A Hot Channel

Tobias Koch   Amos Lapidoth
ETH Zurich, Switzerland
Email: {tkoch, lapidoth}@isi.ee.ethz.ch

Paul P. Sotiriadis
Johns Hopkins University, Baltimore, USA
Email: pps@jhu.edu

***Abstract*— Motivated by on-chip communication, a channel model is proposed where the variance of the additive noise depends on the weighted sum of the past channel input powers. It is shown that, depending on the weights, the capacity can be either bounded or unbounded in the input power. A necessary condition and a sufficient condition for the capacity to be bounded are presented.**

I. Introduction

The channel model we study is motivated by on-chip communication. Data are to be transmitted on a chip, and it is assumed that the signal is corrupted by additive thermal noise. The variance of the thermal noise is determined by the chip's temperature which, in turn, is determined by the ambient temperature and—as the transmissions heat up the chip—by the power in the past transmissions. Thus, the variance of the present noise is a function of the powers used in the previous transmissions.

This channel was studied at low transmit power levels in [1] where it was shown that in the low power limit the heating effect is beneficial. In this paper, we focus on the high transmit power case. When the allowed transmit power is large, then there is a trade-off between optimizing the present transmission and minimizing the interference to future transmissions. Indeed, increasing the transmission power may help to overcome the present ambient noise, but it also heats up the chip and thus increases the noise variance in future receptions. Prima facie it is not clear that, as we increase the allowed transmit power, the capacity tends to infinity. This paper studies conditions under which the capacity is bounded in the transmit power.

A. Channel Model

We consider the communication system depicted in Figure 1. The message M to be transmitted over the channel is assumed to be uniformly distributed over the set $\mathcal{M} = \{1, \ldots, |\mathcal{M}|\}$ for some positive integer $|\mathcal{M}|$. The encoder maps the message to the length-n sequence $X_1, \ldots, X_n$, where n is called the *block-length*. Thus, in the absence of feedback, the sequence X_1^n is a function of the message M, i.e., $X_1^n = \phi_n(M)$ for some mapping $\phi_n : \mathcal{M} \to \mathbb{R}^n$. Here, A_m^n stands for $A_m, \ldots, A_n$, and $\mathbb{R}$ denotes the set of real numbers. If there is a feedback link, then X_k, $k = 1, \ldots, n$, is a function of the message M and, additionally, of the past channel output symbols Y_1^{k-1}, i.e., $X_k = \varphi_n^{(k)}(M, Y_1^{k-1})$ for some mapping $\varphi_n^{(k)} : \mathcal{M} \times \mathbb{R}^{k-1} \to \mathbb{R}$. The receiver guesses the transmitted message M based on the n channel output symbols Y_1^n, i.e., $\hat{M} = \psi_n(Y_1^n)$ for some mapping $\psi_n : \mathbb{R}^n \to \mathcal{M}$.

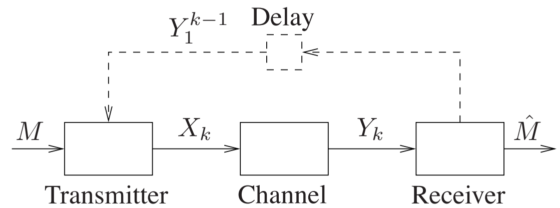

Fig. 1. The communication system.

Let $\mathbb{Z}^+$ denote the set of positive integers. The channel output $Y_k \in \mathbb{R}$ at time $k \in \mathbb{Z}^+$ corresponding to the channel inputs $(x_1, \ldots, x_k) \in \mathbb{R}^k$ is given by

$$Y_k = x_k + \sqrt{\left(\sigma^2 + \sum_{\ell=1}^{k-1} \alpha_{k-\ell} x_\ell^2\right)} \cdot U_k \tag{1}$$

where $\{U_k\}$ are independent and identically distributed (IID), zero-mean, unit-variance random variables, drawn independently of M, and being of finite fourth moment and of finite differential entropy

$$h(U_k) > -\infty. \tag{2}$$

The most interesting case is when $\{U_k\}$ are IID, zero-mean, unit-variance Gaussian random variables, and the reader is encouraged to focus on this case. The coefficients $\{\alpha_\ell\}$ in (1) are non-negative and bounded, i.e., $\alpha_\ell \geq 0$, $\ell \in \mathbb{Z}^+$, and

$$\sup_{\ell \in \mathbb{Z}^+} \alpha_\ell < \infty. \tag{3}$$

The above channel is studied under an average-power constraint on the inputs, i.e.,

$$\frac{1}{n} \sum_{k=1}^{n} \mathsf{E}\left[X_k^2\right] \leq \mathrm{P}. \tag{4}$$

The signal-to-noise ratio (SNR) is defined as

$$\mathrm{SNR} \triangleq \frac{\mathrm{P}}{\sigma^2}. \tag{5}$$

B. Capacity

We define the *rate* R (in nats per channel use) as

$$R \triangleq \frac{\log |\mathcal{M}|}{n} \tag{6}$$

where we use $\log(\cdot)$ to denote the natural logarithm function. A rate is said to be *achievable* if there exists a sequence of mappings ϕ_n (without feedback) or $\varphi_n^{(1)}, \ldots, \varphi_n^{(n)}$ (with feedback) and ψ_n such that the error probability $\Pr\Big(\hat{M} \neq M\Big)$ vanishes as n tends to infinity. The *capacity* C is the supremum of all achievable rates. We denote by $C(\text{SNR})$ the capacity under the input constraint (4) when there is no feedback, and we add the subscript "FB" to indicate that there is a feedback link. Clearly,

$$C(\text{SNR}) \leq C_{\text{FB}}(\text{SNR}) \tag{7}$$

as we can always ignore the feedback link.

For the above channel the *capacities per unit cost* which are defined as [2]

$$\dot{C}(0) \triangleq \sup_{\text{SNR}>0} \frac{C(\text{SNR})}{\text{SNR}} \tag{8}$$

and

$$\dot{C}_{\text{FB}}(0) \triangleq \sup_{\text{SNR}>0} \frac{C_{\text{FB}}(\text{SNR})}{\text{SNR}} \tag{9}$$

were studied in [1] under the additional assumptions that $\{U_k\}$ are IID, zero-mean, unit-variance Gaussian random variables, and that the coefficients fulfill

$$\sum_{\ell=1}^{\infty} \alpha_\ell \triangleq \alpha < \infty. \tag{10}$$

It was shown that in the informationally stable case, irrespective of whether feedback is available or not, the capacity per unit cost is given by

$$\dot{C}(0) = \dot{C}_{\text{FB}}(0) = \frac{1}{2}(1+\alpha). \tag{11}$$

In this paper, we focus on the high SNR case. Specifically, we explore the question whether the capacity is bounded or unbounded in the SNR.

C. The Main Result

We show that whether the capacity is bounded or not depends highly on the decay rate of the coefficients $\{\alpha_\ell\}$. This is stated precisely in the following theorem.

Theorem 1: Consider the above channel model. Then,

$$\text{i)} \quad \varliminf_{\ell\to\infty} \frac{\alpha_{\ell+1}}{\alpha_\ell} > 0 \quad \Longrightarrow \quad \sup_{\text{SNR}>0} C_{\text{FB}}(\text{SNR}) < \infty, \tag{12}$$

$$\text{ii)} \quad \varlimsup_{\ell\to\infty} \frac{\alpha_{\ell+1}}{\alpha_\ell} = 0 \quad \Longrightarrow \quad \sup_{\text{SNR}>0} C(\text{SNR}) = \infty, \tag{13}$$

where we define, for any positive $a > 0$, $a/0 \triangleq \infty$, $0/a \triangleq 0$, and $0/0 \triangleq 0$.

Remark 1: Part i) of Theorem 1 holds also when U_k has an infinite fourth moment. In Part ii) of Theorem 1, the condition on the left-hand side (LHS) of (13) can be replaced by the weaker condition

$$\varliminf_{\ell\to\infty} \frac{1}{\ell} \log \frac{1}{\alpha_\ell} = \infty. \tag{14}$$

A proof of Theorem 1 is given in the next section. In Section III we address the case where neither the LHS of (12) nor the LHS of (13) holds, i.e., when $\varlimsup_{\ell\to\infty} \alpha_{\ell+1}/\alpha_\ell > 0$ and $\varliminf_{\ell\to\infty} \alpha_{\ell+1}/\alpha_\ell = 0$. We show that in this case the capacity can be bounded or unbounded.

II. Proof of Theorem 1

In this section we provide a proof of Theorem 1. Part i) is proven in the next subsection, while the proof of Part ii) can be found in the subsequent subsection.

A. Bounded Capacity

In order to show that

$$\varliminf_{\ell\to\infty} \frac{\alpha_{\ell+1}}{\alpha_\ell} > 0 \tag{15}$$

implies that the feedback capacity $C_{\text{FB}}(\text{SNR})$ is bounded, we derive a capacity upper bound, which is, like in [3, Sec. 8.12], based on Fano's inequality and on an upper bound on $\frac{1}{n}I(M;Y_1^n)$. To simplify notation, we define $\alpha_0 \triangleq 1$.

We first note that, due to (15), we can find an $\ell_0 \in \mathbb{Z}^+$ and a $0 < \rho < 1$ so that $\alpha_{\ell_0} > 0$ and

$$\frac{\alpha_{\ell+1}}{\alpha_\ell} > \rho, \quad \ell \geq \ell_0. \tag{16}$$

We continue with the chain rule for mutual information [3]

$$\begin{aligned}
&\frac{1}{n} I(M;Y_1^n) \\
&= \frac{1}{n}\sum_{k=1}^{\ell_0} I(M;Y_k|Y_1^{k-1}) + \frac{1}{n}\sum_{k=\ell_0+1}^{n} I(M;Y_k|Y_1^{k-1}).
\end{aligned} \tag{17}$$

Each term in the first sum on the right-hand side (RHS) of (17) is upper bound by

$$\begin{aligned}
&I(M;Y_k|Y_1^{k-1}) \\
&\leq h(Y_k) - h(Y_k|Y_1^{k-1},M) \\
&= h(Y_k) - \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=1}^{k-1}\alpha_{k-\ell}X_\ell^2\right)\right] - h(U_k) \\
&\leq \frac{1}{2}\log\left(2\pi e\left(1 + \sum_{\ell=1}^{k}\alpha_{k-\ell}\frac{\mathsf{E}\left[X_\ell^2\right]}{\sigma^2}\right)\right) - h(U_k) \\
&\leq \frac{1}{2}\log\left(2\pi e\left(1 + \Big(\sup_{\ell'\in\mathbb{Z}_0^+}\alpha_{\ell'}\Big)\cdot\sum_{\ell=1}^{k}\frac{\mathsf{E}\left[X_\ell^2\right]}{\sigma^2}\right)\right) - h(U_k) \\
&\leq \frac{1}{2}\log\left(2\pi e\left(1 + \Big(\sup_{\ell'\in\mathbb{Z}_0^+}\alpha_{\ell'}\Big)\cdot n\cdot\text{SNR}\right)\right) - h(U_k)
\end{aligned} \tag{18}$$

where $\mathbb{Z}_0^+$ denotes the set of non-negative integers. Recall that $\sup_{\ell'\in\mathbb{Z}_0^+}\alpha_{\ell'}$ is assumed to be finite. Here, the first inequality follows because conditioning cannot increase entropy; the following equality follows because X_1^k is a function of (M, Y_1^{k-1}) and from the behavior of entropy under translation and scaling [3, Thms. 9.6.3 & 9.6.4] in conjunction with the fact that U_k is independent of (X_1^k, M, Y_1^{k-1}); the subsequent inequality follows from the entropy maximizing property of Gaussian random variables [3, Thm. 9.6.5] and by lower bounding $\mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=1}^{k-1}\alpha_{k-\ell}X_\ell^2\right)\right] \geq \sigma^2$; the next inequality by upper bounding each coefficient $\alpha_\ell \leq \sup_{\ell'\in\mathbb{Z}_0^+}\alpha_{\ell'}$, $\ell = 1,\ldots,k$; and the last inequality follows from

$$\sum_{\ell=1}^{k}\frac{\mathsf{E}\left[X_\ell^2\right]}{\sigma^2} \leq n\cdot\text{SNR}, \quad k \leq n,$$

which is a consequence of the power constraint (4) and of the fact that $\mathsf{E}\left[X_\ell^2\right]/\sigma^2 \geq 0$, $\ell \in \mathbb{Z}^+$.

The terms in the second sum on the RHS of (17) are upper bounded using the general bound for mutual information [4, Thm. 5.1]

$$I(X;Y) \leq \int D(W(\cdot|x)\|R(\cdot))\,\mathrm{d}Q(x) \tag{19}$$

where $D(\cdot\|\cdot)$ denotes relative entropy, $W(\cdot|\cdot)$ is the channel law, $Q(\cdot)$ is the distribution on the channel input X, and $R(\cdot)$ is any distribution on the output alphabet. Thus, any choice of output distribution $R(\cdot)$ yields an upper bound on the mutual information.

For $k = \ell_0+1,\ldots,n$ we upper bound $I(M;Y_k|Y_1^{k-1} = y_1^{k-1})$ for a given $Y_1^{k-1} = y_1^{k-1}$ by choosing $R(\cdot)$ to be of a Cauchy distribution whose density is given by

$$\frac{\sqrt{\beta}}{\pi}\frac{1}{1+\beta y_k^2}, \quad y_k \in \mathbb{R}, \tag{20}$$

where we choose the scale parameter β to be[1] $\beta = 1/(\tilde{\beta} y_{k-\ell_0}^2)$ and

$$\tilde{\beta} = \min\left\{\rho^{\ell_0-1}\cdot\frac{\alpha_{\ell_0}}{\max\limits_{0\leq\ell'<\ell_0}\alpha_\ell}, \alpha_{\ell_0}, \rho^{\ell_0}\right\}. \tag{21}$$

Note that (15) together with the assumption that the coefficients are bounded implies that $0 < \tilde{\beta} < 1$. Applying (20) to (19) yields

$$\begin{aligned}
&I(M;Y_k|Y_1^{k-1} = y_1^{k-1})\\
&\leq \mathsf{E}\left[\log\left(1+\frac{Y_k^2}{\tilde{\beta} y_{k-\ell_0}^2}\right)\right] + \frac{1}{2}\log(\tilde{\beta} y_{k-\ell_0}^2) + \log\pi\\
&\quad - h(Y_k|M, Y_1^{k-1} = y_1^{k-1}),
\end{aligned} \tag{22}$$

and we thus obtain, averaging over Y_1^{k-1},

$$\begin{aligned}
&I(M;Y_k|Y_1^{k-1})\\
&\leq \log\pi - h(Y_k|Y_1^{k-1}, M) + \frac{1}{2}\mathsf{E}\left[\log(\tilde{\beta} Y_{k-\ell_0}^2)\right]\\
&\quad + \mathsf{E}\left[\log(\tilde{\beta} Y_{k-\ell_0}^2 + Y_k^2)\right] - \mathsf{E}\left[\log(Y_{k-\ell_0}^2)\right] - \log\tilde{\beta}.
\end{aligned} \tag{23}$$

We evaluate the terms on the RHS of (23) individually. We begin with

$$h(Y_k|Y_1^{k-1}, M) = \frac{1}{2}\mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=1}^{k-1}\alpha_{k-\ell}X_\ell^2\right)\right] + h(U_k) \tag{24}$$

where we use the same steps as in (18). The next term is upper bounded by

$$\begin{aligned}
&\mathsf{E}\left[\log(\tilde{\beta} Y_{k-\ell_0}^2)\right]\\
&= \mathsf{E}_{X_1^{k-\ell_0}}\left[\mathsf{E}\left[\log\left(\tilde{\beta}(x_{k-\ell_0} + \sigma_{k-\ell_0}U_{k-\ell_0})^2\right)\,\middle|\, x_1^{k-\ell_0}\right]\right]\\
&\leq \mathsf{E}_{X_1^{k-\ell_0}}\left[\log\left(\tilde{\beta}\mathsf{E}\left[(x_{k-\ell_0} + \sigma_{k-\ell_0}U_{k-\ell_0})^2 \,\middle|\, x_1^{k-\ell_0}\right]\right)\right]\\
&= \mathsf{E}\left[\log\left(\tilde{\beta} X_{k-\ell_0}^2 + \tilde{\beta}\sigma^2 + \tilde{\beta}\sum_{\ell=1}^{k-\ell_0-1}\alpha_{k-\ell_0-\ell}X_\ell^2\right)\right]
\end{aligned} \tag{25}$$

where we define

$$\sigma_k^2 \triangleq \sigma^2 + \sum_{\ell=1}^{k-1}\alpha_{k-\ell}x_\ell^2, \quad k \in \mathbb{Z}^+, \tag{26}$$

and where the inequality follows from Jensen's inequality applied to the concave function $\log(x)$, $x > 0$. It can be verified easily that our choice of $\tilde{\beta}$ together with (16) implies

$$\tilde{\beta}\alpha_\ell < \alpha_{\ell+\ell_0}, \quad \ell \in \mathbb{Z}_0^+, \tag{27}$$

so that (25) can be further upper bounded by

$$\mathsf{E}\left[\log(\tilde{\beta} Y_{k-\ell_0}^2)\right] \leq \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=1}^{k-\ell_0}\alpha_{k-\ell}X_\ell^2\right)\right]. \tag{28}$$

Similarly, we use Jensen's inequality along with (27) to upper bound

$$\begin{aligned}
&\mathsf{E}\left[\log(\tilde{\beta} Y_{k-\ell_0}^2 + Y_k^2)\right]\\
&\leq \mathsf{E}\left[\log\left(2\sigma^2 + 2\sum_{\ell=1}^{k-\ell_0}\alpha_{k-\ell}X_\ell^2 + \sum_{\ell=k-\ell_0+1}^{k}\alpha_{k-\ell}X_\ell^2\right)\right]\\
&\leq \log 2 + \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=1}^{k}\alpha_{k-\ell}X_\ell^2\right)\right].
\end{aligned} \tag{29}$$

In order to lower bound $\mathsf{E}\left[\log(Y_{k-\ell_0}^2)\right]$ we need the following lemma:

Lemma 2: Let X be a random variable of density $f_X(x)$, $x \in \mathbb{R}$. Then, for any $0 < \delta \leq 1$ and $0 < \eta < 1$ we have

$$\sup_{c\in\mathbb{R}}\mathsf{E}\left[\log|X+c|^{-1}\cdot I\{|X+c|\leq\delta\}\right] \leq \epsilon(\delta,\eta) + \frac{1}{\eta}h^-(X) \tag{30}$$

where $I\{\cdot\}$ denotes the indicator function; $h^-(X)$ is defined as

$$h^-(X) \triangleq \int_{\{x\in\mathbb{R}:f_X(x)>1\}} f_X(x)\log f_X(x)\,\mathrm{d}x; \tag{31}$$

and where $\epsilon(\delta,\eta) > 0$ tends to zero as $\delta \downarrow 0$.

Proof: A proof can be found in [4, Lemma 6.7]. ■

We write the expectation as

$$\begin{aligned}
&\mathsf{E}\left[\log(Y_{k-\ell_0}^2)\right]\\
&= \mathsf{E}_{X_1^{k-\ell_0}}\left[\mathsf{E}\left[\log(x_{k-\ell_0} + \sigma_{k-\ell_0}U_{k-\ell_0})^2 \,\middle|\, x_1^{k-\ell_0}\right]\right]
\end{aligned} \tag{32}$$

[1]When $y_{k-\ell_0} = 0$ then the density of the Cauchy distribution (20) is undefined. However, this event is of zero probability and has therefore no influence on the mutual information $I(M;Y_k|Y_1^{k-1})$.

and lower bound the conditional expectation by

$$\begin{aligned}
&\mathsf{E}\left[\log\left(x_{k-\ell_0}+\sigma_{k-\ell_0}U_{k-\ell_0}\right)^2 \,\middle|\, x_1^{k-\ell_0}\right] \\
&= \log\sigma_{k-\ell_0}^2 - 2\cdot\mathsf{E}\left[\log\left|\frac{x_{k-\ell_0}}{\sigma_{k-\ell_0}}+U_{k-\ell_0}\right|^{-1} \,\middle|\, x_1^{k-\ell_0}\right] \\
&\geq \log\sigma_{k-\ell_0}^2 - 2\epsilon(\delta,\eta) - \frac{2}{\eta}h^{-}(U_k) + \log\delta^2 \qquad (33)
\end{aligned}$$

for some $0<\delta\leq 1$ and $0<\eta<1$. Here, the inequality follows by splitting the conditional expectation into the two expectations given in (34) (on the top of the next page) and by upper bounding then the first term on the RHS of (34) using Lemma 2 and the second term by $-\log\delta$. Averaging (33) over $X_1^{k-\ell_0}$ yields

$$\begin{aligned}
\mathsf{E}\left[\log(Y_{k-\ell_0}^2)\right] &\geq \mathsf{E}\left[\log\left(\sigma^2+\sum_{\ell=1}^{k-\ell_0-1}\alpha_{k-\ell_0-\ell}X_\ell^2\right)\right] \\
&\quad - 2\epsilon(\delta,\eta) - \frac{2}{\eta}h^{-}(U_k) + \log\delta^2. \qquad (35)
\end{aligned}$$

Note that, since U_k is of unit variance, (2) together with [4, Lemma 6.4] implies that $h^{-}(U_k)$ is finite.

Turning back to the upper bound (23) we obtain from (35), (29), (28), and (24)

$$\begin{aligned}
&I(M;Y_k|Y_1^{k-1}) \\
&\leq \log 2 + \mathsf{E}\left[\log\left(\sigma^2+\sum_{\ell=1}^{k}\alpha_{k-\ell}X_\ell^2\right)\right] + \frac{2}{\eta}h^{-}(U_k) \\
&\quad + 2\epsilon(\delta,\eta) - \log\delta^2 - \mathsf{E}\left[\log\left(\sigma^2+\sum_{\ell=1}^{k-\ell_0-1}\alpha_{k-\ell_0-\ell}X_\ell^2\right)\right] \\
&\quad - \log\tilde{\beta} + \frac{1}{2}\mathsf{E}\left[\log\left(\sigma^2+\sum_{\ell=1}^{k-\ell_0}\alpha_{k-\ell}X_\ell^2\right)\right] + \log\pi \\
&\quad - \frac{1}{2}\mathsf{E}\left[\log\left(\sigma^2+\sum_{\ell=1}^{k-1}\alpha_{k-\ell}X_\ell^2\right)\right] - h(U_k) \\
&\leq \mathsf{E}\left[\log\left(\sigma^2+\sum_{\ell=1}^{k}\alpha_{k-\ell}X_\ell^2\right)\right] \\
&\quad - \mathsf{E}\left[\log\left(\sigma^2+\sum_{\ell=1}^{k-\ell_0-1}\alpha_{k-\ell_0-\ell}X_\ell^2\right)\right] + \mathsf{K} \qquad (36)
\end{aligned}$$

where

$$\mathsf{K} \triangleq \frac{2}{\eta}h^{-}(U_k) - h(U_k) + 2\epsilon(\delta,\eta) + \log\frac{2\pi}{\tilde{\beta}\delta^2} \qquad (37)$$

is a finite constant, and where the last inequality in (36) follows because for any $X_{k-\ell_0+1}^{k-1}=x_{k-\ell_0+1}^{k-1}$ we have $\sum_{\ell=k-\ell_0+1}^{k-1}\alpha_{k-\ell}x_\ell^2\geq 0$. Note that K does not depend on k as $\{U_k\}$ are IID.

Turning back to the evaluation of the second sum on the RHS of (17) we use that for any sequences $\{a_k\}$ and $\{b_k\}$

$$\begin{aligned}
&\sum_{k=\ell_0+1}^{n}(a_k-b_k) \\
&= \sum_{k=n-2\ell_0+1}^{n}(a_k-b_{k-n+3\ell_0}) + \sum_{k=\ell_0+1}^{n-2\ell_0}(a_k-b_{k+2\ell_0}). \qquad (38)
\end{aligned}$$

Using Jensen's inequality followed by the same steps as in (18) we obtain for $k=n-2\ell_0+1,\ldots,n$

$$\begin{aligned}
&\mathsf{E}\left[\log\left(\frac{\sigma^2+\sum_{\ell=1}^{k}\alpha_{k-\ell}X_\ell^2}{\sigma^2+\sum_{\ell=1}^{k-n+2\ell_0-1}\alpha_{k-n+2\ell_0-\ell}X_\ell^2}\right)\right] \\
&\qquad\leq \log\left(1+\big(\sup_{\ell\in\mathbb{Z}_0^+}\alpha_\ell\big)\cdot n\cdot\mathrm{SNR}\right). \qquad (39)
\end{aligned}$$

Thus, applying (39) and (38) to (36) yields

$$\begin{aligned}
&\frac{1}{n}\sum_{\ell=\ell_0+1}^{n}I(M;Y_k|Y_1^{k-1}) \\
&\leq \frac{n-\ell_0}{n}\mathsf{K} + \frac{2\ell_0}{n}\log\left(1+\big(\sup_{\ell\in\mathbb{Z}_0^+}\alpha_\ell\big)\cdot n\cdot\mathrm{SNR}\right) \\
&\quad + \frac{1}{n}\sum_{k=\ell_0+1}^{n-2\ell_0}\mathsf{E}\left[\log\left(\frac{\sigma^2+\sum_{\ell=1}^{k}\alpha_{k-\ell}X_\ell^2}{\sigma^2+\sum_{\ell=1}^{k+\ell_0-1}\alpha_{k+\ell_0-\ell}X_\ell^2}\right)\right] \\
&\leq \frac{n-\ell_0}{n}\mathsf{K} + \frac{2\ell_0}{n}\log\left(1+\big(\sup_{\ell\in\mathbb{Z}_0^+}\alpha_\ell\big)\cdot n\cdot\mathrm{SNR}\right) \\
&\quad + \frac{1}{n}\sum_{k=\ell_0+1}^{n-2\ell_0}\mathsf{E}\left[\log\left(\frac{\sigma^2+\sum_{\ell=1}^{k}\alpha_{k+\ell_0-\ell}X_\ell^2}{\sigma^2+\sum_{\ell=1}^{k+\ell_0-1}\alpha_{k+\ell_0-\ell}X_\ell^2}\right)\right] \\
&\quad - \frac{n-3\ell_0}{n}\log\tilde{\beta} \\
&\leq \frac{n-\ell_0}{n}\mathsf{K} + \frac{2\ell_0}{n}\log\left(1+\big(\sup_{\ell\in\mathbb{Z}_0^+}\alpha_\ell\big)\cdot n\cdot\mathrm{SNR}\right) \\
&\quad - \frac{n-3\ell_0}{n}\log\tilde{\beta} \qquad (40)
\end{aligned}$$

where the second inequality follows by adding $\log\tilde{\beta}$ to the expectation and by upper bounding then $\tilde{\beta}\alpha_\ell<\alpha_{\ell+\ell_0}$, $\ell\in\mathbb{Z}_0^+$ (27); and the last inequality follows because for any given $X_{k+1}^{k+\ell_0-1}=x_{k+1}^{k+\ell_0-1}$ we have $\sum_{\ell=k+1}^{k+\ell_0-1}\alpha_{k+\ell_0-\ell}x_\ell^2\geq 0$.

Combining (40), (18), and (17) we obtain

$$\begin{aligned}
&\frac{1}{n}I(M;Y_1^n) \\
&\leq \frac{n-\ell_0}{n}\mathsf{K} - \frac{n-3\ell_0}{n}\log\tilde{\beta} + \frac{\ell_0}{2n}\log(2\pi e) - \frac{\ell_0}{n}h(U_1) \\
&\quad + \frac{\ell_0}{n}\frac{5}{2}\log\left(1+\big(\sup_{\ell\in\mathbb{Z}_0^+}\alpha_\ell\big)\cdot n\cdot\mathrm{SNR}\right) \qquad (41)
\end{aligned}$$

which converges to $\mathsf{K}-\log\tilde{\beta}<\infty$ as we let n go to infinity. With this, we have shown that $\underline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell>0$ implies that the capacity $C_{\mathrm{FB}}(\mathrm{SNR})$ is bounded.

$$\mathsf{E}\left[\log\left|\frac{x_{k-\ell_0}}{\sigma_{k-\ell_0}}+U_{k-\ell_0}\right|^{-1}\,\middle|\,x_1^{k-\ell_0}\right]=\mathsf{E}\left[\log\left|\frac{x_{k-\ell_0}}{\sigma_{k-\ell_0}}+U_{k-\ell_0}\right|^{-1}\cdot I\left\{\left|\frac{x_{k-\ell_0}}{\sigma_{k-\ell_0}}+U_{k-\ell_0}\right|\leq\delta\right\}\,\middle|\,x_1^{k-\ell_0}\right]$$
$$+\,\mathsf{E}\left[\log\left|\frac{x_{k-\ell_0}}{\sigma_{k-\ell_0}}+U_{k-\ell_0}\right|^{-1}\cdot I\left\{\left|\frac{x_{k-\ell_0}}{\sigma_{k-\ell_0}}+U_{k-\ell_0}\right|>\delta\right\}\,\middle|\,x_1^{k-\ell_0}\right] \tag{34}$$

B. Unbounded Capacity

In order to show that $\overline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell=0$ implies that the capacity $C(\mathrm{SNR})$ is unbounded we propose a coding scheme that achieves an unbounded rate.

We first note that $\overline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell=0$ implies that there is an $\ell_0\in\mathbb{Z}^+$ so that either $\alpha_\ell=0$, $\ell\geq\ell_0$, or that $\alpha_{\ell_0}>0$ and for some arbitrary $0<\delta<1$

$$\frac{\alpha_{\ell+1}}{\alpha_\ell}<\delta,\quad \ell\geq\ell_0. \tag{42}$$

In the former case, we can achieve the (unbounded) rate

$$R=\frac{1}{2L}\log(1+L\cdot\mathrm{SNR}),\quad L\geq\ell_0, \tag{43}$$

by a coding scheme where the channel inputs $\{X_{kL+1}\}$ are IID, zero-mean Gaussian random variables of variance $L\mathsf{P}$, and where the other inputs are deterministically zero. Indeed, by waiting L time-steps, the chip's temperature cools down to the ambient one so that the noise variance is independent of the previous channel inputs and we can achieve—after appropriate normalization—the capacity of the additive white Gaussian noise (AWGN) channel [5].

To show that the capacity is also unbounded in the latter case, we propose the following encoding and decoding scheme. Let $x_1^n(m)$, $m\in\mathcal{M}$, denote the codeword sent out by the transmitter that corresponds to the message $M=m$. We choose some period $L\geq\ell_0$ and generate the components $x_{kL+1}(m)$, $m\in\mathcal{M}$, $k=0,\ldots,\lfloor n/L\rfloor-1$ (where $\lfloor\cdot\rfloor$ denotes the floor function), independently of each other according to a zero-mean Gaussian law of variance $L\mathsf{P}$. The other components are set to zero.[2]

The receiver uses a *nearest neighbor decoder* in order to guess M based on the received sequence of channel outputs y_1^n. Thus, it computes, for any $m'\in\mathcal{M}$, $\|\mathbf{y}-\mathbf{x}(m')\|^2$ and decides on the message that satisfies

$$\hat{M}=\arg\min_{m'\in\mathcal{M}}\|\mathbf{y}-\mathbf{x}(m')\|^2 \tag{44}$$

where ties are resolved with a fair coin flip. Here, $\|\cdot\|$ denotes the Euclidean norm, and $\mathbf{y}$ and $\mathbf{x}(m')$ denote the respective vectors $\{y_{kL+1}\}_{k=0}^{\lfloor n/L\rfloor-1}$ and $\{x_{kL+1}(m')\}_{k=0}^{\lfloor n/L\rfloor-1}$.

We are interested in the average probability of error $\Pr\left(\hat{M}\neq M\right)$, averaged over all codewords in the codebook, and averaged over all codebooks. Due to the symmetry of the codebook construction, the probability of error corresponding to the mth message $\Pr\left(\hat{M}\neq M\,\middle|\,M=m\right)$ does not depend on m, and we thus conclude that $\Pr\left(\hat{M}\neq M\right)=\Pr\left(\hat{M}\neq M\,\middle|\,M=1\right)$. We further note that

$$\Pr\left(\hat{M}\neq M\,\middle|\,M=1\right)\leq\Pr\left(\bigcup_{m'=2}^{n}\|\mathbf{Y}-\mathbf{X}(m')\|^2<\|\mathbf{Z}\|^2\right) \tag{45}$$

where $\mathbf{Z}=\{\sigma_{kL+1}\cdot U_{kL+1}\}_{k=0}^{\lfloor n/L\rfloor-1}$ which is, conditional on $M=1$, equal to $\|\mathbf{Y}-\mathbf{X}(1)\|^2$. In order analyze (45) we need the following lemma.

Lemma 3: Consider the channel described in Section I-A, and assume that the coefficients $\{\alpha_\ell\}$ satisfy (42). Further assume that $\{X_{kL+1}\}$ are IID, zero-mean Gaussian random variables of variance $L\mathsf{P}$. Let the set $\mathcal{D}_\epsilon$ be defined as

$$\mathcal{D}_\epsilon\triangleq\Bigg\{\mathbf{y}\in\mathbb{R}^{\lfloor n/L\rfloor},\mathbf{z}\in\mathbb{R}^{\lfloor n/L\rfloor}:\left|\frac{1}{\lfloor n/L\rfloor}\|\mathbf{y}\|^2-(\sigma^2+L\mathsf{P}+\alpha_L\cdot L\mathsf{P})\right|<\epsilon,\ \left|\frac{1}{\lfloor n/L\rfloor}\|\mathbf{z}\|^2-(\sigma^2+\alpha_L\cdot L\mathsf{P})\right|<\epsilon\Bigg\} \tag{46}$$

with α_L being defined as $\alpha_L\triangleq\sum_{\ell=1}^{\infty}\alpha_{\ell L}$. Then,

$$\lim_{n\to\infty}\Pr((\mathbf{Y},\mathbf{Z})\in\mathcal{D}_\epsilon)=1 \tag{47}$$

for any $\epsilon>0$.

Proof: First note that, since U_k has a finite fourth moment, our choice of input distribution implies that $\mathsf{E}\left[(\sigma_k\cdot U_k)^4\right]<\infty$. This along with (42) yields that the variances $\mathsf{Var}\left(\frac{1}{\lfloor n/L\rfloor}\|\mathbf{Y}\|^2\right)$ and $\mathsf{Var}\left(\frac{1}{\lfloor n/L\rfloor}\|\mathbf{Z}\|^2\right)$ vanish as n tends to infinity. The lemma follows then by computing $\mathsf{E}\left[\frac{1}{\lfloor n/L\rfloor}\|\mathbf{Y}\|^2\right]$ and $\mathsf{E}\left[\frac{1}{\lfloor n/L\rfloor}\|\mathbf{Z}\|^2\right]$ and by Chebyshev's inequality [6, Sec. 5.4]. ∎

In order to upper bound the RHS of (45) we proceed along the lines of [5], [7]. We have

$$\Pr\left(\bigcup_{m'=2}^{n}\|\mathbf{Y}-\mathbf{X}(m')\|^2<\|\mathbf{Z}\|^2\right)\leq\Pr((\mathbf{Y},\mathbf{Z})\notin\mathcal{D}_\epsilon)+\int_{\mathcal{D}_\epsilon}\Pr\left(\bigcup_{m'=2}^{n}\|\mathbf{y}-\mathbf{X}(m')\|^2<\|\mathbf{z}\|^2\,\middle|\,(\mathbf{y},\mathbf{z})\right)\mathrm{d}P(\mathbf{y},\mathbf{z}) \tag{48}$$

and it follows from Lemma 3 that the first term on the RHS of (48) vanishes as n tends to infinity. Note that since the codewords are independent of each other, conditional on $M=1$, the distribution of $\mathbf{X}(m')$, $m'=2,\ldots,|\mathcal{M}|$, does not depend on $(\mathbf{y},\mathbf{z})$. We upper bound the second term on the

[2] It follows from the weak law of large numbers that, for any $m\in\mathcal{M}$, $\frac{1}{n}\sum_{k=1}^{n}x_k^2(m)$ converges to P in probability as n increases.

RHS of (48) by analyzing $\Pr\big(\|\mathbf{y}-\mathbf{X}(m')\|^2 < \|\mathbf{z}\|^2 \,\big|\, (\mathbf{y},\mathbf{z})\big)$ for each $m' = 2, \ldots, |\mathcal{M}|$ and by applying then the union of events bound.

For $(\mathbf{y},\mathbf{z}) \in \mathcal{D}_\epsilon$ and $m' = 2, \ldots, |\mathcal{M}|$ we have

$$\begin{aligned}&\Pr\big(\|\mathbf{y}-\mathbf{X}(m')\|^2 < \|\mathbf{z}\|^2 \,\big|\, (\mathbf{y},\mathbf{z})\big)\\ &\le \exp\Bigg\{ -s\lfloor n/L\rfloor(\sigma^2+\alpha_L\cdot L\mathsf{P}+\epsilon) + \frac{s\|\mathbf{y}\|^2}{1-2sL\mathsf{P}}\\ &\qquad -\frac{1}{2}\lfloor n/L\rfloor \log(1-2sL\mathsf{P})\Bigg\}, \quad s<0, \end{aligned} \tag{49}$$

which follows by upper bounding $\|\mathbf{z}\|^2$ by $\lfloor n/L\rfloor(\sigma^2+\alpha_L\cdot L\mathsf{P}+\epsilon)$ and from the Chernoff bound [6, Sec. 5.4]. Using that, for $(\mathbf{y},\mathbf{z}) \in \mathcal{D}_\epsilon$, $\|\mathbf{y}\|^2 \ge \lfloor n/L\rfloor(\sigma^2+L\mathsf{P}+\alpha_L\cdot L\mathsf{P}-\epsilon)$ it follows from the union of events bound and (49) that (48) goes to zero as n tends to infinity if

$$\begin{aligned}R &< \frac{s}{L}(\sigma^2+\alpha_L\cdot L\mathsf{P}+\epsilon) + \frac{1}{2L}\log(1-2sL\mathsf{P})\\ &\quad -\frac{s}{L}\frac{\sigma^2+L\mathsf{P}+\alpha_L\cdot L\mathsf{P}-\epsilon}{1-2sL\mathsf{P}}\end{aligned} \tag{50}$$

for some $s<0$. Thus, choosing $s = -1/2\cdot 1/(1+\alpha_L\cdot L\mathsf{P})$ yields that any rate below

$$\begin{aligned}&-\frac{1}{2L}\frac{\sigma^2+\alpha_L\cdot L\mathsf{P}+\epsilon}{1+\alpha_L\cdot L\mathsf{P}} + \frac{1}{2L}\log\left(1+\frac{L\mathsf{P}}{1+\alpha_L\cdot L\mathsf{P}}\right)\\ &+\frac{1}{2L}\frac{\sigma^2+L\mathsf{P}+\alpha_L\cdot L\mathsf{P}-\epsilon}{1+\alpha_L\cdot L\mathsf{P}}\frac{1}{1+\frac{L\mathsf{P}}{1+\alpha_L\cdot L\mathsf{P}}}\end{aligned} \tag{51}$$

is achievable. As P tends to infinity this converges to[3]

$$\frac{1}{2L}\log\left(1+\frac{1}{\alpha_L}\right) > \frac{1}{2L}\log\frac{1}{\alpha_L}. \tag{52}$$

It remains to show that given (42) we can make α_L arbitrarily small. We note that (42) implies $\alpha_\ell < \delta^{\ell-\ell_0}\alpha_{\ell_0}$, $\ell \ge \ell_0$. Thus,

$$\alpha_L = \sum_{\ell-1}^{\infty}\alpha_{\ell L} < \sum_{\ell=1}^{\infty}\delta^{\ell L}\frac{\alpha_{\ell_0}}{\delta^{\ell_0}} = \frac{\delta^L}{1-\delta^L}\frac{\alpha_{\ell_0}}{\delta^{\ell_0}} \tag{53}$$

and (52) can therefore be further lower bounded by

$$\frac{1}{2L}\log\frac{(1-\delta^L)\delta^{\ell_0}}{\alpha_{\ell_0}} + \frac{1}{2}\log\frac{1}{\delta}. \tag{54}$$

Letting L tend to infinity then yields that we can achieve any rate below $\frac{1}{2}\log\frac{1}{\delta}$. As this can be made arbitrarily large by choosing δ sufficiently small, we conclude that $\overline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell = 0$ implies that the capacity is unbounded.

[3]The same rate can also be derived by evaluating $\underline{\lim}_{n\to\infty}\frac{1}{n}I(X_1^n;Y_1^n)$ for a distribution on the channel inputs under which $\{X_{kL+1}\}$ are IID, zero-mean, variance-$L\mathsf{P}$ Gaussian random variables while the other inputs are deterministically zero. However, it is not obvious whether there is a coding theorem associated with this quantity.

III. Beyond Theorem 1

Theorem 1 resolves the question whether the capacity is bounded or unbounded when the coefficients satisfy either $\underline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell > 0$ or $\overline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell = 0$. We next address the case where neither condition holds, i.e., when

$$\overline{\lim_{\ell\to\infty}}\,\alpha_{\ell+1}/\alpha_\ell > 0 \quad \text{and} \quad \underline{\lim_{\ell\to\infty}}\,\alpha_{\ell+1}/\alpha_\ell = 0. \tag{55}$$

Example 1 exhibits a sequence $\{\alpha_\ell\}$ satisfying (55) for which the capacity is bounded, and Example 2 provides a sequence $\{\alpha_\ell\}$ satisfying (55) for which the capacity is unbounded.

Example 1: Consider the sequence $\{\alpha_\ell\}$ where all coefficients with an even index are 1 and all coefficients with an odd index are zero. It satisfies (55) because $\overline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell = \infty$ and $\underline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell = 0$. Thus, at even times, the output Y_{2k}, $k\in\mathbb{Z}^+$, only depends on the "even" inputs $\{X_{2\ell}\}_{\ell=1}^k$, while at odd times, the output Y_{2k+1}, $k\in\mathbb{Z}_0^+$, only depends on the "odd" inputs $\{X_{2\ell+1}\}_{\ell=0}^k$. By proceeding along the lines of the proof of Part i) of Theorem 1 while choosing in (20) $\beta = 1/y_{k-2}^2$, it can be shown that the capacity of this channel is bounded.

Example 2: Consider the sequence $\{\alpha_\ell\}$ where $\alpha_0 = 1$, where all coefficients with an odd index are 1, and where all other coefficients (whose index is an even positive integer) are zero. (Again, we have $\overline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell = \infty$ and $\underline{\lim}_{\ell\to\infty}\alpha_{\ell+1}/\alpha_\ell = 0$.) Using Gaussian inputs of power $2\mathsf{P}$ at even times while setting the inputs to be zero at odd times, and measuring the channel outputs only at even times, reduces the channel to a memoryless additive noise channel and demonstrates the achievability of [5]

$$R = \frac{1}{4}\log(1+2\cdot\text{SNR})$$

which is unbounded in the SNR.

References

[1] T. Koch, A. Lapidoth, and P. P. Sotiriadis, "A channel that heats up," in *Proc. IEEE Int. Symposium on Inf. Theory*, Nice, France, June. 24–29, 2007.

[2] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1019–1030, Sept. 1990.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 1991.

[4] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.

[5] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1520–1529, Sept. 1996.

[6] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.

[7] A. Lapidoth and S. Shamai (Shitz), "Fading channels: how perfect need 'perfect side-information' be?" *IEEE Trans. Inform. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.