[5] H. Kobayashi, "Simultaneous adaptive estimation and decision algorithm for carrier-modulated data transmission systems," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 268-280, June 1971.
[6] G. Ungerboeck, "Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems," *IEEE Trans. Commun.*, vol. COM-22, pp. 624-636, May 1974.
[7] D. L. Lyon, "Timing recovery in synchronous equalized data communication," *IEEE Trans. Commun.*, vol. COM-23, pp. 269-274, Feb. 1975.
[8] G. Ungerboeck, "Theory on the speed of convergence in adaptive equalizers for digital communication," *IBM J. Res. Develop.*, vol. 16, pp. 546-555, Nov. 1972.
[9] J. W. Mark and P. S. Budihardjo, "Joint optimization of receive filter and equalizer," *IEEE Trans. Commun.*, vol. COM-21, pp. 264-266, Mar. 1973.
[10] J. Monrolin, H. Nussbaumer, and J. M. Pierret, "Perfectionnements aux systèmes de detection de données distordues," French Patent No. 7131079, Mar. 19, 1973.
[11] L. Guidoux, "Egaliseur autoadaptif à double échantillonnage," *L'Onde Electrique*, vol. 55, pp. 9-13, Jan. 1975.
[12] S. U. H. Qureshi, "Adjustment of the position of the reference tap of an adaptive equalizer," *IEEE Trans. Commun.*, vol. COM-21, pp. 1046-1052, Sept. 1973.
[13] K. H. Mueller and D. A. Spaudling, "Cyclic equalization–A new rapidly converging adaptive equalization technique for synchronous data communication," *Bell Syst. Tech. J.*, vol. 54, pp. 369-406, Feb. 1975.

★

Gottfried Ungerboeck (M'75) was born in Vienna, Austria, in 1940. He received the Dipl. Ing. degree in telecommunications from the Technical University, Vienna, Austria, in 1964, and the Ph.D. degree from the Swiss Federal Institute of Technology, Zurich, Switzerland, in 1970.

After graduating from the Technical University, he was employed by the Wiener Schwachstromwerke, Vienna, Austria. In 1965 he joined IBM Austria, where he worked as a Systems Engineer. Since 1967 he has been with the IBM Zurich Research Laboratory, Rüschlikon, Switzerland. He worked in speech processing, voice and data switching, and data transmission theory. His current interests cover signal processing, detection and estimation theory, and coding.

Nearly Instantaneous Companding for Nonuniformly Quantized PCM

DONALD L. DUTTWEILER, MEMBER, IEEE, AND DAVID G. MESSERSCHMITT, MEMBER, IEEE

***Abstract*—The technique of nearly instantaneous companding (NIC) that we describe processes n-bit μ-law or A-law encoded pulse-code modulation (PCM) to a reduced bit rate. A block of N samples (typically $N \cong 10$) is searched for the sample having the largest magnitude, and each sample in the block is then reencoded to a nearly uniform quantization having $(n - 2)$ bits and an overload point at the top of the chord of the maximum sample. Since an encoding of this chord must be sent to the receiver along with the uniform reencoding, the resulting bit rate is $f_s(n - 2 + 3/N)$ bits/s where f_s is the sampling rate.**

The algorithm can be viewed as an adaptive PCM algorithm that is compatible with the widely used μ-law and A-law companded PCM. Theoretical and empirical evidence is presented which indicates a performance slightly better than $(n - 1)$ bit companded PCM (the bit rate is close to that of $(n - 2)$ bit PCM). A feature which distinguishes NIC from most other bit-rate reduction techniques is a performance that is largely insensitive to the statistics of the input signal.

I. INTRODUCTION

THERE has been much recent interest in reduced-bit-rate waveform encoding of analog signals. Most attention has focused on adaptive delta modulation (ADM) and adaptive differential pulse-code modulation (ADPCM) (see [1] for a good review and extensive bibliography) with lesser attention given to adaptive PCM (APCM) [2], [3].

In this paper we describe a scheme for processing 15-segment μ-law or 13-segment A-law PCM [4] to achieve a reduced bit rate. This scheme, which we call nearly instantaneous companding (NIC), imitates APCM. Specifically, the NIC coder (see Fig. 1) groups the n-bit μ-law or A-law PCM encodings at its input into groups of N and looks for the sample having the largest magnitude in the group. It then reencodes each sample in the group to an approximately $n - 2$ bit uniform quantization with overload at the top of the chord of the maximum-magnitude sample. These $n - 2$ bit uniform encodings are transmitted to a complementary NIC decoder that regenerates n-bit codes in the original companding law. Since three scaling bits giving the chord of the maximum-magnitude sample must also be transmitted to the decoder, the resulting bit rate is $f_s(n - 2 + 3/N)$ where f_s is the sampling rate. This bit rate is comparable to that of $n - 2$ bit PCM for even modestly large N.

The dynamic range of the original PCM companding law is retained through the adaptive overload point, as is the relative PCM insensitivity to signal statistics. To illustrate these points, we show the NIC signal-to-quantization noise ratio (SNR) with $n = 8$ and three sets of signal statistics in Figs. 2, 3, and 4. The methods used to obtain these graphs will be discussed later. Fig. 2 shows the sinusoid performance for both μ255

Paper approved by the Editor for Transmission Systems of the IEEE Communications Society for publication without oral presentation. Manuscript received November 19, 1975; revised January 7, 1976.

The authors are with Bell Laboratories, Holmdel, NJ 07733.

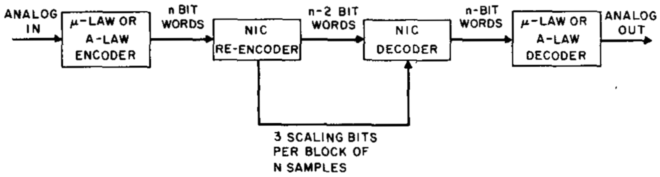


Fig. 1. Processing n-bit PCM with the NIC algorithm.

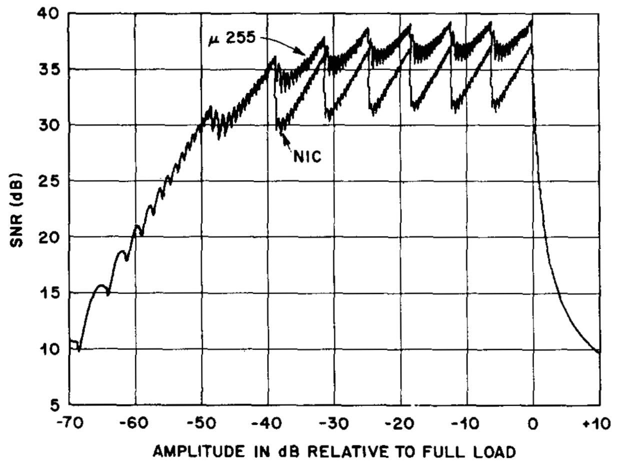


Fig. 2. SNR versus amplitude for μ255 PCM and NIC with a sinusoidal input. Flat noise weighting is assumed. C-message weighting would add 2.9 dB.

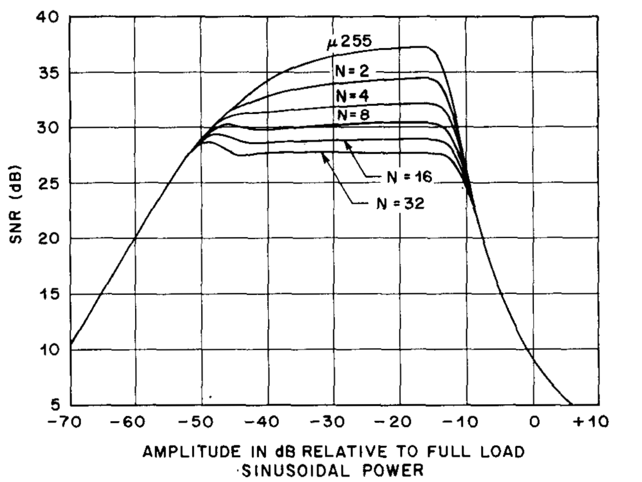


Fig. 3. SNR versus amplitude with independent Laplacian samples. Flat noise weighting is assumed. C-message weighting would add 2.9 dB.

PCM and NIC. The curve for NIC is a lower bound that is a good approximation to the true SNR when the sinusoid frequency is high enough so that the NIC block includes at least half a period (the SNR actually improves slightly at lower frequencies). The maximum NIC degradation relative to μ255 PCM is about 4 dB. Figs. 3 and 4 show the SNR performance of NIC with independent samples having, respectively, Laplacian and Gaussian distributions. The performance now depends more strongly on the block size N. At $N = 8$, the degradation is about 7 dB with a Laplacian distribution and 6 dB with a Gaussian.

The Laplacian distribution assumed for Fig. 3 is characteristic of speech, but speech samples are not independent. Y. C. Ching has simulated NIC with an actual speech input and found a degradation with $N = 8$ of 3.5 dB.

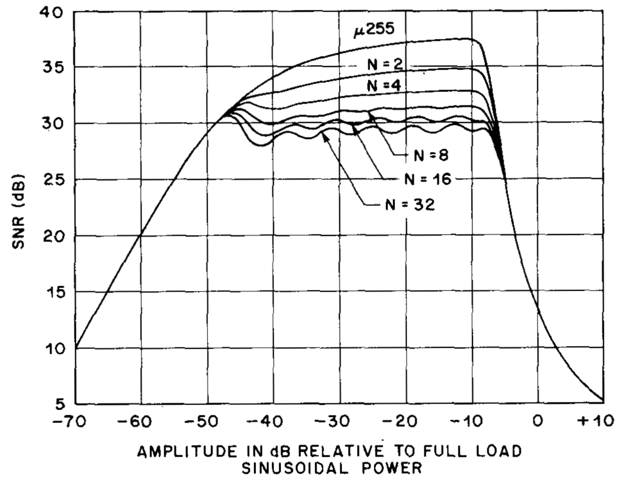


Fig. 4. SNR versus amplitude with independent Gaussian samples. Flat noise weighting is assumed. C-message weighting would add 2.9 dB.

A form of NIC which processes uniformly quantized samples using the same block structure has been previously described by Osborne [5] (see also Croll, Moffat, and Osborne [6]) and Croisier [7]. Our algorithm differs from theirs in that it processes μ-law or A-law encoded samples, making it directly and easily convertible to these widely used companding laws. We adopt the terminology of [6] in calling the algorithm NIC. Croisier used the equally descriptive name "block companding." We have previously described NIC in [8].

The remainder of the paper considers explicitly only μ255 PCM processed by NIC. The NIC algorithm for processing A-law PCM is identical, but the characteristics and performance of the resulting code are slightly different.

II. DESCRIPTION OF NIC

In this section we first describe briefly the μ255 companding law (more detail will be found in [4]). We then describe the NIC translation from a μ255 encoded sample to a uniformly encoded sample given scaling information, and finally the determination of the scaling information.

A. The μ255 Companding Law

We describe here the 15-segment μ255 companding law, where for concreteness we assume 8-bit quantization ($n = 8$).[1] The 8-bit sample is conveniently denoted ($sabcwxyz$) where (s) is the sign bit, (abc) is the binary representation of the "segment" or "chord" L ($0 \leqslant L \leqslant 7$), and ($wxyz$) is the binary representation of the level $V(0 \leqslant V \leqslant 15)$ on that segment. The positive output levels (the companding law is symmetric) are given by [4]

$$\hat{X}(L,V) = 2^L(V + 16.5) - 16.5 \tag{1}$$

with corresponding decision region for sample x of

$$2^L(V + 16) - 16.5 \leqslant x < 2^L(V + 17) - 16.5 \tag{2}$$

[1] In addition, we assume mid-tread bias and decision level assignment. Extensions are straightforward.

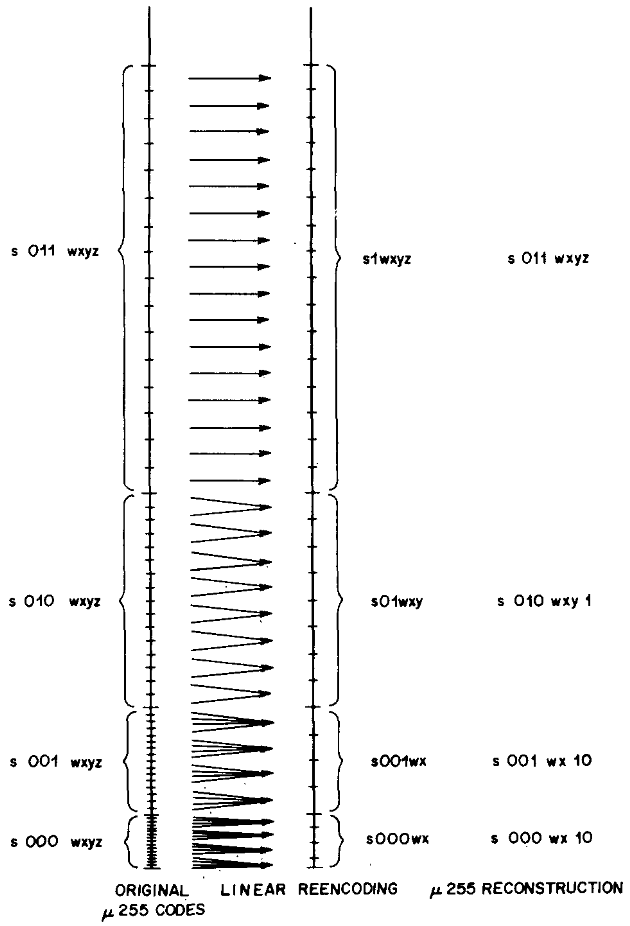


Fig. 5. Translation from 8-bit $\mu 255$ codes to 6-bit uniform codes.

(an exception is $L = V = 0$, where the lower limit is zero). Note from (1) that the 16 output levels on a given segment are equally spaced and that these spacings double from segment L to segment $(L + 1)$. The rapidly increasing spacings (from 1 to 128) provide a wide dynamic range of signal levels for which the SNR is essentially constant. The object of NIC is to reduce the bit rate while retaining this dynamic range.

B. Uniform Reencoding

The idea behind NIC is to transmit the scaling information, comparable to the segment L, only once for each block of N samples rather than with each sample. In this section we describe a natural mapping from an 8-bit $\mu 255$ sample to a 6-bit uniformly quantized sample, where the overload point of the uniform encoding is on the top of segment L_{OVLD}. Similar mappings from $n \neq 8$ bit $\mu 255$ PCM to $n - 2$ bit uniform can be constructed.

The reencoding for $L_{OVLD} = 3$ is illustrated in Fig. 5. The sign bit s is always retained, while the next most significant bit is one for $L = L_{OVLD} = 3$ and zero otherwise. If $L = 3$, the next four bits retain $(wxyz)$, and none of the $\mu 255$ resolution is lost. However, if $L < 3$, the third NIC bit is devoted to specifying whether $L = 2$ or $L < 2$, and the least significant bit z of the $\mu 255$ code is lost. Similarly, if $L < 2$, the fourth NIC bit specifies if $L = 0$ or $L = 1$, and two least significant bits of the $\mu 255$ code are lost. Hence, for $L = 2$ and $L < 2$ the uniformly reencoded samples have a step-size double and quadruple, respectively, that of the $\mu 255$ sample.

In reconstruction, the NIC code and the scaling information L_{OVLD} are used to generate the $\mu 255$ output level closest to the center of the region specified by the NIC code. The resulting output levels, shown on the right side of Fig. 5, are uniformly spaced except for $L = 0$, where the step-size is half as large. The NIC code ($S0001W$) on this segment would eliminate this peculiarity, but would be less efficient in not fully utilizing the 6 NIC bits. It is also interesting to note that this peculiarity does not arise when the original codes are A-law rather than μ-law.

A generalization of the translation of Fig. 5 for other L_{OVLD} is given in Table I. Note that when L_{OLVD} is large, whole segments of $\mu 255$ codes are mapped into a single NIC code. Expressions for the effective decision and output levels after NIC translation are given in Appendix A for $4 \leqslant n \leqslant 8$.

C. Determination of L_{OVLD}

In the introduction we described NIC as using the linear reencoding of the last section with L_{OVLD} for each reencoded sample equal to $L_{\max}$, the largest chord number in the block of that sample. Although by NIC we will always mean the linear reencoding of the last section with $L_{OVLD} = L_{\max}$, this reencoding is potentially useful with other schemes for setting L_{OVLD}. All that is really required is that L_{OVLD} be as small as possible consistent with there being a small probability of overload.

Schemes for determining L_{OVLD} can be divided according to whether they are forward or backward [9]. Forward schemes, as illustrated in Fig. 6, look at the input to the reencoder to determine L_{OVLD}. Backward schemes (Fig. 7) look at its output. Backward schemes have the advantage that L_{OVLD} need not be separately transmitted to the receiver since it is determinable there from the sample encodings also.

Forward acting schemes can be further divided according to whether they determine L_{OVLD} for each sample causally or noncausally. In the latter case, delay must be introduced. Too much delay is undesirable, as will be discussed in Section V-D.

The NIC algorithm, $L_{OVLD} = L_{\max}$, has as disadvantages the fact that it is forward acting, requiring separate transmission of L_{OVLD}, and the fact that it is noncausal, requiring delay. Its advantages are that it is simple and that it is efficient in setting L_{OVLD}. There is never any overload and at least one sample in every N is on segment L_{OVLD}.

D. Intuitive Interpretations of NIC

There are several motivations for NIC which are useful in providing understanding. We list them below.

1) NIC is an adaptive PCM scheme, or digital automatic gain control, in which the quantization is uniform (or nearly so) and the step-size is varied adaptively in accordance with the short-term signal level.

2) NIC is a form of companding midway between instantaneous and syllabic.

3) The chord information for each $\mu 255$ sample is redun-

TABLE I
TRANSLATION BETWEEN μ255 CODES AND NIC CODES

L_{MAX}	ORIGINAL	ENCODING	RECONSTRUCTION
7	s111wxyz	s1wxyz	s111wxyz
7	s110wxyz	s01wxy	s110wxy1
7	s101wxyz	s001wx	s101wx10
7	s100wxyz	s0001w	s100w100
7	s011wxyz	s00001	s0111000
7	s010wxyz	s00000	s0100000
7	s001wxyz	s00000	s0100000
7	s000wxyz	s00000	s0100000
6	s110wxyz	s1wxyz	s110wxyz
6	s101wxyz	s01wxy	s101wxy1
6	s100wxyz	s001wx	s100wx10
6	s011wxyz	s0001w	s011w100
6	s010wxyz	s00001	s0101000
6	s001wxyz	s00000	s0010000
6	s000wxyz	s00000	s0010000
5	s101wxyz	s1wxyz	s101wxzy
5	s100wxyz	s01wxy	s100wxy1
5	s011wxyz	s001wx	s011wx10
5	s010wxyz	s0001w	s010w100
5	s001wxyz	s00001	s0011000
5	s000wxyz	s00000	s0001000
4	s100wxyz	s1wxyz	s100wxyz
4	s011wxyz	s01wxy	s011wxy1
4	s010wxyz	s001wx	s010wx10
4	s001wxyz	s0001w	s001w100
4	s000wxyz	s0000w	s000w100
3	s011wxyz	s1wxyz	s011wxyz
3	s010wxyz	s01wxy	s010wxy1
3	s001wxyz	s001wx	s001wx10
3	s000wxyz	s000wx	s000wx10
2	s010wxyz	s1wxyz	s010wxyz
2	s001wxyz	s01wxy	s001wxy1
2	s000wxyz	s00wxy	s000wxy1
1	s001wxyz	s1wxyz	s001wxyz
1	s000wxyz	s0wxyz	s000wxyz

dant since it is indicative of speech power and speech power varies at a relatively slow syllabic rate. NIC extracts some common characteristic of the entire block, and transmits it only once for the block.

4) It seems unnecessary to finely quantize a small sample when a neighboring sample is quite large and hence more coarsely quantized. NIC quantizes each with similar accuracy, assuming they are in the same block.

III. PERFORMANCE

We give in this section analytic and empirical results on the SNR of NIC. Aside from giving analytical results for a sinusoid (important because it can be conveniently measured) and sequences of independent samples of arbitrary probability density, we will give empirical results for speech and voiceband data signals.

A. Sinusoidal Signal

The minor undulations of the SNR for a sinusoid (Fig. 2) are due to the fine structure of the quantizer in relation to the peak of the sinusoid. The major trend, on the other hand, is pictured in Fig. 8, and is interpreted as follows: When the peak is near the top of a chord, and the NIC block contains at least a half period of the sinusoid, the quantization is nearly 6-bit uniform with overload near the peak, which results in an SNR of 37.9 dB [10]. (We are assuming flat noise weighting in the band $[0, f_s/2]$. The SNR with 8-kHz sampling and C-message noise weighting is 2.9 dB greater.) As the sinusoid peak decreases within the same chord, the quantizer and hence the noise power remain fixed (because the quantization is uniform), so that the SNR decreases decibel for decibel with the signal power. The minimum SNR of 31.9 dB is representative of 5-bit uniform quantization, with again the overload matched to the peak. Eventually, for $L_{\max} = 1$, the SNR is representative of the μ255 law itself.

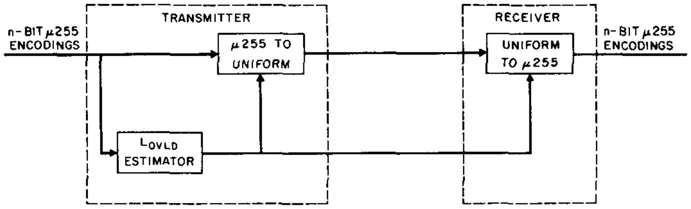


Fig. 6. Forward structure for determining L_{OVLD}.

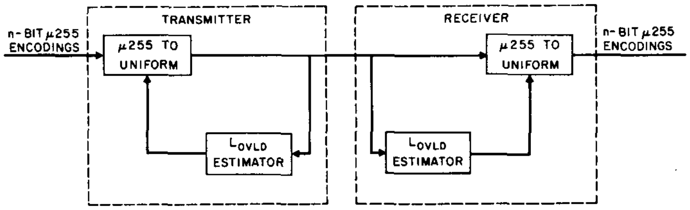


Fig. 7. Backward structure for determining L_{OVLD}.

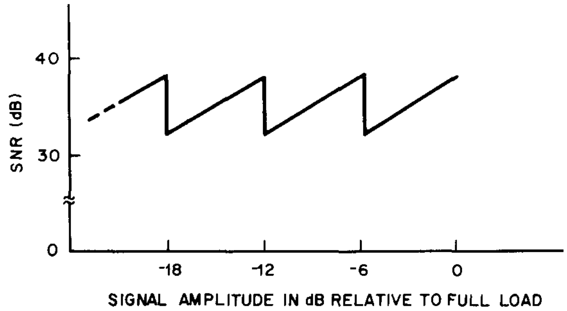


Fig. 8. Approximate SNR curve for 6-bit NIC. C-message noise weighting adds 2.9 dB.

The SNR of Figs. 2 and 8 is worst case in the sense that it is assumed that the maximum chord equals that chord on which the peak of the sinusoid is coded. This is a good approximation except in the extreme case where the sinusoid frequency is so low that there is less than 1/2 cycle in a NIC block.

B. Independent Samples With Arbitrary Density

A sinusoid signal has a well-defined peak value, leading to a relatively simple performance characterization. Such is not the case for densities with tails extending to infinity, such as the Laplacian (characteristic of speech) and Gaussian, for which it can be said with certainty only that $L_{\max} \leqslant 7$.

Let $x_1, x_2, \cdots, x_N$ denote a block of N input samples from a stationary process that is to be PCM encoded, processed by NIC, and finally converted back to analog with a PCM decoder. A general formula for the resulting SNR at the output is

$$\mathrm{SNR} = 10 \log_{10} EX_k{}^2/P \tag{3}$$

where E denotes expectation and

$$P = \min_G \frac{1}{N} \sum_{k=1}^{N} E(GX_k - \hat{X}(X_k, L_{\max}))^2. \tag{4}$$

The quantization $\hat{X}(X_k, L_{\max})$ is the deterministic function of X_k and $L_{\max}(X_1, \cdots, X_N)$ that is defined by Table I and (1) and (2). Miminizing over G in (4) allows for an average gain without penalty. The signal power $EX_k{}^2$ in (3) is independent of k under the stationarity assumption. Averaging over sample position in the calculation of the noise power P is necessary because in general the joint probability density of a sample X and the maximum chord $L_{\max}$ of its block are dependent on the position of that sample in the block.

Evaluating (3) for a speech input is thwarted by a lack of suitable characterization of the joint probability density function of N consecutive speech samples as much as by the cumbersome nature of (4). To proceed analytically, we will assume independent samples. Although this is a poor assumption for a speech input, which is the input of greatest interest, the results obtained still give insight into the performance of NIC as N varies.

With independent samples, averaging over sample position as in (4) is no longer necessary, so that we obtain the simpler formula

$$P = \min_G E(GX - \hat{X}(X, L_{\max}))^2. \tag{5}$$

The key step is then to find the joint density of X and $L_{\max}$, which is accomplished in Appendix B. The resulting SNR has been previously portrayed in Figs. 3 and 4. We would expect the SNR for speech to be slightly greater because of a large sample-to-sample correlation, making the maximum sample somewhat more representative of all the samples in the block. The performance for independent samples is, however, again indicative of the relative insensitivity of NIC performance to signal statistics.

C. SNR Versus N

Another interesting way of presenting SNR data is as a graph of SNR versus the average number of bits per sample as the block size N varies. To make such graphs it is necessary to assume a particular signal level, but as indicated by the flat tops of the plots in Figs. 3 and 4, the result is not particularly sensitive to signal level as long as a level within the dynamic range of the μ-law is choosen. Two plots of this type appear in Fig. 9. One assumes independent Laplacian samples and the other is based on Y.C. Ching's NIC simulations with actual speech. For both, the signal power is 20 dB below full load sinusoidal power. Also plotted in Fig. 9 is a line connecting the SNR for 6-, 7-, and 8-bit μ255 PCM. The maximum advantage of NIC is 3 dB with independent Laplacian samples and 6 dB with actual speech. In both cases the maximum advantage occurs at about $N = 10$.

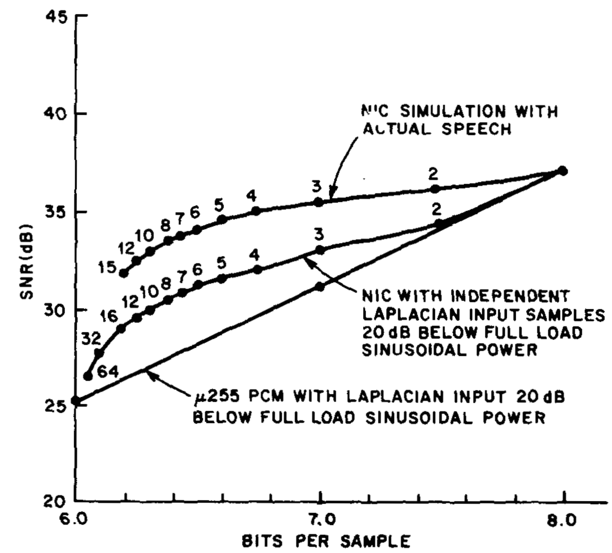


Fig. 9. SNR versus bits/sample. C-message noise weighting adds 2.9 dB.

D. Subjective Effects

There is not necessarily a one-for-one relationship between SNR and the subjective appraisal of an algorithm. Subjective tests of NIC are in progress, but no final results are available yet. Informal listening tests revealed no surprises: the subjective improvement of NIC over PCM at the same bit rate was evident at bit rates below 48 kbits/s where quantization noise is perceptible, and no marked dependence on block size (within the range $2 \leqslant N \leqslant 32$) was observed.

E. Voice-Band Data Performance

The good NIC performance for sinusoidal signals leads to the expectation of similarly good performance for voice-band data signals, which also have a well defined peak level. Measurements of voice-band data error rates were made for a Western Electric 208A data modem at 4800 bits/s by J. H. Fennick of Bell Laboratories. These measurements indicate that both NIC and PCM are well characterized as to their error-rate performance in the presence of various tandem impairments such as delay distortion, harmonic distortion, and Gaussian noise by an equivalent additive thermal noise, which is plotted in Fig. 10 as a function of the average number of transmitted bits per sample.[2] Again we see a superior NIC performance at a given bit rate (typically about 3 dB higher equivalent SNR.)

IV. IMPLEMENTATION

Consider the conversion between parallel 8-bit PCM words and parallel 6-bit NIC words. Implementing the PCM to NIC encoder requires a bank of 8 N-bit shift registers for word storage while the maximum chord is being determined and additional logic to perform the actual code conversion of

[2] 8-bit PCM was difficult to characterize in this manner because of its high quality, while 5-bit PCM reduced to 3 bits by NIC could not support error-free transmission even in the absence of tandem impairments.

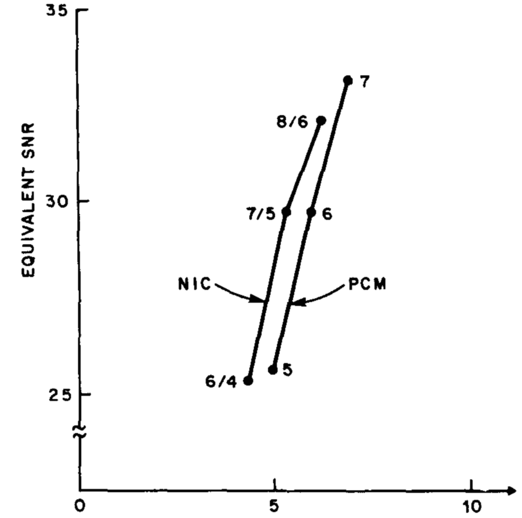


Fig. 10. Voice-band data performance. The points labeled $n/n-2$ on the NIC curve are for n-bit PCM reduced to $n-2$ bits by NIC.

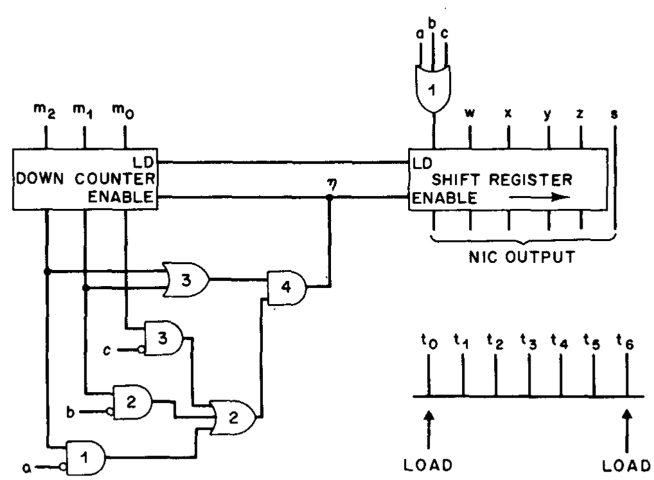


Fig. 11. PCM-to-NIC code translation.

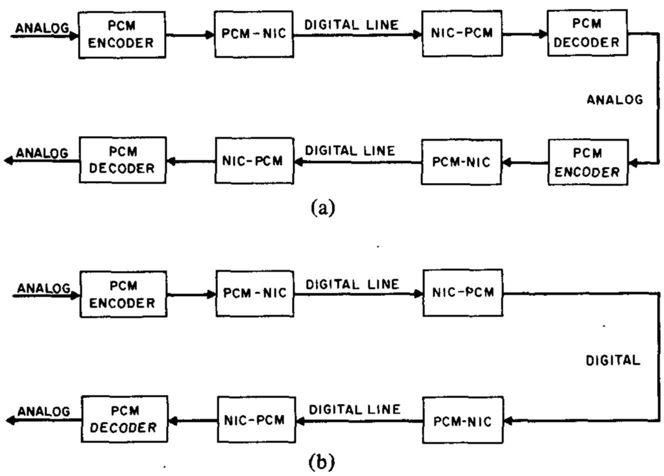


Fig. 12. Two types of tandem NIC encoding. (a) Tandem NIC/A/NIC encoding. (b) Tandem NIC/PCM/NIC encoding.

Table I. The corresponding NIC-to-PCM decoder only requires the code conversion logic. In an encoder/decoder combination the gate count of both code converters together is insignificant in comparison to that of the shift register bank.

While the conversion logic requires insignificant hardware, it is still conceptually the more interesting. Fig. 11 illustrates the logic required for generating a NIC code from an 8-bit PCM encoding (s abc $wxyz$) and the maximum chord ($m_2m_1m_0$). In a complete PCM-to-NIC translator, the PCM encoding would be taken from the output of the shift register bank and the maximum chord information, from a latch scanning the "abc" bits at the input to the shift register bank for a maximum. The code converter consists of a 3-bit down-counter (CTR) and 5-bit shift-register (SR). Initially CTR is loaded with ($m_2m_1m_0$) and SR with ($ewxyz$) where e, generated by OR_1, is zero only if $(abc) = 0$. CTR then counts down until either: 1) its output equals (abc) as detected[3] by AND_1, AND_2, AND_3, and OR_2; or 2) CTR reaches (001) or (000) as detected by OR_3. Counting and shifting are disabled at this point because the output of AND_4 goes low and CTR and SR only count and shift while the enable input (EN) is high. At the end of five clock pulses the NIC code is available in parallel at the SR output. Circuitry for converting from NIC codes to PCM is similarly simple.

V. SYSTEM IMPLICATIONS

A. Effect of Channel Errors

With NIC the important scaling information (i.e., chord information) is transmitted once per block, rather than with every sample as in PCM. In both NIC and PCM every sample contains a sign bit. These four bits are the most vulnerable to errors. The scaling information will be affected less often with NIC, but when it is in error an entire block of samples is affected rather than just one. The question of whether NIC or PCM is affected the most by channel errors therefore comes down to whether it is subjectively more desirable to have fewer "pops" with greater energy as with NIC, or more frequent "pops" with less energy as with PCM. Extensive subjective tests have not been attempted, but laboratory listening suggests strongly that NIC is subjectively more pleasing than PCM at the same error rate for error rates as high as 10^{-3}.

On a more objective basis, experimental measurement of the ratio of voice band data error rate to channel error rate have indicated a ratio of seven for PCM and five for NIC,[4] showing slightly less sensitivity of NIC-encoded voice band data to channel errors. (This difference would be insignificant in practice.)

B. Tandem Encodings

In discussing tandem NIC encodings it is necessary to distinguish the two types shown in Fig. 12. In a tandem NIC/A/NIC encoding, an intermediate analog link with its associated PCM decoder and encoder is involved. In contrast, a tandem NIC/PCM/NIC encoding involves an intermediate

[3] A full comparator is not required since CTR always counts down.

[4] A Western Electric 208A modem at 4800 bits/s and with no tandem analog impairments was used.

PCM link but no conversion to analog. Such a connection would be necessary at the interface of a digital switch that presumed a PCM format.

The noise introduced in tandem NIC/A/NIC encodings is expected to add incoherently, so that the total noise added in decibels has a $10 \log_{10} K$ dependence, where K is the number of tandem encodings. The NIC/PCM/NIC encoding may or may not add any additional impairment, depending on the phase of the blocks assumed by tandem PCM-NIC converters. If they use identical block phases, the NIC/PCM/NIC conversion will be transparent, with no additional impairment. There are, however, $N - 1$ other possible phases the NIC/PCM/NIC conversion could use.

The distortion a particular sample incurs is directly dependent on the maximum chord of its block. In a NIC/PCM/NIC conversion additional distortion will occur when the new block phase has a larger maximum chord than the old block phase. The total distortion for a particular sample is thus dependent on the maximum chord in a new larger block which includes all the samples in the original block and those of the new one as well.

The worst distortion a given sample can incur after any number of NIC/PCM/NIC conversions is determined by the largest sample among the $N - 1$ samples preceding and $N - 1$ samples following, since no block of N samples can include any samples outside this set. The resulting worst-case performance will be equal to (independent samples), or a little better than (dependent samples), what it would be with an original block size of $2N - 1$.

This bound shows that the maximum penalty for any number of NIC/PCM/NIC conversions is small. For instance, we would suffer at most a 1.5 dB penalty with independent Laplacian samples and $N = 8$ (see Fig. 3). This is less than the 3-dB degradation of even a single NIC/A/NIC encoding. Nevertheless, this worst-case bound is rather pessimistic, since it can only be achieved after N tandem encodings, and then only if they each happen to have a different phase. Further quantitative consideration of this issue is given in Appendix C.

It is not necessary to accept any increase in effective block size in tandem NIC/PCM/NIC encodings.[5] By a process similar to statistical framing, the block in a subsequent NIC/PCM/NIC encoding can be lined up with the block in the original encoding by making use of the fact that in each NIC-PCM conversion (Table I) the discarded least significant bits are replaced by (100 ⋯). In a subsequent PCM-NIC conversion, block alignment is indicated when the discarded bits assume that (100 ⋯) pattern. A different method of block alignment will be described in the next section.

C. Delay

When the maximum chord of a block is determined prior to the NIC translation of each sample in the block, there is an implicit delay of one block in the transmitter. For a block size $N = 8$ and an 8-kHz sampling rate, this delay is 1 ms.

The penalty for this delay is that some trunks which fall just below the length where echo suppression is required would be pushed over the threshold and require echo suppression at additional expense. Echo suppressors are required at a length of 1850 mi, where the median incremental round-trip delay for each mile of trunk is 0.013 ms [11]. Thus, 2-ms round-trip processing delay corresponds to about 150 mi of trunk at nominal length of 1850 mi.

This processing delay would thus represent an economic penalty which would likely be more than compensated for by the savings in transmission cost due to the reduced bit rate. However, if tandem NIC/A/NIC or NIC/PCM/NIC translations were allowed, and each involved a processing delay, the total delay might possibly become excessive.

We will now show that the tandem delay problem can be avoided (at the expense of a slight decrease in SNR) in tandem NIC/PCM/NIC encodings. The technique is to generate the first sample in the block in the first and each subsequent NIC-to-PCM reconstruction in a special way, with the result that the maximum chord for the block in the subsequent PCM-to-NIC translation can be determined from the first sample alone without examining the entire block. The inherent processing delay in PCM-to-NIC translation is thereby eliminated in all but the initial encoding. Block alignment is assumed.

The first sample in each block in the first and each subsequent NIC-to-PCM reconstruction is generated according to the modified version of Table I shown in Table II. This modified table is precisely what would be obtained if the least significant bit of the 6-bit NIC code was replaced by a 1, giving a slight SNR penalty (1.4 dB for $N = 8$). At a subsequent PCM-to-NIC translation the maximum chord for the block can be determined from this sample alone; specifically, the maximum chord equals $L + K$, where L is the chord of the first sample and K is the number of least significant zeros in the first sample (under the restriction that $K \leqslant 4$).

Block alignment can also be achieved by exploiting this special encoding of the first sample in the block. The method is to generate the maximum chord as described above from the sample presumed to be the first of the block. If, after examining the remaining chords in the block, an inconsistency is observed, the block phase is incorrect. An inconsistency is a chord larger than and/or no chord equal to the maximum chord inferred from the first sample. Of course, one must account for infrequent inconsistencies due to line errors.

VI. CONCLUSIONS

Reencoding n-bit μ-law or A-law PCM according to the NIC algorithm reduces the bit rate by about 2 bits per sample while only degrading performance for both speech and voice band data by about 1 bit. A block size N of approximately 10 seems optimum. Simple conversion from PCM as well as excellent voice band data performance are advantages of NIC that distinguish it from techniques such as ADPCM (which yields a comparable SNR performance advantage for speech). Among the disadvantages of NIC are a slight increase in multiplexing complexity associated with transmitting the maximum chord information for a block of samples, the increased frame

[5] This possibility together with the potential for eliminating the tandem delay problem (next section) was suggested by J. W. Pan.

TABLE II
MODIFIED NIC TRANSLATION TABLE FOR THE FIRST SAMPLE IN EACH BLOCK (TO ELIMINATE DELAY ACCUMULATION)

L_{MAX}	NIC Encoding	Reconstruction
7	slwxyz	s111wxy1
7	s01wxy	s110wx10
7	s001wx	s101w100
7	s0001w	s1001000
7	s00001	s0110000
7	s00000	s0110000
6	slwxyz	s110wxy1
6	s01wxy	s101wx10
6	s001wx	s100w100
6	s0001w	s0111000
6	s00001	s0100000
6	s00000	s0100000
5	slwxyz	s101wxy1
5	s01wxy	s100wx10
5	s001wx	s011w100
5	s0001w	s0101000
5	s00001	s0010000
5	s00000	s0010000
4	slwxyz	s100wxy1
4	s01wxy	s011wx10
4	s001wx	s010w100
4	s0001w	s0011000
4	s0000w	s0001000
3	slwxyz	s011wxy1
3	s01wxy	s010wx10
3	s001wx	s001w100
3	s000wx	s000w100
2	slwxyz	s010wxy1
2	s01wxy	s001wx10
2	s00wxy	s000wx10
1	slwxyz	s001wxy1
1	s0wxyz	s000wxy1

size due to the block format, and the inherent processing delay.

APPENDIX A

In this Appendix we give without derivation equations for decision and output levels for NIC. We assume an input n-bit μ255 code, $4 \leqslant n \leqslant 8$, which is obtained by truncation of $(8-n)$-bits from an 8-bit code, and similarly an $(n-2)$-bit NIC code which is a truncation of (8-n) bits from the 6-bit NIC code of Table I. Let M be the integer equivalent of the binary NIC code less sign bit, $0 \leqslant M \leqslant 2^{n-3}-1$. We separate the set of all M and $1 \leqslant L_{\max} \leqslant 7$ into two regions Ω and Ω^c, where Ω corresponds to $L = 0$,

$$\Omega = \{M, L_{\max}: 1 \leqslant L_{\max} \leqslant n-3 \text{ and } 0 \leqslant M \leqslant 2^{n-3-L_{\max}}-1\}. \quad \text{(A1)}$$

Then the range of inputs corresponding to NIC code M is

$$X \in \begin{cases} [M2^{L_{\max}+7-n} - 0.5, (M+1)2^{L_{\max}+7-n} - 0.05], & (M,L_{\max}) \in \Omega \\ [M2^{L_{\max}+8-n} - 16.5, (M+1)2^{L_{\max}+8-n} - 16.5], & (M,L_{\max}) \in \Omega^c. \end{cases} \quad \text{(A2)}$$

Finally, let $\hat{L}$ and $\hat{V}$ be the 8-bit chord and level of the 8-bit μ255 code closest to the center of the region specified above. Then

$$\hat{L} = 0, \qquad (M,L_{\max}) \in \Omega \quad \text{(A3)}$$

and otherwise $\hat{L}$ must satisy

$$2^4 \leqslant 2^{\hat{L}_{\max}-L}([2^{5-n}(2M+1)] + 0.5) + 2^5 \quad \text{(A4)}$$

where $[\cdot]$ denotes truncation of fractional part. Finally, $\hat{V}$ is given by

$$\hat{V} = \begin{cases} [2^{L_{\max}-\hat{L}-1}([2^{5-n}(2M+1)] + 0.5)], & (M,L_{\max}) \in \Omega \\ [2^{L_{\max}-\hat{L}}([2^{5-n}(2M+1)] + 0.5) - 16], & (M,L_{\max}) \in \Omega^c. \end{cases} \quad \text{(A5)}$$

These formulas can be used in conjunction with those of Appendix B to calculate the SNR of NIC for independent samples.

APPENDIX B

JOINT DENSITY OF X AND $L_{\max}$

In this Appendix we calculate several probability densities relating to an input sample X and maximum chord $L_{\max}$. These densities are used for calculation of the SNR in Section III-B and the optimum quantization in Appendix C. Since the NIC algorithm depends on input sample magnitudes alone, we consider only positive random variables in the sequel.

Let the block of N statistically independent input-sample magnitudes be denoted by $X, Z_1, Z_2, \cdots, Z_{N-1}$ and have density $p_X(x)$ and cumulative distribution function

$$F_X(x) = \int_0^x p_X(\alpha)\, d\alpha. \quad \text{(B1)}$$

The order of the samples is irrelevant because of their independence.

Let

$$Y = \max\{Z_1, Z_2, \cdots, Z_{N-1}\} \quad \text{(B2)}$$

and

$$W = \max\{X, Y\}$$

where the important observation is that X and Y are independent. Further, W is the maximum sample in the block. The variable Y has the cumulative distribution function

$$\begin{aligned} F_Y(y) &= \Pr\{Y \leqslant y\} \\ &= \Pr\{Z_1 \leqslant y, Z_2 \leqslant y, \cdots, Z_{N-1} \leqslant y\} \\ &= F_X{}^{N-1}(y) \end{aligned} \quad \text{(B3)}$$

and probability density function

$$p_Y(y) = (N-1)F_X{}^{N-2}(y)p_X(y). \tag{B4}$$

The first quantity of interest is the joint density of X and W. First note that

$$\begin{aligned}\Pr\{X \leqslant x, W \leqslant w\} &= \Pr\{X \leqslant x, X \leqslant w, Y \leqslant w\}\\ &= \Pr\{X \leqslant \min(x,w)\}\Pr\{Y \leqslant w\}\\ &= \begin{cases} F_X(x)F_X{}^{N-1}(w), & x \leqslant w\\ F_X(w)F_X{}^{N-1}(w), & x > w.\end{cases}\end{aligned} \tag{B5}$$

Finding the density is just a matter of differentiating (B5). Defining the unit step function

$$u(\alpha) = \begin{cases} 1, & \alpha \geqslant 0\\ 0, & \alpha < 0\end{cases} \tag{B6}$$

we have

$$\frac{\partial}{\partial x}\Pr\{X \leqslant x, W \leqslant w\} = p_X(x)F_X{}^{N-1}(w)u(w-x) \tag{B7}$$

and hence

$$\begin{aligned}p_{X,W}(x,w) &= \frac{\partial^2}{\partial x \partial w}\Pr\{X \leqslant x, W \leqslant w\}\\ &= p_X(x)F_X{}^{N-2}(w)\{(N-1)p_X(w)u(w-x)\\ &\quad + F_X(w)\delta(w-x)\}.\end{aligned} \tag{B8}$$

While (B8) is difficult to interpret, the conditional density

$$\begin{aligned}p_{X|W}(x \mid w) &= \frac{p_{X,W}(x,w)}{p_W(w)}\\ &= \frac{N-1}{N}\frac{p_X(x)}{F_X(w)}u(w-x) + \frac{1}{N}\delta(w-x)\end{aligned} \tag{B9}$$

does have a simple interpretation; namely, with probability $1/N$, $W = X$ (i.e., X is the maximum samples in the block) and with probability $(N-1)/N, X$ is not the maximum sample, in which case its density is $p_X(x)$ constrained to the region $x \leqslant w$ and divided by the probability that $X \leqslant w$.

Let the range of analog samples corresponding to chord L be

$$a_L \leqslant x < a_{L+1}, \qquad 0 \leqslant L \leqslant 7. \tag{B10}$$

Of interest is the joint density of X and the maximum chord L (which is a random variable). We have

$$\begin{aligned}p_{X,L}(x,l) &= \int_{a_l}^{a_{l+1}} p_{X,W}(x,w)\,dw\\ &= \begin{cases} 0, & a_{l+1} \leqslant x\\ F_X{}^{N-1}(a_{l+1})p_X(x), & a_l \leqslant x < a_{l+1}\\ [F_X{}^{N-1}(a_{l+1}) - F_X{}^{N-1}(a_l)]\,p_X(x), & x < a_l.\end{cases}\end{aligned} \tag{B11}$$

This density was used in obtaining the graphs of Figs. 3 and 4.

APPENDIX C

INCREASE OF EFFECTIVE BLOCK SIZE IN TANDEM NIC/PCM/NIC CONVERSIONS

In Section V-B the concept of the effective block size in tandem NIC conversions was introduced. In this Appendix we calculate the distribution of the effective block size in K tandem encodings when each block phase is uniformly distributed. This will show that the $2N-1$ bound is pessimistic.

We focus attention on a single sample, and let its position in the block be $k_1, k_2, \cdots, k_K$ in the K tandem NIC encodings, where $1 \leqslant k_i \leqslant N$. Further, let $N(k_1, \cdots, k_K)$ be the total effective block size (i.e., the number of samples including the one under consideration which are included in determining the maximum chord) after K tandem encodings. Then, as previously stated,

$$N \leqslant N(k_1, \cdots, k_K) \leqslant 2N-1 \tag{C1}$$

and, in particular, a little thought reveals that

$$\begin{aligned}N(k_1, \cdots, k_K) = N &+ \max\{k_1, \cdots, k_K\}\\ &- \min\{k_1, \cdots, k_K\}.\end{aligned} \tag{C2}$$

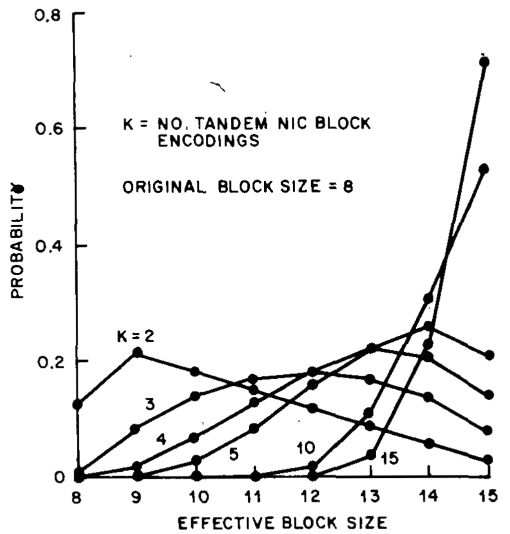
Since we have no preference for any block phase, including that of the original NIC encoding, we let the k_i be independent random variables, each uniformly distributed. Thus,

$$\Pr\{k_i = k\} = 1/N, \qquad i \leqslant k \leqslant N, i \leqslant i \leqslant K.$$

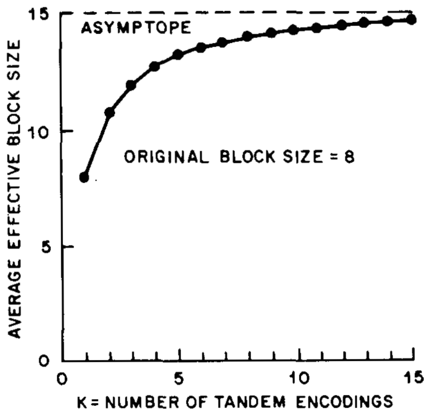
The distribution of $N(k_1, \cdots, k_K)$ is then [12, p. 23]

$$\begin{aligned}&\Pr\{N(k_1, \cdots, k_K) = n\}\\ &= \begin{cases} 1/N^{K-1}, & n = N\\ \dfrac{2N-n}{N^K}\{(n-N+1)^K - 2(n-N)^K + (n-N-1)^K\}, & N+1 \leqslant n \leqslant 2N-1\end{cases}\end{aligned} \tag{C3}$$

which is plotted along with its mean in Fig. 13 for $N = 8$. The mean block-size approaches its asymptope of 15 fairly rapidly, but for the region $K \leqslant 4$ (which in most applications will be the most likely range), the difference is still significant.



(a)



(b)

Fig. 13. (a) Probability density of effective block size. (b) Average effective block size as a function of K.

REFERENCES

[1] N. S. Jayant, "Digital coding of speech waveforms: PCM, DPCM, and DM quantizers," *Proc. IEEE,* vol. 62, pp. 611-632, May 1974.

[2] L. S. Golding and P. M. Schultheiss, "Study of an adaptive quantizer," *Proc. IEEE,* vol. 55, pp. 293-297, Mar. 1967.

[3] D. J. Goodman and A. Gersho, "Theory of an adaptive quantizer," *IEEE Trans. Commun.,* vol. COM-22, pp. 1037-1045, Aug. 1974.

[4] H. Kaneko, "A unified formulation of segment companding laws and synthesis of coders and digital companders," *Bell Syst. Tech. J.,* vol. 49, pp. 1555-1588, Sept. 1970.

[5] D. W. Osborne, "Digital sound signals: Further investigation of instantaneous and other rapid companding systems," BBC Res. Dep., rep. 1972/31.

[6] M. G. Croll, M. E. B. Moffat, and D. W. Osborne, "Nearly instantaneous digital compandor for transmitting six sound-programme signals in a 2.048 Mbits/s multiplex," *Electron. Lett.,* vol. 9, July 1973.

[7] A. Croisier, "Progress in PCM and delta modulation: Block-companded coding of speech signals," in *Proc. 1974 Int. Zurich Seminar Digital Commun.,* Mar. 1974, pp. B1(1)-B1(4).

[8] D. L. Duttweiler and D. G. Messerschmitt, "Nearly instantaneous companding and time diversity as applied to mobile radio transmission," in *Proc. Int. Conf. Commun.,* June 1975, pp. 40-12-40-15.

[9] P. Noll, "Adaptive quantizing in speech coding systems," in *Proc. 1974 Int. Zurich Seminar Digital Commun.,* Mar. 1974, pp. B3(1)-B3(6).

[10] K. W. Cattermole, *Principles of Pulse Code Modulations.* New York: American Elsevier, 1969.

[11] R. W. Hatch and A. E. Ruppel, "New rules for echo suppressors in the DDD network," *Bell Lab. Rec.,* vol. 52, Dec. 1974.

[12] H. A. David, *Order Statistics.* New York: Wiley, 1970.

★

Donald L. Duttweiler (S'68-M'70) was born in Kenmore, NY, on October 2, 1944. He received the B.E.E. degree from Rensselaer Polytechnic Institute, Troy, NY, in 1966, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1967 and 1970 respectively.

Since August 1970 he has been with Bell Laboratories, Holmdel, NJ, where his research interests have been in analog-to-digital encoding and digital communications.

Dr. Duttweiler is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.

★

David G. Messerschmitt (S'65-M'68) was born in Denver, CO, on May 21, 1945. He received the B.S.E.E. degree from the University of Colorado, Boulder, in 1967, and the M.S. and Ph.D. degrees from the University of Michigan, Ann Arbor, in 1968 and 1971, respectively.

Since 1968 he has been with Bell Laboratories, Holmdel, NJ, where he has been studying problems of intersymbol interference and voice coding for digital transmission. He currently supervises a group studying bit-rate reduction techniques.

Dr. Messerschmitt is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. He is Vice-Chairman of the Jersey Shore Chapter and a member of the Communication Theory Committee of the IEEE Communications Society.