Low Computational Complexity of SC Polar Decoder in MIMO Fading Channel

Alaa Hasan*†
*Department of Systems and Computer Engineering
Carleton University
1125 Colonel By Drive, Ottawa, ON, K1S 5B6, Canada
Email: alaahasan@cmail.carleton.ca
†Department of Mechatronics Engineering
Al-Khwarizmi College of Engineering, Baghdad University
Baghdad, Iraq

Ian D. Marsland
Department of Systems and Computer Engineering
Carleton University
1125 Colonel By Drive, Ottawa, ON, K1S 5B6, Canada
Email: ianm@sce.carleton.ca

***Abstract*—Motivated by large capacity gains in multiple antenna systems when ideal channel state information (CSI) is available at both receiver and transmitter and quadrature amplitude modulation (QAM) is applied, we examine the achievable rates of Rayleigh fading channel measurement based optimization techniques. We consider complex-valued noise Gaussian distribution and try to determine the optimal input distribution of fixed signalling points. By using Hermite polynomials and under even-moment constraint, the simulation results show that the information rate is achieved with unique and optimal input distribution. It is also shown that the computational complexity can be reduced by factorizing the optimal distribution into the product of symmetrical distributions.**

Keywords—Polar codes; MIMO fading channel; Hermite polynomials; channel capacity

I. Introduction

A significant amount of work has been reported to analyze the performance of multiple input multiple output (MIMO) systems in combination with other countermeasures, such as error-correcting codes. Among all the error-correcting codes, polar codes can accomplish the bit error rate (BER) performance approaching Shannon limit. In this paper, we discuss a concatenation scheme of polar codes and Alamouti encoder for space-time block codes. The channel capacity $\mathrm{C}_{\text{M-QAM}}$ is defined as the maximum information rate found when quadrature amplitude constellation of size M (M-QAM) is applied. This $\mathrm{C}_{\text{M-QAM}}$ is upper limited by C, the channel capacity, whereas the "reduction" amount in achievable rate between them can be referred due to the the logarithm of the constellation size and the channel SNR. Accordingly, the channel capacity C which grows logarithmically diverges from the bounded $\mathrm{C}_{\text{M-QAM}}$ at high values of SNR while they are close together at low SNRs. First, we try to amplify the transmitted signal amplitude by scaling the points with a factor > 0 in order to approach $\mathrm{C}_{\text{M-QAM}}$ with the channel capacity. Also, we characterize the optimal distribution through using *Hermite polynomials* theory under even-moment constraint.

We show also that the obtained optimal input distribution can be factored into two identical distributions, and then, as a result, apply the optimization algorithm on the new input distribution. This step can hugely lower the complexity of optimization computations along the algorithm. The contributions thus lie in the study of the capacity of this MIMO channel model, in which we show that

- $\mathrm{C}_{\text{M-QAM}}$ is computed with maximum power constraint.
- The capacity-achieving distribution, subject to an even-moment constraint $\mathrm{E}[x^H x] \leq P$ for some $P > 0$, is only optimal, Gaussian and continuous.
- The optimal input symbol distribution, numerically indicated by the results of an optimization procedure based Hermite polynomials theory, can be factored into Cartesian product of symmetrical distributions.

Results evolve along the following lines. A model of the fading MIMO channel model is presented in Section II. Section III formulates our problem then motivate the use of Hermite polynomials theory. The optimal input distribution is then characterized by an iterative and a simplified iterative proposed solutions, see Section IV. In Section V, the main results are provided and, finally, a brief conclusion is given in Section VI.

II. System Model

Consider a complex system of MIMO fading channel, with T transmit antennas under power constraint P and R receive antennas is given by

$$\mathbf{y} = \mathbf{H}\,\mathbf{x} + \mathbf{n}. \tag{1}$$

Where $\mathbf{y} \in \mathbb{C}^R$ represents the received signal, the column vector $\mathbf{x} = (x_1, x_2, \ldots, x_{\mathrm{M}})^{\mathrm{T}} \in \mathbb{C}^T$ is the transmitted signal with M is the constellation size, and $\mathbf{H} \in \mathbb{C}^{T \times R}$ is the channel gain matrix of random coefficients $h_{ij}\ \big(i \in [1, \ldots, T],\ j \in [1, \ldots, R]\big)$ with zero mean and unit variance. The complex noise vector $\mathbf{n} \in \mathbb{C}^R$ has independent and identically distributed (i.i.d) Gaussian samples with mean also equals zero and variance is σ_n^2. Suppose the channel state information (CSI) at the receiver is known and, finally, P is the transmitted signal power.

Fig. 1 shows the signal constellation of quadrature amplitude modulation (QAM) with constellation size M= 16 ($\mathcal{X}_{16\text{-QAM}}$).

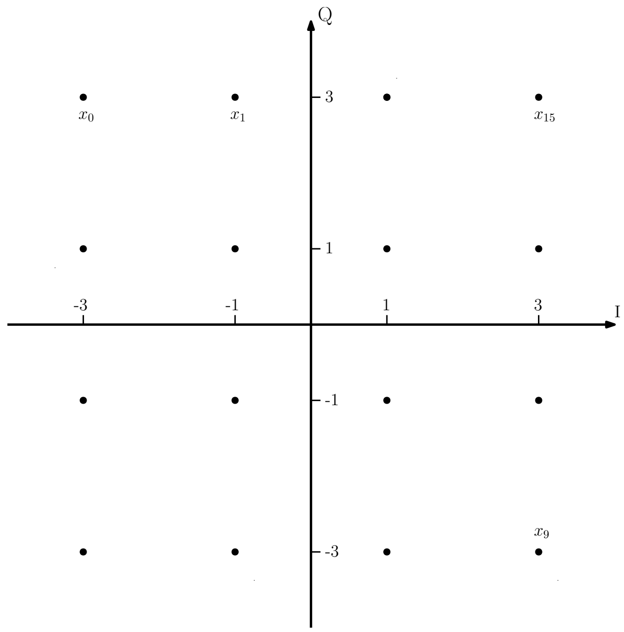

Fig. 1. Signal constellation of $\mathcal{X}_{\text{16-QAM}} = \mathcal{X}_{\text{4-PAM}} \times \mathcal{X}_{\text{4-PAM}}$.

Suppose the conditional probability density function (pdf) is

$$p(y|\mathbf{H}x) = (\frac{1}{\sqrt{2\pi\sigma^2}})^R \, e^{-\frac{||y-hx||^2}{2\sigma_n^2}}, \tag{2}$$

and the channel capacity of the MIMO channel is defined as $\mathrm{C} = \sup_{p} \mathrm{I}\,(\mathbf{x};\mathbf{y})$, where $\mathrm{I}\,(\mathbf{x};\mathbf{y})$ is the mutual information between channel input and output.

$$\mathrm{I}\,(\mathbf{x};\mathbf{y}) = \iint p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y|\mathbf{H})} \,\mathrm{d}h\,\mathrm{d}y \qquad \text{and}$$

$$p(y|\mathbf{H}) = \sum_{x_i \in \mathcal{X}_{\text{16-QAM}}} p(x_i)\, p(y|\mathbf{H}x).$$

Now, let us define the marginal output density of y corresponding to the input distribution Q(x) as $p\big(y\,;\mathrm{Q}(x)\big) = \int p(y|\mathbf{H}x) \;\; \mathrm{dQ}(x)$. Then

$$\mathrm{C} = \sup_{\substack{\mathrm{Q}\in\mathrm{Q} \\ \mathrm{E}[\mathbf{x}^{\mathrm{H}}\mathbf{x}]\le P}} \iint p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x))} \,\mathrm{d}y\,\mathrm{dQ}(x) \tag{3}$$

with Q is the set of all satisfied distributions [1], [2], i.e., $\mathrm{Q} = \mathrm{Q} : \int_{-\infty}^{\infty} x^2 \;\; \mathrm{dQ}(x) \le P$. Calculus may not easily solve the supremum of the last equation. Instead, convex optimization theory is able to show that a unique variable x^* with its input distribution CDF $\mathrm{Q}(x^*)$ attains the capacity C [3]. The required conditions are as the following

$$\mathrm{C} \ge \lambda(P - ||\mathbf{x}||^2) + \iint p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x))} \,\mathrm{d}y\,\mathrm{dQ}(x), \tag{4}$$

for all x, then

$$\mathrm{C} = \mathrm{I}(\mathrm{Q}(x^*))$$

$$= \lambda^*\,(P - ||\mathbf{x}^*||^2) + \int p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))} \,\mathrm{d}y, \tag{5}$$

where $\lambda^* = \lambda(P) \ge 0$ denotes the optimal Lagrange multiplier.

III. The Problem

A. Capacity-Achieving Distribution Characterization

It is necessary to point out here some properties of $\mathrm{I}(\mathrm{Q}(x^*)) = \int p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))} \,\mathrm{d}y$.

Lemma 1. *For a Rayleigh fading channel giving in* (1) *with input distribution* Q(x)*, the integration* $\int p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))} \,\mathrm{d}y$ *satisfies the following properties*

a) $$\lim_{x\to\infty} \int p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))} \,\mathrm{d}y$$

$$\approx -\log_2 \big[p(y\,;\mathrm{Q}(x^*))\big]. \tag{6}$$

b) $$\int p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))} \,\mathrm{d}y \; < -\log_2 \big[p(y\,;\mathrm{Q}(x^*))\big]. \tag{7}$$

Proof: For any finite noise variance σ^2, as $x \longrightarrow \infty$ the conditional pdf $p(y|\mathbf{H}x)$ tends to unity. This leads to

$$\lim_{x\to\infty} \int p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))} \,\mathrm{d}y$$

$$\approx 0 \;-\; \log_2 \big[p(y\,;\mathrm{Q}(x^*))\big] = -\log_2 \big[p(y\,;\mathrm{Q}(x^*))\big].$$

To prove part **b**, and by Jensen's inequality

$$\int p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))} \,\mathrm{d}y$$

$$\ge \; -\log_2 \int p(y|\mathbf{H}x) \, \frac{p(y\,;\mathrm{Q}(x^*))}{p(y|\mathbf{H}x)} \,\mathrm{d}y$$

$$= -\log_2 \int p(y|\mathbf{H}x)\, p(y\,;\mathrm{Q}(x^*)) \,\mathrm{d}y = \; 0$$

$$\Longrightarrow \int p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))} \,\mathrm{d}y \;\ge\; 0$$

Or $$\underbrace{\int p(y|\mathbf{H}x) \log_2 p(y|\mathbf{H}x) \,\mathrm{d}y}_{<\,0}$$

$$< \; -\int p(y|\mathbf{H}x) \log_2 p(y\,;\mathrm{Q}(x^*)) \,\mathrm{d}y.$$

Since the entropy is bounded by the logarithm of the alphabet size $\big(|\mathrm{Q}(x^*)| < \infty\big)$ [4]

$$\Longrightarrow \int p(y|\mathbf{H}x) \log_2 p(y|\mathbf{H}x)\,\mathrm{d}y \, < -\log_2 \big[p(y\,;\mathrm{Q}(x^*))\big]$$

$$\Longrightarrow \int p(y|\mathbf{H}x) \log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))}\,\mathrm{d}y < -\log_2 \big[p(y\,;\mathrm{Q}(x^*))\big].$$

Now

$$\begin{aligned}\mathrm{C} &= \lambda^* \,(P - ||\mathbf{x}^*||^2) \\ &+ \big(\frac{1}{\sqrt{2\pi\sigma_n^2}}\big)^R \int e^{\frac{-||y-hx||^2}{2\sigma_n^2}} \log_2 \frac{\big(\frac{1}{\sqrt{2\pi\sigma_n^2}}\big)^R e^{\frac{-||y-hx||^2}{2\sigma_n^2}}}{p(y\,;\mathrm{Q}(x^*))}\,\mathrm{d}y \\ &= \lambda^* \,(P - ||\mathbf{x}^*||^2) - R \log_2 \sqrt{2\pi\, e\, \sigma_n^2} \\ &- \big(\frac{1}{\sqrt{2\pi\sigma_n^2}}\big)^R \int e^{\frac{-||y-hx||^2}{2\sigma_n^2}} \log_2 p(y\,;\mathrm{Q}(x^*))\,\mathrm{d}y. \end{aligned} \tag{8}$$

The using of Hermite polynomials combination can help in expanding the last integration [5].

B. The Optimal Input

Let $p(y) = p(y\,;\mathrm{Q}(x^*))$ and by assuming that $\sigma_n = 1$ without losing the generality

$$\begin{aligned}\mathrm{C} =& \lambda^* \,(P - ||\mathbf{x}^*||^2) - R \log_2 \sqrt{2\pi\, e} \\ &- \big(\frac{1}{\sqrt{2\pi}}\big)^R \int e^{\frac{-||y||^2}{2}}\, e^{\frac{-||hx^*||^2}{2} + ||hx^*||\, y^{\mathrm{H}}} \log_2 p(y)\,\mathrm{d}y.\end{aligned}$$

According to Hermite polynomial properties [6], the generating function is $e^{\frac{-||hx^*||^2}{2} + ||hx^*||\, y^{\mathrm{H}}} = \sum_{n=0}^{\infty} \mathrm{H}_n(y)\, \frac{||hx^*||^n}{n!}$, and since $p(y)$ is a continuous function of y, so is $\log_2 p(y)$. Then the piecewise continuous function $\log_2\ p(y)$ can be written as $\log_2 p(y) = \sum_{n=0}^{\infty} \mathrm{c}_n\, \mathrm{H}_n(y)$.
By substitution, we get

$$\begin{aligned}&\big(\frac{1}{\sqrt{2\pi}}\big)^R \int e^{\frac{-||y||^2}{2}}\, e^{\frac{-||hx^*||^2}{2} + ||hx^*||\, y^{\mathrm{H}}} \log_2 p(y)\,\mathrm{d}y \\ &= \big(\tfrac{1}{\sqrt{2\pi}}\big)^{R-1} \int \tfrac{1}{\sqrt{2\pi}}\, e^{\frac{-||y||^2}{2}} \textstyle\sum_{n=0}^{\infty} \mathrm{H}_n(y)\, \frac{||hx^*||^n}{n!} \log_2 p(y)\,\mathrm{d}y \\ &= \big(\tfrac{1}{\sqrt{2\pi}}\big)^{R-1} \textstyle\sum_{n=0}^{\infty} \Big[\frac{1}{\sqrt{2\pi}\, n!} \int e^{\frac{-||y||^2}{2}} \log_2 p(y)\, \mathrm{H}_n(y)\,\mathrm{d}y\Big] ||hx^*||^n \\ &= \big(\frac{1}{\sqrt{2\pi}}\big)^{R-1} \sum_{n=0}^{\infty} \mathrm{c}_n\, ||hx^*||^n .\end{aligned}$$

Therefore

$$\begin{aligned}\mathrm{C} =& \lambda^* \,(P - ||\mathbf{x}^*||^2) \,-\, R \log_2 \sqrt{2\pi\, e} \\ &- \big(\frac{1}{\sqrt{2\pi}}\big)^{R-1} \sum_{n=0}^{\infty} \mathrm{c}_n\, ||hx^*||^n .\end{aligned}$$

Or

$$\begin{aligned}\sum_{n=0}^{\infty} \mathrm{c}_n\, ||hx^*||^n =& \big(\sqrt{2\pi}\big)^{R-1} \big[\lambda^* \,(P - ||\mathbf{x}^*||^2) \,-\, \mathrm{C} \\ &- R \log_2 \sqrt{2\pi\, e}\,\big] .\end{aligned}$$

Which yields

$$\mathrm{c}_0 = \big(\sqrt{2\pi}\big)^{R-1} \big[\lambda^*\, P \,-\, \mathrm{C} \,-\, R \log_2 \sqrt{2\pi\, e}\big]\,,\ \mathrm{c}_1 = 0\,,$$
$$\mathrm{c}_2 \leq -\big(\sqrt{2\pi}\big)^{R-1} \frac{\lambda^*}{||h||^2}\,,\ \text{and}\ \ \mathrm{c}_n = 0 \quad \text{for}\ n \geq 3\,.$$

Consequently,

$$\begin{aligned}\log_2 p(y) &= \sum_{n=0}^{\infty} \mathrm{c}_n\, \mathrm{H}_n(y) = \mathrm{c}_0\, \mathrm{H}_0(y) + \mathrm{c}_2\, \mathrm{H}_2(y) \\ &= \mathrm{c}_0 + \mathrm{c}_2(||y||^2 - 1) = (\mathrm{c}_0 - \mathrm{c}_2) + \mathrm{c}_2||y||^2 \\ \text{and hence}\ \ p(y) &= e^{\ln 2\, \big[(\mathrm{c}_0 - \mathrm{c}_2) + \mathrm{c}_2\, ||y||^2\big]} \\ &= \mathrm{k}\, e^{\,\mathrm{c}_2 \ln 2\, ||y||^2},\ \mathrm{k} = e^{\,(\mathrm{c}_0 - \mathrm{c}_2)\ln 2}. \end{aligned} \tag{9}$$

Finally, the appropriate input probability law $p(x^*)$ that induces such output Gaussian distribution could be found as

$$\begin{aligned}p(y|\mathbf{H}x) &= \big(\frac{1}{\sqrt{2\pi}}\big)^R e^{\frac{-||y-hx||^2}{2}} \qquad\qquad (\sigma_n = 1) \\ &= \big(\frac{1}{\sqrt{2\pi}}\big)^R e^{\frac{-||y||^2}{2}}\, e^{\frac{-||hx^*||^2}{2} + ||hx^*||\, y^{\mathrm{H}}} \\ &= \big(\frac{1}{\sqrt{2\pi}}\big)^R e^{\frac{-||y||^2}{2}} \sum_{n=0}^{\infty} \mathrm{H}_n(y)\, \frac{||hx^*||^n}{n!}\,,\end{aligned}$$

then

$$\begin{aligned}p\big(y\,;\mathrm{Q}(x^*)\big) &= \int p(y|\mathbf{H}x)\ \ \mathrm{dQ}(x^*) \\ &= \int \big(\tfrac{1}{\sqrt{2\pi}}\big)^R e^{\frac{-||y||^2}{2}} \textstyle\sum_{n=0}^{\infty} \mathrm{H}_n(y)\, \frac{||hx^*||^n}{n!}\ \mathrm{dQ}(x^*) \\ &= p(y) = \mathrm{k}\, e^{\,\mathrm{c}_2 \ln 2\, ||y||^2} .\end{aligned}$$

Hence

$$\mathrm{k}\, e^{\,\mathrm{c}_2 \ln 2\, ||y||^2} =$$
$$\big(\frac{1}{\sqrt{2\pi}}\big)^{R\text{-}1} \sum_{n=0}^{\infty} \frac{1}{\sqrt{2\pi}\, n!} \int e^{\frac{-||y||^2}{2}}\, ||hx^*||^n\ \ \mathrm{dQ}(x^*)\, \mathrm{H}_n(y).$$

Multiplying by $\mathrm{H}_n(y)$ gives

$$\begin{aligned}&\Big[\mathrm{k}\, e^{\,\mathrm{c}_2 \ln 2\, ||y||^2}\Big]\, \mathrm{H}_n(y) = \big(\frac{1}{\sqrt{2\pi}}\big)^{R\text{-}1} \\ &\times \sum_{n=0}^{\infty} \Big[\frac{1}{\sqrt{2\pi}\, n!} \int e^{\frac{-||y||^2}{2}}\, ||hx^*||^n\, \mathrm{dQ}(x^*)\, \mathrm{H}_n(y)\Big]\, \mathrm{H}_n(y).\end{aligned}$$

The last result could be manipulate as the following:

$$\int \Big[\mathrm{k}\, e^{\,\mathrm{c}_2 \ln 2\, ||y||^2}\Big]\, \mathrm{H}_n(y)\,\mathrm{d}y = \big(\frac{1}{\sqrt{2\pi}}\big)^{R\text{-}1} \int ||hx^*||^n\, \mathrm{dQ}(x^*).$$

The integration on the left hand side could be consider as the expectation operation of $\mathrm{H}_n(z)$ where z is a complex random variable with normal distribution of zero mean and σ_z standard deviation. Therefore,

$$\int \left[\mathrm{k}\, e^{\,\mathrm{c}_2 \ln 2\, ||z||^2}\right] \mathrm{H}_n(z)\ \mathrm{d}z$$

$$= \mathrm{k}\ \sqrt{2\pi\sigma_z^2}\ \int \mathrm{H}_n(z)\, \frac{e^{\frac{-||z||^2}{2\sigma_z^2}}}{\sqrt{2\pi\sigma_z^2}}\ \mathrm{d}z = \mathrm{k}\ \sqrt{2\pi\sigma_z^2}\ \mathrm{E}\left[\mathrm{H}_n(z)\right]$$

$$= \mathrm{k}\,\sqrt{2\pi\sigma_z^2}\ \mu^n\ \mathrm{H}_n(0)\,, \qquad \mu \text{ is the expectation of } \mathrm{H}_n(z)\ [6]$$
$$= \mathrm{k}\,\sqrt{2\pi\sigma_z^2}\ \left(\sqrt{\mathrm{E}(z^2) - \sigma_z^2}\,\right)^n \mathrm{H}_n(0)$$
$$= \mathrm{k}\,\sqrt{2\pi\sigma_z^2}\ \left(\sqrt{(2-1)!! - \sigma_z^2}\,\right)^n \mathrm{H}_n(0)$$
$$= \mathrm{k}\,\sqrt{2\pi\sigma_z^2}\ \left(\sqrt{1 - \sigma_z^2}\,\right)^n \mathrm{H}_n(0)$$
$$= \begin{cases} 0 & \text{if } n \text{ is odd} \\ \mathrm{k}\,\sqrt{2\pi\sigma_z^2}\ \left(\sqrt{1 - \sigma_z^2}\,\right)^n \mathrm{H}_n(0) & \text{if } n \text{ is even}, \end{cases} \tag{10}$$

where $\sigma_z^2 < \dfrac{||h||^2}{(\sqrt{2\pi})^{R\text{-}1}\,\lambda^* \ln 4}$. Hence

$$\int ||hx^*||^n\ \ \mathrm{dQ}(x^*) = \mathrm{k}\,\sqrt{(2\pi)^R\,\sigma_z^2}\ \left(\sqrt{1-\sigma_z^2}\,\right)^n \mathrm{H}_n(0). \tag{11}$$

The last integral is kind of Riemann-Stieltjes integral applies to probability theory. It represents the n-th moment of the probability distribution. Here, the moment generating function (mgf), $\mathrm{M}_{x^*}(t)$, is

$$\mathrm{M}_{x^*}(t) = \int ||hx^*||^n\ \ \mathrm{dQ}(x^*) = \mathrm{E}\left[e^{\,t\,||hx^*||}\right]$$
$$= \sum_{x^*} e^{\,t\,||hx^*||}\ \mathrm{k}\,\sqrt{(2\pi)^R\,\sigma_z^2}\ \left(\sqrt{1-\sigma_z^2}\,\right)^n \mathrm{H}_n(0)$$
$$= \sum_{x^*}\sum_{n=0}^{\infty} \frac{(t\,||hx^*||)^n}{n!}\ \mathrm{k}\,\sqrt{(2\pi)^R\,\sigma_z^2}\ \left(\sqrt{1-\sigma_z^2}\,\right)^n \mathrm{H}_n(0)$$
$$= \mathrm{k}\,\sqrt{(2\pi)^R\,\sigma_z^2}\ \sum_{n=0}^{\infty}\left[\mathrm{H}_n(0)\,\frac{(\sqrt{1-\sigma_z^2}\,t)^n}{n!}\ \ \underbrace{\sum_{x^*}||hx^*||}_{=\,1 \text{ for } n\,=\,0}\ \right]$$
$$= \mathrm{k}\,\sqrt{(2\pi)^R\,\sigma_z^2}\ e^{-\frac{1-\sigma_z^2}{2}\,t^2}\,. \tag{12}$$

Consequently, we can recognize that the pdf of the continuous input x^* is $p(x^*) = \mathrm{k}\,\sqrt{(2\pi)^R\,\sigma_z^2}\ e^{-\frac{1-\sigma_z^2}{2}\,||x^*||^2}$, which is valid whenever $\sigma_z^2 \le 1$ or $\lambda^* \ge \dfrac{||h||^2}{(2\pi)^{R-1}\ln 4}$.

IV. Capacity Computation

A. Capacity under Amplitude-Limited Inputs

We now need to maximize $\mathrm{I}\,(\mathbf{x};\mathbf{y})$ for all $s > 0$. To do this, we must find the *largest* s, referred to s^*, such that the code still converges to zero error rate. Such s^* could be found through binary search, noting that it is possible to limit the search space to $s_{\min} \le s \le s_{\max}$, where $s_{\min}$ and $s_{\max}$ are the minimum and maximum, respectively.

Lemma 2. *There exists a bounded range for* M-QAM *signal constellation gains (s) such that*

$$\sqrt{\frac{P}{2(\sqrt{\mathrm{M}}-1)^2\,T}} = s_{\min} \le s \le s_{\max} = \sqrt{\frac{P}{2T}}\,,$$

where P is transmitting power over T antennas.

At the beginning, it is necessary to show how could the power constraint given by $s^2\,\mathrm{E}[||x_i||^2] \le P, \quad \forall\, x_i \in$ M-QAM be replaced with equality (without loss of optimality). For the two scaling factors $s_1 > 0$ and $s_2 > 0$, assume that $\forall\ x_i \in$ M-QAM

$$s_1{}^2\,\mathrm{E}[||x_i||^2] \le P \quad \text{and} \quad s_2{}^2\,\mathrm{E}[||x_i||^2] = P.$$

It is being adequate to say that $\mathrm{I}\,(\mathrm{x};\mathrm{y}_1) \ge \mathrm{I}\,(\mathrm{x};\mathrm{y}_2)$ with y_1 and y_2 are the corresponding output random variables. Thus,

$\mathrm{I}\,(\mathrm{x};\mathrm{y}_1) = h(\mathrm{x}) - h(\mathrm{x}|\mathrm{y}_1)$
conditioning on an extra variable can only decrease entropy
$\le h(\mathrm{x}) - h(\mathrm{x}|\mathrm{y}_2, \mathrm{y}_1)$
$= h(\mathrm{x}) - h(\mathrm{x}|\mathrm{y}_2) \qquad p_{\mathrm{x|yz}}(x|y,z) = p_{\mathrm{x|y}}(x|y)$
$= \mathrm{I}\,(\mathrm{x};\mathrm{y}_2)$.

Hence, $s^2 \sum\limits_{x_i \in \mathcal{X}_{\text{M-QAM}}} ||x_i||^2 = P.$

Secondly, in many $\mathcal{X}_{\text{M-QAM}}$ constellations, the power requires to transmit the highest-amplitude symbol is $P_{\max} = 2(\sqrt{\mathrm{M}} - 1)^2$, and for T transmit antennas it is $P_{\max} = 2(\sqrt{\mathrm{M}} - 1)^2\,T$, where $\mathrm{M} = 2^m$. While the peak needs power approximately tends to vary with the number of m as $P \propto 2^m$, hence the minimum transmitted power over T antennas is $P_{\min} = 2T$.

Algorithms 1 and 2 present in details an iterative procedure to perform the M-QAM capacity computation problem.

B. Factorizing the Optimal QAM *Distribution*

The system with $(\mathrm{M}^T\!-1)$ optimized variables over $\mathcal{X}_{\text{M-QAM}}$ constellation gets extremely high complexity and needs a solution that simplifies such presented algorithm (Algorithm 1). This solution focuses on the idea that QAM constellation could be performed of signal points placed symmetrically in (I-Q) plane and gives an amplitude-modulation of two orthogonal waveforms. Thus,

$$\mathbf{x} = (x_{1i}\,,\, x_{1q}\,,\, x_{2i}\,,\, x_{2q}\,,\, \ldots\,,\, x_{\mathrm{M}i}\,,\, x_{\mathrm{M}q}) \in \mathcal{X}^2_{\sqrt{\mathrm{M}}\text{-PAM}}\,. \tag{13}$$

We will denote $q(x)$ to be any valid probability function defined over $\mathcal{X}_{\sqrt{\mathrm{M}}\text{-PAM}}$. If all of the elements in (13) are supposed to be i.i.d. with $q(x)$, this means we are actually looking up for the distribution $q(x^*)$ that maximizes C. Accordingly, we make the following important conjectures.

Conjecture 1. *The input distribution $p(x^*)$ that could maximize* C *over $\mathcal{X}_{\text{M-QAM}}$ realizes as*

$$p(x^*) = \prod_{m\in\{i,q\}}\prod_{n=1}^{T} q(x^*_{mn}) \qquad \forall\ x^*_{mn} \in \mathcal{X}_{\sqrt{\mathrm{M}}\text{-PAM}} \tag{14}$$

Algorithm 1 s^*, $p(x^*)$, and $\mathrm{C}_{\text{M-QAM}}$ Calculations.

1: Set $s = s_{\text{min}}$.
2: For all $x \in \mathcal{X}_{\text{M-QAM}}$, compute
- $\mathrm{C} = \sup_{\mathrm{Q}\in\mathrm{Q}} \iint p(y|\mathbf{H}x)\log_2 \frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x))}\,\mathrm{d}y\,\mathrm{dQ}(x)$,
- $p\big(y\,;\mathrm{Q}(x)\big) = \mathrm{k}\,e^{\,\mathrm{c}_2\ln 2\,||y||^2}$.
- $p(x) = \mathrm{k}\sqrt{(2\pi)^R\sigma_z^2}\quad e^{-\frac{1-\sigma_z^2}{2}\,||sx||^2}$, $\mathrm{k} = e^{(\mathrm{c}_0-\mathrm{c}_2)\ln 2}$, $\mathrm{c}_2 \le -\big(\sqrt{2\pi}\big)^{R-1}\frac{\lambda}{||h||^2}$, $\mathrm{c}_0 = \big(\sqrt{2\pi}\big)^{R-1}\big[\lambda P - \mathrm{C} - R\log_2\sqrt{2\pi e}\,\big]$ and $\sigma_z^2 < \dfrac{||h||^2}{(\sqrt{2\pi})^{R\text{-}1}\,\lambda\ln 4}$.

3: λ is chosen to satisfy
- $\lambda > \frac{-\log_2[p(y\,;\mathrm{Q}(x))]-\mathrm{C}}{||sx||^2-P}$ and $\lambda \ge \frac{||h||^2}{(2\pi)^{R-1}\ln 4}$.

4: Repeat steps 2 & 3. When $p(x)$ does converge, call it $p(x^*)$.
5: Set $s = s_{\text{new}}$ (as in Algorithm 2). Go to step 2.
6: Compute $\mathrm{C}_{\text{M-QAM}} = \int p(y|\mathbf{H}x)\log_2\frac{p(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))}\,\mathrm{d}y$.

Algorithm 2 Calculating the optimum value of (s) by binary search optimization method.

1: Sweep $s \in \big[s_{\text{min}}, s_{\text{max}}\big]$.
2: Set $s_1 = s_{\text{min}}$, and $s_2 = s_{\text{max}}$.
3: Do Algorithm 1 so that the probability functions converge using s_1 and do not using s_2.
4: If $s_1 - s_2 < \delta$ (the search accuracy), put $s_{\text{new}} = s_1$. Stop.
5: Calculate $s_3 = \dfrac{s_1+s_2}{2}$. Do Algorithm 1 with $s_{\text{new}} = s_3$. If it converges, let $s_1 = s_{\text{new}}$. Otherwise, let $s_2 = s_3$.
6: Go to step 4.

Conjecture 2. *If* $\boldsymbol{x} = (x_{1i}, x_{1q}, x_{2i}, x_{2q}, \ldots, x_{Mi}, x_{Mq})$ *and* $p(x)$ *has the form of (14), then*

$$\mathrm{E}[x_{mn}^2] \le \frac{P}{2T} \tag{15}$$

Algorithm 3 presents the procedure for accomplishing this optimization task.

Finally, to examine the results, $\mathrm{C}_{\text{M-QAM}}$ is compared with the capacity determined according to independent and uniformly distributed signalling ($\mathrm{C}_{\text{M-QAM-i.u.d.}}$). We consider all signal points are equiprobable and therefore the constellation

Algorithm 3 s^*, $q(x^*)$, $p(x^*)$, and $\mathrm{C}_{\text{M-QAM}}$ Calculations.

1: Set $s = s_{\text{min}}$.
2: For all $x \in \mathcal{X}_{\sqrt{\text{M}}\text{-PAM}} = (x_1, x_2, \ldots, x_{\sqrt{\text{M}}})$ and $(m,n) \in \{1,2,\ldots,T\}\times(i,q)$, compute
- $\mathrm{C} = \sup_{\mathrm{Q}\in\mathrm{Q}} \iint p_{y|\mathbf{H}x_{mn}}(y|\mathbf{H}x)\log_2\dfrac{p_{y|\mathbf{H}x_{mn}}(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x))}\,\mathrm{d}y\,\mathrm{dQ}(x)$, where due Baye's law $p_{y|\mathbf{H}x_{mn}}(y|\mathbf{H}x) = \sum_{x_i\in\mathcal{X}_{\text{M-QAM}}} p(y|\mathbf{H}x)$ and $\mathrm{Q} = \mathrm{Q} : \int_{-\infty}^{\infty}(sx_{mn})^2\;\mathrm{dQ}(x) \le \frac{P}{2T}$.
- $q(x) = \mathrm{k}\sqrt{(2\pi)^R\sigma_z^2}\quad e^{-\frac{1-\sigma_z^2}{2}\,(sx_{mn})^2}$, $\mathrm{k} = e^{(\mathrm{c}_0-\mathrm{c}_2)\ln 2}$, $\mathrm{c}_2 \le -\big(\sqrt{2\pi}\big)^{R-1}\frac{\lambda}{||h||^2}$, $\mathrm{c}_0 = \big(\sqrt{2\pi}\big)^{R-1}\big[\lambda\frac{P}{2T} - \mathrm{C} - R\log_2\sqrt{2\pi e}\,\big]$ and $\sigma_z^2 < \dfrac{||h||^2}{(\sqrt{2\pi})^{R\text{-}1}\,\lambda\ln 4}$.

3: λ is chosen to satisfy
- $\lambda > \frac{-\log_2[q(y\,;\mathrm{Q}(x))]-\mathrm{C}}{(sx_{mn})^2-\frac{P}{2T}}$ and $\lambda \ge \frac{||h||^2}{(2\pi)^{R-1}\ln 4}$.

4: Repeat steps 2 & 3. When $q(x)$ does converge, call it $q(x^*)$.
5: Compute $p(x^*)$ as in (21).
6: Set $s = s_{\text{new}}$ (as in Algorithm 2). Go to step 2.
7: Compute $\mathrm{C}_{\text{M-QAM}} = \int p_{y|\mathbf{H}x_{mn}}(y|\mathbf{H}x)\log_2\frac{p_{y|\mathbf{H}x_{mn}}(y|\mathbf{H}x)}{p(y\,;\mathrm{Q}(x^*))}\,\mathrm{d}y$.

gain s sets to $\sqrt{\dfrac{P}{\frac{\mathrm{M}^2-1}{12}}}$ and $p(x_i) = \dfrac{1}{\mathrm{M}^T}$, $i \in \{1,2,\ldots,\mathrm{M}^{\mathrm{T}}\}$.

V. Numerical Results

We employ Algorithm 3 to compute the achievable rate under Rayleigh fading channel with successive cancellation (SC) polar codes of 1024 code length and rate of $\frac{1}{2}$ coded the 16-QAM input signal constellation for a (2×2) system. For example, in Fig. 2 we can see that a rate of 1.3 bit per channel used is achieved at 9.298 dB when $\mathcal{X}_{\text{M-QAM}}$ constellation is transmitted over each antenna. By using $\mathcal{X}^2_{\sqrt{16}\text{-PAM}}$ constellation, this rate could be achieved at 9.357 dB by utilizing an optimum constellation gain of $s^* = 0.5825$. Thus, $\mathrm{C}_{\text{M-QAM}}$ is within 0.059 dB from channel capacity limit at 1.3 bit per channel used; while $\mathrm{C}_{\text{M-QAM-i.u.d.}}$ is about 0.349 dB from the limit, displaying a gap of 0.29 dB with $\mathrm{C}_{\text{M-QAM}}$.

Now, according to Algorithm 1 there are $16^2-1 = 255$ variables required to fulfill the optimization task; compared

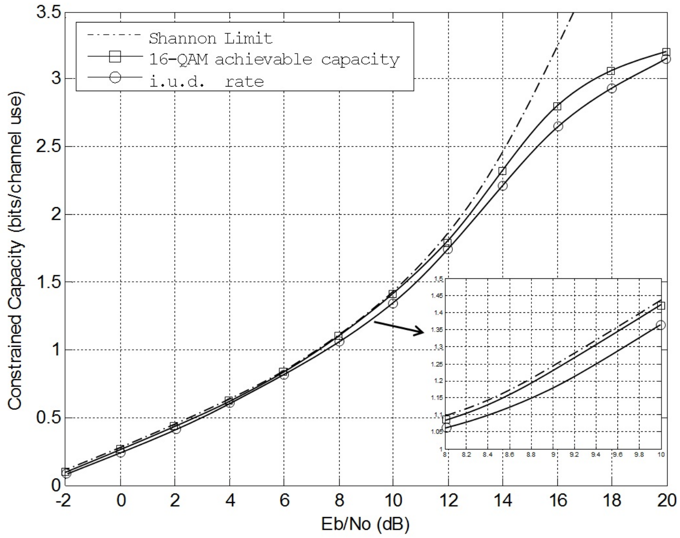

Fig. 2. Achieved capacities under different transmission scenarios. The SC polar codes has code length of 1024 with half rate.

with $\sqrt{16} - 1 = 3$ carried out according to Algorithm 3.

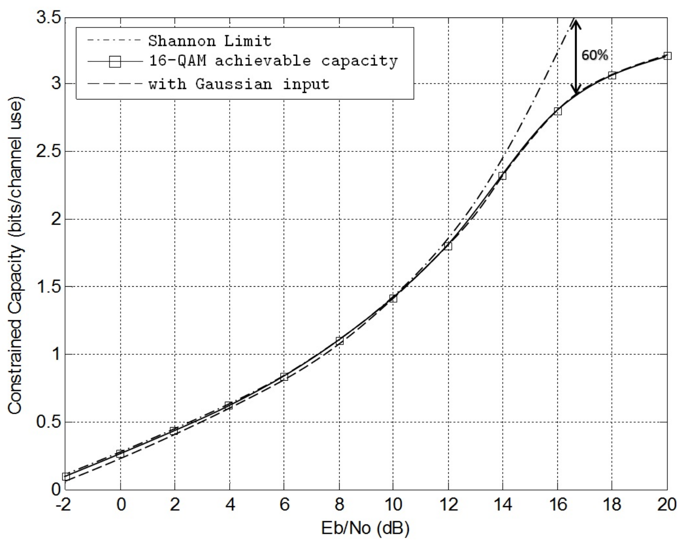

Fig. 3. Achieved capacities under different transmission scenarios. The SC polar codes has code length of 1024 with half rate.

Fig. 3 presents the numerically computed rates when Gaussian distribution is employed. It is shown that the use of the continuous optimal input can performs better than the Gaussian one with the fading channel. Also it is observed the cumulative detrimental effect of such channels linearity yields a rate lower that Shannon limit by 60% at 16.7 dB. Such rates degradation is found to be for low dBs as 4% in maximum.

VI. Conclusion

An optimization method is introduced in this work for approaching the capacity of fading channel with multi-antennas. Under 16-QAM modulation with the condition $\mathrm{E}\left[x^H x\right] \leq P$ and the aid of Hermite polynomials, an optimum Gaussian distribution with its correction factor are determined. Moreover, we have proved that the complexity of the proposed algorithm could be much reduced by factoring the determined Gaussian distribution into a couple of 4-PAM signals.

References

[1] S. Shamai (Shitz) and I. David, "*The capacity of average and peak-power-limited quadrature Gaussian channels*," IEEE Trans. Inf. Theo., vol. 41, no. 4, pp. 1060-1071, July 1995.

[2] I. Faycal and S. Shamai (Shitz), "*The capacity of discrete-time memoryless rayleigh-fading channels*," IEEE Trans. Inf. Theo., vol. 47, no. 4, pp. 1290-1301, May 2001.

[3] J. Smith, "*The information capacity of peak and average power constrained scalar Gaussian channels*," Inf. and Cont., vol. 18, pp. 203-219, 1971.

[4] R. Ash, "*Information Theory*," University of Illinois, Illinois, John Wiley & Sons, 1965.

[5] X. Wang, E. Serpedin and K. Qaraqe, "*A variational approach for assessing the capacity of a memoryless nonlinear MIMO channel*," IEEE Lett. Comm., vol. 18, no. 8, pp. 1315-1318, Aug. 2014.

[6] G. Djordjević, "*On some properties of generalized Hermite polynomials*," University of Niš, Faculty of Technology, Yugoslavia, 1994.