Capacity of Gaussian Channels with Duty Cycle and Power Constraints

Lei Zhang and Dongning Guo
Department of Electrical Engineering and Computer Science
Northwestern University, Evanston, IL 60208, USA

***Abstract*—In many wireless communication systems, radios are subject to *duty cycle* constraint, that is, a radio only actively transmits signals over a fraction of the time. For example, it is desirable to have a small duty cycle in some low power systems; a half-duplex radio cannot keep transmitting if it wishes to receive useful signals; and a cognitive radio needs to listen and detect primary users frequently. This work studies the capacity of scalar discrete-time Gaussian channels subject to duty cycle constraint as well as average transmit power constraint. The duty cycle constraint can be regarded as a requirement on the minimum fraction of nontransmission or zero symbols in each codeword. A unique *discrete* input distribution is shown to achieve the channel capacity. In many situations, numerical results demonstrate that using the optimal input can improve the capacity by a large margin compared to using Gaussian signaling over a deterministic transmission schedule, which is capacity-achieving in the absence of the duty cycle constraint. This is in part because the positions of the nontransmission symbol in a codeword can convey information. The results suggest that, under the duty cycle constraint, departing from the usual paradigm of intermittent packet transmissions may yield substantial gain.**

I. Introduction

In many wireless communication systems, a radio is designed to transmit actively only for a fraction of the time, which is known as its *duty cycle*. For example, the ultra-wideband system in [1] transmits short bursts of signals to trade bandwidth for power savings. The physical half-duplex constraint also requires a radio to stop transmission over a frequency band from time to time if it wishes to receive useful signals over the same band. Thus wireless relays are subject to duty cycle constraint, so do cognitive radios which have to listen to the channel frequently to avoid causing interference to primary users. The *de facto* standard solution under duty cycle constraint is to transmit packets intermittently.

This work studies the question of what is the optimal signaling for a Gaussian channel with duty cycle constraint as well as average transmission power constraint. An important observation is that the signaling in nontransmission periods can be regarded as transmission of the special *zero* signal. To keep the discussion concise, we make a simplistic and ideal assumption that the analog waveform corresponding to each transmitted symbol is exactly of the span of one symbol interval. We restrict our attention to discrete-time scalar additive white Gaussian noise (AWGN) channels for simplicity, where the duty cycle constraint is equivalent to a requirement on the minimum fraction of zero symbols in each transmitted codeword. The mathematical model of the AWGN channel and input constraints is described in Section II. In case a practical pulse shaping filter introduces higher duty cycle in continuous time than its discrete-time counterpart, we can use a smaller duty cycle in discrete time to satisfy the continuous-time constraint. This, however, is out of the scope of this paper.

Determining the capacity of a channel subject to various input constraints is a classical problem. It is well-known that Gaussian signaling achieves the capacity of a Gaussian channel with average input power constraint only. Smith [2] investigated the capacity of a scalar AWGN channel under both peak power constraint and average power constraint. The input distribution that achieves the capacity is shown to be discrete with a finite number of probability mass points. The discreteness of capacity-achieving distributions for various channels, including quadrature Gaussian channels, and Rayleigh-fading channels is also established in [3]–[8]. Chan [9] studied the capacity-achieving input distribution for conditional Gaussian channels which form a general channel model for many practical communication systems.

The impact of duty cycle constraint on capacity-achieving signaling is underexplored in the literature. In this paper, we use a similar approach as in [2] and [9] to show that the capacity-achieving input distribution for an AWGN channel with duty cycle and average power constraints is discrete. The optimal distribution has an infinite number of probability mass points, whereas only a finite number of the points are found in every bounded interval. This allows efficient numerical optimization of the input distribution. The main theorem is reported in Section III and its proof is completed in Section IV.

Numerical results in Section V demonstrate that using optimized discrete signaling may achieve significantly higher rates than using Gaussian signaling over a deterministic transmission schedule. For example, in case the radio is allowed to transmit no more than half the time, i.e., the duty cycle is no greater than 50%, a near-optimal discrete input achieves 50% higher rate at 10 dB signal-to-noise ratio (SNR). This suggests that, compared to intermittently transmitting packets using Gaussian or Gaussian-like signaling, it is more efficient to disperse nontransmission symbols within each packet to form codewords, which results in a form of *on-off* signaling.

This work was initiated during the authors' visit to the Institute of Network Coding at the Chinese University of Hong Kong. The work of D. Guo was partially supported by a grant from the University Grants Committee of the Hong Kong Special Administrative Region, China (Project No. AoE/E-02/08). This work was also partially supported by the NSF under Grant No. 0644344 and by DARPA under Grant No. W911NF-07-1-0028.

978-1-4577-0595-3/11/$26.00 ©2011 IEEE

A key reason for the superiority of on-off signaling is that the positions of nontransmission symbols can be used to convey a substantial amount of information, especially in case of low SNR and low duty cycle. This has been observed in the past. For example, as shown in [10] (see also [11], [12]), time sharing or time-division duplex (TDD) can fall considerably short of the theoretical limits in a relay network: The capacity of a cascade of two noiseless binary bit pipes through a half-duplex relay is 1.14 bits per channel use, which far exceeds the 0.5 bit achieved by TDD and even the 1 bit upper bound on the rate of binary signaling.

Besides that duty cycle constraint is frequently seen in practice, another motivation of this study is a recent work [13], in which on-off signaling is proposed for a clean-slate design of wireless ad hoc networks formed by half-duplex radios. Using this signaling scheme, which is called rapid on-off-division duplex (RODD), a node listens to the channel and receives useful signals during its own off symbols within each frame. Each node can transmit and receive messages at the same time over one frame interval, thereby achieving (virtual) full-duplex communication. Understanding the impact of duty cycle constraint is crucial to characterizing the fundamental limits of such wireless networks.

II. System Model

Consider digital communication systems where coded data are mapped to waveforms for transmission. Usually there is a collection of waveforms, all of one symbol duration, where each waveform represents a symbol (or letter) from a discrete alphabet. Without loss of capacity, we assume some linear modulation of the signal is used. We view nontransmission over a symbol interval as transmitting the all zero waveform. In other words, a symbol interval of nontransmission is simply regarded as transmitting the symbol 0, which carries no energy.

For simplicity, we consider the baseband discrete-time model for the AWGN channel. The received signal over a block of n symbols can be described by

$$Y_i = X_i + N_i \tag{1}$$

where $i = 1, \dots, n$, X_i denotes the transmitted symbol at time i and $N_1, \dots, N_n$ are independent standard Gaussian random variables. The constraint that the duty cycle is no greater than $1-q$, where $q \in (0,1)$, can be considered as a constraint that the fraction of symbol 0 in the input codeword is no less than q. That is, every codeword $(x_1, x_2, \dots, x_n)$ satisfies

$$\frac{1}{n}\sum_{i=1}^{n} 1_{\{x_i \neq 0\}} \leq 1 - q \tag{2}$$

where $1_{\{\cdot\}}$ is the indicator function. In addition, we consider the usual average input power constraint, which is that the SNR of the channel is no greater than P.

In many wireless systems, the transmitter's activity is constrained in the frequency domain as well as in the time domain. In fact the results in this paper also apply to the case where the duty cycle constraint is on the time-frequency plane.

III. The Capacity-Achieving Input

The conditional probability density function (pdf) of the output given the input of the AWGN channel (1) is

$$p_{Y|X}(y|x) = \phi(y-x) \tag{3}$$

where

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} . \tag{4}$$

Let μ denote the distribution of the channel input X. In general, μ is a probability measure defined on the Borel algebra on the real number set, denoted by $\mathcal{B}(\mathbb{R})$. The pdf of the output exists and is:

$$p_Y(y;\mu) = \int p_{Y|X}(y|x)\,\mu(\mathrm{d}x) = \mathsf{E}_\mu\left\{\phi(y-X)\right\} . \tag{5}$$

Denote the relative entropy $D\left(p_{Y|X}(\cdot|x) \| p_Y(\cdot;\mu)\right)$ by $d(x;\mu)$, which is expressed as

$$d(x;\mu) = \int_{-\infty}^{\infty} p_{Y|X}(y|x) \log \frac{p_{Y|X}(y|x)}{p_Y(y;\mu)} \mathrm{d}y . \tag{6}$$

The mutual information $I(\mu) = I(X;Y)$ is then

$$I(\mu) = \int d(x;\mu)\,\mu(\mathrm{d}x) = \mathsf{E}_\mu\left\{d(X;\mu)\right\} . \tag{7}$$

We denote the set of distributions with duty cycle constraint $1-q$ and power constraint P by:

$$\Lambda_P = \left\{\mu : \mu(\{0\}) \geq q,\, \mathsf{E}_\mu\left\{X^2\right\} \leq P\right\} \tag{8}$$

and the finite-power set as $\Lambda = \cup_{0 \leq P < \infty} \Lambda_P$.

Theorem 1: The capacity of the AWGN channel (1) with duty cycle no greater than $1-q$ and SNR constraint P is

$$C(P) = \max_{\mu \in \Lambda_P} I(\mu) . \tag{9}$$

In particular, the following properties hold:

a) the maximum of (9) is achieved by a unique (capacity-achieving distribution) $\mu_0 \in \Lambda_P$;
b) μ_0 is symmetric about 0 and its second moment is exactly equal to P; and
c) μ_0 is discrete with an infinite number of probability mass points, whereas the number of probability mass points in any bounded interval is finite.

Proof: The capacity of the AWGN channel under per-letter duty cycle constraint and power constraint is evidently given by the supremum of the mutual information $I(\mu)$ where $\mu \in \Lambda_P$. The achievability and converse of this result can be established using standard techniques in information theory.

We first establish Property (a). Let $\mathcal{P}$ denote the collection of all Borel probability measures defined on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, which is a topological space with the topology of weak convergence [14]. The following useful lemma is given without proof due to space limitations:

Lemma 1: Λ_P is compact in the topological space $\mathcal{P}$.

Since the mutual information $I(\mu)$ is continuous on $\mathcal{P}$ according to [15, Theorem 9], it must achieve its maximum on the compact set Λ_P. Hence μ_0 exists.

According to [15, Corollary 2], the mutual information $I(\mu)$ is strictly concave. It is easy to see that Λ_P is convex. Hence the maximizer μ_0 must be unique.

Property (b) is established as follows. Since the mirror reflection of μ_0 about 0 is evidently also a maximizer of (9), the uniqueness requires that μ_0 be symmetric. Note that linear scaling of the input to increase its power maintains its duty cycle and cannot reduce the mutual information, as the receiver can add noise to maintain the same SNR. By the uniqueness of the maximizer μ_0, the power constraint must be binding, i.e., the second moment of μ_0 must be equal to P.

We relegate the proof of Property (c) to Section IV. ∎

IV. Discreteness of the Optimal Input

A. Sufficient and Necessary Conditions

The relative entropy $d(x;\mu)$ defined in (6) can be extended to the complex plane $\mathbb{C}$ and has the following property:

Lemma 2: For any $\mu \in \Lambda$ and $z \in \mathbb{C}$,

$$d(z;\mu) = \int_{-\infty}^{\infty} \phi(y-z) \log \frac{\phi(y-z)}{p_Y(y;\mu)} \mathrm{d}y \tag{10}$$

is well defined and it is a holomorphic function of z on $\mathbb{C}$. Consequently, $d(x;\mu)$ is a continuous function of x on $\mathbb{R}$.

The proof of Lemma 2 is omitted due to space limitations. We establish the following result.

Lemma 3: Let

$$f_\lambda(x;\mu) = d(x;\mu) - I(\mu) - \lambda(x^2 - P). \tag{11}$$

Then $\mu_0 \in \Lambda_P$ achieves the capacity if and only if there exists $\lambda \geq 0$ such that $\lambda \mathsf{E}_{\mu_0}\left\{X^2 - P\right\} = 0$ and $\mathsf{E}_\mu\left\{f_\lambda(X;\mu_0)\right\} \leq 0$ for all $\mu \in \Lambda$.

Proof: Define the Lagrangian

$$J(\mu) = I(\mu) - \lambda \mathsf{E}_\mu\left\{X^2 - P\right\} \tag{12}$$

where λ is the Lagrange multiplier. Since Λ is a convex set and $I(\mu) < \infty$ on Λ, μ_0 is capacity-achieving if and only if there exists $\lambda \geq 0$ such that the following conditions hold [16]:

(i) $\lambda \mathsf{E}_{\mu_0}\left\{X^2 - P\right\} = 0$;
(ii) for all $\mu \in \Lambda$, $J(\mu_0) \geq J(\mu)$.

Due to concavity of $I(\mu)$, $J(\mu)$ is also concave. Define the weak derivative of $J(\mu)$ at μ_0 as

$$J'_{\mu_0}(\mu) = \lim_{\theta \to 0^+} \frac{J\left((1-\theta)\mu_0 + \theta\mu\right) - J(\mu_0)}{\theta}. \tag{13}$$

Condition (ii) is then equivalent to that $J'_{\mu_0}(\mu) \leq 0$ for all $\mu \in \Lambda$. The following result, which finds its parallel in [5] and [9], is useful and given without proof.

Lemma 4: Let $\mu_0, \mu \in \Lambda$, the weak derivative of the mutual information function $I(\mu)$ at μ_0 is

$$I'_{\mu_0}(\mu) = \int d(x;\mu_0)\,\mu(\mathrm{d}x) - I(\mu_0). \tag{14}$$

By Lemma 4, the linearity of $\mathsf{E}_\mu\left\{X^2 - P\right\}$ with respect to (w.r.t.) μ and Condition (i), $J'_{\mu_0}(\mu)$ can be easily calculated from (13) as

$$J'_{\mu_0}(\mu) = \mathsf{E}_\mu\left\{f_\lambda(X;\mu_0)\right\}. \tag{15}$$

Therefore, Condition (ii) is equivalent to $\mathsf{E}_\mu\left\{f_\lambda(X;\mu_0)\right\} \leq 0$ for all $\mu \in \Lambda$. Thus Lemma 3 follows. ∎

We call $x \in \mathbb{R}$ a point of increase of a measure μ if $\mu(O) > 0$ for every open subset O of $\mathbb{R}$ containing x. Let S_μ be the set of points of increase of μ. Based on Lemma 3, we derive another sufficient and necessary condition for the optimal input distribution, which will be used to prove Property (c) of Theorem 1 in Section IV-B.

Lemma 5: Let

$$g_\lambda(x;\mu) = q f_\lambda(0;\mu) + (1-q) f_\lambda(x;\mu). \tag{16}$$

Then $\mu_0 \in \Lambda_P$ achieves the capacity if and only if there exists $\lambda \geq 0$ such that for every $x \in \mathbb{R}$,

$$g_\lambda(x;\mu_0) \leq 0\,. \tag{17}$$

Furthermore, $g_\lambda(x;\mu_0) = 0$ for every $x \in S_{\mu_0}\backslash\{0\}$.

Proof: The necessity part is shown as follows. Suppose μ_0 achieves the capacity, then by Lemma 3, there exists $\lambda \geq 0$ such that $\lambda \mathsf{E}_{\mu_0}\left\{X^2 - P\right\} = 0$ and $\mathsf{E}_\mu\left\{f_\lambda(X;\mu_0)\right\} \leq 0$ for all $\mu \in \Lambda$. For any $x \in \mathbb{R}\backslash\{0\}$, choose μ such that $\mu(\{0\}) = q$ and $\mu(\{x\}) = 1-q$, so by the fact that $\mu \in \Lambda$, we have

$$0 \geq \mathsf{E}_\mu\left\{f_\lambda(X;\mu_0)\right\} = q f_\lambda(0;\mu_0) + (1-q) f_\lambda(X;\mu_0). \tag{18}$$

Due to the continuity of $d(x;\mu_0)$ by Lemma 2, $f_\lambda(x;\mu_0)$ is also continuous so that (18) holds for all $x \in \mathbb{R}$, i.e., $g_\lambda(x;\mu_0) \leq 0$ for every $x \in \mathbb{R}$.

To finish proving the necessity, it suffices to show that $g_\lambda(x;\mu_0) = 0$ for all $x \in S_{\mu_0}\backslash\{0\}$. Evidently, $g_\lambda(0;\mu_0) = f_\lambda(0;\mu_0)$ and by (7) and $\lambda \mathsf{E}_{\mu_0}\left\{X^2 - P\right\} = 0$,

$$\int f_\lambda(x;\mu_0)\,\mu_0(\mathrm{d}x) = 0\,. \tag{19}$$

Hence,

$$\begin{aligned}
&\int_{\mathbb{R}\backslash\{0\}} g_\lambda(x;\mu_0)\,\mu_0(\mathrm{d}x) \\
&= \int g_\lambda(x;\mu_0)\,\mu_0(\mathrm{d}x) - g_\lambda(0;\mu_0)\mu_0(\{0\}) \\
&\geq q f_\lambda(0;\mu_0) + (1-q)\int f_\lambda(x;\mu_0)\,\mu_0(\mathrm{d}x) - q f_\lambda(0;\mu_0) \\
&= 0
\end{aligned} \tag{20}$$

Since $g_\lambda(x;\mu_0) \leq 0$ for every $x \in \mathbb{R}$, (20) implies that on $\mathbb{R}\backslash\{0\}$, $g_\lambda(x;\mu_0) = 0$ μ_0-almost surely, so that $g_\lambda(x;\mu_0) = 0$ for all $x \in S_{\mu_0}\backslash\{0\}$ follows immediately.

The sufficiency part of Lemma 5 is established as follows. Suppose $g_\lambda(x;\mu_0) \leq 0$ for every $x \in \mathbb{R}$. By integrating $g_\lambda(x;\mu_0)$ w.r.t. μ_0, we have

$$\begin{aligned}
q g_\lambda(0;\mu_0) &\geq \int g_\lambda(x;\mu_0)\,\mu_0(\mathrm{d}x) \\
&= q g_\lambda(0;\mu_0) - (1-q)\lambda \mathsf{E}_{\mu_0}\left\{X^2 - P\right\} \qquad (21)\\
&\geq q g_\lambda(0;\mu_0) \qquad (22)
\end{aligned}$$

where (21) is due to (7) and $g_\lambda(0;\mu_0) = f_\lambda(0;\mu_0)$, and (22) follows from $\mathsf{E}_{\mu_0}\left\{X^2\right\} \leq P$ since $\mu_0 \in \Lambda_P$. Hence,

$\lambda \mathsf{E}_{\mu_0}\left\{X^2 - P\right\} = 0$ due to the fact that $q < 1$. Furthermore, for any $\mu \in \Lambda$, by integrating $g_\lambda(x;\mu_0)$ w.r.t. μ, we have

$$\begin{aligned} q g_\lambda(0;\mu_0) &\geq \int g_\lambda(x;\mu_0)\,\mu(\mathrm{d}x) \\ &= q f_\lambda(0;\mu_0) + (1-q)\mathsf{E}_\mu\left\{f_\lambda(X;\mu_0)\right\}. \end{aligned} \quad (23)$$

Because $g_\lambda(0;\mu_0) = f_\lambda(0;\mu_0)$, we have $\mathsf{E}_\mu\left\{f_\lambda(X;\mu_0)\right\} \leq 0$. Together with $\lambda \mathsf{E}_{\mu_0}\left\{X^2 - P\right\} = 0$ and Lemma 3, this implies that μ_0 must be capacity-achieving. ■

B. Proof of the Discreteness of μ_0

With Lemma 5 established, we now prove Property (c) in Theorem 1.

Let $\lambda \geq 0$ satisfy condition (17) and $d(z;\mu)$ be defined in (10). We extend functions $f_\lambda(x;\mu)$ in Lemma 3 and $g_\lambda(x;\mu)$ in Lemma 5 to be defined on the whole complex plane $\mathbb{C}$ as (11) and (16), respectively, with x replaced by $z \in \mathbb{C}$. By Lemma 2, $d(z;\mu)$ is a holomorphic function of z on $\mathbb{C}$, hence so is $g_\lambda(z;\mu)$. According to Lemma 5, each element in the set $S_{\mu_0}\backslash\{0\}$ is a zero of the function $g_\lambda(z;\mu_0)$.

Next we show that for any bounded interval L of $\mathbb{R}$, $S_{\mu_0} \bigcap L$ is a finite set. Suppose, to the contrary, $S_{\mu_0} \bigcap L$ is infinite, then it has a limit point in $\mathbb{R}$ by the Bolzano-Weierstrass Theorem [17] and hence, $g_\lambda(z;\mu_0) = 0$ on the whole complex plane $\mathbb{C}$ by the Identity Theorem [18]. Then, by (6), (11) and (16), for every $x \in \mathbb{R}$,

$$\int_{-\infty}^{\infty} \phi(y-x) r(y)\mathrm{d}y = 0 \quad (24)$$

where

$$r(y) = \log p_Y(y;\mu_0) + \lambda y^2 + c \quad (25)$$

and $c = \frac{1}{2}\log(2\pi e) + I(\mu_0) - \frac{q}{1-q} d(0) - \lambda(P+1)$ is a constant.

By (5) and Jensen's inequality, we have

$$1 \geq p_Y(y;\mu) \geq \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\mathsf{E}_\mu\left\{(y-X)^2\right\}} = e^{-\frac{1}{2}y^2 - ay - b} \quad (26)$$

where $a = -\mathsf{E}_\mu\left\{X\right\}$ and $b = \frac{1}{2}\left(\mathsf{E}_\mu\left\{X^2\right\} + \log(2\pi)\right)$. By the assumption that $\mathsf{E}_\mu\left\{X^2\right\} < \infty$, it is easy to see that $a, b \in \mathbb{R}$, so $|\log p_Y(y;\mu)| \leq \frac{1}{2}y^2 + ay + b$. As a result, there exist some $\alpha, \beta > 0$ such that $|r(y)| \leq \alpha y^2 + \beta$. Since the convolution of $r(y)$ and the Gaussian density is equal to the zero function by (24), $r(y)$ must be the zero function according to [9, Corollary 9]. This requires the capacity-achieving output distribution $p_Y(y;\mu_0)$ be Gaussian, which cannot be true unless X is Gaussian, which contradicts the assumption that X has a probability mass at 0. Therefore, $S_{\mu_0} \bigcap L$ must be a finite set for any bounded interval L, which further implies that S_{μ_0} is at most countable.

Finally, we show that S_{μ_0} is countably infinite. Suppose, to the contrary, $S_{\mu_0} = \{x_i\}_{i=1}^N$ is a finite set with $\mu_0(\{x_i\}) = p_i$ and $|x_i| \leq B_1$ for all $i = 1, 2, \ldots, N$. For any $y > B_1$,

$$p_Y(y;\mu_0) = \sum_{i=1}^N p_i \phi(y - x_i) \leq e^{-\frac{(y-B_1)^2}{2}}. \quad (27)$$

For any $\epsilon > 0$, choose $B_2 > 0$ such that $\int_{-B_2}^{B_2} \phi(x)\mathrm{d}x > 1-\epsilon$. By (6), (11), (16) and (17), for any $x > B_1 + B_2$, we have

$$\begin{aligned} 0 &\geq -\int_{-\infty}^{\infty} \phi(y-x) \log p_Y(y;\mu_0)\mathrm{d}y - \lambda x^2 - (c+\lambda) \\ &\geq \int_{x-B_2}^{x+B_2} \phi(y-x) \frac{1}{2}(y-B_1)^2 \mathrm{d}y - \lambda x^2 - (c+\lambda) \\ &= \int_{B_2}^{B_2} \phi(t) \frac{1}{2}(x - B_1 + t)^2 \mathrm{d}t - \lambda x^2 - (c+\lambda) \\ &\geq \frac{1}{2}(x-B_1)^2(1-\epsilon) - \lambda x^2 - (c+\lambda). \end{aligned} \quad (28)$$

For (28) to hold for large x, λ must satisfy $\lambda \geq \frac{1}{2}$.

To finish the proof, it suffices to show that $\lambda < \frac{1}{2}$ for any $P > 0$, so that contradiction arises, which implies that S_{μ_0} must be countably infinite. For fixed $q \in (0,1)$, denote the Lagrange multiplier in (17) as $\lambda(P)$. Denote $C_G(P) = \frac{1}{2}\log(1+P)$, which is the channel capacity of a Gaussian channel with the average power constraint only. By the envelope theorem [16], $\lambda(P)$ is the derivative of $C(P)$ w.r.t. P. Since $C(0) = C_G(0) = 0$ and the derivative of $C_G(P)$ at $P = 0$ is $\frac{1}{2}$, we have $\lambda(0) \leq \frac{1}{2}$, otherwise we could find a small enough P such that $C(P)$ would exceed $C_G(P)$ which is obviously impossible. Next we show that $C(P)$ is strictly concave for $P \geq 0$. Suppose μ_1 and μ_2 are the capacity-achieving input distributions of (9) for different power constraints P_1 and P_2, respectively. Due to Property (b) in Theorem 1, μ_1 and μ_2 must be different. Define $\mu_\theta = \theta\mu_1 + (1-\theta)\mu_2$ for $\theta \in (0,1)$. It is easy to see that μ_θ satisfies that the duty cycle is no greater than $1-q$ and the average input power is no greater than $\theta P_1 + (1-\theta)P_2$. Now we have

$$\begin{aligned} C(\theta P_1 + (1-\theta)P_2) &\geq I(\mu_\theta) \\ &> \theta I(\mu_1) + (1-\theta) I(\mu_2) \quad (29) \\ &= \theta C(P_1) + (1-\theta) C(P_2), \quad (30) \end{aligned}$$

where (29) is due to the strict concavity of $I(\mu)$. Therefore, the strict concavity of $C(P)$ for $P \geq 0$ follows, which implies that $\lambda(P) < \lambda(0) = \frac{1}{2}$ for all $P > 0$.

V. Numerical Results

One implication of Theorem 1 is that directly computing the capacity-achieving input distribution requires solving an optimization problem with infinite variables which is prohibitive. Assuming any upper bound on the number of probability mass points, however, a numerical optimization over the mutual information can yield a suboptimal input distribution and a lower bound on the channel capacity. As we increase the number of mass points, the lower bound can be further refined. We take this approach to numerically compute a good approximation of the channel capacity by optimizing over a sufficient number of probability mass points. Given the duty cycle and power constraints, we first numerically optimize the mutual information by a 3-point input distribution (including a mass at 0), then increase the number of probability mass

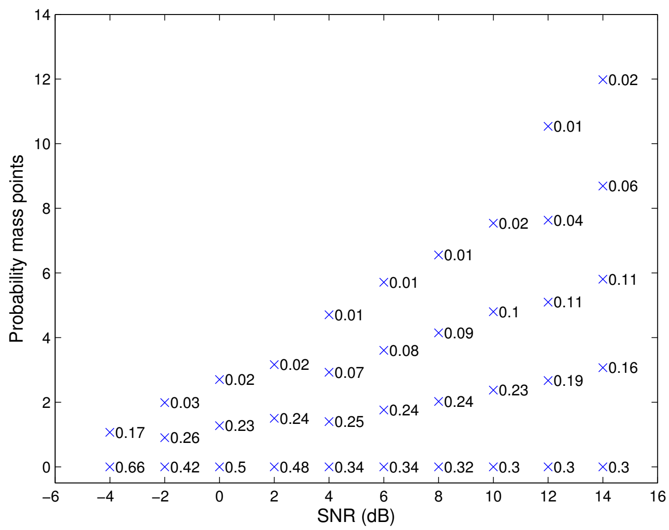

Fig. 1. Suboptimal input distribution for $P(X=0) \geq q = 0.3$.

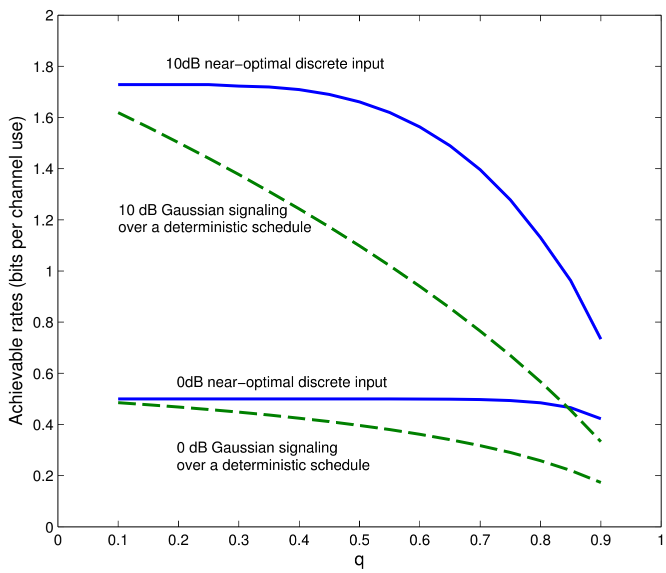

Fig. 2. Capacity under duty cycle constraint for 0 dB and 10 dB SNRs.

points by 2 at a time to improve the mutual information, until the improvement is less than 10^{-3}.

First consider the case that the duty cycle is no greater than 70%, i.e., $P(X=0) \geq q = 0.3$. For different SNRs, the mass points of the near-optimal input distribution with finite support along with the corresponding probability masses are shown in Fig. 1. Due to symmetry, only the positive half of the input distribution is plotted. We can see that as the SNR increases, more masses are put on higher-amplitude points, whereas the probability mass at zero achieves its lower bound 0.3 eventually.

In Fig 2, we compare the rate achieved by the near-optimal input distribution and the rate achieved by a conventional scheme using Gaussian signaling over a deterministic schedule, which is $(1-q)$ times the Gaussian channel capacity without duty cycle constraint. It is shown in the figure that there is substantial gain for both 0 dB and 10 dB SNRs by using discrete input over Gaussian signaling with a deterministic schedule. For example, when the SNR is 10 dB, given the duty cycle is no more than 50%, the discrete input distribution achieves 50% higher rate. Hence departing from the usual paradigm of intermittent packet transmissions may yield significant gains.

Acknowledgement

The authors would like to thank Raymond Yeung and Shuo-Yen Robert Li for hosting them in the Institute of Network Coding at the Chinese University of Hong Kong during the production of the first draft of this paper. The authors would also like to thank Terence Chan for sharing the code for the numerical results in [9] and Yihong Wu and Sergio Verdú for their helpful comments.

References

[1] D. Julian and S. Majumdar, “Low power personal area communication,” in *Proc. Inform. Theory Appl. Workshop*, La Jolla, CA, USA, 2011.
[2] J. G. Smith, “The information capacity of amplitude and variance-constrained scalar Gaussian channels,” *Inf. Contr.*, vol. 18, pp. 203–219, 1971.
[3] S. Shamai (Shitz), “Capacity of a pulse amplitude modulated direct detection photon channel,” *Proc. IEE Communications, Speech and Vision*, vol. 137, pp. 424–430, Dec. 1990.
[4] S. Shamai (Shitz) and I. Bar-David, “The capacity of average and peak-power-limited quadrature Gaussian channels,” *IEEE Trans. Inform. Theory*, vol. 41, pp. 1060–1071, July 1995.
[5] I. C. Abou-Faycal, M. D. Trott, and S. Shamai (Shitz), “The capacity of discrete-time memoryless Rayleigh-fading channels,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 1290–1301, May 2001.
[6] M. Katz and S. Shamai (Shitz), “On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels,” *IEEE Trans. Inform. Theory*, vol. 50, pp. 2257–2270, Oct. 2004.
[7] M. C. Gursoy, H. V. Poor, and S. Verdú, “The noncoherent Rician fading channel-part I: Structure of the capacity-achieving input,” *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2193–2206, Sep. 2005.
[8] J. Huang and S. Meyn, “Characterization and computation of optimal distributions for channel coding,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 2336 –2351, July 2005.
[9] T. H. Chan, S. Hranilovic, and F. R. Kschischang, “Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 2073–2088, June 2005.
[10] T. Lutz, C. Hausl, and R. Kötter, “Coding strategies for noise-free relay cascades with half-duplex constraint,” in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 2385–2389, Toronto, ON, Canada, July 2008.
[11] G. Kramer, “Communication strategies and coding for relaying,” *Wireless Networks*, vol. 143 of The IMA Volumes in Mathematics and its Applications, pp. 163–175, 2007.
[12] T. Lutz, G. Kramer, and C. Hausl, “Capacity for half-duplex line networks with two sources,” in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 2393–2397, Austin, TX, USA, June 2010.
[13] D. Guo and L. Zhang, “Rapid on-off-division duplex for mobile ad hoc neworks,” in *Proc. Allerton Conf. Commun., Control, & Computing*, Monticello, IL, USA, 2010.
[14] D. W. Stroock, *Probability Theory, an Analytic View*. New York: Cambridge Univ. Press, 1993.
[15] Y. Wu and S. Verdú, “Functional properties of mmse,” in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 1453–1457, June 2010.
[16] D. G. Luenberger, *Optimization by Vector Sapce Methods*. New York: Wiley, 1969.
[17] S. Lang, *Complex Analysis*. New York: Springer-Verlag, 1999.
[18] W. Rudin, *Real and Complex Analysis*. McGraw-Hill Science Engineering, 1986.