Achieving Future Space Very Long Baseline Interferometry Gigabits-per-Second Data Rates

J. C. Springett[1]

The next generation of space very long baseline (SVLBI) missions will require a minimum downlink data rate of 1.024 Gb/s, and even more advanced missions propose to utilize 8.196 Gb/s. These high data rates place new demands on downlink spectrum resources and implementation technologies, especially for the space research Ka-band frequency allocation of 37 to 38 GHz which is only 1 GHz wide. To meet this challenge, it becomes necessary to make use of simultaneous left-hand circular and right-hand circular polarized carriers, as well as various forms of bandwidth compressive modulation, along with modulation pulse shaping. Consideration is also given to the National Telecommunications and Information Administration regulations regarding out-of-band spectrum emissions. Other serious problems for future SVLI missions are restrictions on spacecraft effective isotropic radiated power and the cost of attaining the necessary G/T for ground tracking stations. Accordingly, special attention is given to specifying the allowable maximum bit-error rate, along with a statistical downlink design philosophy regarding the three most important temporal parameters of communications slant range, ground antenna elevation, and weather conditions. Additionally, it is proposed that separate data and timing transfer carriers be employed as a means for achieving ground tracking station implementation and operating economies.

I. Introduction

The VLBI Space Observatory Program (VSOP) [or Highly Advanced Laboratory for Communications and Astronomy (HALCA)] space very long baseline interferometry (SVLBI) mission that began in 1997 has a science data rate of 128 Mb/s (megabits per second). A proposed VLBI Space Observatory Program 2 (VSOP2) expects to provide a minimum data rate of 1024 Mb/s or 1.024 Gb/s (gigabits per second). Studies on even more advanced missions project data rates of at least 8.192 Gb/s.

Each of these eight-fold increments places new demands on downlink spectrum resources and implementation technologies. VSOP utilizes 914.2 GHz (Ku-band) and a single quadrature phase-shift-keyed (QPSK) modulated carrier at 14.2 GHz. The VSOP2 baseline plans to transmit 1.024 Gb/s at 37- to

[1] NeoComm Systems, Inc., La Crescenta, California.

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

38-GHz (Ka-band) using unshaped QPSK modulation on simultaneous left circularly polarized (LCP) and right circularly polarized (RCP) carriers. For even higher rates, it will be shown in this article that 4.096 Gb/s on this 1-GHz-wide band can be done provided 16-quadrature amplitude modulation (QAM) bandwidth compressive modulation, along with square-root raised-cosine (SRRC) pulse shaping, is employed.[2] Missions contemplating data rates above 4.096 Gb/s will most likely be forced to higher frequencies (W-band, somewhere in the range of 74 to 84 GHz) and/or the use of greater dimensioned QAM.

In the effort to attain higher data rates, there is the perennial tension between necessary spacecraft effective isotropic radiated power (EIRP) and ground station G/T. Furthermore, these are conditioned on other requirements, including link bit-error rate (BER) and performance allowances (especially for weather). One observation made for VSOP is that there has nearly always been an abundance of link margin, implying a far too conservative design. Therefore, these, along with modulation/detection and spectrum issues, are the subjects addressed by this article.

II. Considerations Regarding Data-Link Maximum BER

The reception of radio telescope signals contends with a number of signal-related degradations or losses. These may be divided into two categories: (1) instrumentation or processing (having to do with implementation choices) and (2) propagation through the atmosphere and ionosphere. Category 1 losses are generally fixed, while category 2 degradations are temporal in nature. These losses will now be further examined to acquaint the reader with their magnitude and to explain a marked difference between ground and space radio telescopes with respect to category 2 influences.

Radio telescope RF signals consist of the sum of weak source and strong receiver-noise components (SNR $<<$ 1). These typically are converted to baseband, where they are Nyquist-rate sampled and quantized to 1 bit (two levels) or 2 bits (four levels) [1, Section 8.3; 2, Chapter 4]. Each quantized sample is said to undergo a pre-quantized SNR loss of η_{QB}, where the subscript value "B" indicates the number of sample bits used. For example, $\eta_{Q1} = 0.638$ (-1.96 dB), and $\eta_{Q2} = 0.881$ (-0.55 dB). When two radio telescope signals observing the same source are identically quantized and cross-correlated, the resulting SNR or sensitivity[3] is likewise degraded by η_{QB}. Relative to non-quantized SNR, 2-bit losses amount to 12 percent, while the 1-bit impairment is a whopping 36 percent. Radio telescope measurements also must deal with several additional processing loss factors (fringe rotation, sideband rejection, discrete delay step, frequency instability) that can collectively further degrade sensitivity by as much as 12 percent [1, Section 9.7].

For Earth radio telescopes, the most significant non-instrumentation losses occur because of troposphere- and ionosphere-induced signal absorption and phase fluctuations [1, Sections 13.1 and 13.2]. In turn, these degrade sensitivity by (1) decreasing SNR due to signal attenuation along with increased system noise temperature, (2) introducing additional random components (rapid phase scintillation), and (3) limiting correlation coherence time (by slow phase meandering). Absorption-related loss factors can range from 0.90 (-0.5 dB) to less than 0.05 (-13 dB) depending on weather conditions and telescope antenna elevation angle. Partial compensations or corrections for these effects may be obtained using calibration techniques, such as water vapor radiometers, phase referencing, and differential frequency

[2] The Advanced Radio Interferometry between Space and Earth (ARISE) mission and spacecraft, as reported in [11], proposed transmitting 8.192 Gb/s using SRRC 4-QAM (see Subsections IV.C and V.A below) on eight carriers per polarization over the band from 37.0 to 39.5 GHz. However, there are two crucial reasons why this is presently impractical. One is that 38.0 to 39.5 GHz is not allocated for space research space-to-Earth transmissions. The second is a workable SRRC implementation for this data rate has not yet been demonstrated.

[3] More fully, the "sensitivity threshold." As used in this article, increased/decreased sensitivity is equivalent to higher/lower SNR.

measurements. However, even with the very best methods, along with relatively benign atmospheric conditions, a net mean sensitivity loss $\eta_{ATMOS-G1}\eta_{ATMOS-G2}$ between telescopes $G1$ and $G2$ of at least a few percent[4] generally will remain. The total ground radio telescope loss becomes $\eta_{QB}\eta_{ATMOS-G1}\eta_{ATMOS-G2}$.

Turning now to the space radio telescope, it is of paramount importance to realize that the spacecraft data differ from the ground radio telescope data sent to the correlator in one crucial sense. As just outlined, a ground observation of a source signal *directly* undergoes induced atmospheric contamination, and it is sometimes difficult to accurately assess the net sensitivity loss due to composite effects. On the other hand, spacecraft-transmitted VLBI data experience an *indirect* atmospheric degradation, manifest only as an increase in ground station received data BER. Within the space radio telescope, the radio source signals are converted, sampled, and quantized (just as is done at the Earth telescopes). In turn, these are transmitted from the spacecraft to a ground tracking station using digital modulation techniques. But due to ground receiver noise, any bit may be received in error with a probability of P_E. It can be shown analytically that the cross-correlation SNR of the spacecraft data (with transmission errors) against ground telescope data (having no quantization bit errors) is degraded by $\eta_{QB}\eta_{ATMOS-G}\,(1-2P_E)$, where the BER loss factor $\eta_{\rm BER} = (1-2P_E)$ is valid for both $B = 1$ and $B = 2$. Generally, it will be found that $\eta_{ATMOS-G} << \eta_{\rm BER}$. Furthermore, the BER and, therefore, $\eta_{\rm BER}$ can be continuously estimated by counting errors in the synchronization word of each data frame,[5] thereby achieving satisfactory temporal calibration. In this respect, the space radio telescope has a distinct advantage against the atmosphere. VSOP is designed for $P_E^{\rm MAX} = 5E-4$, so its factor $\left(1-2P_E^{\rm MAX}\right) = 0.999$ is negligible. This is one reason for the abundance of link margin mentioned in Section I.

Acknowledging this situation gives rise to the question: What is a reasonable criterion for maximum allowable P_E? Consider the following pragmatic proposal. Allow that the worst-case sensitivity loss from bit errors on the spacecraft link be on a par with the least atmospheric sensitivity loss experienced by the most adept ground radio telescope used. By this criterion, the space radio telescope always contributes the least sensitivity degradation to the telescope array. For the sake of the following analysis, this maximum sensitivity loss is taken to be 2 percent, or $\left(1-2P_E^{\rm MAX}\right) = 0.980$ (-0.09 dB). Correspondingly, $P_E^{\rm MAX} = 0.01 = 10^{-2}$. This is a factor of 20 larger than $P_E^{\rm MAX}$ for VSOP and results in a theoretical data SNR (E_b/N_0) savings of 2.6 dB (all other conditions remaining equal). *It is therefore proposed that $P_E^{MAX} = 10^{-2}$ be adopted for all future SVLBI missions.* By way of application, this could reduce spacecraft transmitter power by a factor of 1.82, or spacecraft transmitting antenna diameter by 1.35, or ground station G/T by 1.82, or whatever combination of parameters seems appropriate.

A. Data-Link Design Margin Philosophy

Designing for $P_E^{\rm MAX} = 10^{-2}$ is a strictly worst-case condition. Most of the time the BER will be much lower. There are three temporal conditions that predominantly affect data downlink performance: (1) communications slant range, (2) ground antenna elevation, and (3) atmospheric or weather conditions.

Free space radio field intensity attenuation or loss is a function of the square of the propagation distance between the spacecraft and ground station (slant range). Thus, sufficient capability must be provided to achieve the desired performance at maximum slant range. The Earth's atmosphere attenuates radio waves passing through it and adds to the noise temperature of the ground receiving system. Both of these effects become accentuated as a function of decreasing ground antenna elevation angle. Furthermore, the magnitude of these effects for any given elevation depends on atmospheric conditions—small for clear, dry air and relatively large during storms.

---

[4] A more categorized assessment was requested from the National Radio Astronomy Observatory (NRAO); according to E. Formalont, the NRAO has no more definitive information on the variability of the water vapor content.

[5] This determination is quite accurate for a BER $\geq 10^{-3}$. For a lower BER, the loss may be ignored.

Standard communication-link design practice usually provides sufficient capability to simultaneously accommodate maximum slant range, lowest elevation angle (customarily 5 deg), and very poor weather (heavy overcast with rain), as well as other negative tolerances on certain parameters (transmitter power, circuit losses, etc.). But the three principal extremes are very rarely concurrent, so most of the time there will be a substantial excess of performance (a positive link margin of many dB). In other words, for most ground tracking station passes, P_E will be orders of magnitude less than 10^{-2} for the entire pass.

B. A Practical Weather Model

The key to understanding temporal link performance is the weather model. The following assessment of atmospheric effects at 37.5 GHz has been scaled from 32.0-GHz zenith attenuation and atmosphere noise temperature data [3]. Specific performance is tabulated for Canberra, Australia, and conforms to zone-K on Consultative Committee International Radio (CCIR) rain climate maps (which includes Madrid, Spain, and Usuda, Japan) [4, Section 3.6.1].

Table 1 lists zenith values for the atmospheric attenuation, A, and noise temperature, T, at 37.5 GHz. The left-hand columns indicate cumulative distribution (CD) weather conditions, including qualitative labels. For example, a CD value of 0.900 (90 percent) corresponds to a very cloudy but no rain condition and signifies that, for 90 percent of the time, $A = 0.6119$ dB and $T = 36.602$ K *or lesser* effects will prevail at zenith. The likelihood, expressed as σ (standard deviation) above nominal, also attests to the statistical extremes of the weather. Table 2 provides elevation data for several different weather conditions.

Table 1. Zenith atmosphere attenuation and noise temperature at 37.5 GHz.

Cumulative distribution (exponential)			Frequency = 37.5 GHz	
Weather condition	Likelihood	CD value	A, dB	T, K
Clear and dry	—	0.000	0.1278	7.683
	—	0.100	0.1981	11.881
	—	0.200	0.2446	14.678
Average clear	—	0.250	0.2589	15.555
	—	0.300	0.2732	16.433
	—	0.400	0.3016	18.184
Clear and humid	Nominal weather	0.500	0.3300	19.940
	—	0.600	0.3583	21.701
Thin clouds and very humid	Nominal + 0.3σ	0.700	0.3988	24.170
	—	0.800	0.4638	28.055
Very cloudy and no rain	Nominal + 1.4σ	0.900	0.6119	36.602
Very cloudy and heavy mist	Nominal + 2.0σ	0.950	0.8762	51.019
1.0 mm/h rain	Nominal + 2.7σ	0.980	1.3331	73.932
2.0 mm/h rain	Nominal + 3.0σ	0.990	1.6830	89.905
4.5 mm/h rain	—	0.995	2.3726	117.825
8.5 mm/h rain	Nominal + 4.6σ	0.998	4.7042	185.195

Table 2. Atmosphere attenuation and noise temperature versus antenna elevation.

Elevation, deg	Heavy mist		1.0 mm/h rain		2.0 mm/h rain	
	A95%,dB	T95%, K	A98%,dB	T98%, K	A99%,dB	T99%, K
90	0.8762	51.019	1.3310	73.870	1.6830	89.944
85	0.8795	51.195	1.3361	74.111	1.6894	90.225
80	0.8897	51.728	1.3515	74.841	1.7089	91.076
75	0.9071	52.637	1.3780	76.085	1.7424	92.524
70	0.9324	53.955	1.4164	77.882	1.7910	94.612
65	0.9668	55.729	1.4686	80.294	1.8570	97.406
60	1.0117	58.032	1.5369	83.408	1.9433	101.001
55	1.0696	60.961	1.6249	87.346	2.0545	105.525
50	1.1438	64.657	1.7375	92.275	2.1970	111.154
45	1.2391	69.316	1.8823	98.427	2.3801	118.124
40	1.3631	75.225	2.0707	106.127	2.6183	126.760
35	1.5276	82.807	2.3205	115.839	2.9342	137.510
30	1.7524	92.716	2.6620	128.241	3.3660	150.992
25	2.0732	106.001	3.1494	144.336	3.9823	168.056
20	2.5618	124.434	3.8916	165.623	4.9207	189.804
15	3.3853	151.175	5.1426	194.212	6.5025	217.329
10	5.0457	191.868	7.6649	231.939	**9.6919**	**249.915**
5	**10.0530**	**251.664**	15.2715	271.537	19.3100	276.688

Ground antenna elevation modeling in Table 2 is calculated from Table 1 using the following equations:

$$A\mathrm{CD}\%\,(el) = \frac{A\mathrm{CD}\%\,(\mathrm{zenith})}{\sin\,[el]} \tag{1}$$

$$T\mathrm{CD}\%(el) = [265 + 15 \times \mathrm{CD}] \times \left[1 - 10^{-[A\mathrm{CD}\%(el)/10]}\right] \tag{2}$$

Telecommunication design allowance for weather should jointly consider the CD and fraction of tracking time the ground antenna spends at some minimum operational elevation angle. Experience attests that most antenna sites are able to track the satellite virtually all of the time (subject to local horizon profiles and obstacles) for elevation angles above 10 deg. However, for VSOP, tracking down to an elevation of 5 deg has been routine.[6]

A design control weather allowance for a CD of 99 percent at 10-deg elevation is a reasonable engineering choice and, as shown in Table 2, results in A = 9.7 dB and T = 250 K (bold-face entries in the 99 percent columns). For 5-deg elevation, the same order of degradation (bold-face entries in the 95 percent columns) will be experienced 5 percent of the time. If the weather CD reaches 99.8 percent, elevation must be raised to 28 deg in order to guarantee link performance. Of course, when the slant range is materially less than maximum, the extra signal level thereby obtained will partly assuage weather effects.

[6] For the VSOP mission, the DSN stations tracked below 10 deg only 11 percent of the total tracking time. But for 2.3 percent of all passes, most of the entire track (between 110 and 290 minutes in duration) fell between 5-deg and 10-deg elevation.

The specific criterion, then, is to provide 10 dB of weather margin at an antenna elevation of 10 deg, which guarantees a link P_E of 10^{-2} or less 99 percent of the time at maximum slant range.

C. Example of Statistical Link Design

Table 3 is a design control table for VSOP2. This mission has a 1.024-Gb/s data rate transmitted in two orthogonal polarizations of 512 Mb/s each (see Section III). Ground station parameters are for a minimum capability that might be obtained through modifying an existing 13.7-m radio astronomy antenna.[7] Spacecraft antenna and microwave performance is specified by Japan's Institute of Space and Astronautical Science (ISAS).

Table 3. VSOP2 link budget for 512 Mb/s VLBI data transmission.

VSOP2 Ka-band downlink per polarization	Nominal	Deviation ($\Delta = 2\sigma$)
Ground station—13.7-m antenna		
VLBI data maximum BER	10^{-2}	10^{-2}
Ideal E_b/N_0 (QPSK with differential encoding)	5.2 dB	5.2 dB
Receiving and detection losses	1.5 dB-Hz	2.2 dB-Hz
Required E_b/N_0	6.7 dB	7.4 dB
Bit rate (512 Mb/s)	87.1 dB-b/s	87.1 dB-b/s
Required P_r/N_0	93.8 dB-Hz	94.5 dB-Hz
50 percent efficient 13.7-m antenna, uncooled LNA		
Ground station G/T_{sys}, 10-deg elevation angle	48.4 dB/K	44.2 dB/K
Ground pointing loss	0.3 dB	1.0 dB
Ground power received (P_r)	−182.9 dBW	−177.3 dBW
Propagation medium—10-deg elevation angle	Clear, humid	2 mm/h rain
Atmospheric loss	1.9 dB	7.3 dB
Propagation loss (max range = 36×10^3 km, 38.0 GHz)	215.1 dB	215.1 dB
Polarization mismatch loss	0.2 dB	0.8 dB
Spacecraft		
Required spacecraft EIRP	34.3 dBW	45.9 dBW
Provided spacecraft EIRP	45.1 dBW	41.7 dBW
Spacecraft pointing loss	0.3 dB	0.8 dB
Spacecraft antenna gain (0.6-m dish)	45.5 dBi	44.5 dBi
Transmission circuit loss	5.5 dB	6.5 dB
Zero-margin spacecraft transmitter power	24.6 dBm	38.7 dBm
Provided spacecraft transmitter power	3.5 W	2.8 W
Provided spacecraft transmitter power	35.4 dBm	34.5 dBm
Link direct margin (based on sum of column parameters)	10.8 dB	−4.2 dB
Joint-parameters standard deviation	—	3.6 dB
Nominal link direct margin—3* (joint standard deviation)	—	0.0 dB

[7] Several have been retired and are available.

Most parameters undergo some sort of change as a function of time. A variation may be more or less random (e.g., the weather) or a degradation over time (e.g., spacecraft transmitter power). Accordingly, each parameter is specified in terms of a nominal value and a deviation value. Each deviation, estimated from past experience as well as engineering judgement, is the nominal $- 2\sigma$ value (with a 5 percent possibility of being less). The variations on all parameters are then combined to calculate the standard deviation σ (the joint-parameters standard deviation near the bottom of Table 3).

The objective of the link design is to determine just how much spacecraft transmitter power is needed so that the nominal-link direct margin, minus three times the joint-parameters standard deviation, results in a 0-dB margin. This statistical methodology reflects the fact that temporal variation of the parameters does not result in simultaneous worst-case extremes. Accordingly, the likelihood at 10-deg elevation of being above the 0-dB margin (where $P_E \leq 10^{-2}$) is better than 99 percent.[8]

III. Use of Orthogonal Polarization and Associated Spectra

Transmission of 1.024-Gb/s (1024-Mb/s) data using rudimentary QPSK modulation on the 37- to 38-GHz band requires that the 1.024-Gb/s data stream be split (if not already so) and transferred using two 512-Mb/s QPSK carriers radiated in orthogonal circular polarizations.

In addition to communicating VLBI data, spacecraft telemetry critical to the science aspect of the mission must also be delivered to the ground. For VSOP, this was accomplished by encoding the telemetry within the frame synchronization headers of the VLBI data frames.

Another function performed is the transmission of a phase-stable frequency reference from the ground tracking station to the spacecraft. Known as phase transfer, the integrity of this process requires that the uplink frequency be transponded and sent back at a different frequency to the ground tracking station.

For VSOP, one downlink carrier frequency was used for all functions. But the fact that the VLBI data, spacecraft telemetry, and phase transfer were all combined into a single signal resulted in some undue complications that future ground tracking stations should avoid. As mentioned in the previous section, a modified 13.7-m radio astronomy (or other suitable) antenna can serve for VLBI data reception. During ongoing studies of future SVLBI missions, it also has been recognized that the prerequisites for the phase transfer and telemetry data functions are, for all practical purposes, independent of VLBI data transmission requirements. For one thing, the phase transfer and telemetry data functions require far less downlink EIRP. Also, phase transfer and telemetry are inherently narrowband, while VLBI data are very wideband. Additionally, it seems certain that the cost of retrofitting existing antennas (such as the 13.7-m antenna) to perform phase transfer will be unduly high. For these reasons, the use of a functionally separate ground station for phase transfer and telemetry is favored.

Partitioning the data reception and phase transfer services requires individual downlink carriers, which may be expressed mathematically in the forms

$$s_{\mathrm{LCP}}(t) = \sqrt{2P_{\mathrm{data}}}\left[d_{IL}(t)\cos\left(\omega_{dL}t\right) + d_{QL}(t)\sin\left(\omega_{dL}t\right)\right] + \sqrt{2P_{\mathrm{pilot}}}d_{tlm}(t)\sin\left(\omega_{pL}t\right) \tag{3}$$

$$s_{\mathrm{RCP}}(t) = \sqrt{2P_{\mathrm{data}}}\left[d_{IR}(t)\cos\left(\omega_{dR}t\right) + d_{QR}(t)\sin\left(\omega_{dR}t\right)\right] + \sqrt{2P_{\mathrm{pilot}}}d_{tlm}(t)\sin\left(\omega_{pR}t\right) \tag{4}$$

where subscripts containing L and R respectively designate LCP and RCP components. The VLBI data modulations d_{IL}, d_{QL}, d_{IR}, andd_{QR} represent essentially independent ± 1 amplitude symbol streams

[8] Perhaps even this represents over-design.

(256 Ms/s). Signal parameters, using VSOP2 link budget values (Table 3), a phase transfer tracking SNR of 30 dB in 1 kHz, and a telemetry bit rate of 8 kb/s are

$$P_{\text{data}} = 3.5 \text{ W}$$
$$P_{\text{pilot}} = 80 \text{ mW}$$
$$f_{pL} = 37.116 \text{ GHz}$$
$$f_{dL} = 37.372 \text{ GHz}$$
$$f_{dR} = 37.628 \text{ GHz}$$
$$f_{pR} = 37.884 \text{ GHz}$$

Figure 1 shows the RF spectra for both equations. The signal equations and associated spectra are ideal. Data modulation employs rectangular QPSK, and telemetry modulation is rectangular binary phase-shift keying (BPSK). Carrier static (offset) phase relationships between the various components are arbitrary (not systematized by design or implementations). The carriers placed at the data spectra nulls will be referred to as "pilots." Offsetting the carrier frequencies by the symbol clock of 256 MHz (one-half the 512-Mb/s bit clock) minimizes cross-polarization interference.

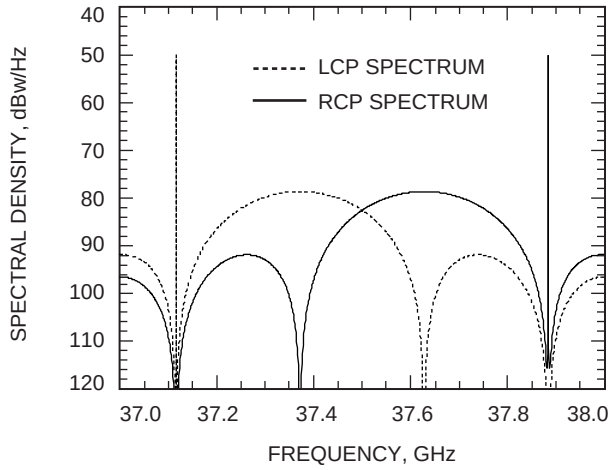

Fig. 1. VSOP2 LCP and RCP RF spectra.

A. Meeting Spectrum Regulation Requirements

In the United States, the National Telecommunications and Information Administration (NTIA) establishes regulations regarding emission spectrum and, in particular, the determination of necessary bandwidth. To meet such regulations, some form of spacecraft transmission filtering must be employed to minimize out-of-band (<37- and >38-GHz) spectral levels.

NTIA guidelines tend to be general, with engineering formulas addressing only the more common types of modulations. There is a necessary bandwidth formula for QPSK, but not one that considers the placement of pilot carriers in the spectral nulls.

The *NTIA Red Book* [5] in part provides the following.

1. Terminology: Desired Relationship of Occupied Bandwidth to Necessary Bandwidth. The emission designator(s) associated in the authorization for any particular frequency assignment specifies the value of the necessary bandwidth of emission for the particular type(s) of transmission permitted.

The values of necessary bandwidth are generally idealized. All reasonable effort shall be made in equipment design and operation by Government agencies to maintain the occupied bandwidth of the emission of any authorized transmission as close to the necessary bandwidth as is reasonably practicable.

2. Unwanted Emission Standards for Earth and Space Stations Operating at 960 MHz and Above. For all systems operating in this frequency range, the emissions radiated outside of the necessary bandwidth shall roll off at a rate equal to or greater than 40 dB/decade (12 dB/octave) from the attenuation level at the limit of necessary bandwidth. The emissions power shall roll off to a level of at least 60 dB below the transmission's maximum peak spectral power density (SPD) or less. (Figure 2 depicts these requirements.) The requirements in this standard specify the upper bounds on unwanted emissions from space and Earth stations associated with the space services. This standard cannot be used alone for planning and evaluation purposes; it is emphasized that the modulation type, emission spectrum, power output, frequency tolerance, and maximum expected Doppler shift should be considered.

The necessary bandwidth, B_n, is determined by ([5], Annex J):

$$B_n = \frac{2\kappa R_b}{\log_2 (\text{sigstates})} \tag{5}$$

where R_b is the data bit rate, "sigstates" represents the number of signaling states (four for QPSK), and κ is a factor chosen to meet some allowable, but not specifically defined, signal distortion criterion. Since $\log_2 (4) = 2$, then $B_n = \kappa R_b$. If $\kappa = 1$ (taken for the example given in [5]), the necessary bandwidth becomes equal to the data bit rate, which is the spectral null-to-null bandwidth of the QPSK modulation spectrum. But in this case the attenuation level at the limit of necessary bandwidth theoretically is $-\infty$ dB, which presents a problem with regard to the requirements of [5, paragraph 5.6.3], as previously stated. So with this situation, some further qualification is needed.

An important consideration is that the filtering provided to meet NTIA emission standards does not significantly attenuate the pilot carrier levels or distort the telemetry data modulated onto the pilots. This requires that the necessary bandwidth be somewhat wider than the null-to-null bandwidth. One solution is to define the total 37- to 38-GHz band as the necessary bandwidth and let the peaks of the second spectral lobes be the attenuation level at the limit of necessary bandwidth. (These peaks fall just within the band edge limits.) With this interpretation, all spectral levels <37 GHz and >38 GHz can easily be made to meet the NTIA emission standards.

There is no single "best" solution for the bandpass filter (BPF) type required. Nonetheless, as the following example demonstrates, a high-order BPF is not needed. Figure 3 shows the upper sideband

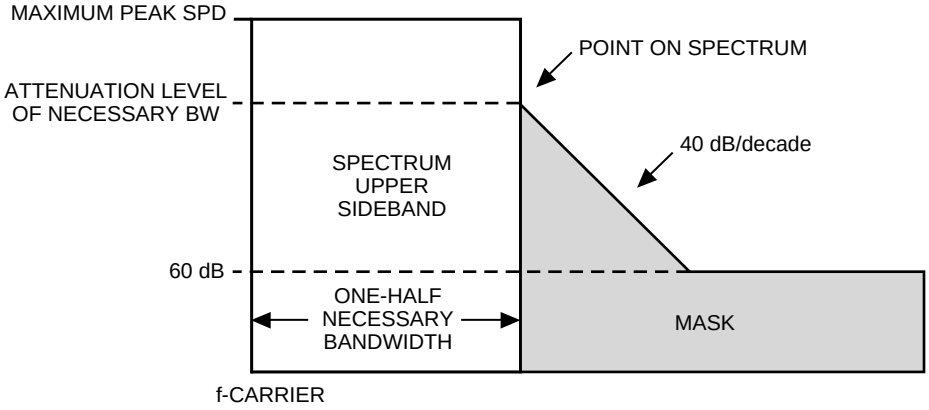

Fig. 2. Unwanted emissions standards (as per Fig. 5.6.1 in [5]).

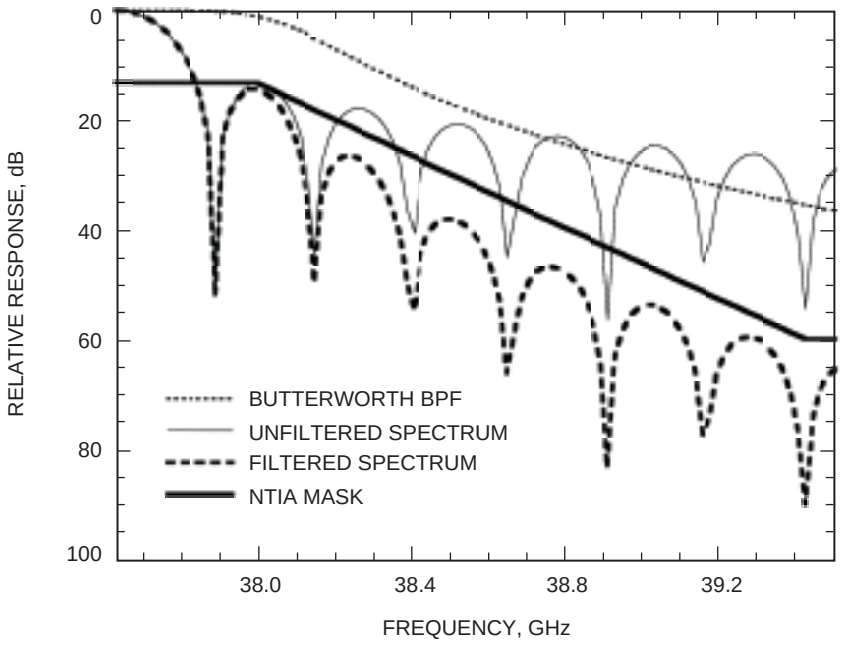

Fig. 3. VSOP2 RCP upper sideband spectrum filtering by three-pole Butterworth BPF.

(USB) result of applying a three-pole Butterworth BPF having a center frequency of 37.628 GHz and a −3 dB bandwidth of 0.9 GHz to the RCP signal. Clearly, the BPF has negligible effect at the first null where the pilot carrier is located. Moreover, above 38 GHz, the attenuation of the unfiltered spectrum (solid thin plot) by the BPF (dashed thin plot) produces an outcome (dashed thick plot) that falls below the NTIA mask.

Filtering invariably introduces some corrupting consequences in the forms of data waveform distortion and cross-channel interference, as well as power reduction due to BPF insertion loss. But these effects should be small, if not inconsequential.

B. Supplying Carrier and Pilot Power

There are two basic ways to combine the data and pilot carriers: (1) after individual carrier power amplification or (2) prior to joint-carrier power amplification. In the first case, two separate power amplifiers are involved, along with their individual efficiencies, and some power insertion loss in the combiner. For the second case, there is a single power amplifier (PA), and the combiner is before its input, where the losses sustained at the milliwatt level may be ignored. However, there will be notable harmonic and intermodulation (IM) loss in the PA.

Studies of both approaches indicate that the joint-carrier power amplification method results in the least power loss when the ratio of output pilot-to-data power is less than 0.1. For example, the VSOP2 data power (see Table 3) is 3.5 W, while the pilot power needed to perform with a 3-m phase transfer antenna[9] is 0.08 W, resulting in a pilot-to-data power ratio of 0.023.

The stipulated quantities are usually the data and pilot powers and the pilot-to-data power ratio. From these, it is essential to determine the PA input pilot-to-data power ratio and the total power (useful and wasted) that will be generated by the PA. To obtain estimates for these values, the PA is modeled as a

[9] This ground station, referred to as the phase transfer antenna (PTA), is a receiving site physically isolated from the antenna used to receive the data. This concept is discussed in [12].

power hard-limiter (because the real PA characteristics are unknown, and this is something easily handled analytically). What this model assumes is that the PA is driven well into the output power saturation region where power supply clipping occurs. The model also assumes that the PA has infinite bandwidth, so that all output harmonic and IM products are produced without attenuation. This latter qualification, of course, is unrealistic since there are always active device and circuit operating frequency limitations, so most harmonics and IM products are not generated at their theoretical power levels.[10] Consequently, the value for the total harmonics and IM product's power obtained from the model represents an upper bound, with any actual PA generating a lesser amount.

Figure 4 plots the two derived quantities as a function of PA output pilot-to-data power ratio. Continuing the example for the ratio of 0.023 calculated above, the PA input pilot-to-data power ratio becomes 0.40, and the total power that will be generated by the PA is 1.26×3.5 W = 4.41 W. Thus, 4.41 W − 3.5 W − 0.08 W = 0.83 W wasted. In other words, 19 percent of the total power is lost to harmonics and IM products. But again, for a real PA, the total power (as well as the wasted power percentage) should be substantially smaller (probably around 10 percent).

IV. Achieving Higher Data Rates

The strategy presented in the previous sections has concentrated on providing 1.024-Gb/s data transfer on the Ka-band 37- to 38-GHz allocation. This is accomplished by generating two 512 Mb/s QPSK carriers and respectively placing them in left-circular and right-circular polarizations.

In order to go to higher bit rates on the same 37- to 38-GHz band, it becomes necessary to make use of higher-order, or bandwidth compressing, modulations. At the same time, it is also very desirable to retain the spectral structure illustrated by Fig. 1 for any data rate, so that switching between bit rates does not require changing components, such as the NTIA BPFs, or the location of the pilots. In order to achieve this, the use of M-ary signal constellations, as well as modulation pulse shaping, are now examined.

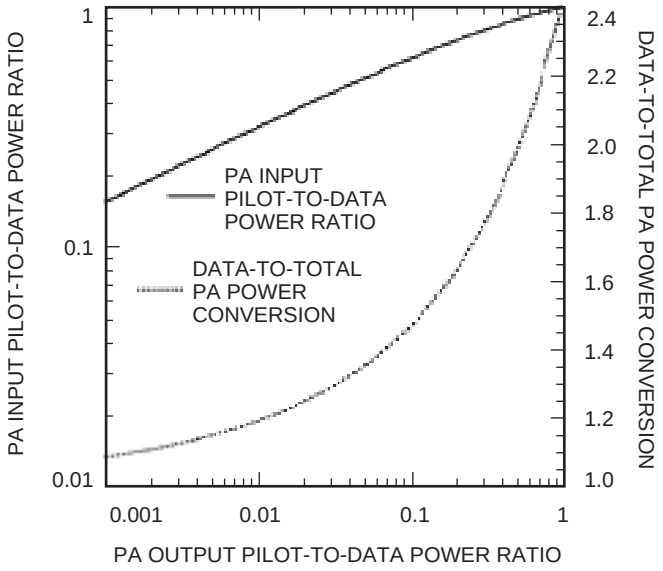

Fig. 4. Performance of single power amplifier for data and pilot carriers.

[10] For a perfect squarewave, the harmonic power is 23 percent of the fundamental power. But if only, say, the third and fifth harmonics are produced at their theoretical power, they have 15 percent of the fundamental power.

As a point of reference, the QPSK signal used for the baseline VSOP2 mission involves bi-phase modulation of quadrature carriers. Or, stated another way, 2 data bits at a time specify one of four discrete and equally spaced carrier phases (0, 90, 180, 270 deg). This 4-ary form of modulation is also referred to as 4-phase-shift keying (4-PSK). If the modulation rate is R_b bits/s, then the carrier spectrum null-to-null bandwidth is R_b Hz.

A. Use of *M*-PSK

Consider next 8-PSK, where three data bits at a time are used to specify one of eight discrete carrier phases (0, 45, 90, 135, 180, 225, 270, 315 deg). For a modulation rate of R_b bits/s, 8-PSK has a carrier spectrum null-to-null bandwidth of $(2/3)R_b$ Hz. But then, this spectral narrowing can be utilized to allow a higher bit rate of $(3/2)R_b$ that produces a carrier spectrum null-to-null bandwidth of R_b Hz. In other words, the spacecraft can transmit 1.024 Mb/s using QPSK or 1.536 Gb/s using 8-PSK, both with the same spectral structure shown in Fig. 1.

However, there are additional considerations when employing M-ary modulations. With increasing M, the distance between points making up the signal constellation decreases. In the presence of a fixed noise level, this is tantamount to increasing the BER for a given E_b/N_0. To counteract this effect, it becomes necessary to raise E_b/N_0.[11] Accordingly, to keep $P_E = 10^{-2}$, the theoretical E_b/N_0 needed for QPSK must be increased by 3 dB for 8-PSK.

Table 4 compiles the performance of M-PSK for M up to 256, and shows the bit-rate increase possible while maintaining the spectral structure of Fig. 1. What is most striking is the marked increase in E_b/N_0 needed to maintain a maximum BER of 10^{-2} for the 3σ worst-case conditions. Clearly, for an SVLBI spacecraft that is already raw-power limited, any such enhancement on the part of the PA alone, except perhaps for 8-PSK, seems implausible. But since it is EIRP that counts, some of the increase may be obtained by using a larger transmitting antenna.

Another method for effectively obtaining a higher E_b/N_0 is to cash in on excess link margin when it is available (which is most of the time). For example, going from a weather CD of 99 percent to 70 percent (thin clouds and very humid) automatically adds a minimum of 5.6 dB to the received E_b/N_0 at 10-deg elevation, more than enough to support 8-PSK (and approaching enough for 16-PSK) without any spacecraft EIRP increase. Or, if a weather CD of 99 percent prevails, a 3.0-dB excess may be obtained for an antenna elevation angle of 15 deg. The 7.3-db needed for 16-PSK is available for a weather CD of

Table 4. *M*-PSK performance.

M-PSK modulation	Bits per symbol	Bit rate, Gb/s	E_b/N_0 increase needed at $P_E = 10^{-2}$, dB
4-PSK or QPSK	2	1.024	0.0
8-PSK	3	1.536	3.0
16-PSK	4	2.048	7.1
32-PSK	5	2.560	11.7
64-PSK	6	3.072	16.6
128-PSK	7	3.584	21.6
256-PSK	8	4.096	26.7

[11] For all of the E_b/N_0 values that follow, especially in Tables 4 and 5, no differential encoding and decoding of the data symbols are assumed. This will require some additional E_b/N_0, typically less than 1 dB, but its inclusion goes beyond the treatment possible in this article. For further information, see [9] and [6, Section 4.7].

70 percent and elevation above 15 deg. And, finally, higher E_b/N_0 is also obtained for less than maximum slant range, increasing by 6 dB for each halving of the maximum range.

Briefly, then, in addition to the VSOP2 basic data rate of 1.024 Gb/s, higher rates of 1.536 Gb/s and 2.048 Gb/s may reasonably be attained through respective use of 8-PSK and 16-PSK modulation, plus some mix of increased EIRP and/or operational (weather, elevation, range) restrictions. However, the use of higher-order M-PSK ($M > 16$) appears precluded because the payoffs from the types of trade-offs just outlined lead to unacceptable operating constraints. Additionally, hardware implementation of higher-order M-PSK modulators and receivers imposes further limitations (as discussed in Section V).

B. Use of *M*-QAM

M-QAM (multiple-level quadrature-amplitude modulation) is a modulation form theoretically more efficient than M-PSK. It is beyond the scope of this article to do more than touch briefly on a few essentials as they apply to SVLBI. A broad treatment of the subject may be found in [6], [7, Chapter 8], and [8, Section 10.1.2].

The dilemma of M-PSK is the inordinate increase of E_b/N_0 needed for any value of M larger than 16. The reason for this is that, as more points are added to the signal constellation, a fixed distance between points must be maintained in order to preserve the desired P_E. This should be expected. But the principal drawback is that M-PSK inefficiently uses the signaling plane; all of the signal points lie around the circumference of a circle, as shown in Fig. 5(a). For each doubling of M, the distance between points halves, necessitating that the radius (or phasor amplitude) of the circle grow to compensate. But this is very poor "packing" of the amplitude–phase plane. Suppose instead that every alternate point around the circle of Fig. 5(a) is placed along the same radial line, but around another circle having a radius 0.63 that of the large radius. This will separate adjacent (closest) points by a factor of 1.23 times the distance between the points around the original circle, allowing a corresponding reduction of increased E_b/N_0 needed. But this still is not optimum packing.

A far better packing is depicted by the square[12] signal constellation in Fig. 5(b), where the distance to the outer corner points is the same as the radius of Fig. 5(a). Designated 16-quadrature-amplitude shift keying (QASK), the array is obtained by multi-level amplitude modulation of quadrature carriers. In Fig. 5, the distance between the 16-QASK points is about 1.21 the distance between the 16-PSK points. Also, the *average* E_b/N_0 for 16-QASK is substantially less than that for 16-PSK (which has identical average and peak E_b/N_0).

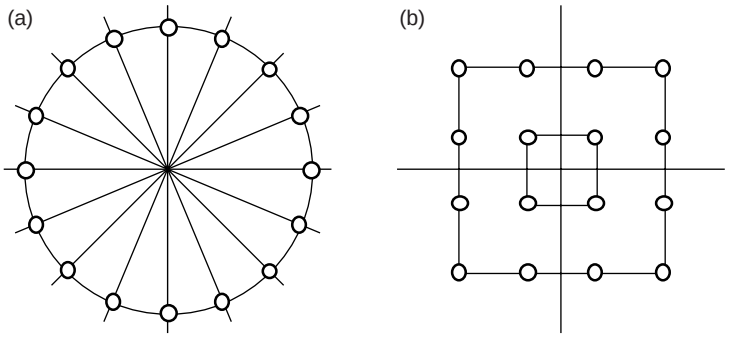

Fig. 5. *M*-PSK and *M*-QASK constellations for *M* = 16: (a) 16-PSK and (b) 16-QASK.

[12] A more general designation is "rectangular." When M is both a power of two and an integer-squared (4, 16, 64, 256), the signal point array will be square with dimensions $\sqrt{M} \times \sqrt{M}$. Otherwise, the array will be rectangular, with dimensions $\sqrt{2M} \times \sqrt{2M}/2$.

Table 5 compiles M-QASK average and peak E_b/N_0 requirements[13] as a function of M, along with the M-PSK conditions repeated from Table 4. An apt glance shows that M-QASK provides a substantial power economy over M-PSK (except, perhaps, for $M = 8$, where a higher peak power is needed). However, there is an additional qualifying consideration, namely that M-QASK does involve linear amplitude modulation, which will have an impact on actual raw power savings (see Subsection V.A).

Table 5. Comparison of M-PSK and M-QASK E_b/N_0 increases.

Modulation states (M)	M-PSK E_b/N_0 increase needed at $P_E = 10^{-2}$, dB	M-QASK average E_b/N_0 increase needed at $P_E = 10^{-2}$, dB	M-QASK peak E_b/N_0 increase needed at $P_E = 10^{-2}$, dB
4 (M-QASK 2×2 array)	0.0	0.0	0.0
8 (M-QASK 4×2 array)	3.0	1.7	3.9
16 (M-QASK 4×4 array)	7.1	3.6	6.2
32 (M-QASK 8×4 array)	11.7	5.5	9.0
64 (M-QASK 8×8 array)	16.6	7.6	11.3
128 (M-QASK 16×8 array)	21.6	10.8	14.9
256 (M-QASK 16×16 array)	26.7	12.1	16.3

C. Use of Pulse Shaping

Again, only a brief discussion will address some salient aspects of this topic, and [6], [7], and [8] should be consulted for further details.

In all of the foregoing, the form of the modulating symbol waveform has been rectangular, i.e., constant amplitude over each symbol period T_S, as depicted by Figs. 6(a) and 6(b), where two successive "1" symbols[14] are shown. In the frequency domain, the Fourier transform of a symbol pulse becomes the typical $\sin^2(x)/x^2$ spectral shape displayed by Figs. 6(a) and 6(b) (and also represented by Fig. 1). This spectrum, which theoretically is infinite in extent, has a null-to-null bandwidth of $2R_S$.

Alternatively, consider Figs. 6(c) and (d), which show the pulse shape and corresponding Fourier spectrum for what is known as the Nyquist waveform. This pulse has infinite extent over time but has a strictly limited rectangular spectrum of bandwidth R_S Hz.

To illustrate the importance of the Nyquist pulse, it already has been noted that going from QPSK to some form of 16-ary modulation effectively halves the required RF bandwidth. But this same savings may also be obtained by staying with "QPSK" (now actually 4-QAM) and using Nyquist, rather than rectangular, symbol waveforms. It is also essential to realize that the same theoretical E_b/N_0 is needed for rectangular QPSK and Nyquist 4-QAM. Thus, by employing 4-QAM instead of 16-PSK, the 7.1-dB penalty is seemingly avoided. Practically speaking, however, this will be partially offset by the need for a linear, and therefore less efficient, power amplifier (see Subsection V.A), plus the fact that ideal Nyquist waveforms cannot be generated.

---

[13] The formulas for calculations were taken from J. G. Smith, *Bandwidth-Conserving Digital Signals*, JPL internal document, Jet Propulsion Laboratory, Pasadena, California, 1975.

[14] In this article, non-shaped symbol data waveforms are represented as having a constant (flat) amplitude level over each symbol interval, which for binary modulations is referred to as a non-return-to-zero (NRZ) characteristic. For multiple-amplitude-level modulations, the designation M-NRZ is used.

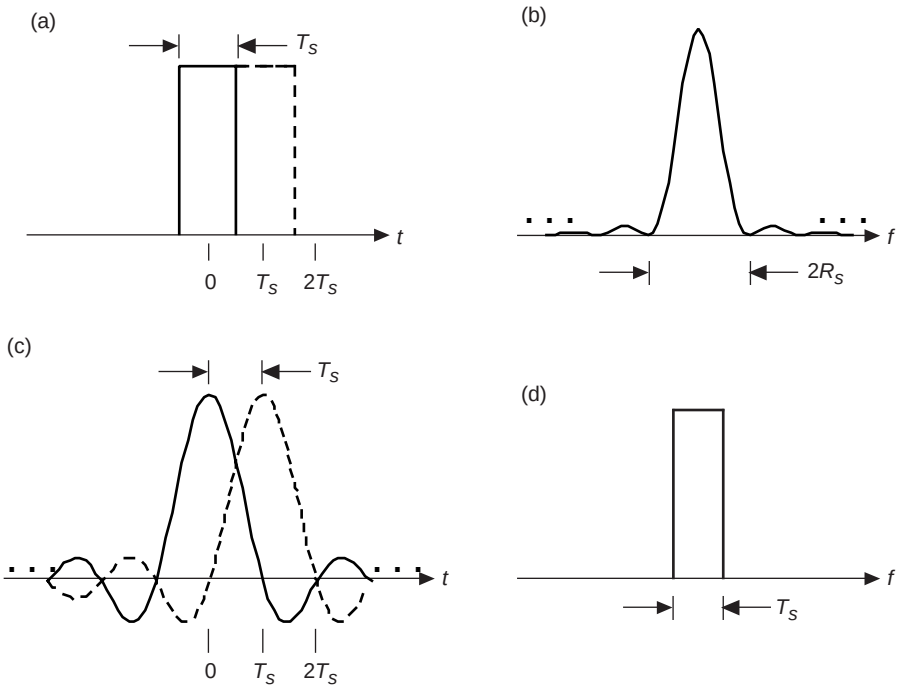

Fig. 6. Pulses with associated spectra: (a) rectangular pulses, (b) rectangular pulse spectrum, (c) Nyquist pulses, and (d) Nyquist pulse spectrum, $R_S = 1/T_S$.

A strictly rectangular bandwidth, and its associated Nyquist pulse, is unrealizable in practice, but may be approximated by various spectral shapes known by the names Butterworth, Gaussian, raised cosine (RC), and square-root raised-cosine (SRRC). The corresponding pulse waveforms may be generated by analog (filtering of the NRZ or M-NRZ symbol stream) or by digital (synthesis) methods. Most theoretical discussions defer to the filter method, with the Nyquist transfer function being denoted as $G_N(f)$. In actual systems, filtering must take place at the transmitter (the NRZ or M-NRZ waveform filter) and at the receiver (the SNR optimizing matched filter). Usually the transmit filter and receive filter transfer functions are identical, so that $G_T(f) = G_R(f) = \sqrt{G_N(f)}$ (and thus the distinction between SRRC and RC).

Some of these pulses are further qualified by a sharpness or roll-off factor, denoted by $\alpha (0 \leq \alpha \leq 1)$. The product αR_S is known as the excess bandwidth relative to the minimum bandwidth R_S. Thus, the nominal spectral bandwidth becomes $(1 + \alpha)R_S$. For $\alpha = 1$, the resulting spectrum has a bandwidth of $2R_S$, but there are no lobes outside this range (as there are for M-PSK). Realistic values for α are generally between $\alpha = 0.25$ and $\alpha = 0.5$. As a result, there will be a slight spectrum variance with respect to the VSOP2 baseline spectrum shown in Fig. 1. If it were possible to generate $\alpha = 0$ ideal Nyquist pulses for 4-QAM operating at 1.024 Gb/s (per polarization), the main spectral lobes (null-to-null) of Fig. 1 would be replaced by ideal rectangular spectra having the same width (512 MHz). But realistically, when $\alpha = 0.25$, each SRRC spectrum will have a width of 640 MHz (see Fig. 7 for the SRRC pulse and its spectrum). This places the theoretical upper frequency limit of the RCP spectrum at 37.948 GHz (52 MHz short of 38.0 GHz). (The lower frequency limit of the LCP spectrum will likewise be 52 MHz above 37.0 GHz.) Moreover, the pilot carriers no longer reside in spectral nulls (a disparity requiring further analysis).

To sum up, M-QAM with SRRC pulse shaping potentially appears to be the preeminent technique for placing the highest possible bit rate in the 37- to 38-GHz band. With it, and dual polarization, 4-QAM can provide 2.048 Gb/s, and ultimately it may become possible to communicate 4.096 Gb/s

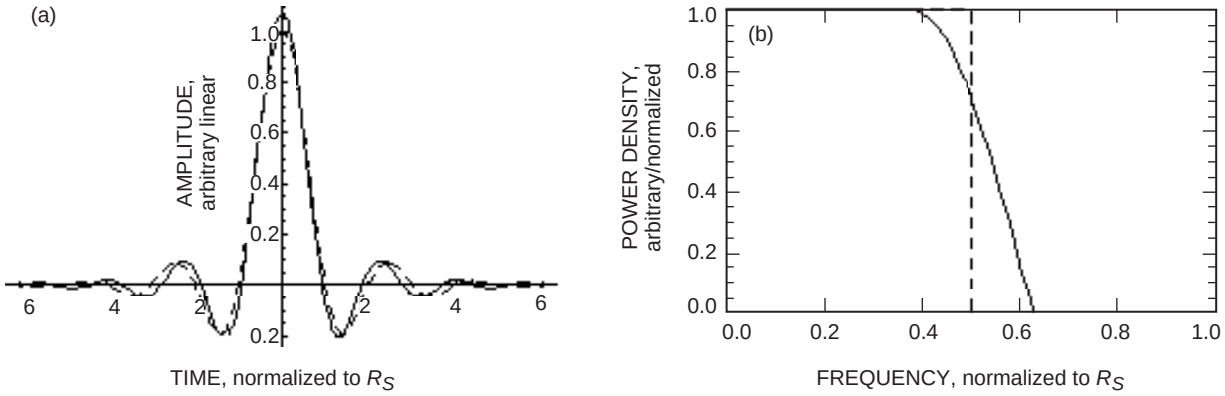

Fig. 7. Ideal Nyquist (dashed) and 125 percent SRRC (solid) Fourier transform pairs: (a) time and (b) one-sided (lowpass) frequency functions.

using 16-QAM. On the practical side, a number of problems require solutions. With the use of SRRC pulses, especially for large M, the nonlinearity and distortion problems that plague M-QASK (see Subsection V.A) will to an even greater extent afflict M-QAM. And, unfortunately, adequate Gb/s M-QAM modulators and linear power amplifiers may not become available in the foreseeable future.

V. Practical Modulators and Receivers

The following two subsections briefly examine practical implementations, especially some of the problems and limitations that currently beset M-ary transmitters.

A. Transmitters (Waveform Generators, Modulators, and Power Amplifiers)

1. *M*-PSK and *M*-QASK Transmitters. Very high data rate QPSK modulators typically have been constructed using analog circuits (quadrature hybrids, balanced mixers, splitter/combiners, etc.). Invariably, the result is that the RF QPSK signal has some in-phase (I) and quadrature-phase (Q) amplitude imbalance (≈0.5 dB), as well as quadrature error (≈5 deg). Errors of this magnitude have little impact on QPSK performance (generally <0.5 db in E_b/N_0 degradation).

M-PSK modulators may be constructed along similar lines. For example, 8-PSK can be realized by combining the outputs of two QPSK modulators that are respectively referenced to 45-deg phase-shifted carriers. But analog circuit imperfections also increase, and, because the points in the signal constellation are now much closer to each other in phase, imperfection errors (especially phase) will begin to manifest measurable performance degradation (around 1 dB). Conceptually, 16-PSK can be obtained by combining two 8-PSK modulators (four QPSK modulators), but the effects of analog circuit errors become really significant. Although there are other architectures for realizing 16-PSK (e.g., a series of dynamically switched phase shifters), as long as analog circuits are used, there will always be serious errors. The bottom line is that high-quality M-PSK modulators for $M > 8$ are quite difficult to implement using analog devices, especially at Gb/s data rates. Even when the best possible design and construction care is taken, equivalent E_b/N_0 losses of considerably more than 1 dB can be expected due to imperfections.

On the horizon are all-digital realizations for M-PSK and M-QASK modulators. These will be free of the amplitude and phase errors that plague analog modulators. Some already exist, but not for hundreds of Mb/s—into Gb/s—data rates.

A more pessimistic assessment exists for M-QASK transmitters (modulator plus PA). For M-PSK transmitters, the PA functions at a constant power level and, therefore, introduces negligible effects. However, when a PA is operated at multiple power levels (essentially linearized), AM-to-AM, AM-to-PM,

and PM-to-PM conversions produce serious signal constellation distortion (even with perfect modulator-generated input). So, the signal array fidelity of the combined M-QASK modulator–PA can be quite poor.

One solution to PA-induced distortions is to place what is known as a pre-distortion network (a linearizer) between the modulator and PA. But not only is this complex, it suffers from degeneration over temperature and with component aging. To combat this problem, a combination of feed-forward and feed-back loops has been designed to adjust the pre-distortion network parameters. This may be done in a pre-transmission mode, where phasors may be directly measured, or during transmission by measuring irregularities such as excessive out-of-band power. Either way, the pre-distortion method is formidable to implement, especially for $M > 16$.

Another solution to the PA nonlinearity problem is to employ a number of separate saturated power amplifiers and to phase-sum their outputs to obtain the total array. For example, two QPSK modulations, and therefore two amplifiers, one delivering four-times the power (twice the amplitude) of the other, may have their individual phasors added to produce all 16 points shown in Fig. 5(b). This method requires that the amplifier outputs be added using a hybrid combiner. Not only does this necessitate precise phasing between the two amplifiers (at Ka-band) under the conditions of temperature variation and component aging, the combiner itself has an insertion loss of 1 dB or so. Moreover, since both amplifiers must continuously produce power somewhat in excess of the array peak power, the potential power savings (Table 5) of 16-QASK over 16-QPSK becomes negated. In fact, it is probably reasonable to conclude that any solution to the PA distortion problem will involve substantial complications and high development costs and thereby tip the balance in favor of 16-QPSK over 16-QASK (at least for the near future).

An additional concern when considering M-QASK versus M-QPSK must be the DC (or prime) power efficiency of the design. Because M-QPSK works with a PA operating in its RF power saturation region, efficiencies can run as high as 35 percent at 8 GHz (X-band). For Ka-band, the efficiency with currently available technologies is somewhat less than 30 percent (principally because several lower-power field-effect transistors (FETs) must be combined to realize output power levels of a few watts). For M-QASK (and even more so for M-QAM, discussed below), a high degree of linearity is required, necessitating that the PA be some form of Class-A design, which typically has an efficiency less than 20 percent. Further information on transmitter efficiency may be found in [10].

2. *M*-QAM Shaped-Pulse Transmitters. The following discusses the generation of SRRC and its linearity requirement, plus other practical caveats pertaining to the implementation of a serviceable SRRC transmission system. Specifics are confined to 4-QAM architectures, but all results apply in principle (if not degree) for larger M's.

For reference, Fig. 7 shows the time and one-sided (lowpass) frequency functions for ideal Nyquist and $\alpha = 0.25$ (or 125 percent) SRRC pulses. (Note the time functions and spectra are normalized to R_S.) The SRRC spectrum is zero for $f \geq 0.625$.

A 4-QAM transmitter basically consists of an SRRC pulse generator, IF carrier modulator, frequency upconverter (U/C), and RF power amplifier. As much as possible is implemented at some IF, with the 4-QAM SRRC-modulated IF then being translated to RF. Three basic SRRC pulse waveform generator and carrier modulator designs seem possible: (1) all digital, (2) hybrid digital–analog, and (3) all analog. Each of these will be briefly examined.

A high-level block diagram for a transmitter employing an all-digital waveform generator and carrier modulator is shown in Fig. 8. Digital blocks are drawn with a thin line, and analog items are rendered with a thick line. The digital-to-analog converter (DAC) is a part-digital, part-analog device. The data symbol streams, designated $S_I(t)$ and $S_Q(t)$, have a symbol rate R_S. For practical reasons, it is assumed that the digital hardware "samples" the symbols and operates at a clock frequency of $4R_S$. (How the symbol streams and the $4R_S$ clock are generated is not of immediate concern.)

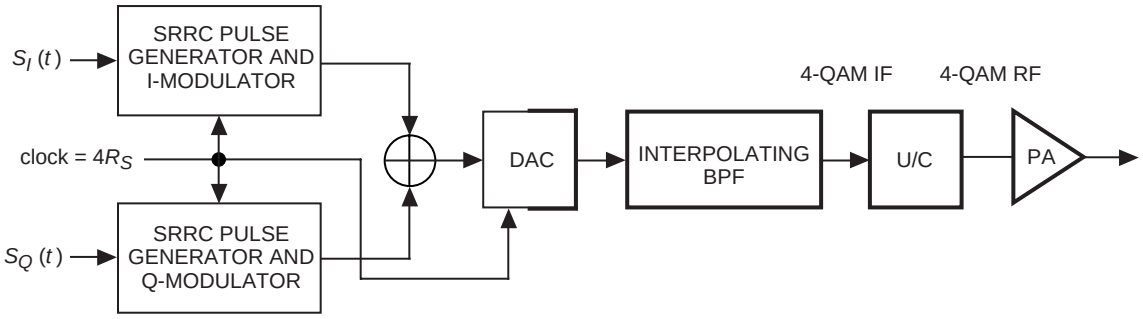

Fig. 8. Transmitter with all-digital waveform generator and modulator.

Each SRRC pulse generator and modulator functions to produce precise amplitude samples of a 125 percent SRRC time function, multiplied by samples of the IF carrier amplitude, at the $4R_S$ rate. The modulator is easily implemented provided the IF frequency is $f_{\mathrm{IF}} = R_S$ and its quadrature sample's amplitudes are 1-bit (logically, $\cdots, 1, 1, 0, 0, \cdots$ for the I carrier and $\cdots, 1, 0, 0, 1, \cdots$ for the Q carrier). Thus, modulation simply amounts to sign reversal of the appropriate SRRC time function samples. There are various ways the latter may be obtained, but a look-up scheme appears quite straightforward. In fact, because each SRRC pulse corresponding to each symbol will, of necessity, be time limited (finite; see below), the entire process of I and Q SRRC pulse generation, modulation, and I and Q addition might be accomplished using a look-up table [read-only memory (ROM)] that contains all of the possible combinations.

The DAC must take each modulation number and convert it to a linearly quantized voltage or current value at the effective rate of $4R_S$ (which becomes a major burden for the DAC at very high data rates, as discussed below). Finally, the interpolating BPF passes the fundamental IF spectrum (extending from $0.375R_S$ to $1.625R_S$), adequately suppressing all harmonics. A two-stage upconverter (U/C) is necessary to translate the IF signal to RF (Ka-band), followed by a linear PA.

The next option for the transmitter involves a digital waveform generator and an analog carrier modulator, as shown in Fig. 9. This differs from the all-digital scheme in that the QAM modulator is constructed from analog components. This appears to have little practical advantage over the all-digital scheme, because two (rather than a single) DACs are needed and obtaining the necessary linearity for the modulators (multipliers) is very difficult. On the plus side, upconversion to RF may be accomplished in a single stage because the IF carrier can be much higher in frequency than R_S. Also, the lowpass filters (LPFs) should be much easier to realize than the interpolating BPF. Nevertheless, it is difficult to envision the hybrid digital–analog architecture having undeniable advantages over the all-digital approach.

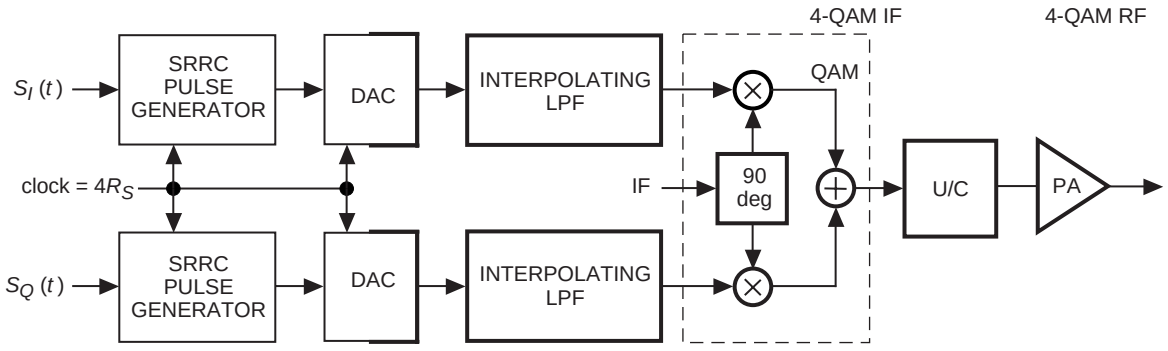

Fig. 9. Transmitter with hybrid digital—analog waveform generator and modulator.

Lastly, the all-analog architecture is shown in Fig. 10. Ahead of the U/C and PA, it consists of two principal parts: SRRC pulse filters and the QAM modulator. Conceptually, the all-analog approach is relatively simple, but it also suffers from several drawbacks, especially the inability to produce preciseness and long-term constancy for the SRRC pulse waveforms. Because of the multitude of problems that plague analog implementations, an all-analog 4-QAM transmitter appears to be out of the question for Gb/s rates.

It seems reasonable to conclude that *in principle* the most attractive approach for constructing an M-QAM transmitter is to make use of the all-digital waveform generator and modulator. Meeting all fidelity requirements, however, especially those that impact the actual out-of-band spectral levels, is not an easy matter. Furthermore, there are acute PA and DAC problems that must be solved. These issues are addressed in the following subsections.

3. *M*-QAM Transmitter Imperfections. Theoretically, the $\alpha = 0.25$ normalized SRRC spectrum is zero for $f \geq 0.625$ (Fig. 7). But the effects of (1) finite-pulse waveform duration, (2) the number of DAC bits employed, and (3) amplitude nonlinearities can give rise to significant spectral content for $f \geq 0.625$. This becomes important in terms of adjacent channel interference and allowable (legislated) out-of-band emission levels.

4. Effect 1—Finite-Pulse Waveform. A finite-pulse waveform means that the SRRC pulse function is generated over the limited time period NT_S. Figure 11 shows three spectra for $N = 64, 128, 512$, where the vertical dashed line denotes the ideal terminus at $f \geq 0.625$, and the vertical solid line is

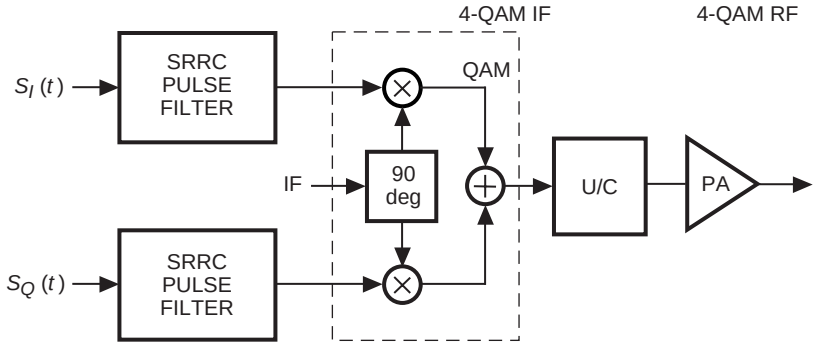

Fig. 10. Transmitter with all-analog waveform generator and modulator.

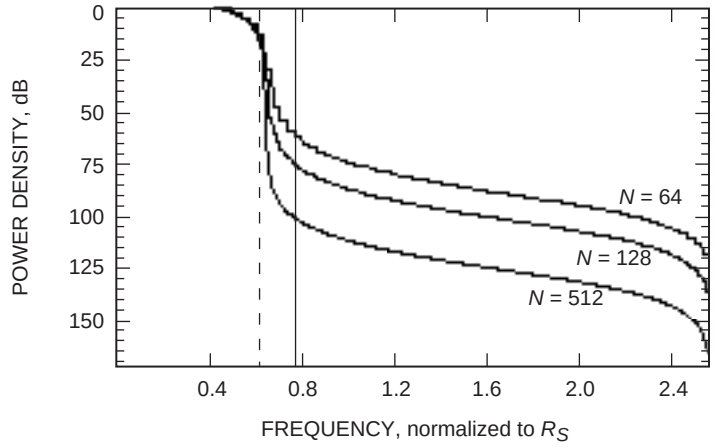

Fig. 11. Spectra for finite-length SRRC pulses.

synonymous with 38 GHz (the Ka-band upper limit, as discussed in Subsection III.A). Note that the vertical axis is in logarithmic units, rather than the absolute axis used for Fig. 7. The out-of-band level of the spectrum drops by about 12 dB for every doubling of N.

5. Effect 2—Number of DAC Bits. Figure 12 indicates the sensitivity of an $N = 512$ SRRC pulse when converted to analog form using bipolar DACs of 5, 8, 10, and 12 bits. The out-of-band "hash" arises from quantization noise and decreases about 6 dB for every additional DAC bit.

6. Effect 3—Amplitude Distortion. Small amplitude distortions are unavoidable in the analog circuits after the DAC output, and especially through the PA. Large amplitude distortions must be expunged through design. The most common nonlinearity is that of gradual amplitude saturation or compression, where large voltages are gracefully "clipped." This is typical of linear amplifiers, most particularly the PA (even with linear equalization). Another type of nonlinearity arises from "under-drive," wherein very small voltages are suppressed. This phenomenon is indigenous to some types of amplifiers, as well as analog modulators (especially passive ring-bridge-type devices). Figure 13 shows both generic characteristics.

One way of characterizing distortions is by the percentage of total harmonics produced when a test sinusoid (in lieu of the actual signal) having equivalent amplitude and frequency is passed through the nonlinearity. For microwave amplifiers, usually only the third harmonic is considered or specified. With the characteristics in Fig. 13, the harmonic distortion produced is 4.4 percent (−27 dBr) for gradual saturation and 2.5 percent (−32 dBr) for under-drive. These levels are considered quite conservative for a complex transfer chain (the U/C and PA). But note in Fig. 14 how drastically they raise the out-of-band spectrum levels for an $N = 512$ SRRC pulse.

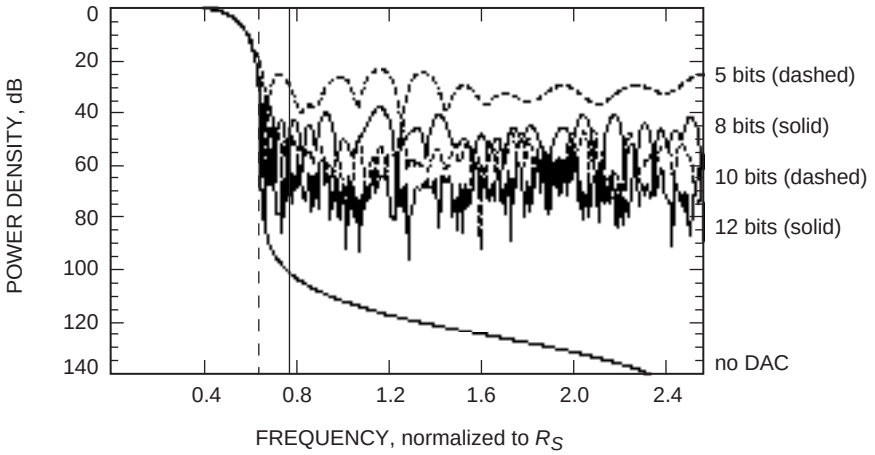

Fig. 12. Spectra for various DAC bits.

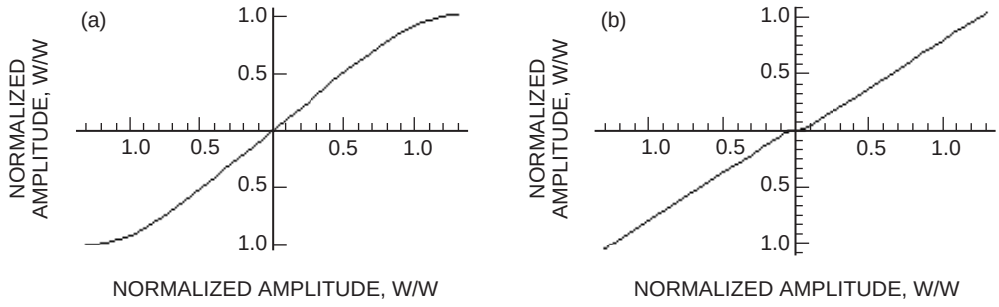

Fig. 13. Two types of amplitude distortion: (a) saturation and (b) under-drive.

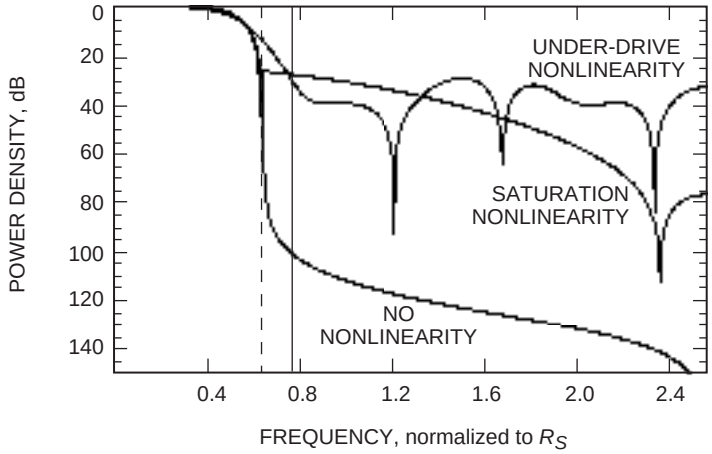

Fig. 14. Spectra for an amplitude-distorted SRRC pulse.

Clearly, linearity becomes a critical issue in order to prevent spectral out-of-band power from rising to unacceptable levels. It remains to be determined whether even the small nonlinearities portrayed in the above example can be achieved in practice.

7. Composite Effects. Figure 15 shows the composite result for an $N = 512$, $\alpha = 0.25$ SRRC pulse, using a 12-bit DAC, and allowing 4.4 percent harmonic distortion due to gradual amplitude saturation. (Under-drive is unlikely to occur for the all-digital architecture.) Notice that the out-of-band contribution appears dominated by the amplitude saturation effect for f up to about 1.6, and by the number of DAC bits for larger f.

Thus far, all of the examples have dealt with a single SRRC pulse. In reality, the real-time waveform is made up of a sequence of SRRC pulses separated by T_S, the sign of each pulse being determined by the data symbols. Figure 16 shows just such a sequence over 41 symbols for $T_S = 1$. It is worth noting that the amplitude peak of the SRRC sequence is about 50 percent greater than that of a single pulse. The corresponding spectra are portrayed in Fig. 17. The additional "hash" is due to the finite sequence of random data symbols, as is the slightly higher out-of-band level for "$N = 512$ only" as compared with its counterpart in Fig. 15. However, the out-of-band level due to the composite effects appears to be on a par with that for the single pulse.

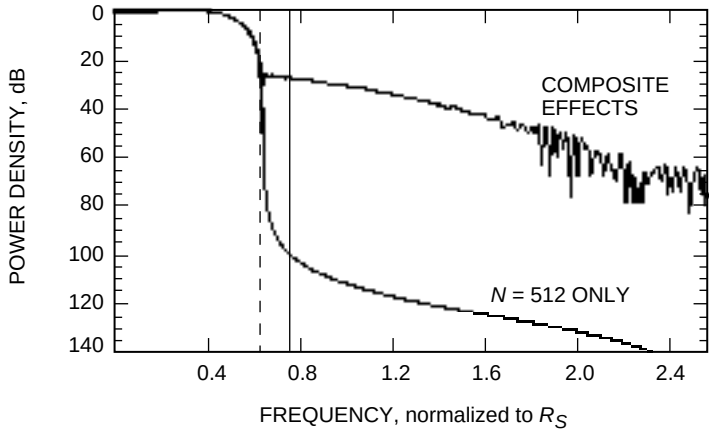

Fig. 15. Composite 12-bit DAC and 4.4 percent gradual distortion effects.

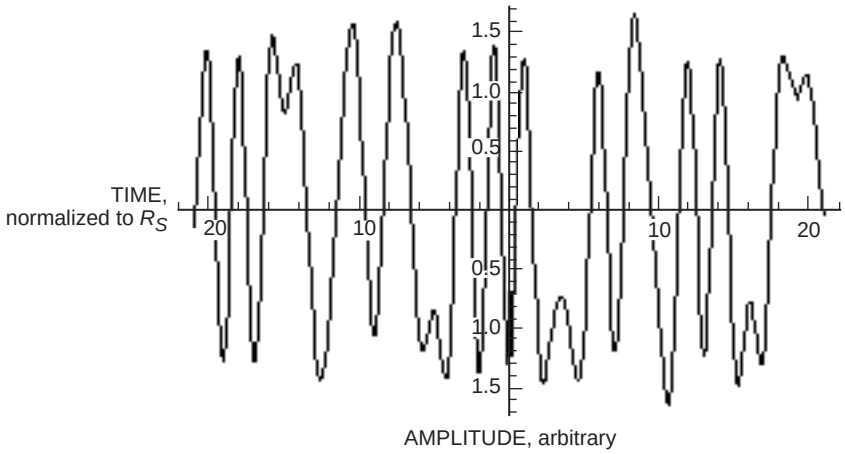

Fig. 16. Sequence of data-symbol-modulated SRRC pulses.

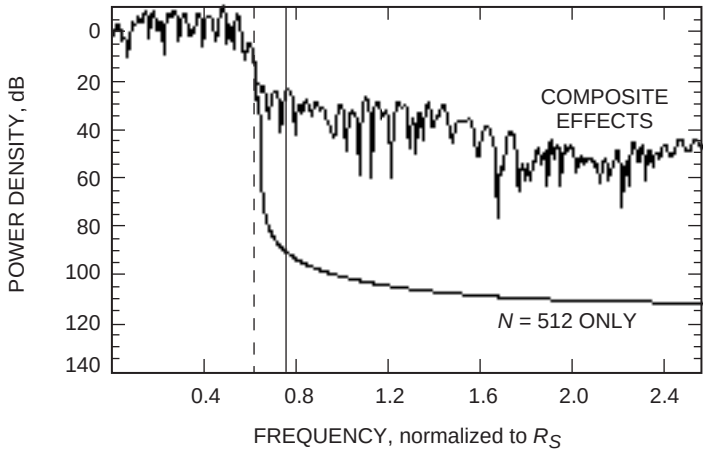

Fig. 17. Sequence composite 12-bit DAC and 4.4 percent gradual distortion effects.

8. DAC Problems. When calculating the preceding results, the DAC was simply represented as a quantized number. But a real DAC outputs some form of non-ideal rectangular pulse over each interval $T_S/4$. Such high output rates require the use of a video-class DAC. For these, the output pulses are distinguished by substantial (non-trivial) onset delay, finite rise (or transition) time, settling time, and glitching. Video-class DACs typically provide a unipolar (and biased) output current (rather than voltage), capable of driving low impedance loads (e.g., 50Ω). De-glitching is often effected by following the DAC with a track-and-hold amplifier, which holds during the glitch interval. The bottom line here is that the DAC's output will be far from ideal. What must be done, therefore, is to pursue a design which ensures that the output pulse, or waveform step, is highly consistent from sample to sample. Then, common irregularities can be accommodated, the most important being to provide suitable correction for the aperture distortion caused by the real pulse shape so that accurate SRRC pulse synthesis results.

Now, although the foregoing has talked in terms of a single DAC, there are no video DACs available (at least commercially) capable of operating at a clock frequency of $4R_S$ (an example being 2.048 GHz for the 1.024-Gb/s 4-QAM system discussed in Subsection IV.C). Pragmatically, the synthesis of the SRRC pulse sequence at this rate is ultimately limited by the slowest component in the chain of generation functions, which may or may not be the DAC. Commercial video DACs (not space qualified) are currently able to support throughput rates on the order of 200 mega-samples per second (e.g., Analog Devices AD9732).

Video DACs seem to be either 8 bit or 10 bit. Thus, from a practical standpoint, more than ten DACs could be necessary for each symbol stream.[15] This means that at least 20 DACs are needed for each 1.024-Gb/s 4-QAM modulator unit.

Of course, implied by the use of multiple DACs is the fact that the logic scheme which generates the composite I and Q SRRC modulation will have some sort of pipeline or parallel processing structure. Also required is an analog multiplexer to sequentially multiplex each DAC's outputs (at an analog switching rate of 2.048 GHz) into a common signal path for input to the interpolating BPF. At this speed, some arrangement using PIN diode switches, having matched delays and minimum switching transients, seems likely. Clearly, the design and implementation of this complex structure requires new developments. One can only hope that video DAC technology will advance over the next few years, or perhaps that a specialized DAC might be developed, to allow a drastic reduction of the number of DACs needed.

Summing up the findings for M-QAM transmitters, it seems that acceptable 4-QAM is possible with $N = 512$, $\alpha = 0.25$ SRRC pulses, using a 12-bit DAC, and allowing no more than 4.4 percent harmonic distortion due to gradual amplitude saturation in the analog circuits (especially the PA). Realizing this in practice, however, remains a monumental challenge. Additionally, a big uncertainty for any M-QAM system (including M-QASK) is whether its DC power efficiency can exceed that of M-PSK for the same M.

B. Receivers

Since good M-PSK receiver architectures have been known for decades, the principal receiver challenge is hardware implementation. For bit rates above 100 Mb/s, most contemporary designs are for QPSK (4-PSK) and employ a mixture of discrete and integrated analog and digital circuits. The upper bit-rate limit of these designs is about 500 Mb/s, and they are expensive ($30,000 to $60,000). It is believed that, as with existing analog modulators, analog design methods for such receivers have reached their end. Only all-digital methods will permit efficient gigabit receivers to be produced at reasonable costs. An emerging design (and the only one presently being considered for advanced SVLBI missions) is predicated on the use of a first-generation Goddard Space Flight Center (GSFC)-/JPL-developed application-specific integrated circuit (ASIC) capable of acquiring, tracking, demodulating, and detecting a BPSK symbol stream of up to 300 Ms/s. (The next-generation ASIC is expected to be capable of 600 Ms/s.) Two such ASICs operating from a common signal sample stream can receive the VSOP2 baseline 512-Mb/s QPSK data (256-Ms/s I data in one ASIC, and 256-Ms/s Q data in the other ASIC). A complete receiver incorporating the ASICs is shown in Fig. 18.

The receiver input consists of a downconverted IF at the QPSK symbol rate of R_S. A high-quality (good symmetry, linear phase) anti-aliasing BPF with a bandwidth of $0.9R_S$ and sharp cutoff is required. In order to keep the IF signal properly scaled into the analog-to-digital converter (ADC), an automatic gain control (AGC) loop is provided. The GaAs technology flash ADC samples the input IF at $4R_S$,[16] producing 8-bit samples in parallel (8 wide). Following the ADC, another GaAs digital converter further parallelizes the samples via a 1:8 demultiplexer (DEMUX) to produce 8 parallel 8-bit samples (64 wide) that pass into the ASICs. Internally, each ASIC stacks the 64-wide samples into a 128-wide array before demodulation processing occurs. This allows the samples to be processed using CMOS technology.

Once the data are converted into a 128-wide stream, each ASIC (one operating on the I component, and the other on the Q) carries out all of the demodulation and detection operations, including complex baseband mixing, lowpass filtering, detection filtering, symbol synchronization, and carrier tracking

[15] It is always wise to back away somewhat from the supposedly maximum throughput rate.

[16] Actually, the ADC operates at a fixed sample rate of 1.2 GHz, and the ASICs run at a fixed clock frequency of 75 MHz. These clocks are generally asynchronous with respect to the input symbol rate. Internal to the ASICs are tracking loops and clock algorithms that account for this disparity. However, for the sake of discussion, it is easier to conceptualize the sampling as four times the symbol rate (or four samples per symbol).

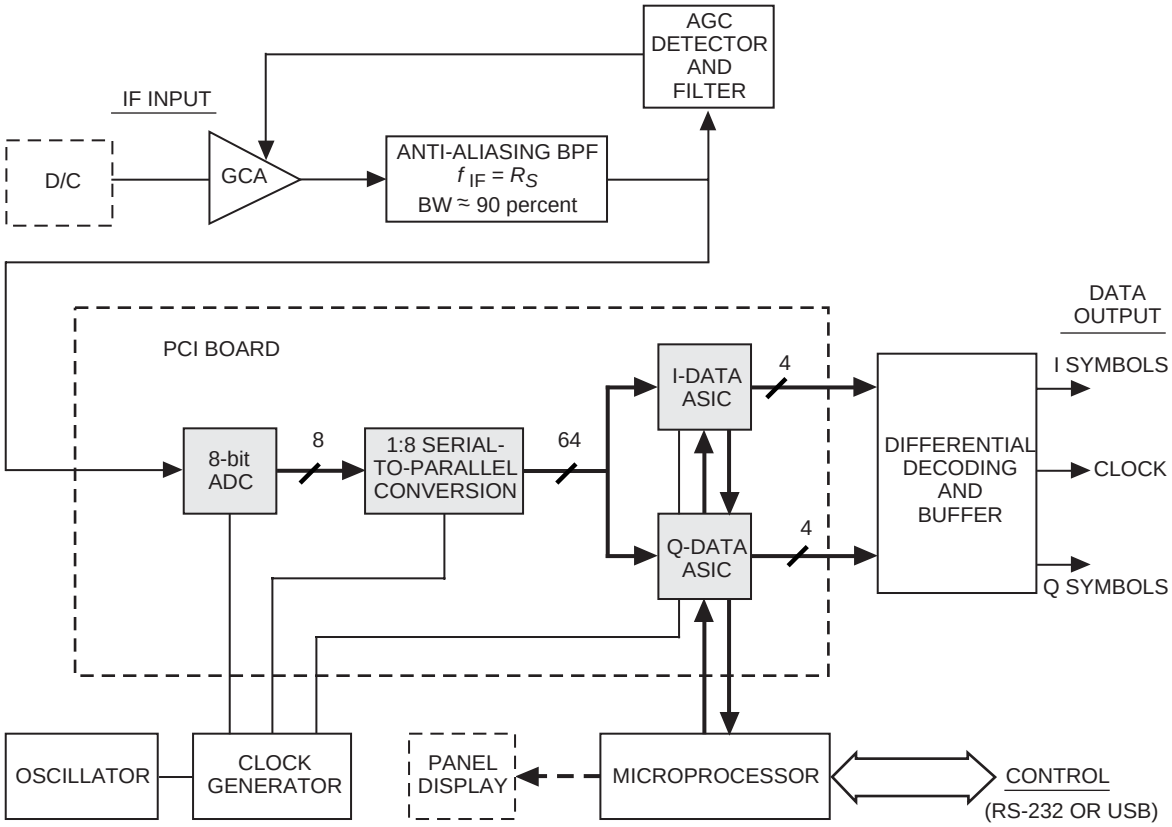

Fig. 18. QPSK digital receiver.

(including carrier Doppler removal). To accomplish this, the samples are simultaneously operated on by an array of mixers (demodulators), following which a discrete Fourier transform (DFT) is performed. Then finite-impulse response (FIR) filtering is applied, followed by an inverse discrete Fourier transform (IDFT) to reconstruct a filtered sample time sequence. The FIR filtering constitutes the symbol detection matched filter. Each ASIC also contains algorithms for symbol synchronization (optimum timing) and Costas-type carrier regeneration. The symbol synchronizer is of the digital transition tracking loop (DTTL) type, having a threshold E_b/N_0 of 2 dB (more than good enough for SVLBI). The ultimate output of this processing is 4 data symbols at a time from each ASIC.

In general, the ASICs are capable of demodulating and detecting BPSK, QPSK, offset QPSK, unbalanced QPSK, and other I/Q modulation schemes. But in their present form, the ASICs cannot directly accommodate either 8-PSK or 16-PSK, or any form of M-QAM. However, 8-PSK can be handled through the addition of a field programmable gate array (FPGA) processor to perform the eight-phase carrier regeneration and symbol detection functions. In the planning stages, the next-generation ASIC will directly incorporate 8-PSK and 16-QAM receiving algorithms.

The present generation ASICs, along with the ADC and DEMUX, are configured on a PCI board. A more complete breadboard receiver has been constructed by GSFC and includes the anti-aliasing BPF and an AGC circuit. It has been successfully used to track and demodulate data from several Earth satellites, including Landsat 7 (150 Mb/s). However, a complete stand-alone receiver that contains all of the blocks shown in Fig. 18 and is capable of being configured to handle a range of data rates and several modulation types does not yet exist. Nor has any commercial manufacturer adopted the technology for its product line, although a few companies have shown interest, and in the near future one or more may develop an off-the-shelf unit.

VI. Summary and Conclusions

This article has covered a wide range of topics connected with attaining SVLBI data rates of 1.024 Gb/s and higher. A minimum data rate of 1.024 Gb/s for VSOP2 is a factor of eight increase over the VSOP mission, and its transmission on the 37- to 38-GHz band using dual-polarization QPSK modulation with pilot carriers is another distinct advance.

It has been demonstrated that the link design for VSOP was unduly conservative and that its performance margin was excessive. In Section II, a more lenient statistical assessment, especially regarding weather margin, was presented. It was consequently established that spacecraft EIRP requirements can be significantly reduced from earlier assessments and that the link allowable maximum BER should be established at 10^{-2}.

In Section III, the signal design for VSOP2 was presented, including a way to meet NTIA out-of-band spectrum regulation requirements without requiring sophisticated filtering. The advantages of separate data and phase transfer/telemetry carriers were introduced, plus how to allocate data and pilot carrier powers from a single power amplifier.

A formidable problem is providing data rates above 1.024 Gb/s, up to 4.096 Gb/s, without compromising the basic spectral structure of Fig. 1. Section IV covered a number of modulation types that permit transmission up to 4.096 Gb/s on the 37- to 38-GHz band. Various forms of bandwidth compressive modulation were examined. One fundamental price paid for their use is a significantly higher E_b/N_0 than that for QPSK. M-PSK requires the largest increase in E_b/N_0. M-QASK needs less E_b/N_0 increase, but at the expense of greater implementation complexity and less efficient use of DC power. For this reason, it was determined that for $M = 8$ (1.536 Gb/s) and $M = 16$ (2.048 Gb/s), M-PSK was the choice over M-QASK. On the other hand, M-QAM gains considerable advantage for $M = 4$ (2.048 Gb/s) and $M = 16$ (4.096 Gb/s), but its usefulness critically depends on SRRC pulse-generation fidelity and the linearity achievable for its PA.

Lastly, Section V addressed practical problems with the construction of various hardware elements for Gb/s data rates. Waveform generators, modulators, and PAs are the most challenging devices, requiring substantial advanced developments if they are to provide adequate signal transfer fidelity for M-QASK and M-QAM. On the other hand, digital hardware Gb/s receivers are currently under development and will generally become available way before the transmitters that communicate to them.

What can be done if M-QAM cannot be successfully brought to fruition at Gb/s rates? One answer is to go to W-band and obtain a wide enough allocation to use QPSK. The alternative is analog transmission. Production of 4.096 Gb/s, using Nyquist rate sampling and 2 bits per sample, presupposes an equivalent analog bandwidth of 1.024 GHz. This easily fits into the 37- to 38 GHz band by subdividing[17] and placing 512-MHz-wide analog signals in each of the orthogonal polarizations. Digitization is then accomplished at the ground tracking station. Analog transmission, as compared to 16-QAM, requires no innovative designs and components to implement, but it does have some drawbacks, especially its being subject to atmospheric effects from which digital transmission is virtually immune. However, there are reasonable solutions to the analog system problems, which seem to make analog transmission a viable and far more cost-effective alternative to M-QAM. Details of the analog architecture and link will be presented in a future article in this publication.

[17] This may not even be necessary if the radio telescope simultaneously receives 512-MHz-wide LCP and RCP signals.

References

[1] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy*, 2nd ed., New York: John Wiley and Sons, 2001 (1st ed. 1986).

[2] R. A. Perley, F. R. Schwab, and A. H. Bridle, eds., *Synthesis Imaging in Radio Astronomy*, Astronomical Society of the Pacific Conference Series, Bookcrafters Inc., 1989.

[3] *DSN/Flight Project Interface Design Handbook*, rev. D, vol. I, DSN Document 810-5, "Telecommunications Interfaces Atmospheric and Environmental Effects," TCI-40, rev. C, Jet Propulsion Laboratory, Pasadena, California, May 1, 1992.

[4] L. J. Ippolito, *Propagation Effects Handbook for Satellite Systems Design (A Summary of Propagation Impairments on 10 to 100 GHz)*, NASA Reference Publication 1082(04), Fourth Edition, NASA, 1989.

[5] *NTIA Manual of Regulations and Procedures for Federal Radio Frequency Management (Red Book)*, NTIA, January 2000 edition with May/September 2000 revisions.

[6] W. Webb and L. Hanzo, *Modern Quadrature Amplitude Modulation*, New York: IEEE Press and John Wiley and Sons, 1994.

[7] F. Xiong, *Digital Modulation Techniques*, Dedham, Massachusetts: Artech House Publishers, 2000.

[8] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques*, Upper Saddle River, New Jersey: PTR Prentice Hall, 1995.

[9] W. J. Weber III, "Differential Encoding for Multiple Amplitude and Phase Shift Keying Systems," *IEEE Transactions on Communications*, vol. COM-26, no. 3, pp. 385–391, March 1978.

[10] J. G. Smith, "Transmitter Efficiency Considerations in MAPSK Signal Selection," *NTC Record*, December 1974.

[11] A. B. Chmielewski, M. Noca, and R. D. Wietfeldt, eds., *Advanced Radio Interferometry between Space and Earth (ARISE) Mission and Spacecraft*, 2nd. ed., JPL Publication 99-14, Jet Propulsion Laboratory, Pasadena, California, October 1999.

[12] J. C. Springett, "Space Very Long Baseline Interferometry Ground-Station Segmented Architecture," *The Interplanetary Network Progress Report, January–March 2002*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–17, May 15, 2002.
http://ipnpr.jpl.nasa.gov/progress_report/42-149/149G.pdf