On information rates for faster than Nyquist signaling

Rusek, Fredrik; Anderson, John B

Published in:
Proc., 2006 IEEE Global Telecommunications Conference

DOI:
10.1109/GLOCOM.2006.52

2006

On Information Rates for Faster than Nyquist Signaling

Fredrik Rusek and John B. Anderson
Dept. of Information Technology
Lund University, Lund, Sweden
Email: {fredrikr,anderson}@it.lth.se

***Abstract*— In this paper we consider the information rates of faster than Nyquist (FTN) signaling schemes. We consider binary, quaternary and octal schemes that use root raised cosine pulses. Lower and upper bounds to the information rates are given. The main result is that the lower bounds are often above the information rates for standard Nyquist signaling schemes. This implies that FTN must be superior to Nyquist signaling in some cases. Test results for one coding scheme are given; these show that high throughput communication based on FTN is indeed practical.**

I. Introduction and System Model

The concept of Faster Than Nyquist (FTN) signaling is well established. If a PAM system $\sum a[n]v(t-nT)$ is based on orthogonal pulse $v(t)$, the pulses can be packed closer than the Nyquist rate $1/T$ without suffering a loss in minimum square distance in the signaling system.

The result is a much more bandwidth-efficient coding system. Mazo showed [1] that for ideal sinc pulses the symbol time can be reduced to $.802T$ without loss in minimum Euclidean distance. We refer to this value as the Mazo limit. More recently the limits for root raised cosine pulses with nonzero excess bandwidth were derived in [2]. Efficient receivers for FTN signaling were also presented in this paper for the first time. Methods of computing the minimum distance of FTN signaling can be found in [3] and [4]. Mazo–type limits can also be derived for other pulse shapes [5]. Mazo limit phenomena turn up in other places as well; see [6] and references therein. A generalized version of FTN is given in [7]. More FTN results appear in [8].

The capacity calculation for a channel involves a maximization over all probability distributions on the input symbols. In the AWGN channel in this paper, the ultimate capacity is

$$\int_0^\infty \log_2[1+\frac{2|\Phi(f)|^2}{N_0}]df$$

where $N_0/2$ is the noise density and $\Phi(f)$ is the signal power spectral density (PSD). This rate is shown as a heavy line in plots that follow. Coding schemes with the same $\Phi(f)$ but additional constraints on the symbols or signal generation must perform worse. In the case of a given distribution or generation form, the limiting data rate will be called the information rate.

The information rate of an FTN system with a standard alphabet is still an open question and it is not even known whether FTN is superior, inferior or equivalent to ordinary multilevel Nyquist signaling. Moreover, the throughput problem is very different from the distance problem. Here we will compute upper and lower bounds to the information rate of FTN. These will show that FTN can indeed have rate superior to multilevel Nyquist signaling.

Consider a baseband PAM system based on a T-orthogonal root raised cosine (RC) pulse $\psi(t)$ where α denotes the excess bandwidth. When $\alpha = 0$ a sinc pulse is obtained. The one sided bandwidth of $\psi(t)$ is then $W = (1+\alpha)/(2T)$. The signal transmitted over the channel is

$$s_{\mathbf{a}}(t) = \sum_{n=-\infty}^{\infty} a[n]\psi(t-n\tau T), \quad \tau \leq 1 \tag{1}$$

where $a[n]$ are independent identically distributed (i.i.d.) data symbols from an alphabet $\mathcal{A}$ and $1/\tau T$ is the signaling rate. We assume $\psi(t)$ to be unit energy, i.e. $\int_{-\infty}^{\infty}|\psi(t)|^2dt = 1$. Since the pulse is T-orthogonal the system will not suffer from intersymbol interference (ISI) when $\tau = 1$. For $\tau < 1$ we say that we have FTN signaling, and ISI is unavoidable for i.i.d. input symbols.

Signals of the form (1) with i.i.d. data symbols have PSD [6]

$$\Phi(f) = \frac{\sigma_a^2}{\tau T}|\Psi(f)|^2 \tag{2}$$

where $\sigma_a^2 = \mathcal{E}\{a[n]a[n]^*\}$. Note that the shape of the PSD is not affected by FTN.

The channel is assumed to be the AWGN channel with one sided PSD $N_0/2$; the signal presented to the decoder is then $r(t) = s_{\mathbf{a}}(t) + n(t)$. Forney has shown [10] that a set of sufficient statistics to estimate $a[n]$ is the sequence

$$y[n] = \int_{-\infty}^{\infty} r(t)\psi^*(t-n\tau T)dt \tag{3}$$

Inserting the expression for $r(t)$ into (3) yields

$$y[n] = \sum_{m=1}^{N} a[m]g_\psi[n-m] + \eta[n] \tag{4}$$

where

$$g_\psi[n-m] = \int_{-\infty}^{\infty} \psi(t-n\tau T)\psi^*(t-m\tau T)dt \tag{5}$$

and

$$\eta[n] = \int_{-\infty}^{\infty} n(t)\psi^*(t-n\tau T)dt \tag{6}$$

Eq. (4) is the so called Ungerboeck observation model. The autocorrelation of the noise sequence $\eta[n]$ is

$$\mathcal{E}\{\eta[n]\eta^*[m]\} = \frac{N_0}{2} g_\psi[n-m] \tag{7}$$

In matrix notation (4) can be written as $\mathbf{y}^N = \mathbf{G}_\psi^N \mathbf{a}^N + \eta^N$. The matrix $\mathbf{G}_\psi^N$ is a $N \times N$ Toeplitz matrix formed from $\{g_\psi[0], g_\psi[1], \ldots, g_\psi[N]\}$, and $\mathbf{a}^N$ denotes the column vector formed from $\{a[1], \ldots, a[N]\}$; later we will use the notation $\mathbf{a}_{n_1}^{n_2}$ which denotes the column vector $\{a[n_1], \ldots, a[n_2]\}^{tr}$.

Since the noise variables are correlated it is convenient to work with the whitened matched filter (WMF) model instead of the Ungerboeck model. By filtering $y[n]$ with a whitening filter the sequence encountered by the decoder becomes

$$x[k] = \sum_{l=1}^{k} b[k-l]a[l] + w[k] \tag{8}$$

where $b[n]$ is a causal ISI tap sequence such that $b[n] \star b^*[-n] = g_\psi[n]$ and $w[k]$ are independent Gaussian variables having variance $\sigma^2 = N_0/2$. Since the whitening filter is invertible $x[k]$ also form a set of sufficient statistics. In matrix notation we write $\mathbf{x}^N = \mathbf{B}^N \mathbf{a}^N + \mathbf{w}^N$ The generator matix $\mathbf{B}^N$ is a $N \times N$ lower triangular matrix constructed from the ISI response $b[n]$.

II. Bounds on the Information Rate

In this section we bound the information rate of an FTN system. More bounds and reviews of what is known about capacity of ISI channels can be found in [9], [12], [13]. Methods to closely approximate the information rate are found in [14]. The bounds here are extensions of bounds in [9] and are therefore not new in themselves. However, our emphasis is their application to FTN.

We start by defining the information rate when pulse $\psi(t)$ is being used:

$$\begin{aligned} I_\psi &\triangleq \lim_{N\to\infty} \frac{1}{N} I(\mathbf{y}^N; \mathbf{a}^N) = \lim_{N\to\infty} \frac{1}{N} I(\mathbf{x}^N; \mathbf{a}^N) \\ &= \lim_{N\to\infty} \frac{1}{N} [H(\mathbf{x}^N) - H(\mathbf{x}^N | \mathbf{a}^N)] \end{aligned} \tag{9}$$

where $H(U)$ denotes the differential entropy of the random variable U and $I(V;U)$ is the mutual information between U and V. The data $\mathbf{a}^N$ take values from $\mathcal{A}$ according to a given probability distribution. The WMF and Ungerboeck models are assumed to be derived from $\psi(t)$.

The fact that $\psi(t)$ has infinite support generates a problem; it is hard to find its WMF model from the Ungerboeck model. To solve that we seek another pulse $\psi_1(t)$ that has a finite time discrete model and has an information rate I_{ψ_1} that can be related to I_ψ. Construct the pulse $\psi_1(t)$ as

$$\psi_1(t) = \psi(t) + \psi_2(t) \tag{10}$$

where $\psi_2(t)$ has a Fourier transform that satisfies

$$\Psi_2(f) = 0, \quad |f| \le W \tag{11}$$

This implies that $\psi(t)$ and $\psi_2(t)$ are non overlapping in frequency. Now, take $\psi_2(t)$ to be a pulse that has τT-sampled autocorrelation

$$g_{\psi_2}[n] = \begin{cases} \sum_{|m|>L} |g_\psi[m]|, & n = 0 \\ 0, & 0 < |n| \le L \\ -g_\psi[n], & |n| \ge L \end{cases} \tag{12}$$

Existence of the pulse $\psi_2(t)$ follows from the positivity of the Fourier transform of $g_{\psi_2}[n]$, see [8], [15]. Since $\psi_2(t)$ and $\psi(t)$ are mutually orthogonal we have

$$g_{\psi_1}[n] = g_\psi[n] + g_{\psi_2}[n] \tag{13}$$

and consequently $g_{\psi_1}[n] = 0, \ n \ge L$; i.e. $\psi_1(t)$ has a finite time discrete model.

If a signaling scheme is based on pulse $\psi_1(t)$ the receiver filter should be matched to $\psi_1(t)$; its output samples are $\mathbf{y}_1^N$. However, $\mathbf{y}_1^N$ could be obtained by having two matched filters $\psi_2(t)$ and $\psi(t)$; after τT-sampling the decoder then sees $\mathbf{y}_2^N$ and $\mathbf{y}^N$. Since $\psi_2(t)$ and $\psi(t)$ are mutually orthogonal it is clear that $\mathbf{y}_1^N = \mathbf{y}^N + \mathbf{y}_2^N$ and thus $\mathbf{y}_2^N$ and $\mathbf{y}^N$ also form a set of sufficient statistics for estimating $\mathbf{a}^N$. Furthermore, since ML decoding can be done either with $\mathbf{y}_1^N$ or with the pair $(\mathbf{y}^N, \mathbf{y}_2^N)$ we have

$$I(\mathbf{y}_1^N; \mathbf{a}^N) = I((\mathbf{y}_2^N, \mathbf{y}^N); \mathbf{a}^N) \tag{14}$$

By a standard result I_{ψ_1} and I_ψ can be related as

$$I((\mathbf{y}_2^N, \mathbf{y}^N); \mathbf{a}^N) \le I(\mathbf{y}_2^N; \mathbf{a}^N) + I(\mathbf{y}^N; \mathbf{a}^N) \tag{15}$$

Consequently,

$$I_\psi \ge I_{\psi_1} - I_{\psi_2} \tag{16}$$

We will later show how to lower bound I_{ψ_1}; since the Ungerboeck model is finite the method can be based on the Forney model which can be found exactly. However, I_{ψ_2} still has an infinite time discrete model. This difficulty we avoid by using a very loose upper bound: For signals with given PSD the so called Gaussian upper bound [9] for any Ψ is

$$I_\psi \le I_{G,\psi} \triangleq \int_0^\infty \log_2[1 + \frac{2\sigma_a^2 |\Psi(f)|^2}{\tau T N_0}] df \tag{17}$$

Inserting (17) into (16) gives

$$I_\psi \ge I_{\psi_1} - I_{G,\psi_2} \tag{18}$$

The spectrum $|\Psi_2(f)|^2$ can be expressed in terms of $|\Psi_1(f)|^2$ and $|\Psi(f)|^2$. The spectra $|\Psi_1(f)|^2$ and $|\Psi(f)|^2$ can be found exactly, thus $|\Psi_2(f)|^2$ can be given exactly even though its sampled autocorrelation is infinite.

Now we lower bound I_{ψ_1}. To simplify notation we omit the subscript and write $\mathbf{x}^N$ instead of $\mathbf{x}_1^N$ etc. The last term of (9) can be directly computed as

$$H(\mathbf{x}^N | \mathbf{a}^N) = H(\mathbf{w}^N) = \frac{N}{2} \log_2(2\pi e \sigma^2) \tag{19}$$

By the chain rule we have

$$H(\mathbf{x}^N) = \sum_{l=1}^{N} H(x[l] | \mathbf{x}^{l-1}) \tag{20}$$

Since conditioning does not increase entropy we get

$$H(x[l]\big|\mathbf{x}^{l-1}) \geq H(x[l]\big|\mathbf{x}^{l-1},\mathbf{a}^{l-k}) \tag{21}$$

This implies that

$$I_\psi \geq \lim_{N\to\infty}\frac{1}{N}\sum_{l=1}^{N} H(x[l]\big|\mathbf{x}^{l-1},\mathbf{a}^{l-k}) - \frac{N}{2}\log_2(2\pi e\sigma^2) \tag{22}$$

In [9] k is 1; here we calculate the bound for $k > 1$. By the chain rule

$$\begin{aligned} H(\mathbf{x}_{l-k+1}^{l}\big|\mathbf{x}^{l-k},\mathbf{a}^{l-k}) &= H(\mathbf{x}_{l-k+1}^{l-1}\big|\mathbf{x}^{l-k},\mathbf{a}^{l-k}) \\ &+ H(x[l]\big|\mathbf{x}^{l-1},\mathbf{a}^{l-k}) \end{aligned} \tag{23}$$

The last term of (23) is the entropy we are seeking in (21); rearranging gives

$$\begin{aligned} H(x[l]\big|\mathbf{x}^{l-1},\mathbf{a}^{l-k}) &= H(\mathbf{x}_{l-k+1}^{l}\big|\mathbf{x}^{l-k},\mathbf{a}^{l-k}) \\ &- H(\mathbf{x}_{l-k+1}^{l-1}\big|\mathbf{x}^{l-k},\mathbf{a}^{l-k}) \end{aligned} \tag{24}$$

Since $\mathbf{w}^N$ are i.i.d. Gaussian and the ISI response is causal the conditioning on $\mathbf{x}^{l-k}$ in (24) can be removed, i.e.

$$H(\mathbf{x}_{l-k+1}^{l}\big|\mathbf{x}^{l-k},\mathbf{a}^{l-k}) = H(\mathbf{x}_{l-k+1}^{l}\big|\mathbf{a}^{l-k}) \tag{25}$$

Now, since the data $\mathbf{a}^N$ are i.i.d.

$$H(\mathbf{x}_{l-k+1}^{l}\big|\mathbf{a}^{l-k}) = H(\tilde{\mathbf{x}}_{l-k+1}^{l}) \tag{26}$$

where

$$\tilde{x}[s] = \sum_{m=0}^{s-l+k-1} a[s-m]b[m]+w[s], \quad l-k+1 \leq s \leq l \tag{27}$$

That is, $\tilde{x}[s]$ is formed by subtracting the interference to $\mathbf{x}_l^{l-k+1}$ caused by $\mathbf{a}^{l-k}$. To summarize, the entropy in (26) is equal to the entropy of the variable $\mathbf{z}^k$, where $\mathbf{z}^k = \mathbf{B}^k\mathbf{a}^k + \mathbf{w}^k$. For l larger than the ISI length the terms to be summed in (22) are all identical and the limit can be removed. If we express the information rate in bits/s we have shown that

$$I_\psi \geq I_{LB,\psi}^k \triangleq \frac{I(\mathbf{z}^k;\mathbf{a}^k) - I(\mathbf{z}^{k-1};\mathbf{a}^{k-1})}{\tau T} \tag{28}$$

We have not found any closed form expression for (28) when $k > 1$. Instead we calculate it by direct evaluation of the integral $H(\mathbf{z}^k)$ in (28). Inserting (28) into (16) the lower bound is complete and reads

$$I_\psi \geq I_{LB,\psi_1}^k - I_{G,\psi_2}, \quad k \geq 1 \tag{29}$$

It should be noted that as $N_0 \to 0$ we have $I_{G,\psi} \to \infty$ and the lower bound therefore tends to $-\infty$. However, by choosing L large, the power in $\psi_2(t)$ is very small which leads to $I_{G,\psi} \approx 0$ for reasonable SNRs.

The lower bound (29) was derived for infinite root RC pulses. In the bounding technique it was necessary to work with finite time discrete models which led to the construction of pulses $\psi_1(t)$ and $\psi_2(t)$. However, a finite model can be obtained by truncating the root RC pulse. Altough the results for truncated pulses probably would be extremely close to our results they would not be valid bounds for the infinite root RC pulse.

Next we derive upper bounds. Since $\psi_1(t)$ was constructed from $\psi(t)$ by adding an orthogonal part it is clear that

$$I_\psi \leq I_{\psi_1} \tag{30}$$

A more intuitive explanation that (30) holds is that the decoder can choose between either matching against $\psi(t)$ or $\psi_1(t)$ when decoding a system based on $\psi_1(t)$.

We now derive the upper bound for I_{ψ_1}. As with the lower bound we omit the subscript and and write $\mathbf{x}^N$ instead of $\mathbf{x}_1^N$ etc. Replace (21) by

$$H(x[l]\big|\mathbf{x}^{l-1}) \leq H(x[l]\big|\mathbf{x}_{l-k}^{l-1}) \tag{31}$$

Similarily to the lower bounding technique we apply the chain rule and obtain

$$H(x[l]\big|\mathbf{x}_{l-k}^{l-1}) = H(\mathbf{x}_{l-k}^{l}) - H(\mathbf{x}_{l-k}^{l-1}) \tag{32}$$

Since

$$\begin{aligned} H(\mathbf{x}_{l-k}^{l}) &= I(\mathbf{x}_{l-k}^{l};\mathbf{a}^l) + H(\mathbf{x}_{l-k}^{l}\big|\mathbf{a}^l) \\ &= I(\mathbf{x}_{l-k}^{l};\mathbf{a}^l) + H(\mathbf{w}_{l-k}^{l}) \end{aligned} \tag{33}$$

we have shown that

$$I_\psi \leq \lim_{N\to\infty}\frac{1}{N}\sum_{l=1}^{N}\frac{I(\mathbf{x}_{l-k}^{l};\mathbf{a}^l) - I(\mathbf{x}_{l-k}^{l-1};\mathbf{a}^{l-1})}{\tau T} \tag{34}$$

For l such that $l-k$ is larger than the ISI memory L, the terms in (34) are statistically equivalent and we can remove the limit. A computational problem here is that $\mathbf{x}_{l-k}^{l-1}$ depends on L data symbols $a[n]$. Therefore L has to be taken quite small; for large alphabets L has to be very small which makes the bound loose.

Finally we remark that we have derived a more advanced and better construction that can replace (10). This construction is based on solving a linear program to get a pulse that can be shown to be finite and to have superior/inferior information rate compared to $\psi(t)$. The main benefit of this construction is that it eliminates the loose Gaussian bound from the final lower bound. For the upper bound the main benefit is that the pulse found can have very short discrete model, while still having an information rate quite close to $\psi(t)$. The improvements in the actual bounds are not very significant for low SNRs. The improved constructions have not been used in following sections.

III. Numerical Results and Comparisons

Before we give numerical results we discuss some competing systems. For Nyquist systems higher throughput is obtained by expanding the signaling alphabet. Here we will consider 8PSK, 16QAM, 32CR and 128CR; these alphabets are defined in [6]. Although slightly better alphabets exist these are standard alphabets in the literature as well as in practice. They are all two dimensional; when we compute their information rates we compute them per dimension, i.e. the 16QAM system's maximal throughput is $2/T$ bits/s per

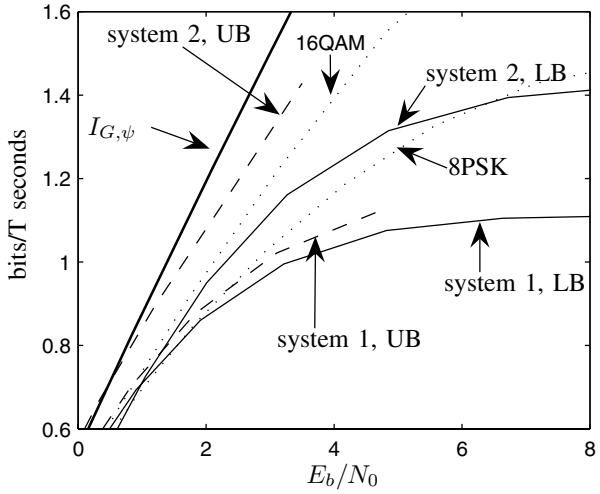

Fig. 1. Information rates for binary FTN transmission. System 1 has signaling parameters $\tau = .9$ and $\alpha = .3$. System 2 has signaling parameters $\tau = .7$ and $\alpha = .3$. The solid lines are lower bounds and dashed are upper. Dotted lines are information rates for equiprobable 8PSK and 16QAM. The heavy solid line is the Gaussian upper bound (17) for root RC pulses with $\alpha = .3$.

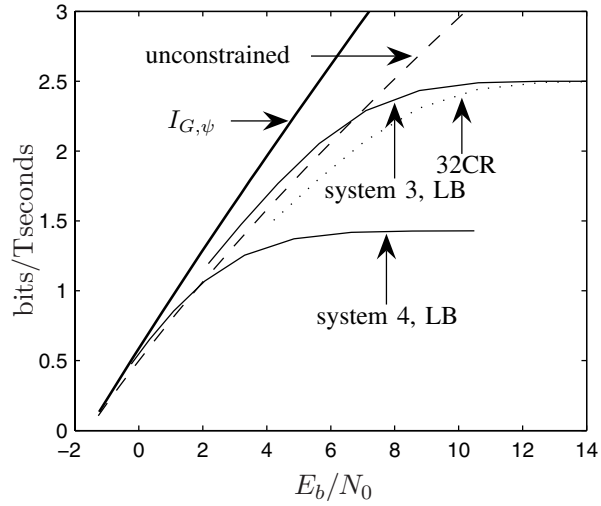

Fig. 2. Lower bounds on information rate for binary and quaternary FTN. System 3 is quaternary signaling with $\tau = .8$ and $\alpha = .5$. System 4 is binary signaling with $\tau = .7$ and $\alpha = .5$. The dotted line is the information rate of equiprobable 32CR. The dashed line is the capacity of an unconstrained Nyquist system. The heavy solid line is the Gaussian upper bound (17) for root RC pulses with $\alpha = .5$.

dimension. When comparing FTN to Nyquist systems we demand that both have same PSD shape. This implies that both systems use same $\psi(t)$ and α. The data symbols form an i.i.d. sequence and all outcomes are equiprobable. In some cases we will compare the FTN schemes against Nyquist schemes with unconstrained alphabets; in that case the optimal data distribution is Gaussian.

Nyquist systems suffer from a major weakness: They cannot benefit from excess bandwidth. When α increases, the information rate of a Nyquist system remains constant. For FTN the story is different. When α grows the ISI pattern becomes milder, i.e. energy concentrates in the first ISI taps. This makes the information rate grow and it eventually eclipses that of the Nyquist system.

For FTN we consider 2-, 4- and 8-ary signaling, i.e. $\mathcal{A} = \{-1, 1\}$ or $\{-3, -1, 1, 3\}$ or $\{-7, -5, -3, -1, 1, 3, 5, 7\}$, with appropriate energy normalization.

In figure 1 lower and upper bounds for binary FTN and $\alpha = .3$ are given. Although not strongly bandwidth efficient, $\alpha = .3$ is a common value in practice. Lower and upper bounds are calculated with $k = 2$. It is seen that the FTN scheme with $\tau = .7$ outperforms 8PSK Nyquist signaling over a wide range of SNRs. For $E_b/N_0 < 1$ dB, system 1 outperforms 16QAM Nyquist as well. The gap between the lower and upper bounds is less than 2 dB for $\tau = .7$ (system 2) and virtually closed for $\tau = .9$ (system 1).

In figure 2 we show lower bounds calculated with $k = 2$ for binary FTN with $\alpha = .5$ and $\tau = .7$ and quaternary FTN with $\alpha = .5$ and $\tau = .8$. The comparison system is here a Nyquist system with unconstrained alphabet. This is the best rate that any Nyquist system of the form (1) can achieve. Even so, the FTN system outperforms the Nyquist system.

Comparing system 3 against the Nyquist 32CR system we observe the following. First, the maximal throughput of both systems is 2.5 bits/T seconds (per dimension), thus it is fair to compare them. Second, system 3 has always superior information rate; in $E_b/N_0 = 5$–8 dB the gain of FTN is ≈ 1.5 dB.

Third, one can argue that $\alpha = .5$ is too large, but FTN systems can profit from large α. The bandwidth efficiency of both schemes is $2I_\psi/(1+\alpha)$. For a given value of E_b/N_0 the throughput of the Nyquist scheme is significantly lower than for the FTN system. Thus, in order to equalize *both* the bandwidth and power efficiencies the Nyquist system has to choose a smaller value for α. If we choose $E_b/N_0 = 5.5$dB as an example, the Nyquist system needs $\alpha \approx .3$. Thus the FTN scheme can work with a much larger value of α, resulting in lower implementation complexity. Recall that the shown information rate of the FTN system is only a lower bound.

In figure 3 a lower bound with $k = 2$ for octal FTN with $\alpha = .2$ and $\tau = .8517$ is given. The value for τ is chosen to give the FTN system the maximal information rate of 3.5 bits/T seconds which is equal to the information rate of a 128CR Nyquist system. Even for such a low α as .2 the FTN system has higher information rate than the 128CR system. In fact the FTN system beats even the Nyquist system with Gaussian symbol distribution. This we find remarkable for such a small value of α. Equalizing both the bandwidth and power efficiencies of the systems, we find that the 128CR Nyquist system requires $\alpha \approx .11$ for $E_b/N_0 \lessapprox 12$dB. Also shown is a lower bound with $k = 2$ for quaternary FTN with $\alpha = .2$ and $\tau = .9$. The bound lies above the information rates for both 16QAM and 32CR Nyquist systems at low SNRs.

IV. A Test Coding Scheme

In this section we give the results of a test coding scheme to illustrate that the information rates for FTN are indeed

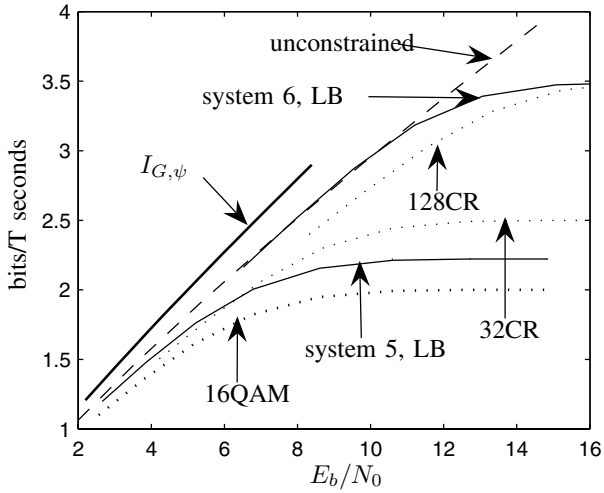

Fig. 3. Lower bounds on information rate for FTN. System 5 is quaternary signaling, $\tau = .9$ and $\alpha = .2$. System 6 is octal signaling, $\tau = .8517$ and $\alpha = .2$. The dashed line is Nyquist signaling with an unconstrained alphabet. The heavy solid line is the Gaussian upper bound (17) for root RC pulses with $\alpha = .2$.

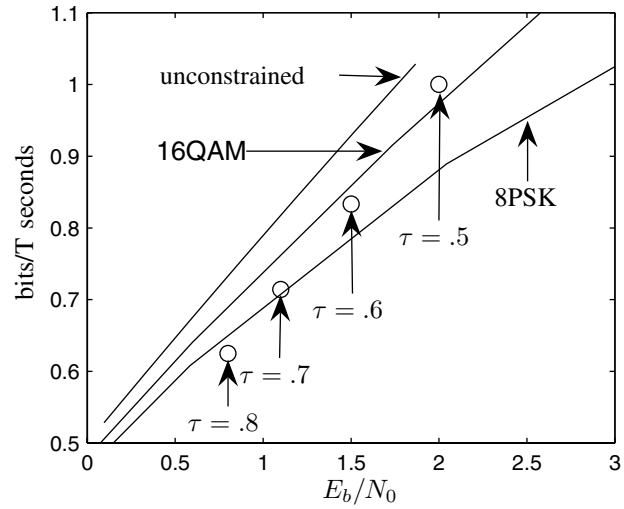

Fig. 4. Receiver test results for a coding scheme based on FTN signaling with $\alpha = .3$. Solid lines are information rates of 8PSK, 16QAM and unconstrained alphabets (Gaussian). The circles are the needed E_b/N_0 to obtain BER 10^{-5} for FTN schemes with various τ.

practically obtainable. The coding scheme used is described in [16], and we review it briefly here. A block of 50000 binary information bits is encoded by a rate $1/2$ repetition code, then an S-random interleaver with blocksize 100000 bits is applied to the output of the repetition code. A rate 1 recursive precoder follows and finally an FTN system with $\alpha = .3$ and signaling rate τT is used as modulation over the AWGN channel. Decoding is done with standard iterative decoding. The bit rate of this scheme is $1/2\tau T$ bits/s. Receiver details and tests for interleaver length 10000 are found in [16].

In figure 4 the bit rate is plotted against the needed E_b/N_0 to obtain BER 10^{-5}, for various values of τ.[1] It is seen that the FTN systems are very competitive compared to the Nyquist systems; for $\tau = .5$ the FTN system outperforms the best performance of any coding system based on equiprobable 16QAM scheme.

V. Conclusions

In this paper we have given lower and upper bounds of the information rate for FTN. First we constructed a finite time discrete model from the infinite time support root RC pulse. This was done by finding pulses that have superior/inferior information rate than the root RC pulse has. The actual bounds must be computed explicitly.

We find that FTN indeed has superior information rate to Nyquist signaling in many cases. This seems to be especially true when the signaling alphabet grows. For example, system 6 above has excess bandwidth of only 20% but is still superior to any form of Nyquist signaling with the same PSD shape. When compared with standard signaling methods such as 128CR the FTN system is better.

[1]Technically the information rates should be computed against BER 10^{-5}, but the difference is invisible in the figure.

The power gains for FTN can also be converted into increased α values. The Nyquist systems are able to match the FTN systems both in power and bandwidth efficiency but only by selecting a smaller, less practical value of α. For example, we gave an example where Nyquist systems with $\alpha = .3$ have the same power and bandwidth efficiency as FTN with $\alpha = .5$.

References

[1] J.E. Mazo, "Faster–than–Nyquist signaling," *Bell Syst. Tech. J.*, vol. 54, pp. 1451-1462, Oct. 1975.

[2] A.D. Liveris and C.N. Georghiades, "Exploiting Faster–than–Nyquist Signaling," *IEEE Trans. Commun.*, vol. 51, pp. 1502-1511, Sep. 2003.

[3] D. Hajela, "On Computing the Minimum Distance for Faster–Than–Nyquist Signaling," *IEEE Trans. Inform. Theory*, vol. 36, pp. 289–295, Mar. 1990.

[4] J.E. Mazo and H.J. Landau, "On the Minimum Distance Problem for Faster–Than–Nyquist Signaling," *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 1420–1427, Nov. 1988.

[5] F. Rusek and J.B. Anderson, "M-ary Coded Modulation by Butterworth Filtering," *Proc.*, IEEE Int. Symp. Inform. Theory, Yokohama, p. 184, June 2003.

[6] J.B. Anderson and A. Svensson, *Coded Modulation Systems*, Plenum, New York, 2003.

[7] F. Rusek and J. B. Anderson, "The Two Dimensional Mazo Limit," *Proc.* IEEE Int. Symp. Inform. Theory, pp. 970-974, Sept. 2005 Adelaide.

[8] G.J. Foschini, "Contrasting Performance of Faster Binary Signaling with QAM," *Bell Syst. Tech. J.*, vol. 63, pp. 1419-1445, Oct. 1984.

[9] S. Shamai, L.H. Ozarow and A.D. Wyner, "Information Rates for a Discrete-Time Gaussian Channel with Intersymbol Interference and Stationary Inputs," *IEEE Trans. Information Theory*, vol. 37, pp. 1527–1539, Nov. 1991.

[10] G. D. Forney, Jr., "Maximum Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," *IEEE Trans. Inform. Theory*, vol. 18, pp. 363–378, May 1972.

[11] G. Ungerboeck, "Adaptive Maximum-Likelihood Receiver for Carrier-Modulated Data-Transmission Systems," *IEEE Trans. Commun.*, vol. 22, pp. 624–636, May 1974.

[12] W. Hirt, "Capacity and Information Rates of Discrete-Time Channels with Memory," Doctoral dissert. (no. ETH 8671), Swiss Federal Inst. Technol., Zurich, 1988.

[13] L.H. Ozarow, A.D. Wyner and Z. Ziv, "Achievable Rates for a Constrained Gaussian Channel," *IEEE Trans. Information Theory*, vol. 34, pp. 365–370, May. 1988.

[14] H. D. Pfister, J. B. Soriaga, and P. H. Siegel, "On the Achievable Information Rates of Finite State ISI Channels," *Proc.* IEEE Global Commun. Conf., San Antonio, TX, Nov. 2001, pp. 2992-2996

[15] A. Said and J.B. Anderson, "Design of Optimal Signals for Bandwidth–Efficient Linear Coded Modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 701–713, Mar. 1998.

[16] F. Rusek and J.B. Anderson, "Serial and Parallel Concatenations Based on Faster than Nyquist Signalnig," *Proc*. IEEE Int. Symp. Inform. Theory, pp. 1993-1997, July 2006, Seattle.