On the Optimal Power Allocation for Additive Color Noise Parallel Channels with Limited Access Constraint

Shih-Wei Wang, Po-Ning Chen, Chung-Hsuan Wang, and Wen-Chieh Chang
Dept. of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan 30010, R.O.C.
Emails: shihwei.wang@ieee.org, poning@faculty.nctu.edu.tw, chwang@mail.nctu.edu.tw, wenchieh.cm98g@nctu.edu.tw

***Abstract*—In this paper, we consider an (N, K)-limited access system consisting of N parallel additive noise channels with spatial dependency, where the receiver starts to decode the information being transmitted when at least K out of N channel outputs are received. We investigate the optimal power allocation that maximizes the minimum mutual information among all possible cases of partial reception. A universal guideline is then obtained for a group of permutation-invariant channels, in which the system mutual information remains unchanged when permuting the parameters that characterize the partial reception and signal-to-noise power ratio (SNR) of channels, that a channel with less noise power should have larger SNR. When all N channels belong to a permutation-invariant group, we also have that the optimal power allocation problem can be transformed to an equivalent problem for K parallel channels without limited access constraint via a water-filling noise-power-redistribution process. The merit of this transformation can be more evidently seen when the channel input-noise pairs are reduced to be spatially independent with distributions scaled from a common random vector, for which the optimal power allocation solution can be simply obtained by a two-phase water-filling process.**

I. Motivation and System Model

Power allocation over multiple access channels is a classical research problem in information theory, and the most known result in the literature is perhaps the water-filling scheme that maximizes the capacity of parallel Gaussian channels by using Gaussian inputs [4]. It has recently been generalized to a mercury/water-filling scheme when the inputs are additionally constrained to be symbols of a discrete modulation such as BPSK, QPSK and QAM, which is referred to as *arbitrary inputs* in [9]. Based on this new result, the optimal power allocation respectively for multi-user orthogonal frequency-division multiplexing (OFDM) channels [10] and multiple-input-multiple-output (MIMO) channels [12] with arbitrary inputs constraint is subsequently obtained.

Although Gaussians are generally appropriate noise models for physical additive channels, experimental measurement indicates that the noises in certain environments are by no means Gaussian distributed [2], [15], [16]. In some situations, dependence among channels is even suggested by measurements [5], [11]. As such, in this paper we consider a system that consists of N additive color noise parallel channels with

This work was supported by the National Science Council of Taiwan under Grants NSC 98-2221-E-009-060-MY3 and NSC 99-2221-E-009-076-MY3.

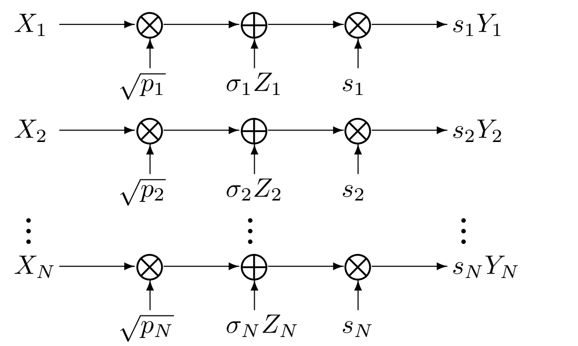

Fig. 1. System model for an (N, K)-limited access additive color noise channel with $E[|X_i|^2] = E[|Z_i|^2] = 1$, $s_i \in \{0, 1\}$ for $1 \le i \le N$, $\sum_{i=1}^N s_i \ge K$, and $\sum_{i=1}^N p_i \le P$.

maximal total transmission power P, as shown in Fig. 1, in which the channel inputs and outputs are characterized by

$$Y_i = \sqrt{p_i} X_i + \sigma_i Z_i \quad \text{for } 1 \le i \le N \tag{1}$$

where Y_i is the channel output, X_i and Z_i are respectively the channel input and additive noise with $E[|X_i|^2] = E[|Z_i|^2] = 1$, and p_i and σ_i^2 are parameters specifying the transmission and noise powers . It is assumed that channel inputs $\{X_i\}_{i=1}^N$ and noises $\{Z_i\}_{i=1}^N$ are independent of each other, but now statistical dependency could exist among channel inputs and also among channel noises.

When all the channel outputs can be perfectly received, the power allocation problem with respect to the system defined in (1) might be solved by whitening technique. However, suppose only a portion of channel outputs can be successfully received without providing the statistical distribution of the accessible channels, and the receiver begins to recover the transmitted information only at least K out of N channel outputs are available. Additional care should be taken in order to provide an optimal power allocation solution since the transmitter does not know whcih channels are accessible in advance. We emulate the situation by a binary process $\{s_i\}_{i=1}^N$, where the ith channel output is blocked when multiplied by $s_i = 0$, and remains accessible when the multiplicative constant s_i is equal to 1 as depicted in Fig. 1. One straightforward scenario for such a binary channel-state model is a packet-switched network, where packets may be lost during transmission [1]. In certain cases, the receiver may still be required to recover

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

the transmitted information from these partial receptions [8], [13], [14].

Since the receiver begins to recover the information only when at least K out of N parallel channels are accessible,[1] our problem becomes to maximize the mutual information of a compound channel [7]. Involved with the channel dependency, it seems difficult to provide a general principle for the power allocation over such a compound channel. Even back to the case of the additive white Gaussian noise channels with arbitrary inputs considered in [9], which corresponds to $K = N$ and independent Gaussian $\{Z_i\}_{i=1}^N$ in our setting, the optimal power allocated to a noisier channel (i.e., with a larger σ_i^2) may not be always smaller (cf., for example, Fig. 5 in [9]). Hence, a general rule regarding the optimal solution appears to be somewhat unlikely. We however consider a group of permutation-invariant channels, in which the system mutual information remains unchanged when simultaneously permuting their power parameters and multiplicative constants (i.e., switching (p_i, σ_i^2) and (p_j, σ_j^2), also s_i and s_j, for channels i and j). A universal guideline for power allocation over the (N, K)-limited access channel that suffers additive color noises can then be addressed in that the optimal signal-to-noise power ratio (SNR) cannot be smaller for a less noisy channel (cf. Theorem 1). This general guideline can be used to supplement what has been demonstrated in [9] that even though a larger power is optimally allocated to a noisier channel within a group of the same arbitrary input such as QPSK, the resulting SNRs are always monotonically nonincreasing with respect to the degree of channel noisiness.

In addition, if all N parallel channels belong to the same permutation-invariant group, then at least $(N-K+1)$ channels achieve the maximal SNR when maximizing the system mutual information (cf. (6) in Theorem 1). Accordingly, the power allocation problem can be solved in two phases. In the first phase, the maximal SNR channels are directly identified based on their degrees of noisiness via a *water-filling noise-power-redistribution* procedure, which reduces the problem to an equivalent one over a (non-compound) multiple access system with K parallel channels but no limited access constraint. In the second phase, the optimal power allocation is obtained via a one-to-one mapping from the power allocation solution of the K equivalent channels according to the Karuch-Kuhn-Tucker (KKT) condition [3]. The merit of this approach can be more clearly seen (cf. Corollary 1) when the channel input-noise pairs are reduced to be spatially independent with distributions scaled from a common random vector, for which the optimal power allocation solution can be simply obtained from a *two-phase water-filling* procedure. Details will be given in the following sections.

Throughout the entire paper, we assume without loss of generality that

$$\sigma_1^2 \ge \sigma_2^2 \ge \cdots \ge \sigma_N^2. \tag{2}$$

[1]Note that our concern in this paper is the maximal rate that yields a vanishing decoding error given $\sum_{i=1}^N s_i \ge K$ for some positive K that is known to the transmitter, not the decoding error with respect to an unconstrained $\boldsymbol{s}$ (namely, $K = 0$) or a statistically distributed $\boldsymbol{s}$.

II. Analysis of the Optimal Power Allocation

Based on our system model, the optimal power allocation $\boldsymbol{p}^* \triangleq (p_1^*, p_2^*, \ldots, p_N^*)$ can be determined by solving:

$$\max_{\sum_{i=1}^N p_i \le P} \ \min_{\sum_{i=1}^N s_i \ge K} I(\sqrt{\boldsymbol{p}} \circ \boldsymbol{X}; \boldsymbol{s} \circ \boldsymbol{Y}) \tag{3}$$

where $\sqrt{\boldsymbol{p}} \triangleq (\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_N})$, $\boldsymbol{s} \triangleq (s_1, s_2, \ldots, s_N)$, $\boldsymbol{X} \triangleq (X_1, X_2, \ldots, X_N)$, $\boldsymbol{Y} \triangleq (Y_1, Y_2, \ldots, Y_N)$, $I(\cdot;\cdot)$ is the mutual information function, and operator "$\circ$" denotes the Hadamard product [6]. We can reformulate (3) as

$$\max_{\sum_{i=1}^N p_i \le P} \ \min_{\sum_{i=1}^N s_i \ge K} I(\sqrt{\boldsymbol{\gamma}} \circ \boldsymbol{X}; \boldsymbol{s} \circ \tilde{\boldsymbol{Y}}) \tag{4}$$

where $\tilde{\boldsymbol{Y}} \triangleq (Y_1/\sigma_1, Y_2/\sigma_2, \ldots, Y_N/\sigma_N)$, and $\boldsymbol{\gamma} \triangleq (\gamma_1, \gamma_2, \ldots, \gamma_N) = (p_1/\sigma_1^2, p_2/\sigma_2^2, \ldots, p_N/\sigma_N^2)$ is the corresponding SNR vector.

It is obvious that $I(\sqrt{\boldsymbol{\gamma}} \circ \boldsymbol{X}; \boldsymbol{s} \circ \tilde{\boldsymbol{Y}})$ is nondecreasing with respect to $\sum_{i=1}^N s_i$; hence, the internal minimization in (4) can be achieved by considering only those $\boldsymbol{s}$ satisfying $\sum_{i=1}^N s_i = K$. Similarly, as the mutual information function is anticipatedly nondecreasing with respect to P, we can further replace the inequality condition of the external maximization by its equality counterpart. However, to solve (4), we require an additional assumption below, which is usually satisfied for channels of general interest such as the additive noises are circularly symmetric dependent Gaussians.

Assumption 1: For any $\boldsymbol{s}$ satisfying $\sum_{i=1}^N s_i = K$, mutual information $I(\sqrt{\boldsymbol{p}} \circ \boldsymbol{X}; \boldsymbol{s} \circ \boldsymbol{Y})$ is a continuous, strictly increasing concave function in $\boldsymbol{s} \circ \boldsymbol{p}$, and its first derivation with respect to p_i exists and is a continuous, strictly decreasing function in p_i for every i with $s_i = 1$.

By this assumption, we are ready to present the first theorem in this paper.

Theorem 1: Let $U_{i,j}$ be the $N \times N$ permutation matrix that exchanges components i and j when applied to a vector. If for a fixed (i, j) pair with $i < j$ and for every $\boldsymbol{s}$ satisfying $\sum_{k=1}^N s_k = K$,

$$\begin{aligned} &I\big(\sqrt{\boldsymbol{\gamma}}\, U_{i,j} \circ \boldsymbol{X}; \boldsymbol{s} U_{i,j} \circ (\sqrt{\boldsymbol{\gamma}}\, U_{i,j} \circ \boldsymbol{X} + \boldsymbol{Z})\big) \\ &\qquad = I\big(\sqrt{\boldsymbol{\gamma}} \circ \boldsymbol{X}; \boldsymbol{s} \circ (\sqrt{\boldsymbol{\gamma}} \circ \boldsymbol{X} + \boldsymbol{Z})\big) \text{ for every } \boldsymbol{\gamma}, \end{aligned} \tag{5}$$

then the SNRs γ_i^* and γ_j^* corresponding to the optimal power allocation satisfy

$$\gamma_i^* \le \gamma_j^*.$$

Moreover, if (5) holds for every (i, j) pair with $1 \le i < j \le N$ and for every $\boldsymbol{s}$ satisfying $\sum_{k=1}^N s_k = K$, then

$$\gamma_1^* \le \gamma_2^* \le \cdots \le \gamma_K^* = \gamma_{K+1}^* = \cdots = \gamma_N^*. \tag{6}$$

It can be concluded from Theorem 1 that, although the optimal allocated powers may be larger or smaller for noisier channels, their SNRs remain monotonic. Two examples are then provided to demonstrate this phenomenon: one with independent Gaussian noises (and $K = N$), and the other with dependent Gaussian noises (and $K < N$).

Example 1: Suppose $\{X_i\}_{i=1}^N$ are independent and identically distributed (i.i.d.) BPSK inputs with uniform marginal distribution, and $\{Z_i\}_{i=1}^N$ are i.i.d. real-valued Gaussian distributed. Let $N = K = 2$ and $(\sigma_1^2,\ \sigma_2^2) = (5,\ 2)$. Then (5) clearly holds for $\boldsymbol{s} = (1, 1)$ and for every i, j with $1 \le i < j \le 2$. By $\boldsymbol{s} = (1, 1)$, (4) is reduced to a maximization problem with a single power-sum constraint; hence, we can apply the Lagrange multiplier technique to find that the optimal allocated powers satisfy $p_1^* = 1.077 < p_2^* = 2.923$ when $P = 4$, but they become $p_1^* = 5.141 > p_2^* = 4.859$ as P increases to 10. It can be verified that $\gamma_1^* < \gamma_2^*$ in both cases, as indicated by Theorem 1. □

Example 2: In this example, $\{X_i\}_{i=1}^N$ remain as uniform i.i.d. BPSK inputs, but $\{Z_i\}_{i=1}^N$ are changed to real-valued dependent Gaussians satisfying $E[Z_iZ_j] = 0.6$ for every $i \ne j$. Take $(N, K) = (3, 2)$ and $(\sigma_1^2,\ \sigma_2^2,\ \sigma_3^2) = (5,\ 1.5,\ 0.5)$. Again, it can be verified that (5) holds for every $\boldsymbol{s}$ with $\sum_{k=1}^3 s_k = 2$ and for every i, j with $1 \le i < j \le 3$. We can then apply Theorem 3, introduced later, to obtain $(p_1^*,\ p_2^*,\ p_3^*) = (1.432,\ 1.926,\ 0.642)$ for $P = 4$ and $(p_1^*,\ p_2^*,\ p_3^*) = (5.456,\ 3.408,\ 1.136)$ for $P = 10$. Consequently, $(\gamma_1^*,\ \gamma_2^*,\ \gamma_3^*)$ is equal to $(0.286,\ 1.308,\ 1.308)$ and $(1.091,\ 2.272,\ 2.272)$ for $P = 4$ and $P = 10$, respectively. The SNRs of the last two channels are thus equal and maximal among the three channels, which conform to Theorem 1. □

With Theorem 1, we know that if condition (5) is valid for every (i, j) pair with $1 \le i < j \le N$ and for every $\boldsymbol{s}$ satisfying $\sum_{i=1}^N s_i = K$, then assumption (2) implies the validity of (6), which in turns implies that the condition below is valid for exactly one value of ℓ in $\{1, 2, \cdots, K\}$:

$$\max_{1\le i<\ell} \gamma_i^* < \gamma_\ell^* = \gamma_{\ell+1}^* = \cdots = \gamma_N^*. \tag{7}$$

In case the optimal ℓ^* that validates (7) is *a priori* known, the max-min problem in (4) can be equivalently simplified to

$$\max_{\boldsymbol{p}\in\mathcal{P}_{\ell^*}} I(\sqrt{\boldsymbol{\gamma}} \circ \boldsymbol{X}; \boldsymbol{s}^* \circ \tilde{\boldsymbol{Y}}) \tag{8}$$

where $s_i^* = 1$ for $1 \le i \le K$, and 0, otherwise, and

$$\mathcal{P}_{\ell^*} \triangleq \left\{ \boldsymbol{p} \in \Re^N : \begin{array}{l} \sum_{i=1}^N p_i = P \text{ with } p_i \ge 0 \\ p_i/\sigma_i^2 < p_N/\sigma_N^2 \text{ for } 1 \le i < \ell^* \\ p_i/\sigma_i^2 = p_N/\sigma_N^2 \text{ for } \ell^* \le i \le N \end{array} \right\}. \tag{9}$$

In other words, the optimizer for (8) is exactly the optimal allocated power $\boldsymbol{p}^*$ for (4). We then notice that $I(\sqrt{\boldsymbol{\gamma}} \circ \boldsymbol{X}; \boldsymbol{s}^* \circ \tilde{\boldsymbol{Y}})$ is only a function of $p_1, p_2, \ldots, p_K$ because the last $(N - K)$ elements of $\boldsymbol{s}^*$ are all equal to zero; hence, we can for convenience rewrite (8) as

$$\max_{\boldsymbol{p}\in\mathcal{P}_{\ell^*}} g\left(\frac{p_1}{\sigma_1^2}, \frac{p_2}{\sigma_2^2}, \cdots, \frac{p_K}{\sigma_K^2}\right) \tag{10}$$

where

$$g\left(\frac{p_1}{\sigma_1^2}, \frac{p_2}{\sigma_2^2}, \cdots, \frac{p_K}{\sigma_K^2}\right) \triangleq I(\sqrt{\boldsymbol{\gamma}} \circ \boldsymbol{X}; \boldsymbol{s}^* \circ \tilde{\boldsymbol{Y}}). \tag{11}$$

Note that (10) is derived based on the presumed knowledge of the optimal ℓ^*. Since ℓ^* is defined by the optimal SNRs (cf. (7)), it seems illogical to assume ℓ^* is known without knowing $\boldsymbol{p}^*$ in advance. Therefore, some preliminary work must be done in order to argue that ℓ^* can actually be identified directly using a water-filling principle (cf. Theorem 3).

Consider

$$\sup_{\boldsymbol{p}\in\mathcal{P}_m} g\left(\frac{p_1}{\sigma_1^2}, \frac{p_2}{\sigma_2^2}, \cdots, \frac{p_K}{\sigma_K^2}\right) \tag{12}$$

where $\mathcal{P}_m$ is defined the same as (9) except that ℓ^* is replaced with m. Here, we use "supremum" instead of "maximum" because, unlike $\boldsymbol{p}^*$ is guaranteed to be in $\mathcal{P}_{\ell^*}$ by its definition from (7), the optimizer of (12) may be outside of $\mathcal{P}_m$ since m is generally not equal to ℓ^*. Then the next lemma shows that (12) can be equivalently transformed to a system with K parallel channels, of which the effective noise variances are $\{\tilde{\sigma}_{m,i}^2\}_{i=1}^K$.

Lemma 1: Define

$$\tilde{\sigma}_{m,i}^2 = \begin{cases} \sigma_i^2 & \text{for } 1 \le i < m \\ \frac{1}{(K-m+1)} \sum_{j=m}^N \sigma_j^2 & \text{for } m \le i \le K. \end{cases} \tag{13}$$

Then,

$$\sup_{\boldsymbol{p}\in\mathcal{P}_m} g\left(\frac{p_1}{\sigma_1^2}, \frac{p_2}{\sigma_2^2}, \cdots, \frac{p_K}{\sigma_K^2}\right) = \sup_{\boldsymbol{q}\in\mathcal{Q}_m} g\left(\frac{q_1}{\tilde{\sigma}_{m,1}^2}, \frac{q_2}{\tilde{\sigma}_{m,2}^2}, \cdots, \frac{q_K}{\tilde{\sigma}_{m,K}^2}\right) \tag{14}$$

where

$$\mathcal{Q}_m \triangleq \left\{ \boldsymbol{q} \in \Re^K : \begin{array}{l} \sum_{i=1}^K q_i = P \text{ with } q_i \ge 0 \\ q_i/\tilde{\sigma}_{m,i}^2 < q_K/\tilde{\sigma}_{m,K}^2 \text{ for } 1 \le i < m \\ q_i/\tilde{\sigma}_{m,i}^2 = q_K/\tilde{\sigma}_{m,K}^2 \text{ for } m \le i \le K \end{array} \right\}.$$

In addition, there exists $\boldsymbol{q}_m^\diamond \triangleq (q_{m,1}^\diamond, q_{m,2}^\diamond, \ldots, q_{m,K}^\diamond)$ satisfying

$$\sup_{\boldsymbol{q}\in\mathcal{Q}_m} g\left(\frac{q_1}{\tilde{\sigma}_{m,1}^2}, \frac{q_2}{\tilde{\sigma}_{m,2}^2}, \cdots, \frac{q_K}{\tilde{\sigma}_{m,K}^2}\right) = g\left(\frac{q_{m,1}^\diamond}{\tilde{\sigma}_{m,1}^2}, \frac{q_{m,2}^\diamond}{\tilde{\sigma}_{m,2}^2}, \cdots, \frac{q_{m,K}^\diamond}{\tilde{\sigma}_{m,K}^2}\right) \tag{15}$$

if, and only if, there exists $\boldsymbol{p}_m^\diamond \triangleq (p_{m,1}^\diamond, p_{m,2}^\diamond, \ldots, p_{m,N}^\diamond)$ satisfying

$$\sup_{\boldsymbol{p}\in\mathcal{P}_m} g\left(\frac{p_1}{\sigma_1^2}, \frac{p_2}{\sigma_2^2}, \cdots, \frac{p_K}{\sigma_K^2}\right) = g\left(\frac{p_{m,1}^\diamond}{\sigma_1^2}, \frac{p_{m,2}^\diamond}{\sigma_2^2}, \cdots, \frac{p_{m,K}^\diamond}{\sigma_K^2}\right) \tag{16}$$

where

$$p_{m,i}^\diamond = \begin{cases} q_{m,i}^\diamond & \text{for } 1 \le i < m \\ \frac{\sigma_i^2/\tilde{\sigma}_{m,K}^2}{(K-m+1)} \sum_{j=m}^K q_{m,j}^\diamond & \text{for } m \le i \le N. \end{cases} \tag{17}$$

Next, we simplify $\mathcal{Q}_m$ by removing the last two constraints so that only the power-sum constraint is left.

Theorem 2: Let $\tilde{\boldsymbol{q}}_m \triangleq (\tilde{q}_{m,1}, \tilde{q}_{m,2}, \ldots, \tilde{q}_{m,K})$ be the maximizer of

$$\max_{\boldsymbol{q}\in\tilde{\mathcal{Q}}} g\left(\frac{q_1}{\tilde{\sigma}_{m,1}^2}, \frac{q_2}{\tilde{\sigma}_{m,2}^2}, \cdots, \frac{q_K}{\tilde{\sigma}_{m,K}^2}\right) \tag{18}$$

with

$$\tilde{\mathcal{Q}} \triangleq \left\{ \boldsymbol{q} \in \Re^K : \sum_{i=1}^K q_i = P \text{ with } q_i \ge 0 \right\}.$$

Then if $\boldsymbol{q}_m^\diamond$ satisfying (15) *lies in $\mathcal{Q}_m$, then $\boldsymbol{q}_m^\diamond = \tilde{\boldsymbol{q}}_m$.*

In light of Theorem 2 and the fact that $\boldsymbol{q}_{\ell^*}^\diamond$ validating (15) lies in $\mathcal{Q}_{\ell^*}$ (since $\boldsymbol{p}^* = \boldsymbol{p}_{\ell^*}^\diamond$ that validates (16) with $m = \ell^*$ is guaranteed to lie in $\mathcal{P}_{\ell^*}$), we have $\boldsymbol{q}_{\ell^*}^\diamond = \tilde{\boldsymbol{q}}_{\ell^*}$. Thus, the desired optimizer $\boldsymbol{q}_{\ell^*}^\diamond$ must be one of $\tilde{\boldsymbol{q}}_1$, $\tilde{\boldsymbol{q}}_2$, ..., $\tilde{\boldsymbol{q}}_K$; hence, under a single power-sum constraint, we can solve (18) using the Lagrange multiplier technique for each $1 \le m \le K$, and compare their resulting mutual information. The optimal $\boldsymbol{q}_{\ell^*}^\diamond$, as well as the optimal ℓ^*, can then be identified, which immediately gives $\boldsymbol{p}^* = \boldsymbol{p}_{\ell^*}^\diamond$ through (17).

Along this idea, we can subsequently prove that the optimal ℓ^* can actually be determined directly through a *water-filling* procedure without performing the aforementioned K maximizations as summarized in the next theorem. Hence, the optimal power allocation for an (N, K)-limited access additive color noise channel satisfying (5) for every $1 \le i < j \le N$ and for every $\boldsymbol{s}$ with $\sum_{i=1}^N s_i = K$ can be obtained by solving (18) by directly assigning $m = \ell^*$.

Theorem 3: The optimal ℓ^* in the sense of (7) *is given by*

$$\ell^* \triangleq \min\left\{ i \,\middle|\, 1 \le i \le K \text{ and } \sigma_i^2 \le \hat{\sigma}_K^2 \right\} \tag{19}$$

where $\hat{\sigma}_i^2 \triangleq \sigma_i^2 + [\lambda - \sigma_i^2]^+$ for $1 \le i \le K$ with λ chosen to satisfy $\sum_{i=1}^K [\lambda - \sigma_i^2]^+ = \sum_{i=K+1}^N \sigma_i^2$ and $[y]^+ \triangleq \max\{0, y\}$. The optimal power allocation $\boldsymbol{p}^*$ can therefore be obtained from $\tilde{\boldsymbol{q}}_{\ell^*}$ (equivalently, $\boldsymbol{q}_{\ell^*}^\diamond$) through (17).

Theorem 3 immediately gives that the optimal ℓ^* can be determined via a water-filling noise-power-redistribution procedure. As an example in subfigures (a) and (b) in Fig. 2, the noise powers of channels 5 and 6 are poured as "noise water" into a tank with $K = 4$ unit-width vessels, each of which has base height σ_i^2 for $1 \le i \le 4$. By noting that the resulting noise water level is strictly below σ_2^2, we determine $\ell^* = 3$. Then the optimal power allocation for channels 3, 4, 5 and 6 should all achieve the maximal SNR; as a result, their allocated powers will be proportional to the respective noise variances as marked by the vertical dotted lines in Fig. 2(f).

After determining $\ell^* = 3$, Fig. 2(b) also gives $\tilde{\sigma}_{3,i}^2 = \hat{\sigma}_i^2$ for $1 \le i \le K$. Then, the original problem is reduced to one that solves (18) with $m = \ell^* = 3$, i.e., a power allocation problem over K parallel channels without limited access constraint, as illustrated in Fig 2(c).

Under a single power-sum constraint, the KKT condition for (18) with $m = \ell^*$ implies that for $1 \le i \le K$,

$$\frac{\partial}{\partial q_i} g\left(\frac{q_1}{\hat{\sigma}_1^2}, \frac{q_2}{\hat{\sigma}_2^2}, \cdots, \frac{q_K}{\hat{\sigma}_K^2}\right)\Bigg|_{\boldsymbol{q}=\tilde{\boldsymbol{q}}_{\ell^*}} \begin{cases} = \nu & \text{if } \tilde{q}_{\ell^*,i} > 0 \\ \le \nu & \text{if } \tilde{q}_{\ell^*,i} = 0 \end{cases} \tag{20}$$

where ν is the Lagrange multiplier chosen to satisfy the power-sum constraint [3]. Without further knowledge of function g, (20) can only be solved numerically in a case-by-case fashion.

In a special case where function g is additive itself, i.e.,

$$g\left(\frac{q_1}{\hat{\sigma}_1^2}, \frac{q_2}{\hat{\sigma}_2^2}, \cdots, \frac{q_K}{\hat{\sigma}_K^2}\right) = \sum_{i=1}^K g\left(\frac{q_i}{\hat{\sigma}_i^2}\right),$$

which, e.g., holds when $\{(X_i, Z_i)\}_{i=1}^N$ are i.i.d., the optimal $\tilde{\boldsymbol{q}}_{\ell^*}$ can also be determined by water-filling. Here, we abuse the notation to reuse g as the component function for convenience. The result is summarized in Corollary 1.

Corollary 1: When function g is additive, the optimal power allocation $\tilde{\boldsymbol{q}}_{\ell^*}$ that verifies (20) *satisfies that for $1 \le i \le K$,*

$$\tilde{q}_{\ell^*,i} = \begin{cases} \hat{\sigma}_i^2 \cdot g'^{(inv)}(\nu\hat{\sigma}_i^2) & \text{if } g'(\infty) < \nu\hat{\sigma}_i^2 < g'(0) \\ 0 & \text{if } \nu\hat{\sigma}_i^2 \ge g'(0) \end{cases} \tag{21}$$

where $g'^{(inv)}$ is the inverse function of the first derivative g' of function g.[2]

With Corollary 1, we remark that the optimal power allocation $\boldsymbol{p}^*$ can be obtained from $\tilde{\boldsymbol{q}}_{\ell^*}$ through a similar relation to (17). Furthermore, additivity of function g implies that the resulting $\tilde{\boldsymbol{q}}_{\ell^*}$ can be graphically interpreted through a variation of water-filling procedure. Continuing the example in Fig. 2(d), we first adjust the base heights of the four vessels according to function g'. The new heights $\{L_i(\nu)\}_{i=1}^K$ are given by

$$L_i(\nu) \triangleq \hat{\sigma}_i^2 \cdot G(\nu\hat{\sigma}_i^2) \quad \text{for } 1 \le i \le K$$

where

$$G(\zeta) \triangleq \begin{cases} \frac{1}{\zeta} - g'^{(\mathrm{inv})}(\zeta) & \text{if } g'(\infty) < \zeta < g'(0) \\ \frac{1}{g'(0)} & \text{if } \zeta \ge g'(0). \end{cases} \tag{22}$$

Notably, the new heights of some channels are higher, while some can be lower, than the original base levels $\{\hat{\sigma}_i^2\}_{i=1}^K$, as indicated in Fig. 2(d). It is however noteworthy when $\{Z_i\}_{i=1}^N$ are Gaussians, it is guaranteed that $L_i(\nu) \ge \hat{\sigma}_i^2$ for $1 \le i \le K$; hence, a *mercury-filling* scheme before water pouring has been proposed to materialize the lifting of vessel bases of all channels [9]. In such a case, (22) returns to what has been defined and identically denoted in [9, Eq. (43)]. The pre-adjustment of base heights before filling water can be regarded as a preparation for these vessels to be capable of supporting the water that will be poured in.

After this adjustment of base heights, the next step is the usual water-filling power allocation as shown in Fig. 2(e). We thus obtain $\tilde{\boldsymbol{q}}_{\ell^*}$. The last step, according to the equal-SNR requirement (or equivalently, by (17)), is to subdivide the last $(K - \ell^* + 1) = 2$ vessels into $(N - \ell^* + 1) = 4$ vessels with their widths proportional to σ_i^2 for $\ell^* \le i \le N$. Then the amount of poured water inside each of the six vessels is the optimal transmission power. This is concluded in Fig. 2(f). Details of Fig. 2 are given in the next example.

Example 3: Figure 2 corresponds to the two-phase water-filling process for a $(6, 4)$-limited access additive noise channel with uniform i.i.d. QPSK inputs $\{X_i\}_{i=1}^6$ and i.i.d. complex Laplacian noises $\{Z_i\}_{i=1}^6$. Consider $P = 15$ and $(\sigma_1^2, \sigma_2^2, \ldots, \sigma_6^2) = (12, 10, 4, 2, 2, 2)$. Following Theorem 3, we obtain

[2]For notational convenience, we denote $g'(\infty) \triangleq \lim_{\rho\uparrow\infty} g'(\rho)$ and $g'(0) \triangleq \lim_{\rho\downarrow 0} g'(\rho)$ in (21), for which the existences of these limits are guaranteed by Assumption 1. From (21), a natural question to ask is, what if $0 < \nu\hat{\sigma}_i^2 \le g'(\infty)$. A quick answer to this question is that for most channels of practical interest, such as channels with finite input alphabet, $g'(\infty) = 0$. A more precise answer is that $0 < \nu\hat{\sigma}_i^2 \le g'(\infty)$ can never occur even if $g'(\infty) > 0$ because the KKT condition in (20) ensures that $g'(\tilde{q}_{\ell^*,i}/\hat{\sigma}_i^2) \le \nu\hat{\sigma}_i^2$; thus by the strict decreasingness of g' from Assumption 1, we have $\nu\hat{\sigma}_i^2 \ge g'(\tilde{q}_{\ell^*,i}/\hat{\sigma}_i^2) \ge g'(P/\hat{\sigma}_i^2) > g'(\infty)$.

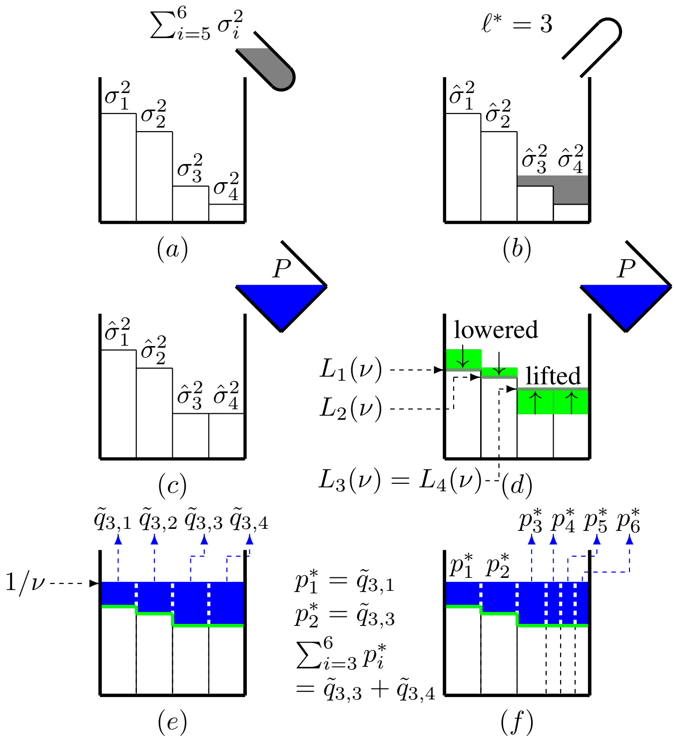

Fig. 2. The exemplified graphical interpretation of the optimal two-phase water-filling power allocation policy for a $(6, 4)$-limited additive noise access channel with uniform i.i.d. QPSK inputs and i.i.d. additive Laplacian noises.

$\ell^* = 3$ and $(\hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\sigma}_3^2, \hat{\sigma}_4^2) = (12, 10, 5, 5)$. Corollary 1 gives $\nu = 0.811$, $L_1(\nu) = 9.766$, $L_2(\nu) = 9.089$, $L_3(\nu) = L_4(\nu) = 7.7325$, and $\tilde{\boldsymbol{q}}_{\ell^*} = (2.564, 3.241, 4.5975, 4.5975)$. Then, the optimal power allocation is given by $\boldsymbol{p}^* = (2.564, 3.241, 3.678, 1.839, 1.839, 1.839)$, and the optimal SNRs are $\boldsymbol{\gamma}^* = (0.214, 0.324, 0.918, 0.918, 0.918, 0.918)$. □

We close this section by the following two remarks.

- When $\sigma_1^2 \leq \frac{1}{(K-1)} \sum_{i=2}^N \sigma_i^2$, we have $\ell^* = 1$. Then $\boldsymbol{p}^*$ can be determined without any maximization labor since we immediately have for $1 \leq i \leq N$

$$p_i^* = \frac{\sigma_i^2}{\sum_{k=1}^N \sigma_k^2} \left(\sum_{j=1}^K q_j^*\right) = \frac{\sigma_i^2}{\sum_{k=1}^N \sigma_k^2} P.$$

- The validity of the above theorems and corollary is not restricted to channels with additive noises, but can be extended to any mutual information function that satisfies Assumption 1. For example, our result can also be applicable to a system characterized by

$$Y_i = (\beta_i H_i)(\sqrt{p_i} X_i) + \sigma_i Z_i \quad \text{for } 1 \leq i \leq N$$

where channel state $\{H_i\}_{i=1}^N$ is i.i.d. with unit second moment and is independent of the channel inputs and additive noises, provided the channel state information is known at the receiver end.

III. Conclusion

In this paper, we consider the (N, K)-limited access additive color noise channel with arbitrary inputs and obtain that the optimal SNRs should be monotonic with respect to noise powers within each group of channels, in which the mutual information is permutation-invariant in the parameters that characterize each channel. For a system where all channels belong to the same permutation-invariant group, we further show that the power allocation max-min problem is reduced to a maximization problem with a single power-sum constraint via a water-filling noise-power redistribution procedure. Based on this result, we proceed to prove that if the mutual information function is additionally assumed to be additive, then the determination of the optimal allocated power can be interpreted by a water-filling process with pre-adjustment.

A possible future work would be to weaken the strict concavity assumption such that the derivative of the mutual information function is only nonincreasing. Another follow-up research work could be to examine the relationship of SNRs among channels in different permutation-invariant groups. Thus, a general guideline for the power allocation for a system containing, e.g., a mixture of different modulations such as BPSK, QPSK and QAM may be addressed.

References

[1] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoded transmission," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1737–1744, Nov. 1996.

[2] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 11, no. 7, pp. 991–1001, Sep. 1993.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, U.K.: Cambridge Univ. Press, 2004.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[5] B. F. Cron and C. H. Sherman, "Spatial correlation functions for various noise models," *J. Acoust. Soc. Amer.*, vol. 34, pp. 1732–1736, Nov. 1962.

[6] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, U.K.: Cambridge Univ. Press, 1985.

[7] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2148–2177, Oct. 1998.

[8] I-C. Lee, C.-S. Chang, and C.-M. Lien, "On the throughput of multicasting with incremental forward error correction," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 900–918, Mar. 2005.

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

[10] A. Lozano, A. M. Tulino and S. Verdú, "Optimum power allocation for multiuser OFDM with arbitrary signal constellations," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 828–837, May 2008.

[11] J. Meyer and K. U. Simmer, "Multi-channel speech enhancement in a car environment using Wiener filtering and spectral subtraction," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, Munich, Germany, Apr. 1997.

[12] F. Pérez-Cruz, M. R. D. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1070–1084, Mar. 2010.

[13] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions—part I: (n, k) source-channel erasure codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 47–61, Jan. 2004.

[14] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions—part II: an achievable rate-distortion region," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1377–1392, Apr. 2005.

[15] M. G. Sánchez, A. V. Alejos, and I. Cuiñas, "Urban wide-band measurement of the UMTS electromagnetic environment," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1014–1022, Jul. 2004.

[16] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Trans. Electromagn. Compat.*, vol. 44, no. 1, pp. 249–258, Feb. 2002.