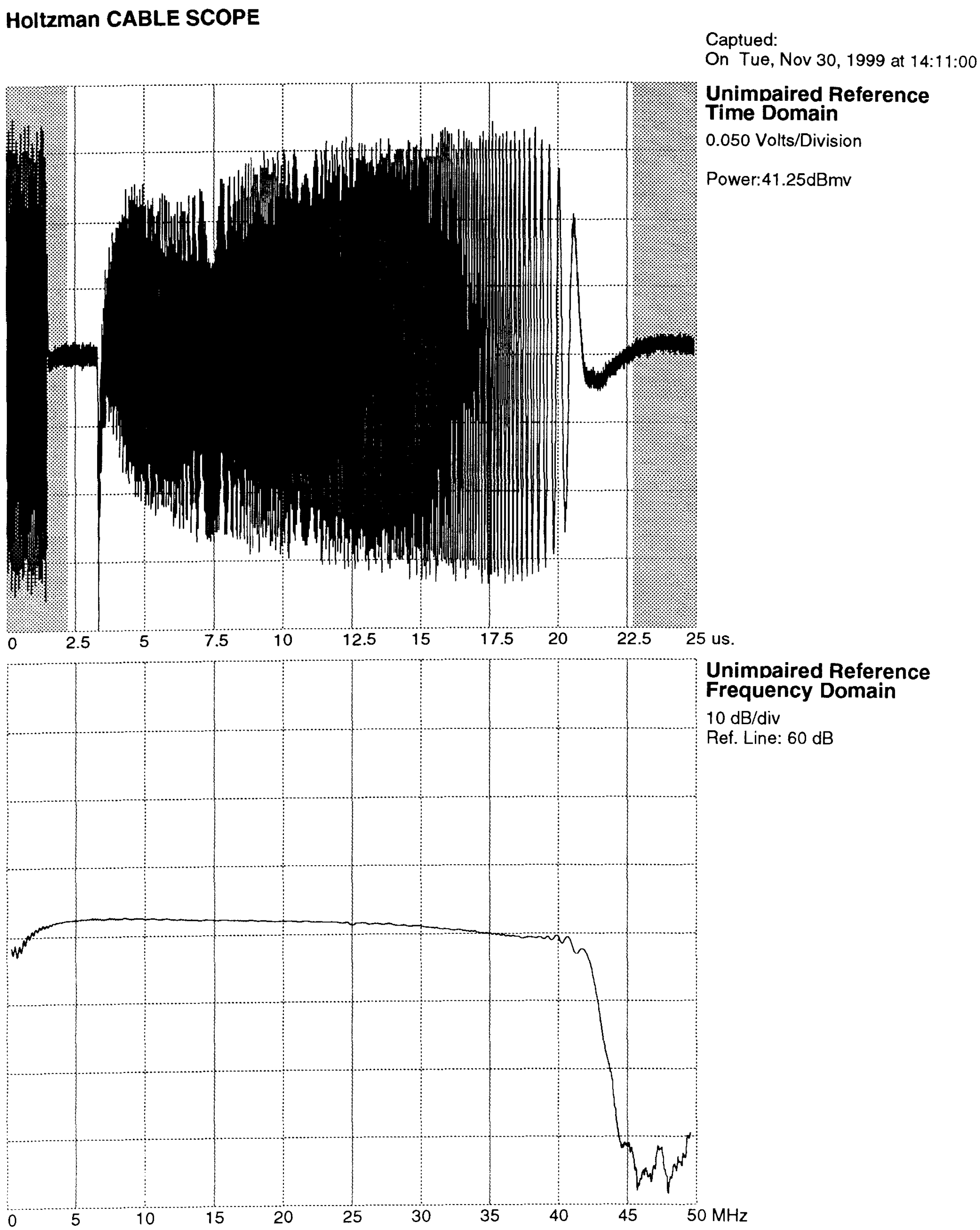

Fig.2 Unimpaired Reference Signal

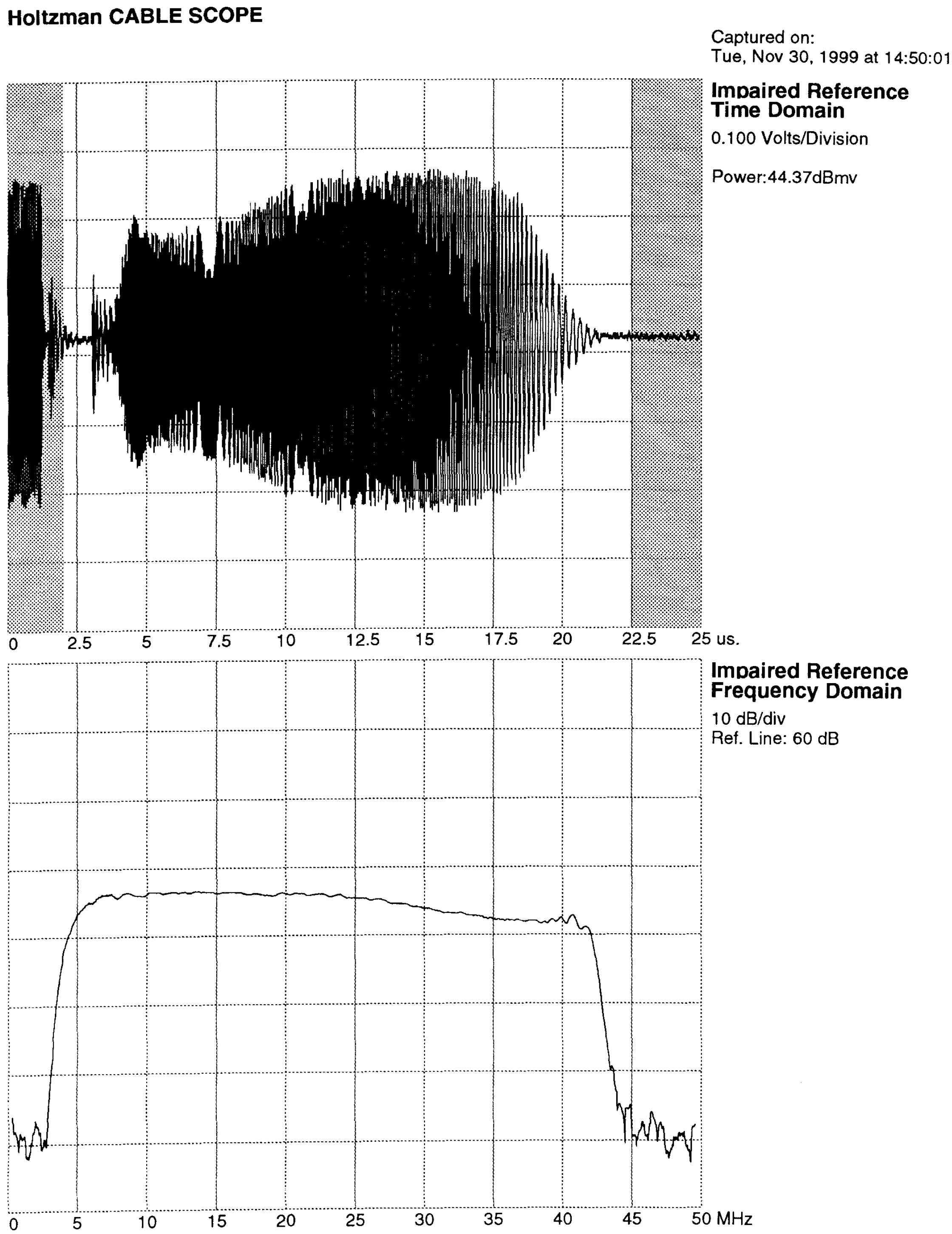

Fig. 3 An Impaired Reference Signal

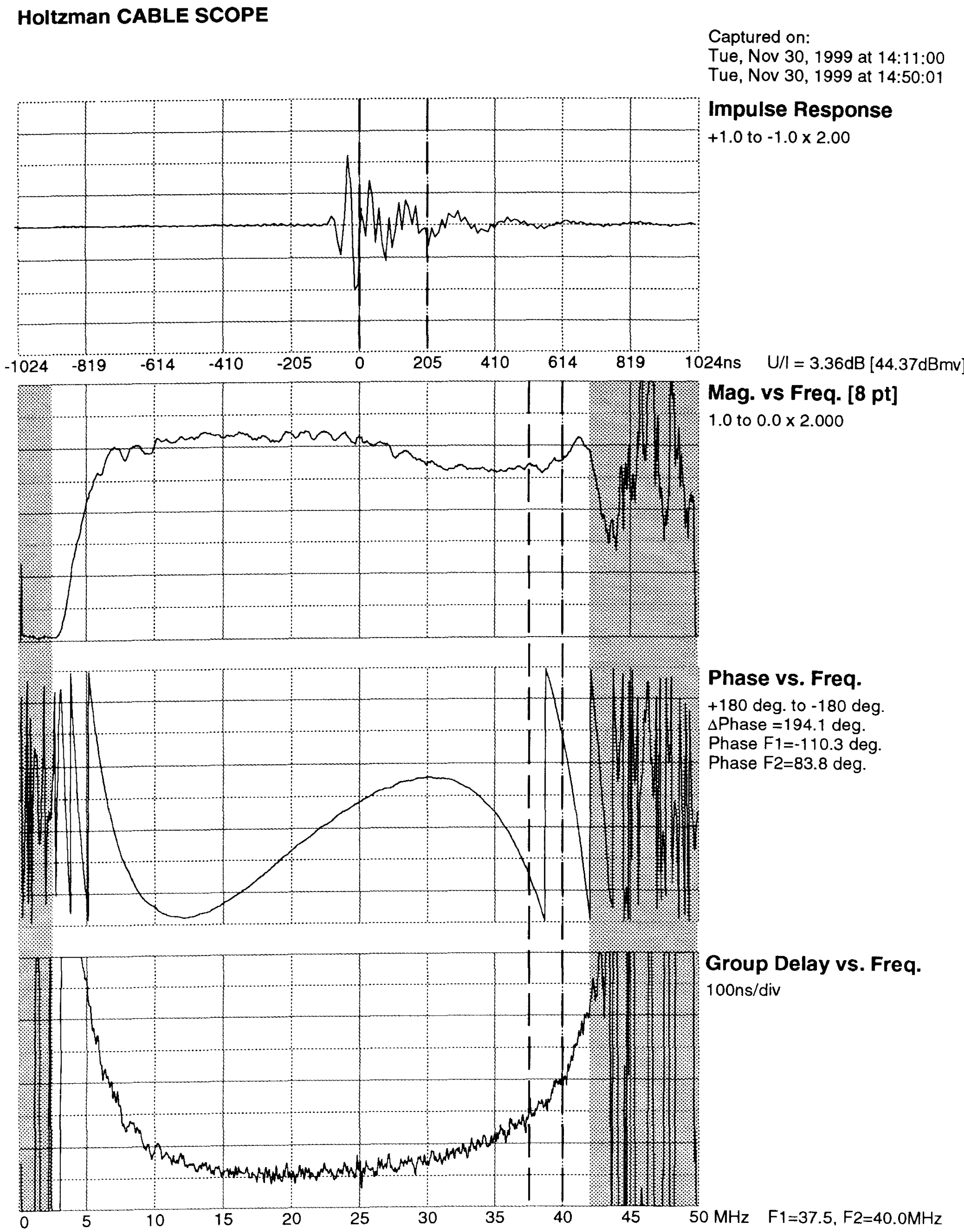

Fig. 4 Frequency Response System B

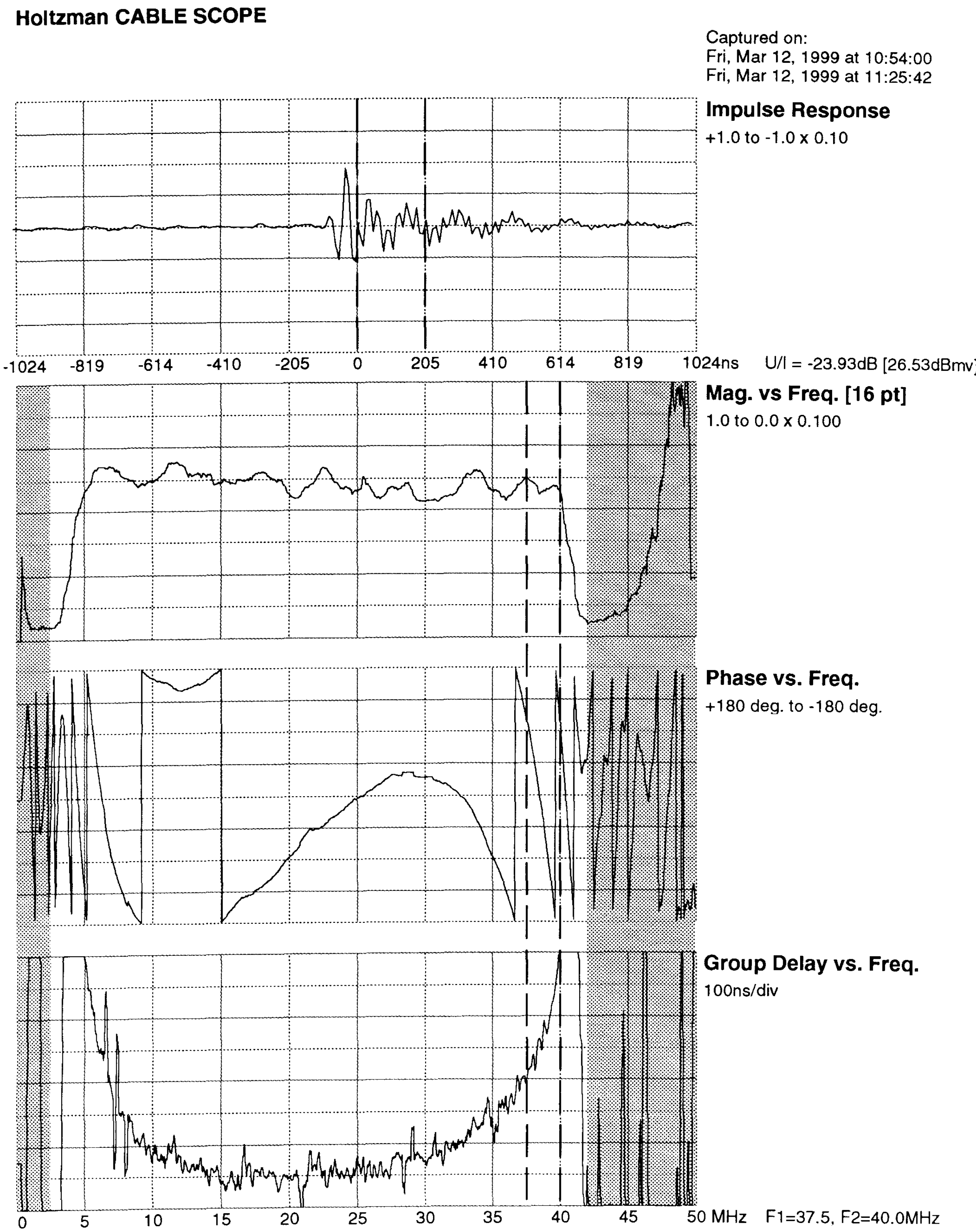

Fig. 5 Frequency Response System A

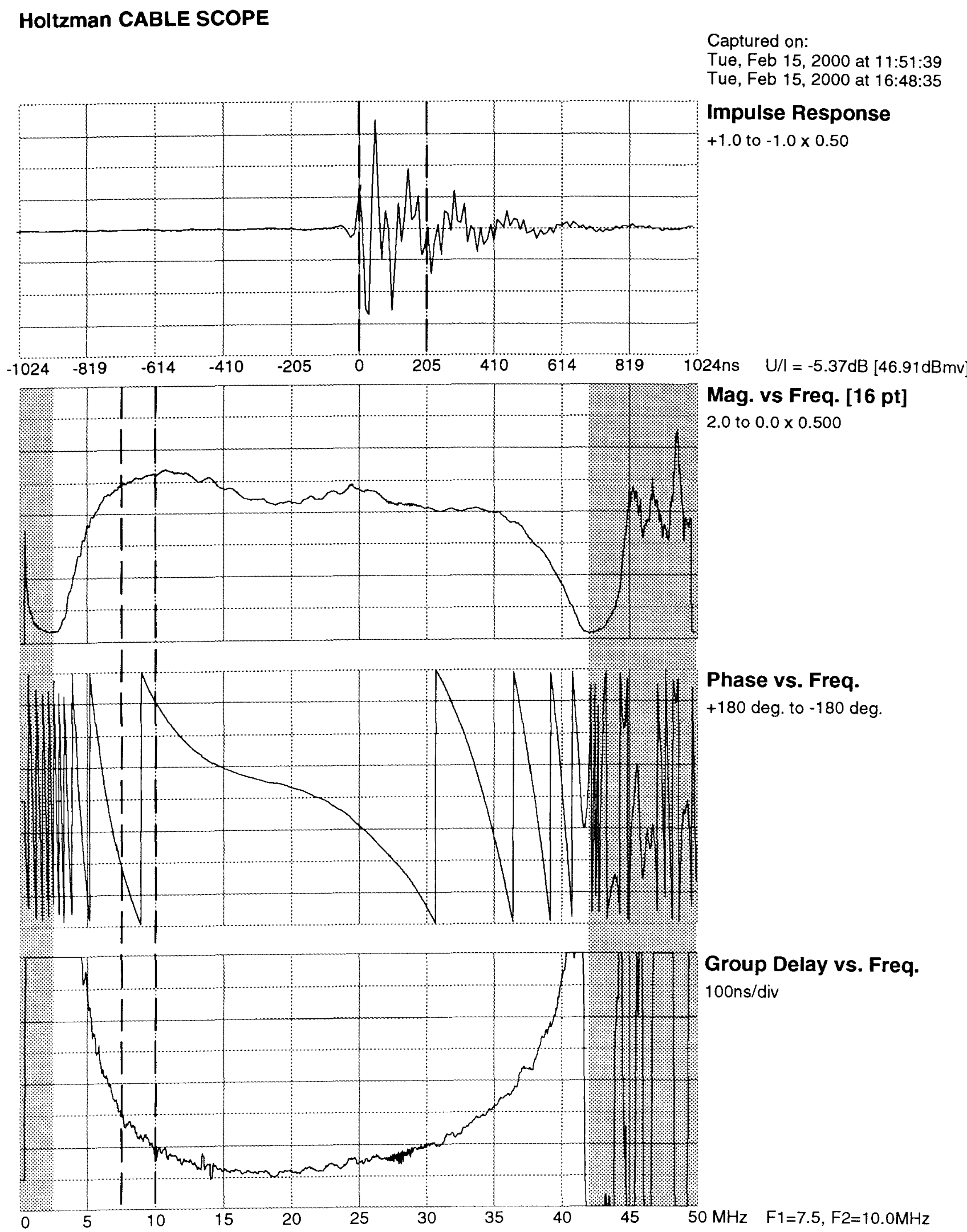

Fig. 6 Frequency Response System C