Finite Wordlength Implementation of a Megachannel Digital Spectrum Analyzer

E. H. Satorius, M. J. Grimm, G. A. Zimmerman, and H. C. Wilck
Communications Systems Research Section

This article presents the results of an extensive system analysis of the megachannel spectrum analyzer currently being developed for use in various applications of the Deep Space Network. The intent of this analysis is to quantify the effects of digital quantization errors on system performance. The results of this analysis provide useful guidelines for choosing various system design parameters to enhance system performance.

I. Introduction

Development of a million channel, Fast Fourier Transform (FFT)-based spectrum analyzer is currently under way at JPL for use in various applications of the Deep Space Network. The system is being designed to provide contiguous output spectra at a real-time throughput rate of 40 MHz. Although the basic system architecture is reasonably well established as described in Ref. 1, i.e., based on a Radix-2 decimation-in-frequency (DIF) FFT algorithm, a number of issues relating to system performance tradeoffs remain to be addressed prior to hardware development.

To this end, we present in this article the results of an extensive computer-aided systems analysis aimed at quantifying the effects of digital quantization noise on system performance. In general, there are three specific sources of quantization error in the implementation and operation of a digital spectrum analyzer: input data quantization; computational errors resulting from finite precision arithmetic operations; and coefficient quantization errors. All of these errors degrade system performance in terms of dynamic range and sensitivity. The results of our analysis not only serve to quantify these degradations, but also provide guidelines for choosing various system design parameters to enhance system performance.

In the remainder of this article, we first briefly describe the general system architecture in Section II. Then, in Section III, we discuss front-end design considerations focusing on image rejection capability. Finally, Section IV contains the results of an extensive quantization error analysis of the basic Radix-2 DIF architecture. The intent of this analysis is a determination of system dynamic range and sensitivity as a function of various input data and wordlength parameters.

II. System Overview

A block diagram of the digital spectrum analyzer system is depicted in Fig. 1 including an IF stage, postprocessing and preprocessing stages, and the FFT processor. As summarized in Ref. 1, the digital spectrum analyzer (FFT processor) is based on the Radix-2 DIF algorithm and will consist of 20 Radix-2 butterfly stages to provide up to 2^{20} output spectral points. Initially (Ref. 1), it was envisioned that 22-bit floating

point arithmetic would be utilized exclusively throughout the FFT processor. However, this is currently being re-evaluated for two reasons. First, due to the excessive memory requirements for the first butterfly stages, it becomes important to minimize wordlengths and simplify the arithmetic in the first stages to meet a 1-rack system packaging goal (this is especially important for a 40-MHz system). Based on the results of the analysis presented in Section IV, it is seen that 16-bit fixed point arithmetic can be used to implement the first butterfly stages (6–8 stages) without sacrificing system performance. Consequently, a hybrid fixed/floating point architecture is being considered as a viable candidate for system implementation. In Section IV, results are presented which quantify hybrid system performance as a function of the number of front-end fixed point stages.

A second reason for re-evaluating a 22-bit floating point architecture also stems from the quantization error analysis as well as the recent advent of fast (10 MHz and faster) 32-bit IEEE floating point chips for performing multiplication and addition operations. In particular, as discussed in Section IV, the dynamic range constraint imposed by a 22-bit floating point architecture is less than that imposed by an 8-bit input quantizer. Thus, dynamic range can not be extended for a 22-bit floating point-based architecture by simply increasing the input quantizer resolution. Such an extension, however, is possible with a 32-bit floating point system. Thus, with the advent of the new chips, 32-bit floating point arithmetic is being considered for hardware implementation.

In addition to re-evaluating the FFT arithmetic, two other critical system design areas are being re-examined prior to system development: (1) system bandwidth, and (2) input IF signal conditioning. A 20-MHz system bandwidth was the original design goal. However, by reducing the FFT front-end size requirements as discussed above, it is possible to package a 40-MHz spectrum analyzer system within one rack. This is highly desirable from the viewpoint of ultimately developing a super wideband (approximately 300 MHz wide), high-resolution system. This can be done with little risk by simply replicating the prototype narrowband system. However, it is certainly more feasible to replicate a 40-MHz system 8 times, for example, than it is to replicate a 20-MHz system 16 times to achieve the same final bandwidth.

Finally, reconsideration of the IF input signal conditioner is also being carried out. In particular, as discussed in Ref. 1, complex mixing was initially considered for translating the input IF signal to baseband. However, only 30-dB image rejection can be realistically provided by complex base-banding due to inherent phase and amplitude mismatching between the in-phase and quadrature channels. Alternatively, a single real channel, sampled at twice the signal bandwidth, can be used to provide the same spectral information. Furthermore, as discussed in Section III, the real baseband approach is not limited to 30-dB image rejection. The only requirement for this approach is an AD converter which operates at twice the input signal bandwidth. Such a requirement can be met, for example, with existing 8-bit AD converters that operate in excess of the 80-MHz sample rate required for a 40-MHz wide system. Thus, a real baseband, 40-MHz wide system is feasible and is currently being evaluated for system implementation.

III. Front End Design Considerations

With reference to Fig. 1, the system front end prior to digital spectrum analysis consists of the IF conditioning stage as well as input buffer memory and preprocessing stages. Signal basebanding and AD conversion are performed at the IF stage, whereas data reformatting, buffering, and windowing are typically performed in the remaining front-end stages. Critical areas in the design of the system front end include dynamic range constraints imposed by the AD converter as well as the generation of spurious noise components (images) by the basebanding processor. The former design area is discussed below in Section IV, whereas the latter is discussed in this section.

In particular, consider a complex baseband front end such as is used in existing digital spectrum analyzer systems at JPL (Refs. 1 and 2). In this case, a single-channel analog input with signal bandwidth B is converted into two in-phase and quadrature channels each with bandwidth $B/2$. Each channel is sampled at B samples/second and is treated as a complex signal throughout the rest of the system. This approach requires two AD converters each operating at the signal IF bandwidth as opposed to twice the IF bandwidth as would be required for real basebanding.

One limitation with this approach, however, is the generation of images as depicted in Fig. 2. These images are a consequence of phase and amplitude mismatching between the in-phase and quadrature local oscillators. In practice, it is difficult to exceed a 30-dB image rejection ratio (IRR). Consequently, strong RFI components which exceed the desired signal by 25–30 dB contaminate twice the RFI band.

An alternative approach is to convert the analog input into a single baseband channel with bandwidth B. This channel is sampled at $2B$ samples/second and the resulting samples are then split into even and odd pairs in the preprocessor stage. The even-odd sample pairs are treated as complex data by the FFT processor and the resulting output is then recombined in a special, "real adjust" FFT stage to produce the spectrum (Ref. 3, pp. 167–169). Since only one baseband channel is

involved, images created by amplitude and phase mismatch are eliminated.

The only limit to image rejection with this approach is numerical precision, which is most critical in the real adjust FFT stage. This stage effectively uses coherent digital subtraction to cancel images as depicted in Fig. 3. The complex signal, X_{RFI}, denotes an RFI spectral component and the two noise components, n_1 and n_2, arise from digital multiplication and addition operations in previous FFT stages. These components represent the limiting factors on image rejection.

Assuming that only the last FFT stage contributes significantly to these noise components, then the variance of n_1 and n_2 can be approximated by (Ref. 8):

$$\sigma_{n1}^2 = \sigma_{n2}^2 \cong \frac{2^{-2(b-1)}}{3} A_{RFI}^2$$

where A_{RFI} denotes the magnitude of X_{RFI} and b is the number of bits (including sign) used to represent the mantissa of the digital words in the FFT. Thus, the digital IRR is approximately:

$$2\,\frac{\sigma_{n1}^2}{A_{RFI}^2} \cong \frac{2}{3}\,2^{-2(b-1)}$$

For a 16-bit mantissa, image rejection is approximately 90 dB as opposed to 30 dB provided by complex basebanding. This is also borne out by computer simulation experiments. Based on these results, a real baseband IF stage is being seriously considered for system implementation.

IV. Digital Quantization Error Analysis

The intent of this analysis is to quantify the effects of quantization errors on system sensitivity and dynamic range. In this section, we describe the basic system model used in our analysis (Section IV-A). We then present some approximate analytical results (Section IV-B), which provide a basis for at least understanding trends in system performance over a wide class of input signal conditions. Finally, in Section IV-C, we present a summary of results stemming from an extensive simulation study.

A. System Model

In performing this analysis, we have utilized the system model depicted in Fig. 4. The additive noise components, n_Q and n_{FFT}, correspond to input AD quantization noise (n_Q) and roundoff noise (n_{FFT}) which originates from the FFT butterfly computations. The factor, 1/2, arises from scaling the input data to the FFT down one bit from the MSB to prevent overflows which might arise within the first butterfly stage (subsequent stages also perform scaling by 1/2).

In quantifying system dynamic range and sensitivity, we have used the following simplified complex input signal model:

$$x(t) = I(t) + s(t) + n(t)$$

where

$$I(t) = A_I e^{2\pi i f_I t} \equiv \text{interference}$$

$$s(t) = A_s e^{2\pi i f_s t} \equiv \text{desired signal}$$

$$n(t) \equiv \text{system noise}$$

The system noise component is assumed to be zero-mean, complex Gaussian noise with total power: $\langle n^2 \rangle = \sigma_0^2 = \alpha k T_s F_s$, where α is a system gain constant (which in practice will vary across the IF bandwidth), k is Boltzman's constant, T_s is the system temperature, and F_s is the sampling rate. It is also assumed that the interference and signal frequencies, f_I and f_s, are at cardinal frequencies of the FFT, and thus the effects of spectral leakage, which is really a separate issue, are not addressed here. (Reference 4 contains an extensive summary of window effects on spectral leakage.)

Even though this is a simplified input signal model, it permits us to develop several important parameters characterizing system performance. First, the FFT input signal-to-noise ratio prior to quantization is given by:

$$SNR_{in} = A_s^2/\sigma_0^2$$

In the absence of either FFT computational noise or input quantization errors, the output SNR from the FFT is given by:

$$SNR_o = N\,SNR_{in}$$

where $N = 2^L$ is the length of the FFT. With quantization errors, the actual output SNR, SNR_o', can be expressed as:

$$SNR_o' = N\,\frac{\sigma_s^2}{\sigma_0^2 + N(\sigma_{QN}'^2 + 4\,\sigma_{FFT}^2)}$$

where $\sigma_{QN}'^2$ denotes the variance of the AD quantization noise at the FFT output, and σ_{FFT}^2 denotes the variance of the FFT computational noise. (Approximate expressions for $\sigma_{QN}'^2$ and σ_{FFT}^2 will be given below in Section IV-B.) The ratio

of SNR_o to SNR_o', which is a good measure of degradation in system sensitivity due to quantization errors, is given by:

$$\rho \equiv \frac{SNR_o}{SNR_o'} = 1 + \frac{N}{\sigma_0^2}\left(\sigma'^2_{QN} + 4\sigma^2_{FFT}\right)$$

Another useful parameter characterizing system performance is the interference-to-input system noise ratio (INR) which is defined as:

$$INR \equiv A_I^2/\sigma_0^2$$

A related parameter is the interference-to-desired signal ratio (ISR):

$$ISR \equiv A_I^2/A_s^2$$

Both of these parameters can be used as measures of system dynamic range. For instance, the INR corresponding to a degradation in system sensitivity of 1 dB (ρ = 1.26) can be defined as the input system dynamic range. Similarly, the ISR corresponding to a 1-dB degradation can be defined as the system output or "two-tone" dynamic range.

Finally, we utilize the output noise standard deviation-to-mean ratio, σ:

$$\sigma = \frac{\sqrt{VAR(X_n^2)}}{\langle X_n^2 \rangle}$$

where X_n^2 denotes the magnitude squared of the complex FFT coefficient in a noise bin and $VAR(\cdot)$ denotes variance. In the absence of quantization errors (assuming no averaging over multiple bins or transforms):

$$\sigma = 1$$

As discussed in Section IV-B, quantization errors not only increase the average noise level in the output spectrum, thereby increasing ρ, but also increase the noise fluctuations due to the generation of narrowband noise "spurs." Furthermore, these spurs can not generally be reduced by incoherently averaging transforms. They can only be reduced by increasing the numerical precision used in the FFT arithmetic. Consequently, depending on the FFT wordlengths, these spurs may represent the limiting factor on system performance.

B. Input Quantization/Computational Noise Approximations

Regardless of the simplicity of our input signal model, an exact quantization error analysis for this class of inputs would be extremely complex due to the nonlinear interactions between the various noise components and would provide little insight into system performance. The problem is further complicated by the complex nature of the hybrid fixed/floating point FFT processor under consideration. This processor consists of P fixed point stages followed by $L - P$ floating point stages. Further, the floating point stages utilize floating point addition operations but may only use fixed point multiplication operations due to the fixed point representation of the twiddle factors (such a "quasi" floating point stage model has been used to evaluate existing candidate FFT hardware architectures).

Clearly, an exact analysis of such a complicated system would be very difficult. Consequently, we consider here a simplified analytical statistical model for the quantization errors which arises from a large body of existing literature (Refs. 5-10).[1] This model provides simple, approximate analytical expressions for the various system performance parameters described above in Section IV-A. These expressions are useful for at least predicting trends in system performance as evidenced by the simulation results presented in Section IV-C.

We first discuss the input quantization noise. In particular, it has been shown (Refs. 5-8 and Footnote 1) that for an input signal class consisting of sinewaves plus Gaussian noise, the spectral density of the quantization noise consists of various harmonics and intermodulation products related to the input sinewaves as well as a white (flat) spectral noise component. The amplitudes of the sinewave-related components are a function of both the number of quantizer bits and the levels of the input sinewaves. Typically these components are diminished by the presence of the additive system noise at the input to the quantizer (Footnote 1). Consequently, we assume that the dominant quantization noise component is the uncorrelated, spectrally flat noise component with variance given by (Ref. 8):

$$\sigma^2_{QN} = q_I^2/12$$

where

$$q_I = 2^{-(BI-1)}$$

and BI is the number of bits (including sign) used in the AD converter.

[1] Also see: Martin, D. R., and Secor, D. J., *High Speed Analog-to-Digital Converters in Communications Systems*, TRW internal report, Nov. 1981.

By virtue of this uncorrelated assumption, the variance of the quantization noise at the output of the FFT, σ'^2_{QN}, is reduced by a factor of $1/N$ from the input:

$$\sigma'^2_{QN} = \sigma^2_{QN}/N$$

Thus, the signal-to-noise ratio, ρ, defined in Section IV-A can be expressed in terms of σ^2_{QN} as follows:

$$\rho = 1 + \frac{1}{\sigma_0^2}(\sigma^2_{QN} + 4N\sigma^2_{FFT})$$

Alternatively, as a function of the interference-to-noise ratio, we have:

$$\rho = 1 + \frac{INR}{A_I^2}(\sigma^2_{QN} + 4N\sigma^2_{FFT})$$

As is seen, the levels of both quantization and FFT computational noise increase as INR increases for fixed A_I. Furthermore, for a given value of σ^2_{FFT}, the effects of computational noise become worse as the FFT length, N, increases whereas input quantization effects are independent of N. Consequently, for shorter transforms the limiting factor on system performance tends to be input quantization noise (regardless of the implementation) whereas roundoff noise starts to dominate as the transform size is increased depending on the implementation. These trends are demonstrated in the simulation results presented in Section IV-C.

The variance of the FFT computational noise, σ^2_{FFT}, has been derived for both fixed and floating point arithmetic (Refs. 8–10) although none of these derivations exactly model the combined effects of quantization/computation/coefficient noise. Nevertheless, we have used these results to derive approximate expressions which characterize system performance trends. Specifically, for a hybrid fixed/floating point architecture, we split the computational noise into two components: one arising from the first P fixed point stages and the other arising from the last L–P floating point stages, i.e.,

$$\sigma^2_{FFT} = \sigma^2_{FXP} + \sigma^2_{FLP}$$

where σ^2_{FXP} and σ^2_{FLP} denote the fixed and floating point noise variances, respectively. Utilizing methods summarized in Ref. 8, we have:

$$\sigma^2_{FXP} \cong \frac{4}{3} 2^{-2(BF-1)}\left(\frac{1}{2}\right)^{L-P}$$

where BF denotes the total number of bits (including sign) used to represent the fixed point FFT words.

Expressions for the floating point noise variance for a wide class of inputs (but assuming floating point multiplication operations) have been derived in Ref. 9. There it is shown that the floating point noise variance can be approximately split into two components: one related to the input system noise and the other related to the large level interference, i.e.,

$$\sigma^2_{FLP}(k) \cong \frac{2}{3} 2^{-2(BM-1)}(L-P)\sigma_0^2/4N + \sigma^2_{sp}(k)$$

where k denotes the FFT spectral bin number, BM is the number of bits (including sign) in the mantissa of the FFT floating point words, and σ^2_{sp} is the variance of the interference-related noise spur components. The system noise-related component is negligible compared to the other quantization noise components (being scaled by σ_0^2). The spur noise components are narrowband with spectral centers related to the interference frequency.

Based on the analysis presented in Ref. 9, it can be shown that the strongest noise spur occurs at either $k_0 + N/2$ if $k_0 < N/2$ (k_0 = interference spectral bin number) or at $k_0 - N/2$. The two next largest noise spurs (3 dB down) are separated by $N/4$ bins from the maximum spur, etc. The maximum spur level depends on the interference level and numerical precision via:

$$\sigma^2_{sp} \leqslant \frac{1}{3} 2^{-2(BM-1)}(A_I/2)^2$$

Using the above analytical expressions, we can now establish dynamic range constraints imposed by the various quantization noise components. To do this, we first expand ρ in terms of all the quantization noise variances given above, i.e.,

$$\rho = 1 + \frac{INR}{A_I^2}\sigma^2_{QN} + 4N\,INR\frac{\sigma^2_{FXP}}{A_I^2} + 4N\,INR\frac{\sigma^2_{sp}(k)}{A_I^2}$$

$$\equiv 1 + INR\,(E_1 + E_2 + E_3)$$

As discussed in Section IV-A, system input dynamic range can be defined as the INR which corresponds to a 1-dB degradation in system sensitivity, i.e., ρ=1.26. Thus the input dynamic range constraint imposed by any quantization noise source is given by:

$$INR_i = 0.26/E_i, \quad i = 1,2,3$$

Additionally, we can compute system output dynamic range constraints from:

$$ISR_i = INR_i\,(N/SNR_o)$$

A summary of dynamic range constraints is presented in Table 1 corresponding to 8-bit input quantization, 16-bit fixed point stages, and both 22-bit (16-bit mantissa, 6-bit exponent) and 32-bit (24-bit mantissa, 8-bit exponent) floating point stages. Also, we have fixed $SNR_o = 10$ (ideal output SNR) and $A_I = 0.75$. As is seen, AD quantization dominates all noise sources (smallest dynamic range constraint) for $N = 2^{14}$. As N increases to $N = 2^{20}$, AD quantization noise continues to dominate fixed point noise from the first butterfly stages. However, for 22-bit floating point stages, spur noise establishes the overall system dynamic range constraint for larger transform sizes ($N = 2^{18}$ or 2^{20}). For 32-bit floating point stages, AD quantization noise remains the dominant noise source for all transform sizes up to 2^{20}. These trends are observed in the simulation results presented in the next section.

C. Simulation Results

We have carried out extensive computer simulations of various hybrid arithmetic FFT systems with transform sizes up to 2^{18}. In addition, we have also obtained some limited hardware simulation results utilizing a 2^{14} point, 22-bit floating point FFT spectrum analyzer. In general, all of these simulation results support the analytical results presented in Section IV-B. As an example, consider the 32-bit floating point simulation results presented in Fig. 5. Here we display a segment (2000 bins) from a 2^{18} point power spectrum (no averaging) containing both interference and desired signal components. Two such segments are presented corresponding to 8-bit input quantization (Fig. 5(b)) and no input quantization, i.e., 32-bit floating point input words (Fig. 5(a)). For both cases, INR = 60 dB, $A_I = 0.75$, and $SNR_o = 10$ dB. Referring to Section IV-B (Table 1), we would expect that for this case input quantization will be the limiting factor on system performance inasmuch as FFT computational noise is relatively negligible. This is indeed the case as evidenced by the much higher noise floor (approximately 10 dB higher) in Fig. 5(b). (Note also the addition of noise spurs in Fig. 5(b) evidently due to input quantization effects.)

For hybrid 16-bit fixed point/22-bit floating point systems, the effects of computational noise can be clearly observed in the computer simulated data. For instance, consider the simulation results presented in Fig. 6 corresponding to a hybrid FFT architecture consisting of six 16-bit fixed point stages followed by twelve 22-bit floating point stages. Here we present a 2000-bin spectral segment resulting from incoherently averaging five successive transforms (no frequency averaging). This particular segment was chosen to contain the desired signal component as well as the dominant floating point noise spur (based on the analytical model presented in Section IV-B). Two such segments are presented corresponding to INR = 40 dB with 8-bit input quantization (Fig. 6(a)) and INR = 60 dB with 12-bit input quantization (Fig. 6(b)). For both cases, $A_I = 0.75$ and $SNR_o = 10$ dB. Based on the discussion in Section IV-B (Table 1), we would expect that even at INR = 40 dB, floating point noise spurs will represent the limiting factor on system performance. This is indicated in Fig. 6(a), although unambiguous signal detection at least over the 2000 bins displayed is still possible. However, the situation degrades significantly in Fig. 6(b) where noise spurs larger than the signal spectral level would prevent unambiguous signal detection.

A summary of further 2^{18} point FFT computer simulation experiments is presented in Table 2 corresponding to the hybrid 16-bit fixed point/22-bit floating point architecture again with $A_I = 0.75$ and $SNR_o = 10$ dB. The system parameters tabulated in Table 2, SNR_o' and σ, were obtained by frequency averaging over one transform. Specifically, SNR_o' was computed by averaging over 64 redundant signals uniformly spaced throughout the 2^{18} output spectral points and does not correspond to the worse-case output SNR at the noise spurs. Indeed, degradation in system performance is only evidenced by the computed values of σ which were obtained by averaging over all the 2^{18}-64-1 noise bins (the interference bin is excluded from computing σ). For INR = 60 dB and 12-bit input quantization, $SNR_o' \cong 9.3$ dB, indicating satisfactory system performance but $\sigma \cong 2.4$–3.0 dB revealing the influence of the noise spurs. Note that, for this case, increasing the number of fixed point butterfly stages from 4 to 6 leads to a 0.6-dB increase in σ, possibly due to interaction between fixed point and floating point noise components. A more ideal behavior is indicated for INR = 40 dB and 8-bit quantization; however, even in this case low level noise spurs are present in the output spectrum as noted above (Fig. 6(a)).

Finally, we present in Fig. 7 an output spectrum resulting from a hardware simulation of a 2^{14} point, 22-bit floating point FFT spectrum analyzer currently under development. This spectrum resulted from first generating a simulated complex input data set (on a MASSCOMP computer) consisting of a sinewave (amplitude = 0.375 at spectral bin 250) plus a low level of Gaussian noise with standard noise deviation approximately equal to the quantization level. This data was quantized to 15 bits and memory-mapped to the FFT hardware which consists of 14 butterfly stages and a data unscrambler stage to perform the bit reversal operation. The resulting spectral output (16,384 bins) was then decimated

by only keeping the largest spectral level in every 4 bins and then plotted.

As is seen in Fig. 7, numerous spectral noise peaks are present at or above the maximum levels predicted by either the general computer FFT simulator program or analysis. It should be noted, however, that a separate, special-purpose computer program written to exactly model this FFT hardware system did agree precisely with the hardware results. Clearly, these results indicate the presence of computational noise spurs resulting from the 22-bit floating point arithmetic although more detailed experiments utilizing multiple transform averages will be performed to further evaluate this system.

V. Conclusions

As a result of this systems analysis, several important conclusions emerge. First, spurious images due to the front-end baseband processor can be significantly reduced by utilizing real basebanding in combination with an additional real-adjust stage at the end of the FFT processor. Second, for 8-bit input quantization, which will most likely be used in implementing the prototype megachannel spectrum analyzer, system input dynamic range will be at most 45 dB and possibly less depending on whether 22-bit or 32-bit floating point arithmetic is used. Certainly the results of this analysis would support a 32-bit implementation. Third, 16-bit fixed point arithmetic can be used to implement the first butterfly stages without sacrificing system performance, provided the number of such stages is 8 or less. If the number of fixed point stages exceeds 8, then the resulting roundoff noise from these stages will begin to dominate the 8-bit input quantization noise. Finally, although the effects of window leakage on system performance were not addressed in this study, it is an important issue in designing a very large FFT. For instance, an input INR = 40 dB corresponds to a spectral dynamic range of 100 dB when $N = 2^{20}$. Thus, window sidelobes should be at least 100–110 dB down near the desired signals or else spectral leakage from the large level interference will prevent signal detection. Clearly, the particular window used for the megachannel system must be carefully chosen, i.e., using the results of previous detailed analysis on FFT window design (e.g., as summarized in Ref. 4).

References

1. Quirk, M. P., Wilck, H. C., and Grimm, M. J., "A Wide-Band, High-Resolution Spectrum Analyzer," *TDA Progress Report 42-83*, pp. 180–190, Jet Propulsion Laboratory, Pasadena, CA, 1985.

2. Morris, G. A., and Wilck, H. C., "JPL 2^{20} Channel 300 MHz Bandwidth Digital Spectrum Analyzer," *1978 IEEE Int. Conf. Acoust., Speech, and Signal Processing*, April 10-12, 1978, Tulsa, OK, pp. 808–811.

3. Brigham, E. O., *The Fast Fourier Transform*, Prentice-Hall, 1974.

4. Harris, F. J., "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proc. IEEE*, Vol. 66, pp. 51–83, Jan. 1978.

5. Bennett, W. R., "Spectra of Quantized Signals," *Bell Sys. Tech. J.*, Vol. 27, pp 446–472, July 1948.

6. Widrow, B., "Statistical Analysis of Amplitude-Quantized Sampled-Data Systems," *Trans. AIEE*, Vol. 79, pp. 555–568, Jan. 1961.

7. Hurd, W. J., "Correlation Function of Quantized Sine Wave Plus Gaussian Noise," *IEEE Trans. Infor. Theory*, Vol. 13, pp. 65–68, Jan. 1967.

8. Oppenheim, A. V., and Schafer, R. W., *Digital Signal Processing*, Chapter 9, Prentice-Hall, 1975.

9. Kaneko, T., and Liu, B., "Accumulation of Round-Off Error in Fast Fourier Transforms," *J. Assoc. Comput. Machin.*, Vol. 17, pp. 637–654, Oct. 1970.

10. Weinstein, C. J., "Roundoff Noise in Floating Point Fast Fourier Transform Computation," *IEEE Trans. Audio, Electroacoust.*, Vol. 17, pp. 209–215, Sept. 1969.

Table 1. Quantization noise contributions to system dynamic range

Noise Source	N	INR, dB	ISR, dB
AD Quantization, σ^2_{QN} (8-bit ADC)	2^{14}	45	77
	2^{18}	45	89
	2^{20}	45	95
Fixed Point Roundoff, σ^2_{FXP} (16 bit)	2^{14}	62 (P=4)	94 (P= 4)
	2^{18}	62 (P=4)	106 (P=4)
	2^{20}	56 (P=6)	106 (P=6)
	2^{20}	50 (P=8)	100 (P=8)
Floating Point Roundoff, $(\sigma^2_{sp})_{max}$ (22 bit)	2^{14}	47	79
	2^{18}	35	79
	2^{20}	29	79
Floating Point Roundoff, $(\sigma^2_{sp})_{max}$ (32 bit)	2^{14}	95	127
	2^{18}	83	127
	2^{20}	77	127

Table 2. Summary of hybrid 16-bit fixed point/22-bit floating point computer simulation results for $N = 2^{18}$

Number of AD Bits	INR, dB	P	SNR'_o, dB	σ, dB
8	40	6	9.0	0
12	60	4	9.3	2.4
12	60	6	9.4	3.0

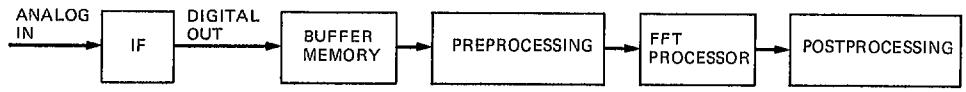

Fig. 1. System block diagram

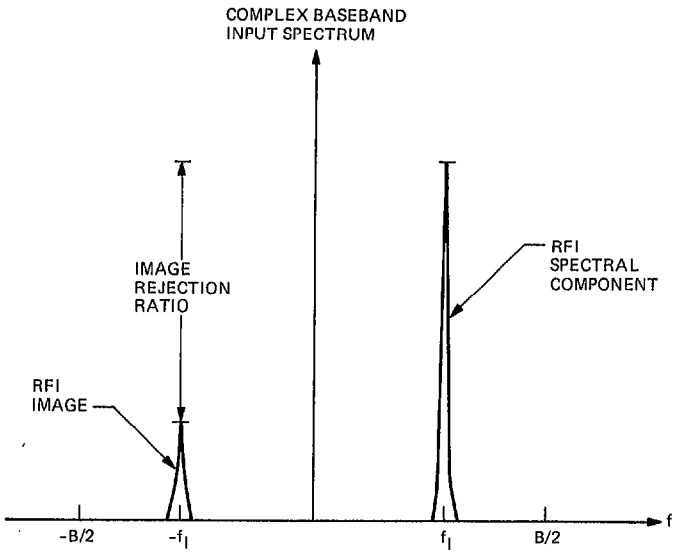

Fig. 2. Image generated by complex basebanding

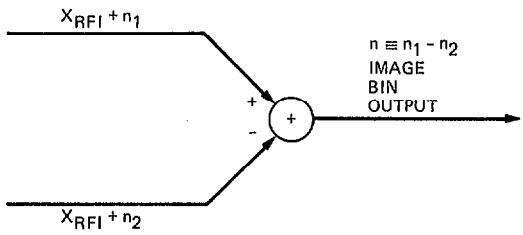

Fig. 3. Model of image cancellation in the real adjust stage

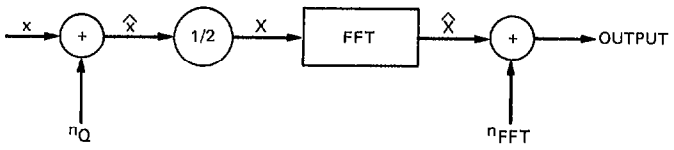

Fig. 4. Basic system model

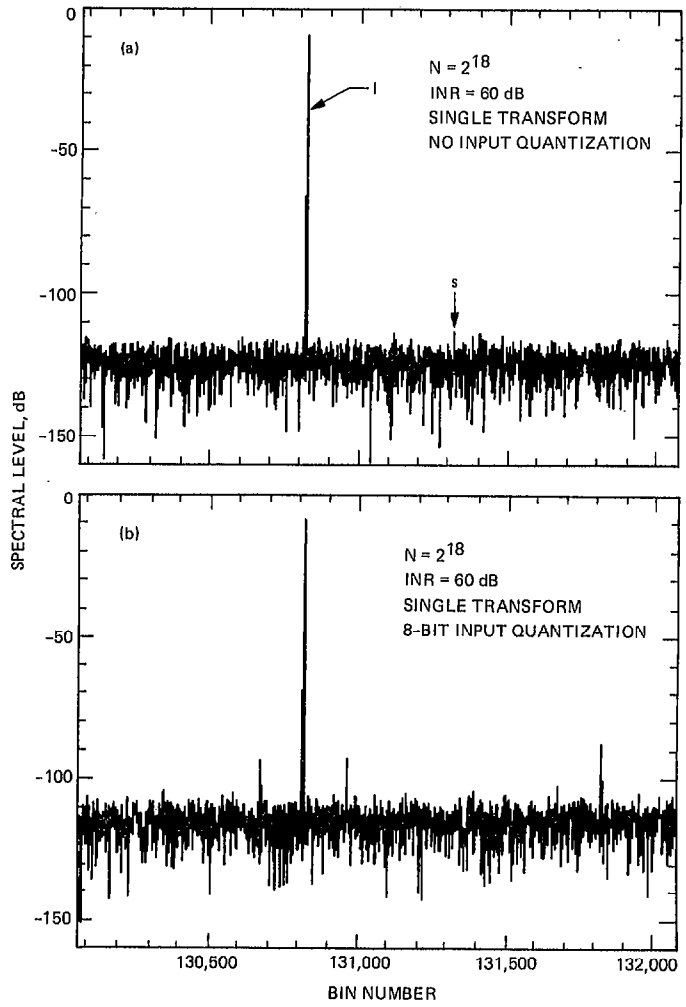

Fig. 5. Computer simulation results for 32-bit floating point FFT arithmetic

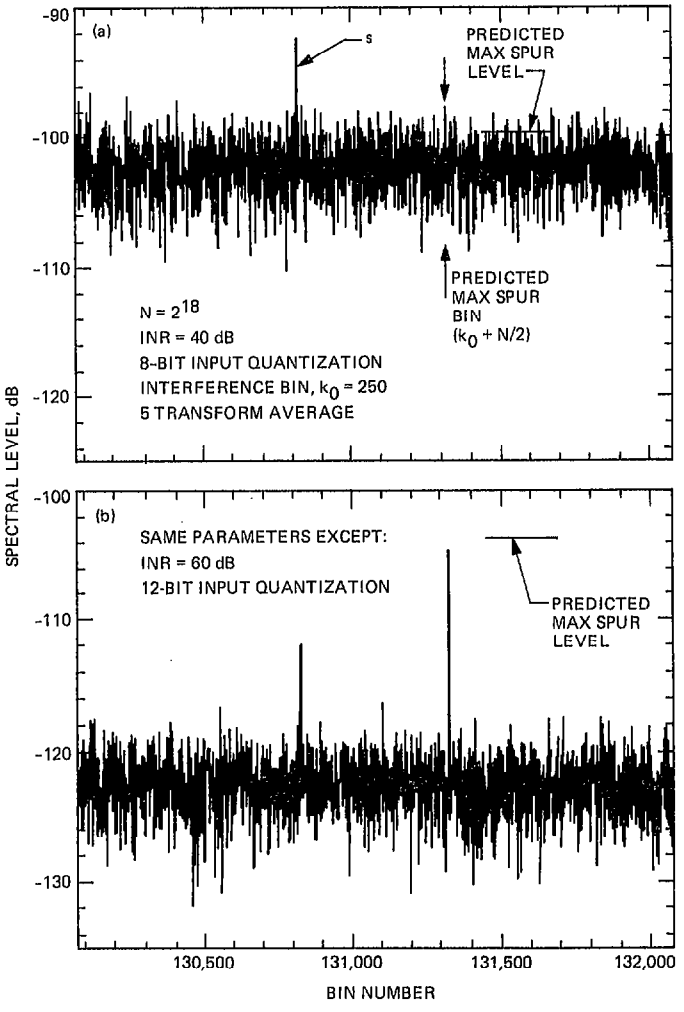

Fig. 6. Computer simulation results for hybrid 16-bit fixed point/22-bit floating point FFT arithmetic with 6 front end fixed point stages

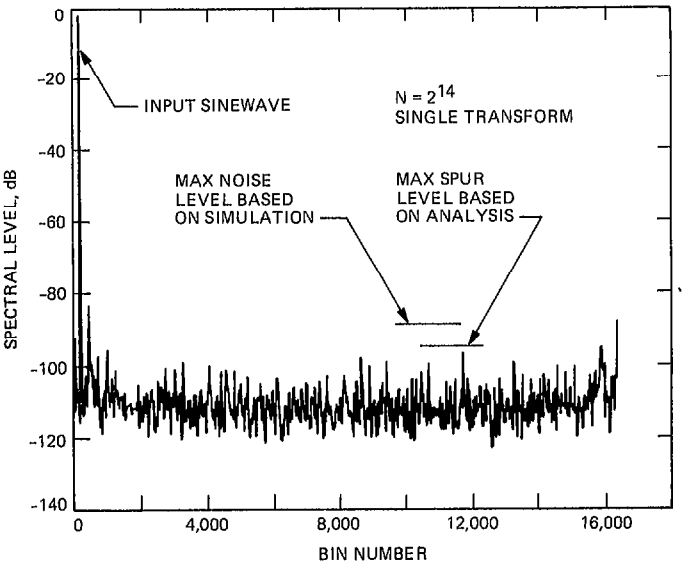

Fig. 7. Hardware simulation for 22-bit floating point system