Capacity of Faster than Nyquist Signaling

Mehdi Ganji, *Student Member, IEEE*, Xun Zou, *Student Member, IEEE*, and Hamid Jafarkhani, *Fellow, IEEE*

Abstract—**This letter considers the capacity computations of faster-than-Nyquist (FTN) signaling. It calculates the theoretical capacity of FTN signaling which is obtained by a correlated input. The capacity-achieving power spectral density (PSD) is derived and its superiority over the independent input is shown. The practical issue imposed by the capacity-achieving PSD, i.e., out-of-band (OOB) emission, is shown. To solve this issue, an upper-bound is introduced for the input PSD to limit the OOB emission. The new optimization problem is solved and the constrained PSD is obtained. The introduced PSD captures the trade-off between the obtained capacity and the OOB emission.**

I. Introduction

In this letter, we consider the problem of transmitting data over a bandlimited additive white Gaussian noise (AWGN) channel by means of a set of signals that are generated by linear modulation. Linear modulation signals have the form $x(t) = \sum a[n]p(t - n\delta T)$ where $\{a[n]\}$ are the symbols to be transmitted and can be independent or correlated. The pulse shape $p(t)$ usually has useful detection properties, e.g., orthogonality, which is commonly mentioned as No-ISI Nyquist condition. An example of such pulses is called raised cosine (RC) pulse and is extensively used in communication standards including the DVB-S2X in satellite communication [1].

Shannon's classical result states that the highest transmission rate over the AWGN channel is $W \log_2(1 + P/WN_0)$ where W is one-sided bandwidth, P is the average power, and $N_0/2$ is the power spectral density (PSD) of the white noise [2]. Such a capacity is achieved by the *sinc* pulse and $T = \frac{1}{2W}$. However, the sinc pulse has serious realization problems and disadvantages in practice. Therefore, smoother orthogonal pulses are utilized which introduce the excess bandwidth. This excess bandwidth remains unutilized by an orthogonal signaling. However, the excess bandwidth can be utilized by the Faster than Nyquist (FTN) Signaling [3].

FTN was first introduced in [4] where its minimum Euclidean distance property was analyzed. FTN has been extensively studied in past few years [5]. For example, asymptotic optimality of binary FTN signaling is shown in [6], non-binary FTN and the bit error rate (BER) performance are studied in [7], [8], the detection and the receiver design are examined in [9], [10] and the application of FTN in multiuser broadcast channel is considered in [11], [12]. The information rates of FTN signaling and the provided gain by faster signaling is analyzed in [3], [13]. In [3], an independent input with a constant PSD that satisfies the power constraint is considered. In [13], a more general set of inputs including independent and correlated inputs is considered. Denoting the input PSD as

$S_a(f), f \in [-1/2, 1/2]$, the power constraint is expressed as $\int_{-1/2}^{1/2} S_a(f)df \le P$ and the capacity-achieving input is found by using the well-known water-filling algorithm. Although the mentioned power constraint is correct for Nyquist signaling, the correlated input in conjunction with FTN signaling will alter the power constraint and thus change the capacity expressions. The erroneous power constraint does not fully reveal the potential of FTN signaling on increasing the capacity.

Our goal in this work is to comprehensively study the power constraint and the resulting capacity expression. We derive the capacity-achieving PSD and show its enhancement over the independent input. To alleviate the practical issues of the capacity-achieving input, we modify the input PSD constraint to strike a trade-off between the independent and capacity-achieving inputs.

II. System model

We consider the fundamental single-carrier communication model. The transmitted linearly modulated baseband signal has the form

$$x(t) = \sum_{n=1}^{N} a[n]p(t - n\delta T), \tag{1}$$

where $\{a[n]\}$ is a stationary Gaussian process with power spectral density (PSD) of $\{S_a(f), f \in [-1/2, 1/2]\}$ which is periodic in f with period of 1. The symbol interval is denoted by δT where $\delta \in (0, 1]$ is the compression factor in FTN signaling and $\delta = 1$ yields the traditional Nyquist signaling with signaling rate of $1/T$. Also, $p(t)$ is a root-Nyquist pulse shaping filter with unit energy, i.e., $\int_{-\infty}^{\infty} |p(t)|^2 dt = 1$. Denoting $g(t) = p(t) \star p^*(-t)$, the Nyquist No-ISI condition is expressed as: $g(nT) = \begin{cases} 1 & n = 0 \\ 0 & n \ne 0 \end{cases}$. The Nyquist No-ISI condition can be equivalently expressed as $G_T(f) = \sum_{k=-\infty}^{\infty} G\left(f - \frac{k}{T}\right) = T$. The PSD of the transmitted signal $x(t)$ can be calculated as $S_x(f) = S_a(\delta T f)G(f)$. The transmission power is calculated as $P_x = \int_{-\infty}^{\infty} S_x(f)df$. Because $S_a(\delta T f)$ is periodic in f with period of $1/\delta T$, the transmission power can be written as:

$$P_x = \int_{-1/2\delta T}^{1/2\delta T} S_a(\delta T f)G_{\delta T}(f)df = \frac{1}{\delta T} \int_{-1/2}^{1/2} S_a(f)G_{\delta T}(\frac{f}{\delta T})df, \tag{2}$$

where $G_{\delta T}(f) = \sum_{k=-\infty}^{\infty} G\left(f - \frac{k}{\delta T}\right)$. Note that for the Nyquist signaling, i.e., $\delta = 1$, $G_T(f) = T$, and thus, the transmit power reduces to $P_x = \int_{-1/2}^{1/2} S_a(f)df$ which is commonly used in the literature. However, for $\delta < 1$, the above equality does not hold. Also note that for independent symbols, i.e., $S_a(f) = \sigma_a^2$, the transmission power reduces
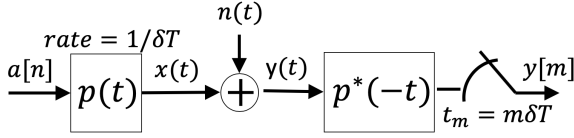
Fig. 1: FTN system model.

to $P_x = \frac{\sigma_a^2}{\delta T} \int_{-1/2}^{1/2} G_{\delta T}(\frac{f}{\delta T}) df = \sigma_a^2$. Assuming the AWGN channel, the received signal at the decoder is denoted as: $y(t) = \sum_{n=1}^{N} a[n]p(t - n\delta T) + n(t)$ where $n(t)$ is the white noise with the PSD of $\sigma_n^2 = N_0/2$. At the decoder, the set of sufficient statistics is obtained by a matched filter followed by a sampler which samples at time instants $t_m = m\delta T$ as shown in Fig. 1. Thus, the samples can be written as:

$$y[m] = \sum_{n=1}^{N} a[n]g((n-m)\delta T) + \eta[m], \quad (3)$$

where $\eta[m] = \int_{-\infty}^{\infty} n(t)p^*(t - m\delta T)dt$ and the correlation of the noise sample can be denoted as $\mathcal{E}\{\eta[m]\eta^*[n]\} = \sigma_n^2 g((n-m)\delta T)$. The input-output relation can be written in matrix form as:

$$\boldsymbol{y}^N = \boldsymbol{G}^N \boldsymbol{a}^N + \boldsymbol{\eta}^N, \quad (4)$$

where $\boldsymbol{y}^N = \{y[1], \cdots, y[N]\}$, $\boldsymbol{a}^N = \{a[1], \cdots, a[N]\}$, $\boldsymbol{\eta}^N = \{\eta[1], \cdots, \eta[N]\}$ and $\boldsymbol{G}^N$ is an $N \times N$ Toeplitz matrix whose elements are defined as:

$$[\boldsymbol{G}^N]_{m,n} = g((n-m)\delta T). \quad (5)$$

Note that with Nyquist signaling, the input-output relation is independent for each transmitted symbol denoted as $y[m] = a[m] + \eta[m]$.

III. EXISTING RESULTS

In [3], the capacity of FTN signaling for independent symbols with power constraint of P is calculated as:

$$C(P, \delta T) = \int_0^{1/2\delta T} \log_2\left(1 + \frac{P}{\sigma_n^2} G_{\delta T}(f)\right) df. \quad (6)$$

In [13], the capacity with correlated symbols is calculated as:

$$C(P, \delta T) = \sup_{S_a(f)} \int_0^{1/2\delta T} \log_2\left(1 + \frac{1}{\sigma_n^2} S_a(\delta T f) G_{\delta T}(f)\right) df$$

$$s.t. \quad \int_{-1/2}^{1/2} S_a(f) df = P, \quad (7)$$

where the the optimization problem is solved by the well-known water-filling strategy. However, in these results, the effect of FTN signaling on the power transmission is ignored which resulted in erroneous results. Hence, we revisit the capacity expression with the correct transmission power and reveal the main advantages of FTN signaling.

IV. FTN CAPACITY EXPRESSIONS REVISITED

For completeness, we derive the capacity expressions for FTN signaling, following the steps taken in [14] and [15]. Based on [14], the capacity of the discrete-time system in (4), which resembles an ISI channel, is equal to $C_{DT} = \lim_{N\to\infty} \frac{1}{N} I(\boldsymbol{y}^N; \boldsymbol{a}^N)$ given the input symbols satisfy the power constraint in (2). The differential entropy of a stationary Gaussian n-vector, $\boldsymbol{x}^N$, with covariance matrix of $\boldsymbol{\Sigma}_x^N$ is equal to [16]:

$$h(\boldsymbol{x}^N) = \frac{N}{2} \log_2\left(2\pi e \left[det(\boldsymbol{\Sigma}_x^N)\right]^{1/N}\right). \quad (8)$$

In addition, according to the Toeplitz distribution theorem [17], as $N \to \infty$, the asymptotic differential entropy approaches:

$$\lim_{N\to\infty} \frac{1}{N} h(\boldsymbol{x}^N) = \frac{1}{2} \int_{-1/2}^{1/2} \log_2\left(2\pi e \sigma_x(f)\right) df, \quad (9)$$

where $\sigma_x(f)$ is the generating function of the Toeplitz matrix $\boldsymbol{\Sigma}_x$ defined as $\sigma_x(f) = \sum_{n=-\infty}^{\infty} [\boldsymbol{\Sigma}_x]_n e^{-j2\pi fn}$ and $[\boldsymbol{\Sigma}_x]_n$ represents the nth diagonal element of $\boldsymbol{\Sigma}_x$.

Using (8) and (9), the asymptotic information rate of $C_{DT} = \lim_{N\to\infty} \frac{1}{N} I(\boldsymbol{y}^N; \boldsymbol{a}^N)$ can be calculated in the following steps:

$$C_{DT} = \lim_{N\to\infty} \frac{1}{N} h(\boldsymbol{y}^N) - \lim_{N\to\infty} \frac{1}{N} h(\boldsymbol{\eta}^N)$$

$$= \frac{1}{2} \int_{-1/2}^{1/2} \log_2\left(\sigma_n^2 \hat{g}(f) + \hat{g}^2(f) S_a(f)\right) - \log_2\left(\sigma_n^2 \hat{g}(f)\right) df$$

$$= \frac{1}{2} \int_{-1/2}^{1/2} \log_2\left(1 + \frac{1}{\sigma_n^2} \hat{g}(f) S_a(f)\right) df, \quad (10)$$

where $\hat{g}(f) = \sum_{n=-\infty}^{\infty} g(n\delta T) e^{-j2\pi fn}$ and after some calculations, it can be shown that $\hat{g}(f) = G_{\delta T}(f/\delta T)$. Then, after normalizing by the signaling rate, the capacity of FTN signaling can be formulated as:

$$C(P, \delta T) = \sup_{S_a(f)} \frac{1}{2\delta T} \int_{-1/2}^{1/2} \log_2\left(1 + \frac{1}{\sigma_n^2} G_{\delta T}(\frac{f}{\delta T}) S_a(f)\right) df$$

$$s.t. \quad \frac{1}{\delta T} \int_{-1/2}^{1/2} S_a(f) G_{\delta T}(\frac{f}{\delta T}) df \leq P. \quad (11)$$

Note that the power constraint in the new capacity formulation differs from the power constraint mentioned in (7). Considering Jensen's inequality, the capacity is achieved by:

$$S_a(f) = \begin{cases} \frac{\delta T P}{\mu G_{\delta T}(f/\delta T)} & G_{\delta T}(f/\delta T) \neq 0 \\ 0 & o.w. \end{cases}, \quad (12)$$

where $\mu = \mathcal{L}\left(f \in [-1/2, 1/2] | G_{\delta T}(f/\delta T) \neq 0\right)$ and $\mathcal{L}(I)$ is the Lebesgue measure of the set I. By simple calculations, the capacity can be written as:

$$C(P, \delta T) = \frac{\mu}{2\delta T} \log_2\left(1 + \frac{\delta T P}{\mu \sigma_n^2}\right). \quad (13)$$

Remarks 1: For the sinc function with any signaling rate, the achieved capacity is $C(P, \delta T) = \frac{1}{2T} \log_2\left(1 + \frac{TP}{\sigma_n^2}\right)$.

Remarks 2: For Nyquist signaling, i.e., $\delta = 1$ with any No-ISI satisfying pulse shape, i.e., $G_T(f/T) = T$, the capacity-achieving input and the achieved capacity are $S_a(f) = P$

and $C(P, \delta T) = \frac{1}{2T} \log_2 \left(1 + \frac{TP}{\sigma_n^2}\right)$, respectively. This verifies dissipation of the excess bandwidth with Nyquist signaling.

Remarks 3: For any non-orthogonal signaling, i.e., $\delta < 1$, the capacity-achieving input and the capacity expression are calculated by inserting $\mu = \min\{1, f_{BW}\delta T\}$ in (12) and (13), respectively, where f_{BW} is the frequency support of $g(t)$.

Example 1: Consider the RC pulse shape with a roll-of-factor of β whose Fourier transform is denoted as:

$$G(f) = \begin{cases} T & |f| \leq \frac{1-\beta}{2T} \\ \frac{T}{2}\left[1 + \cos\left(\frac{\pi T}{\beta}\left(|f| - \frac{1-\beta}{2T}\right)\right)\right] & \frac{1-\beta}{2T} < |f| \leq \frac{1+\beta}{2T} \\ 0 & o.w. \end{cases}$$

The folded-scaled spectrum $G_{\delta T}\left(\frac{f}{\delta T}\right) = \sum_{k=-\infty}^{\infty} G\left(\frac{f}{\delta T} - \frac{k}{\delta T}\right)$ is depicted in Fig. 2 for $T = 1$ and different values of δ and β. Observe that for $\delta = 1$, all the pulse shapes with different roll-of-factors result in a constant folded-scaled spectrum. Decreasing δ alters the shape of the folded-scaled spectrum.
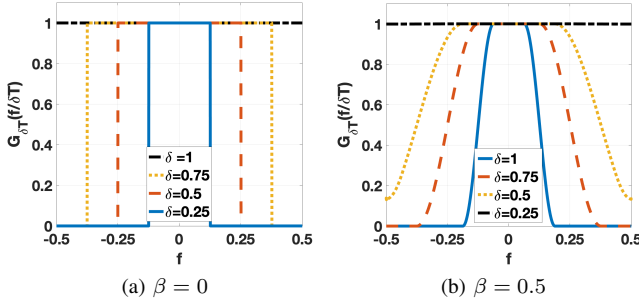

(a) $\beta = 0$ (b) $\beta = 0.5$

Fig. 2: Illustration of the folded-scaled spectrum for different values of β and δ.

It can be calculated that $\mu = \min\{1, (1+\beta)\delta\}$ and as a result:

$$C(P, \delta T) = \begin{cases} \frac{1}{2\delta T} \log_2 \left(1 + \frac{\delta TP}{\sigma_n^2}\right) & \frac{1}{1+\beta} < \delta \leq 1 \\ \frac{1+\beta}{2T} \log_2 \left(1 + \frac{TP}{(1+\beta)\sigma_n^2}\right) & \delta \leq \frac{1}{1+\beta} \end{cases},$$

which is obtained by the correlated input expressed in (12). On the other hand, with independent input, i.e., $S_a(f) = P$, the achievable rate is equal to $C_{ind}(P, \delta T) = \int_0^{1/2\delta T} \log_2 \left(1 + \frac{P}{\sigma_n^2} G_{\delta T}(f)\right) df$ also derived in [3] .
A comprehensive comparison of C and C_{ind} is illustrated in Fig. 3 for $P = 20$ dB. Note that for the compression factor of $\frac{1}{1+\beta} < \delta < 1$, the excess bandwidth is exploited by FTN and the achievable rate is increased. However, as $\delta < \frac{1}{1+\beta}$, further compression provides no extra gain. Using FTN, the capacity-achieving input and the independent input can improve the performance of the Nyquist signaling up to 40% and 65%, respectively.

These results state that, in terms of capacity, Nyquist signaling is optimal only for the sinc pulse. For other pulse shapes, FTN can exploit the excess bandwidth and improve the capacity. Although FTN signaling with independent input can improve the achievable rate, the correlated input achieves the capacity and further boosts the advantage of
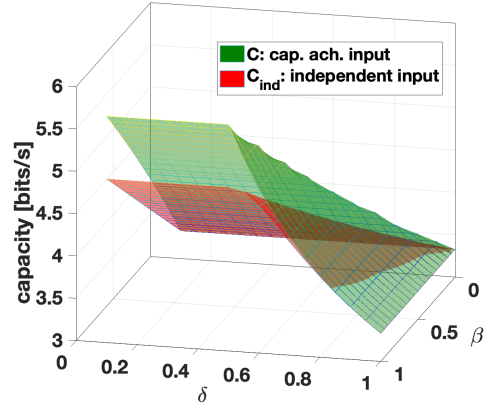

Fig. 3: Comparison of C and C_{ind} for $P = 20$ dB

FTN signaling. However, the capacity-achieving input imposes hardware-related problems, stability issues and out-of-band (OOB) emission in practical settings. This is because the PSD of the capacity-achieving input is proportional to the inverse of the folded-scaled spectrum. Thus, the capacity-achieving PSD can take very large values due to the small values of the folded-scaled spectrum.

V. NEW CONSTRAINED CAPACITY

To solve the mentioned practical issues, we introduce an upperbound constraint on the input PSD, i.e., $S_a(f) \leq \theta P$ where $\theta \in [1, \infty)$ can be chosen based on the hardware capabilities and system sensitivity to OOB emission. Intuitively, looser upperbound results in higher capacity and on contrary more stringent upperbound results in lower capacity in exchange for less hardware complications and lower OOB emission. The constrained capacity for the FTN signaling with PSD upperbound-constraint can be stated as:

$$C_\theta(P, \delta T) = \sup_{S_a(f)} \frac{1}{2\delta T} \int_{-1/2}^{1/2} \log_2 \left(1 + \frac{1}{\sigma_n^2} G_{\delta T}(\frac{f}{\delta T}) S_a(f)\right) df$$

$$s.t. \quad \frac{1}{\delta T} \int_{-1/2}^{1/2} S_a(f) G_{\delta T}(\frac{f}{\delta T}) df \leq P$$

$$S_a(f) \leq \theta P. \tag{14}$$

The KKT conditions can be written as:

$$\lambda_1 \left(\frac{1}{\delta T} \int_{-1/2}^{1/2} S_a(f) G_{\delta T}(\frac{f}{\delta T}) df - P\right) = 0,$$

$$\lambda_2(f)\left(S_a(f) - \theta P\right) = 0,$$

$$\frac{G_{\delta T}(\frac{f}{\delta T})}{\sigma_n^2 + G_{\delta T}(\frac{f}{\delta T}) S_a(f)} - 2\lambda_1 G_{\delta T}(\frac{f}{\delta T}) - 2\delta T\lambda_2(f) = 0,$$

where $\lambda_1 \geq 0$ and $\lambda_2(f) \geq 0$ are the Lagrange multipliers. Therefore, $S_a(f)$ can be calculated as: $S_a(f) =$

$$\begin{cases} \frac{1}{2\lambda_1 G_{\delta T}(f/\delta T) + 2\delta T\lambda_2(f)} - \frac{\sigma_n^2}{G_{\delta T}(f/\delta T)} & G_{\delta T}(f/\delta T) \neq 0 \\ 0 & o.w. \end{cases}.$$
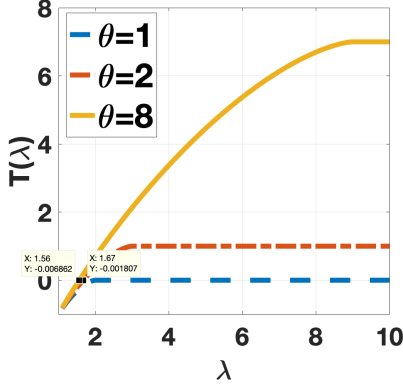
Considering the KKT conditions and with some calculations, $S_a(f)$ can be further simplified as:

$$S_a(f) = \begin{cases} \theta P & \lambda > \theta P G_{\delta T}(f/\delta T) + \sigma_n^2 \\ \frac{\lambda - \sigma_n^2}{G_{\delta T}(f/\delta T)} & \theta P G_{\delta T}(f/\delta T) + \sigma_n^2 \geq \lambda > \sigma_n^2 \\ 0 & o.w. \end{cases}, \tag{15}$$
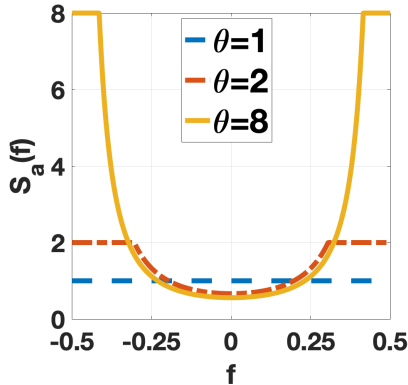
where $\lambda = 1/2\lambda_1$. The procedure to find λ is similar to the well-known water-filling algorithm. Defining $\mathcal{U} = \{f \in [-1/2, 1/2] | \lambda > \theta P G_{\delta T}(f/\delta T) + \sigma_n^2\}$ and $\mathcal{V} = \{f \in [-1/2, 1/2] | \theta P G_{\delta T}(f/\delta T) + \sigma_n^2 \geq \lambda\}$, then the power constraint can be written as:

$$T(\lambda) = \frac{\theta P}{\delta T} \int_{f \in \mathcal{U}} G_{\delta T}(\frac{f}{\delta T}) df + (\lambda - \sigma_n^2)\frac{\mu_{\mathcal{V}}}{\delta T} - P = 0,$$

where $T(\lambda) \leq (\theta - 1)P$ is a monotonically increasing function

(a) Illustration of $T(\lambda)$ with respect to λ



(b) θ-limited PSD for different values of θ

Fig. 4: Analysis of the proposed algorithm for finding θ-limited PSD.

VI. NUMERICAL RESULTS

In this section, we analyze the derived results using the RC pulse. First, we examine the average transmit power with Nyquist and FTN signaling. We assume that the inputs are drawn from either an independent Gaussian process or a Gaussian process with capacity-achieving PSD of $S_a(f) =$

$\begin{cases} \frac{\delta T P}{\mu G_{\delta T}(f/\delta T)} & G_{\delta T}(f/\delta T) \neq 0 \\ 0 & o.w. \end{cases}$. To generate the capacity-achieving symbols, a random realization of white noise is convolved by a sequence with PSD of $S_a^{1/2}(f)$. Next, the symbols are modulated by a root raised cosine filter. Realizations of such sequences are shown in Fig. 5a for independent and capacity-achieving symbols with Nyquist rate, i.e., $\delta = 1$ and a FTN rate, i.e., $\delta = 0.6$. Then, the average power $P_x = \mathcal{E}\left\{\int_t |x(t)|^2 dt\right\}$ is shown in Fig. 5b for different values of δ. Fig. 5 shows that the capacity-achieving PSD combined
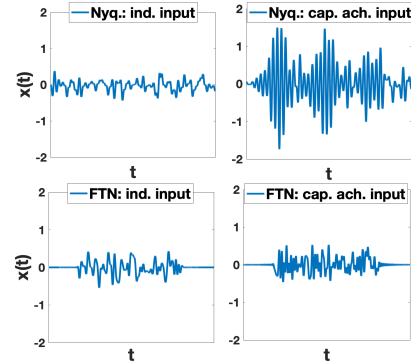
(a) The time domain modulated signal for Nyquist and FTN signaling with $\delta = 0.6$
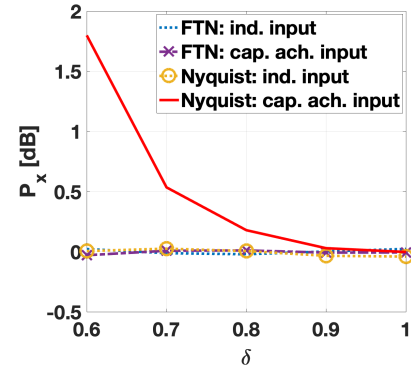
(b) P_x versus δ for Nyquist and FTN signaling

Fig. 5: Illustration of the effect of FTN on transmit power.

with FTN signaling satisfies the power constraint. On the other hand, the same PSD combined with Nyquist signaling, results in large instantaneous and average power. Also, note that independent PSD satisfies the power constraint and results in the same average power for all values of δ.

In Fig. 6, the effect of FTN signaling on the capacity is illustrated. As the signaling rate increases, the capacity increases. However, as δ drops below $\frac{1}{1+\beta}$, the capacity does not improve anymore. In addition, for $\beta \geq \frac{1}{\delta} - 1$, change of β has no effect on the capacity which shows the waste of the excess bandwidth.

As mentioned before, the capacity-achieving PSD imposes some practical issues like OOB emission. The OOB emission is calculated as the total power leaked out of the allocated sub-band, and shown in Fig. 7 for $\beta = 1, \delta = 0.5, P = 10\ dB$. As θ increases, the OOB emission increases and for large values of θ, the OOB emission approaches an unaccepted level of $5\ dB$. On the other hand, by increasing θ, the achievable

in λ as shown in Fig. 4a for $\beta = 1, \delta = 0.5, P = 1$ and different values of θ. Therefore, the optimal λ can be found by a simple bisection search. The optimal values of λ in Fig. 4a are 1.57 and 1.67 for $\theta = 8$ and $\theta = 2$, respectively. The optimal PSD can be calculated by inserting the optimal λ in (15). Using the proposed algorithm, the optimal PSDs are illustrated in Fig. 4b for $\beta = 1, \delta = 0.5, P = 1$ and different values of θ. In summary, the proposed method limits the very large values of the original PSD to θP and then adjusts the rest of the PSD in order to satisfy the power constraint.
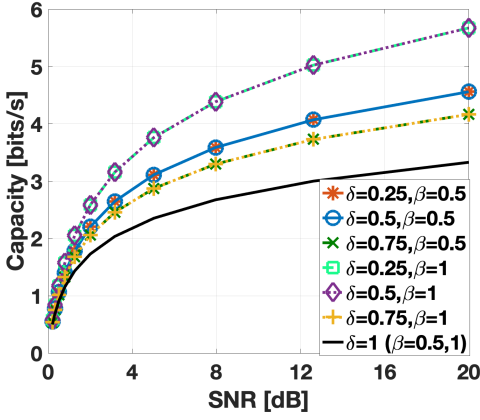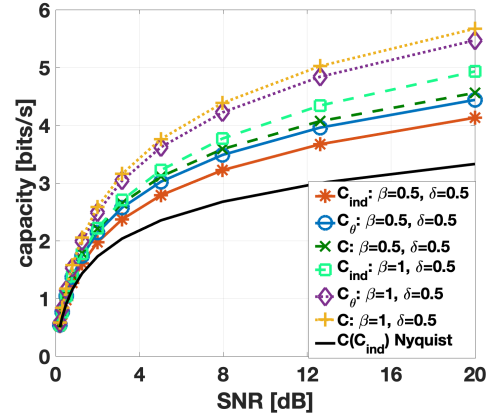
Fig. 6: Illustration of the effect of FTN on capacity.

Fig. 8: Comparison of C, C_θ and $C_{ind.}$.

rate increases. In fact, we have $C_1 = C_{ind}$ and $C_\infty \to C$. Therefore, θ provides a flexible mechanism to capture the trade-off between an acceptable OOB emission and a desired rate. For example, $\theta = 10$ can result in an acceptable OOB emission of $\sim -30\ dB$ while it provides $\sim 2.5\ bits/s$ rate (i.e., 96% of the FTN capacity).
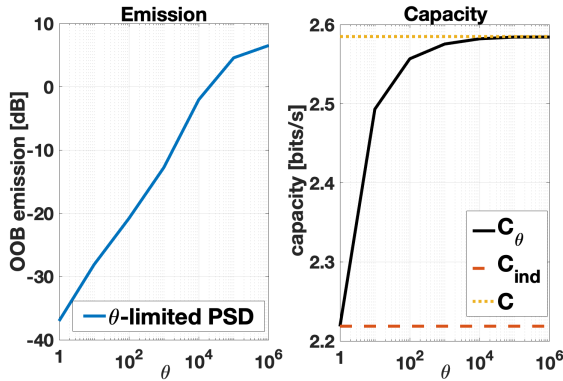

Fig. 7: Illustration of the effect of θ on OOB emission and capacity.

In Fig. 8, the capacity of θ-limited PSD with $\theta = 10$ is compared with that of independent and capacity-achieving PSDs for $\beta = 0.5, 1$ and $\delta = 0.5$. Not only does the proposed PSD maintain an acceptable OOB emission, but also it provides $\sim 10\%$ gain compared with the independent PSD.

VII. CONCLUSION

In this work, we analyzed the capacity of FTN signaling. We showed that an independent PSD does not achieve the capacity of FTN signaling. Instead, the capacity-achieving PSD is derived which is proportional to the inverse of the folded-scaled spectrum of the pulse shape. However, the capacity-achieving input causes practical issues such as OOB emission. To solve these issues, a new optimizing problem is introduced with an additional constraint on the input PSD. The introduced θ-limited input PSD can strike a trade-off between the independent input and the capacity-achieving input. The

value of θ can be engineered to achieve an acceptable OOB emission and a desirable capacity.

REFERENCES

[1] P.-D. Arapoglou, A. Ginesi, S. Cioni, S. Erl, F. Clazzer, S. Andrenacci, and A. Vanelli-Coralli, "DVB-S2X-enabled precoding for high through-put satellite systems," *International Journal of Satellite Communications and Networking*, vol. 34, no. 3, pp. 439–455, May 2016.
[2] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
[3] F. Rusek and J. B. Anderson, "Constrained capacities for faster-than-Nyquist signaling," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 764–775, Feb. 2009.
[4] J. E. Mazo, "Faster-than-Nyquist signaling," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1451–1462, 1975.
[5] D. Dasalukunte, V. Owall, F. Rusek, and J. B. Anderson, *Faster than Nyquist Signaling*. Springer, 2016.
[6] Y. G. Yoo and J. H. Cho, "Asymptotic optimality of binary faster-than-Nyquist signaling," *IEEE Commun. Lett.*, vol. 14, no. 9, pp. 788–790, Aug. 2010.
[7] A. D. Liveris and C. N. Georghiades, "Exploiting faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1502–1511, Sep. 2003.
[8] F. Rusek and J. B. Anderson, "Non binary and precoded faster than Nyquist signaling," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 808–817, May 2008.
[9] H. Sasahara, K. Hayashi, and M. Nagahara, "Symbol detection for faster-than-Nyquist signaling by sum-of-absolute-values optimization," *IEEE Signal Process. Lett.*, vol. 23, no. 12, pp. 1853–1857, Nov. 2016.
[10] E. Bedeer, M. H. Ahmed, and H. Yanikomeroglu, "A very low complexity successive symbol-by-symbol sequence estimator for faster-than-Nyquist signaling," *IEEE Access*, vol. 5, pp. 7414–7422, Mar. 2017.
[11] Y. J. D. Kim, J. Bajcsy, and D. Vargas, "Faster-than-Nyquist broadcasting in Gaussian channels: Achievable rate regions and coding," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1016–1030, Jan. 2016.
[12] D. Spano, M. Alodeh, S. Chatzinotas, and B. Ottersten, "Faster-than-Nyquist signaling through spatio-temporal symbol-level precoding for the multiuser MISO downlink channel," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 5915–5928, Jul. 2018.
[13] D. Kapetanovic and F. Rusek, "The effect of signaling rate on information rate for single carrier linear transmission systems," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 421–428, Jan. 2012.
[14] W. Hirt and J. L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 34, no. 3, pp. 380–388, May 1988.
[15] S. Verdu, "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 35, no. 4, pp. 733–751, Jul. 1989.
[16] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
[17] U. Grenander and G. Szego, *Toeplitz forms and their applications*. University of California Press, 1958.