Power Allocation for Parallel Channels with One-bit Analog-to-digital Conversion

Zhide Li, Zhou Lv,Xuedan Zhang

Department of Electronic Engineering,Tsinghua University,Shenzhen,518055,China

Keywords: Channel capacity, power allocation, precoding, quantization.

Abstract. In this paper, the problem of maximizing the mutual information of independent parallel Gaussian-noise channels is studied under an average power constraint with one-bit analog-to-digital (A/D) convertor. It is shown that the power allocation problem is a convex optimization problem, which can be solved efficiently. Meanwhile, the waterfilling interpretation for the power allocation problem is also given. Further, the combined 1bit/dimension vector quantizer and precoder design are studied in the 2 parallel independent channels. Simulation results demonstrate that substantial gain can be achieved compared to power allocation scheme especially when the difference of channel gains is large and the range of SNR is low to medium.

Introduction

The problem of power allocation among a bank of independent parallel channels has drawn a great deal of attention. For a given independent parallel channels under a total power constraint, it is well known that Gaussian inputs with the waterfilling power allocation scheme will achieve the capacity [1]. However, Gaussian inputs can never be realized in practice due to its infinite peak-to-average ratios. As a consequence, discrete constellations such as phase shift keying (PSK) and quadrature amplitude modulation (QAM) are adopted in practical communication systems. In [2], the power allocation for parallel Gaussian channels with arbitrary input distributions is investigated. The optimum is ensured to be found and has a good interpretation named mercury/waterfilling. In addition, it has been demonstrated that a linear precoder may increase the mutual information between the source and channel's output. Thus the optimal precoder is designed in parallel Gaussian channels with arbitrary inputs [3].

The above works perform the design by assuming that the precision of the analog-to-digital converter (ADC) is infinite. Traditionally, the precision of the ADC is often high enough (e.g., 10-12 bits or more), which makes the assumption of the infinite precision of the ADC practical. Consequently, considerable effort is spent on the optimization of modulation, demodulation and decoding [4,5]. However, as the data rates increases, high-speed high-precision ADC is often not available. Even it is available, the cost and power dissipation are large [6]. One approach is to decrease the resolution of an ADC to a few bits by reducing the number of comparators. As a result, the receiver blocks such as synchronization, channel estimation and demodulation need to be redesigned [7],[8],[9]. In addition, the quantizer must be designed carefully in order to reduce the performance loss compared to the case of high precision ADC.

There has been a lot of work focusing on the performance limits of low precision quantization ranging from signal processing to communication theory. The problem of estimating a frequency of a single complex-valued sinusoid in additive Gaussian noise for 1bit quantization is studied, which demonstrates that the performance degradation is not large [10]. In [11], a scheme is proposed to estimate the impulse response of a dispersive channel when a low precision ADC utilized at the receiver. It is shown that the scheme provides performance close to that of the full precision sampling. The capacity of the additive white Gaussian noise (AWGN) channel with low-precision (e.g., 1-4 bits) receiver is studied in [12]. It is demonstrated that at high enough signal to noise ratio (SNR), the low-precision receiver incus a small loss compared to unquantized observations. This motivates us to study the capacity of the independent parallel channels with low-precision ADC.

In this paper, we first consider the capacity of the independent parallel channels under a total power constraint with one bit ADC. Then a precoder is proposed to a specific two independent

parallel channels, and a 1bit/dimension vector quantizer is utilized. It is shown that this precoder scheme works better than the power allocation method, especially when the SNR is medium and the difference of the channel gain is large.

Preliminary

In this paper, we first review the classical n parallel independent channels model

$$y_i = \sqrt{\mathrm{snr}} h_i \sqrt{p_i} x_i + n_i, i = 1, \cdots, n,\qquad(1)$$

where y_i, h_i, p_i and x_i denote respectively the ith output, channel coefficient, power and input, snr is the signal to noise ratio, $n_i \sim N(0,1)$ denotes the i.i.d. additive Gaussian noise. We assume that each x_i is independent of each other and is a zero-mean unit variance random variable, and $\{p_i\}_{i=1}^n$ satisfy $\sum_{i=1}^n p_i = 1$. Although Gaussian inputs will achieve the channel capacity, we assume that the input $\{x_i\}_{i=1}^n$ is always drawn equiprobable from an M-ary alphabet in this paper. Consequently, the mutual information expression $I(\mathbf{x}, \mathbf{y} \mid \{p_i\}_{i=1}^n)$ which involves Gaussian kernel integration can be obtained. With the established relationship between the mutual information and the minimum mean square error (MMSE), one can obtain the optimal power allocation scheme [2].

Linear precoder can be utilized to increase the mutual information between the input and output with finite alphabet input, i.e., by allowing the transmitter and receiver to cooperate in communicating and decoding. By defining $\mathbf{H} = \mathrm{diag}(h_1, \cdots, h_n)$ as the channel matrix, $\mathbf{y} = [y_1, \cdots, y_n]^\mathrm{T}$ as the n outputs, $\mathbf{P} \in R^{n \times n}$ as the precoder which is designed, the mathematical model is

$$\mathbf{y} = \sqrt{\mathrm{snr}} \mathbf{H} \mathbf{P} \mathbf{x} + \mathbf{n}.\qquad(2)$$

To find the optimal precoder $\mathbf{P}^*$, the fixed point iteration $\mathbf{P}_{k+1} = \Pi\left(\mathbf{P}_k + \nu \mathrm{snr} \mathbf{H}^\mathrm{T} \mathbf{H} \mathbf{P}_k \Sigma_e\right)$ is adopted [3], where k denotes the iteration index, ν is a fixed small step size, Σ_e is a MMSE matrix and $\Pi(\cdot)$ is an operation that enforces the variable to have unite Frobenius norm. Note that both the MMSE matrix Σ_e and mutual information $I(\mathbf{x}, \mathbf{y} \mid \mathbf{P})$ are computed using Gaussian Hermite quadrature rules [13].

Power allocation under 1bit quantization

We first consider the case that power allocation is only adopted with the receiver using one-bit analog-to-digital (A/D) convertor. For the power allocation problem, model (1) reduces to be

$$y_i = \mathrm{sign}\left(\sqrt{\mathrm{snr}} h_i \sqrt{p_i} x_i + n_i\right), \quad i = 1, \cdots, n.\qquad(3)$$

The input-output mutual information on the ith channel is denoted as $I_i(p_i) = I_i(x_i; y_i)$. The capacity of the ith channel is the supreme of the mutual information taken over all possible input distributions

$$C_i = \sup_{q_i(x_i)} I_i(x_i; y_i),\qquad(4)$$

where $q_i(x_i)$ denotes the distribution of x_i. The problem is to determine both the optimal source distribution $\{q_i(x_i)\}_{i=1}^n$ and the power allocation $\{p_i\}_{i=1}^n$ so that the mutual information is maximized, i.e.,

$$(\{p_i^*\}_{i=1}^n, \{q_i(x_i)\}_{i=1}^n) = \arg\max_{\{p_i\}_{i=1}^n} \sup_{q_i(x_i)} \sum_{i=1}^n I_i(x_i; y_i)\qquad(5a)$$

$$\text{subject to } \sum_{i=1}^n p_i \leq 1, \quad p_i \geq 0, \forall i.\qquad(5b)$$

For the i th channel, it is an ideal discrete-time real baseband Additive White Gaussian Noise (AWGN) channel imposed by one-bit ADC. Given p_i, it can be shown that the objective function of (5) is separated with respect to i. For the i th channel, the optimal distribution $\{q_i(x_i)\}_{i=1}^n$ is established, and the capacity is also given [12]. The result is summarized in the following lemma.

Lemma 1. *[12] Given the power p_i of the i th channel, the capacity C_i is attained by BPSK signalling, and the corresponding capacity is*

$$C_i = 1 - h\left(Q\left(\sqrt{p_i h_i^2 \mathrm{snr}}\right)\right), \qquad (6)$$

where $h(\cdot)$ denotes the binary entropy function given as

$$h(x) = -x\log(x) - (1-x)\log(1-x) \qquad (7)$$

and $Q(\cdot)$ denotes the complementary Gaussian distribution function

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt. \qquad (8)$$

As a consequence, it can be shown that the original optimization problem (5) is equivalent to the following optimization problem

$$\underset{\{p_i\}_{i=1}^n}{\text{minimize}} \qquad \sum_{i=1}^n h\left(Q(\sqrt{p_i\gamma_i})\right) \quad (9a)$$

$$\text{subject to} \qquad \sum_{i=1}^n p_i \leq 1, \, p_i \geq 0, i = 1, \cdots, n, \qquad (9b)$$

where γ_i is defined as $\gamma_i = \mathrm{snr} h_i^2$. In [12], it has shown that $h\left(Q(\sqrt{z})\right)$ is strictly convex in z. Besides, the above constraint is a linear inequality. Thus, the power optimization problem is a convex optimization problem, which can be solved efficiently by numerical methods.

Obviously, the whole power should be allocated to all the channels. We may define $f(p_i) \triangleq h\left(Q(\sqrt{p_i\gamma_i})\right)$. Then the Karush-Kuhn-Tucker (KKT) conditions is used to characterize the optimal solutions. By introducing the Lagrange multipliers $v \in \square$ for the equality constraint $\sum_{i=1}^n p_i = 1$ and a multiplier $\lambda^* \in \square^n$ for the inequality constraint $p_i \geq 0, i = 1, \cdots, N$, the KKT conditions are

$$\begin{cases} \dfrac{\partial f(p_i^*)}{\partial p_i} + v^* - \lambda_i^* = 0, \forall i, (10a) \\[2mm] \sum_{i=1}^n p_i = 1, \qquad (10b) \\[2mm] p_i \geq 0, \quad \lambda_i^* p_i^* = 0, \forall i, \quad (10c) \\[2mm] \lambda^* \succeq 0. \qquad (10d) \end{cases} \xrightarrow[\text{variable } \lambda^*]{\text{eliminate the slack}} \begin{cases} \dfrac{\partial f(p_i^*)}{\partial p_i} + v^* \geq 0, \forall i, \qquad (11a) \\[2mm] \sum_{i=1}^n p_i = 1, \qquad (11b) \\[2mm] p_i \geq 0, \forall i, \qquad (11c) \\[2mm] \left(\dfrac{\partial f(p_i^*)}{\partial p_i} + v^*\right) p_i^* = 0, \forall i.(11d) \end{cases}$$

It follows by straightforward calculation that

$$g(p_i^*) \triangleq \frac{\partial f(p_i^*)}{\partial p_i} = \frac{\sqrt{\gamma_i}}{2\sqrt{2\pi p_i}} e^{-\frac{1}{2}p_i\gamma_i} \log\left(\frac{Q(\sqrt{p_i\gamma_i})}{1 - Q(\sqrt{p_i\gamma_i})}\right). \qquad (12)$$

We introduce a new variable $s_i = \sqrt{p_i\gamma_i}$, then $\kappa(s_i) \triangleq g(p_i)/\gamma_i$ and

$$\kappa(0) = \lim_{s_i \to 0} g(s_i) = -\frac{1}{\pi \ln 2}. \qquad (13)$$

Since $f(p_i)$ is strictly convex, $g(p_i)$ is strictly increasing with respect to p_i, thus $g(p_i) \geq -\frac{\gamma_i}{\pi \ln 2}$. If $v^* \geq \frac{\gamma_i}{\pi \ln 2}$, according to (11d), one has $p_i^* = 0$. If $v^* < \frac{\gamma_i}{\pi \ln 2}$, one must have $p_i^* > 0$ and the optimal nonzero power is $p_i^* = g^{-1}(-v^*)$. Otherwise if $p_i^* = 0$, the condition (11a) will be violated. In a nutshell, the optimal power allocation is

$$p_i^* = \begin{cases} \left(\kappa^{-1}(-\frac{v^*}{\gamma_i}) \right)^2 / \gamma_i, & \text{if} \quad v^* < \frac{\gamma_i}{\pi \ln 2}; \\ 0, & \text{otherwise,} \end{cases} \qquad (14)$$

where $\kappa^{-1}(\cdot)$ denotes the inverse function of $\kappa(\cdot)$. Substituting the above expression (14) into the condition (11a), one can solve v^* by bisection methods. Note that the above solution is similar to the power allocation problem in AWGN channel, which has a visualized interpretation named waterfilling. It can be seen that we may view $\pi \ln 2 / \gamma_i$ as the ground level above patch i, and then we pour the water into the region with a depth $\left(\kappa^{-1}(-v^*/\gamma_i) \right)^2 / \gamma_i + \pi \ln 2 / \gamma_i$, and the total amount of water is 1.

Precoding for one bit quantizer

In this section, a precoding scheme is adopted to improve the capacity of the parallel channels. We just consider the 2 parallel channel case. It has been shown that the BPSK signaling is optimal in the scenario of one-bit ADC. As a consequence, we assume that BPSK signalling is utilized.

For the precoding scheme, we use a 1bit/dimension vector quantizer. The output $\mathbf{z} = [z_1, z_2]^T$ of the vector quantizer is

$$\mathbf{z} = Q(\mathbf{y}), \qquad (15)$$

where $Q(\cdot)$ denotes the vector quantizer. The vector quantizer is a partion of $\square^2$ into 4 regions $R_i, i = 1, 2, 3, 4$, where the i th region is defined as

$$R_i = \{ \mathbf{y} \in \square^k : Q(\mathbf{y}) = \mathbf{z}_i \}. \qquad (16)$$

Next we will design the vector quantizer. We use the nearest neighbor rule and the square error $d(\mathbf{x}, \mathbf{y}) = \| \mathbf{x} - \mathbf{y} \|^2$ are adopted as the distortion measures. Then we obtain the nearest neighbor vector quantizer whose partions are

$$R_i = \{ \mathbf{x} : d(\mathbf{x}, \mathbf{y}_i) \leq d(\mathbf{x}, \mathbf{y}_j), \forall j \in J \}. \qquad (17)$$

We want to know the gain of the precoding scheme compared to the power allocation scheme. In fact, it is difficult to evaluate the mutual information $I(\mathbf{x}; \mathbf{z} | \mathbf{P})$ without knowing the precoder $\mathbf{P}$. We just use a greedy precoding scheme, i.e., we find a good precoder $\mathbf{P}^*$ by assuming no quantization, then the noiseless output is used as the represent points for the vector quantizer, and the nearest neighbor rule is adopted to implement the quantization. To find a good precoder, we use the the mutual information $I(\mathbf{x}; \mathbf{y} | \mathbf{P})$ as a criterion, and an optimal precoder can be found as given by [3].

For our numerical experiments, the constellation vectors are $\mathbf{s}_1 = [1,1]^T$, $\mathbf{s}_2 = [1,-1]^T$, $\mathbf{s}_3 = [-1,1]^T$, $\mathbf{s}_4 = [-1,-1]^T$, respectively. Since the number of constellation vectors are 4, the number of receiver constellations $\mathbf{z}$ are 4 assuming no noise, which are used as represent points for the vector quantizer. Thus we have

$$\mathbf{z}_i = \mathbf{H}\mathbf{P}\mathbf{s}_i. \qquad (18)$$

We now give a specific example to illustrate the above ideas. Suppose that the channel matrix is $\mathbf{H} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}$, and the optimal precoder $\mathbf{P}^* = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ for high SNR is [3]. With the optimal precoder $\mathbf{P}^*$, the noise-free received constellation vectors are $\mathbf{z}_1 = [\sqrt{6},0]^T$, $\mathbf{z}_2 = [-\sqrt{6},0]^T$, $\mathbf{z}_3 = [0,\sqrt{2}]^T$, $\mathbf{z}_4 = [0,-\sqrt{2}]^T$. Consequently, the corresponding quantizer regions are described in Figure 1.
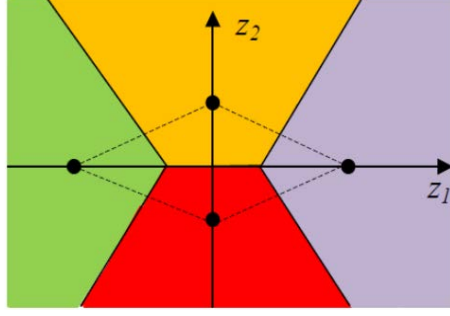
Fig.1 : For the given channel $\mathbf{H}$, the two dimensional 1bit/sample vector quantizer for the precoder $\mathbf{P}^*$ at $\text{snr}=0\text{dB}$. The black circles are the noiseless channel outputs, and the colors represent the quantization regions.

Once the quantizer is designed , we need to compute the mutual information $I(\mathbf{x};\mathbf{z}\,|\,\mathbf{P})$ given as

$$I(\mathbf{x};\mathbf{z}\,|\,\mathbf{P})=H(\mathbf{z})-H(\mathbf{z}\,|\,\mathbf{x}), \tag{19}$$

where $H(\mathbf{z})=-\sum_{j=1}^{4}p(z_{j})\log p(z_{j})$ bit, $H(\mathbf{z}\,|\,\mathbf{x})$ can be computed as

$$H(\mathbf{z}\,|\,\mathbf{x})=-\sum_{i=1}^{4}\sum_{j=1}^{4}\Big(p(z_{i}\,|\,x_{j})p(x_{j})\log p(z_{i}\,|\,x_{j})\Big). \tag{20}$$

We will use Monte Carlo simulation to evaluate $p(z_{i}\,|\,x_{j})$. Meanwhile, the mutual information $I(\mathbf{x};\mathbf{y}\,|\,\mathbf{P})$ based on unquantized measurements are calculated by Gaussian Hermite quadrature methods.

Numerical Simulation

In this section, two numerical simulations are performed. The channel vectors of the first and the second simulation are $[1.4,0.2]^{\mathrm{T}}$ and $[\sqrt{3/2},\sqrt{1/2}]^{\mathrm{T}}$, respectively. Besides, BPSK signalling is always adopted. The first simulation shows that when the difference of the channel gains are large and the SNR is in medium-to-high regions, significant improvement can be obtained with precoding scheme and vector quantizer, compared to the power allocation scheme and scalar quantizer. The second simulation shows that when the difference of the channel gains are small, the improvement is indeed small.

From Figure 2(left) of the first simulation, it can seen that the precoder scheme works significantly better than the power allocation scheme at SNR 5~16dB when no quantization is used. As the SNR increases (SNR≥2dB), the precoder scheme with 1bit vector quantizer works better than the power allocation scheme in the case of no quantization. In addition, the power allocation scheme with 1bit quantization works better than the uniform power allocation scheme at -10~5dB. For the second simulation where Figure 2(right) is generated, it can be seen that the gain of precoder scheme and power allocation scheme compared to the uniform power allocation scheme is small.
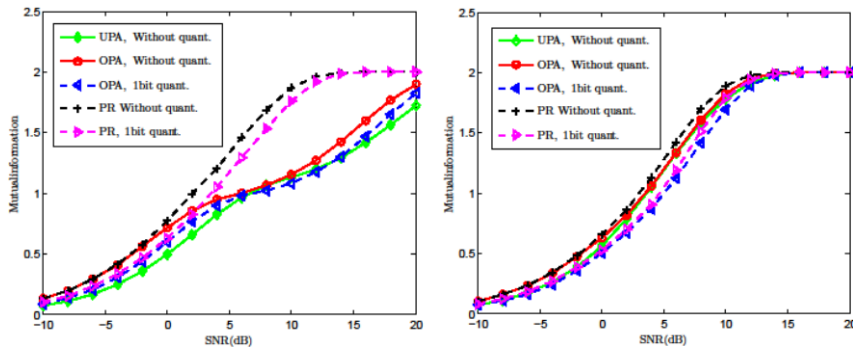

Fig.2 : The first and second simulation, the mutual information of two independent Gaussian channels versus SNR. Note that UPA, OPA and PR denote the uniform power allocation, optimal power allocation and precoder design scheme, respectively.

Conclusion

This paper investigated the power allocation scheme for parallel independent channels under 1bit quantization. The optimal power allocation can be found by bisection methods, and has a good interpretation named waterfilling. For BPSK signalling, we also give a heuristic precoder scheme and 1bit/dimension vector quantizer for 2 parallel independent channels. Simulation results show that significant gains can be achieved for the heuristic precoder scheme compared to power allocation method when the difference of the channel gains is large and the SNR is not high.

References

[1] E. I. Telatar, "Capacity of multi-attenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585-595, Nov. 1999.

[2] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitary input distributions", *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033-3051, Jul. 2006.

[3] F. Pérez-Cruz, M. R. D. Rodrigues, and S. Verdú, "Optimal precoding for digital suscriber lines", in *Proc. Int. Conf. Commun.*, Beijing, China, May 2008, pp. 1200-1204.

[4] Q. Wang, Q. Xie, Z. Wang, S. Chen, and L. Hanzo," A universal low-complexity symbol-to-bit soft demapper", *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 119-130, Jan. 2014.

[5] Q. Wang, Z. Wang, S. Chen, and L. Hanzo, "Enhancing the decoding performance of optical wireless communication systems using receiver-side predistortion", *Optics Express*, vol. 21, no. 25, pp. 30295-30395, Dec. 2013.

[6] R. H. Walden,"Analog-to-digital converter survey and analysis", *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539-550, Apr. 1999.

[7] M. Stein, S. Theiler and J. A. Nossek, "Overdemodulation for high performance receivers with low-resolution ADC," Avaliable at http://arxiv.org/abs/1409.1437.

[8] M. Stein, F. Wendler, A. Mezghani, J. A. Nossek, "Quantization-loss reduction for signal parameter estimation," *Proc. IEEE Intern. Conf. Acoust., Speech and Signal Process.*, 2013, pp. 5800-5804.

[9] S. Krone, G. Fettweis, "Capacity of communications channels with 1-bit quantization and oversampling at the receiver," *IEEE Sarnoff Symp.(SARNOFF)*, May 2012.

[10] A. Host-Madsen, P. Handel, "Effects of sampling and quantization on single-tone frequency estimation", *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 650-662, 2000.

[11] O. Dabeer and U. Madhow, "Channel estimation with low-precision analog-to-digital conversion", *2010 IEEE Intern. Conf. Commun.*, pp. 1-6, May 2010.

[12] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver", *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629-3639, Dec. 2009.

[13] C. Xiao and Y. R. Zheng, "On the mutual information and power allocation for vector Gaussian channels with finite discrete inouts", *IEEE Global Communications Conference*, pp. 1-5, New Orleans, LO, Nov. 30-Dec. 4, 2008.