On Optimal Power Allocation for Modulation-Constrained Gaussian Channels

Yu Han, Maria Urlea and Sergey Loyka

2021 IEEE International Symposium on Information Theory (ISIT) | 978-1-5386-8209-8/21/$31.00 ©2021 Crown | DOI: 10.1109/ISIT45174.2021.9518107

Abstract—**The problem of optimal power allocation for parallel Gaussian channels under modulation order constrains, in addition to the total transmit power constraint, is considered. It is motivated by coded-modulation systems using powerful capacity-approaching codes. While only analytically-intractable solution is known to this problem, an explicit closed-form solution is obtained here using a sphere-packing-based approximation for modulation-constrained rates. It can be interpreted as waterfilling with variable water level, which is also expressed in a closed-form. The obtained power allocation also solves the dual problem of minimizing the total transmit power subject to the sum rate and modulation order constraints. More insightful analytical solutions are obtained in some special cases. While the new power allocation is similar to the well-known waterfilling procedure at low SNR, it is dramatically different at moderate to high SNR. Proportional cardinality allocation is shown to be optimal at high SNR under the uniform power allocation.**

I. Introduction

Parallel Gaussian channel model appears in many areas of modern communication systems in space, frequency or time domain, including multiple-input multiple-output (MIMO) systems (after unitary precoding), OFDM-based wideband systems or wavelength division multiplexing (WDM) optical fiber systems [1]-[7]. For such systems, independent Gaussian signaling is optimal (capacity-achieving) and, under the total (sum) transmit power constraint (TPC), optimal power allocation is given by the well-known waterfilling (WF) procedure, which originates back to Shannon [8]-[11]. The WF procedure plays a prominent role in many areas of information/communication theory, signal processing and control.

However, in many systems modulation format is constrained to have a constellation of given cardinality due to e.g complexity/implementation constraints. If capacity-approaching codes are used, an achievable information rate can be expressed as the input-output mutual information (MI) of an extended channel, which also includes modulator/demodulator [10][14][15]. This approach has recently gained wide acceptance to characterize the performance of optical transmission systems [7][12][13]. While the expressions for the MI of 1-D (e.g. M-PAM) and 2-D (e.g. M-QAM) constellations in the AWGN channel are available [10][14], they are not in a closed form and are analytically-intractable. This makes it difficult to find an optimal power allocation for parallel Gaussian channels under the added modulation constraints, which is important for applications. To overcome this difficulty, an MMSE-based approach was used in [16][17] and an optimal power allocation was obtained, termed ”mercury/waterfilling” (MWF) thus extending the standard WF to modulation-constrained inputs.

However, unlike the standard WF, no closed-form solution is available for the MWF in general since it requires the inverse MMSE functions for which no closed-form expressions are available; their numerical evaluation can be computationally-expensive and, sometimes, can result in numerical instabilities, especially at high SNR [18]. The underlying problem is that the exact MI and MMSE expressions include multiple integrals over infinite intervals for which no closed-from solutions are available, and which have to be evaluated numerically. This can be computationally-expensive, especially for high-order modulation formats (which are now widely used to achieve high spectral efficiency), and also in an iterative optimization process, where such expressions have to be numerically evaluated over many iterations. The lack of closed-form solutions also limits significantly available insights, from which design guidelines can be developed.

To address these issues, a number of approximations to the exact modulation-constrained MI have been developed [19][20]. Here, we make use of the approximation in [20] due to several reasons: it is sufficiently accurate for the whole SNR range and for constellations of various cardinalities; it compares favourably ro recent experimental results for fiber optics systems (being closer to those results than the exact MI); and it is in an explicit closed form, which possess a number of analytical properties making it suitable for optimization. In addition, it was obtained via the sphere-packing method, which has long information-theoretic roots [8]-[11].

We obtain a closed-form optimal power allocation (OPA) to maximize the sum rate (MI) of modulation-constrained Gaussian channels with different channel gains and different constellation cardinalities, all under the total Tx power constraint (TPC). Using the approximation in [20], we formulate the OPA problem as a convex optimization problem, for which a closed-form globally-optimal solution is obtained via Karush-Kuhn-Tucker (KKT) conditions. Unlike the MMSE-based MWF in [16][17], our solution is in an explicit closed form, which is computationally-efficient and from which a number of insights follow. It can be interpreted as waterfilling with variable water level for each channel, for which a closed-form expression is obtained with explicit dependence on channel gain and modulation order. When modulation order increases, our solution gradually converges to the standard WF. Numerical experiments demonstrate that the true sum rate (MI) of this explicit closed-form solution is very close to that of the (non-explicit) MWF. We further show that the OPA also solves the dual problem of minimizing the total transmit power subject to the sum rate constraint.

To obtain an optimal cardinality allocation (OCA), we turn the OPA problem around and ask a question: what is a cardinality allocation so that the OPA is uniform? An explicit solution to this problem is given, which can be used in systems with uniform power allocation (UPA).

II. Channel Model and Information Rates

Let us consider a set of n parallel (independent) AWGN channels, which, after after matched filtering and sampling, can

The authors are with the School of Electrical Engineering and Computer Science, University of Ottawa, Ontario, Canada, K1N 6N5, e-mail: sergey.loyka@uottawa.ca

be expressed as

$$y_k = \sqrt{g_k} x_k + \xi_k,\ k = 1...n, \tag{1}$$

where x_k and y_k are transmitted and received symbols respectively at time k, ξ_k is the additive white Gaussian noise of zero mean and unit variance, and g_k is the channel power gain (if noise variance is not the same in all channels, this can be accounted for in g_k). This model can represent spatial parallel channels available in MIMO systems after unitary precoding based on channel singular value decomposition, channels corresponding to different sub-carriers in OFDM-based wireless systems or WDM-based optical fiber systems [1]-[7].

With no modulation constraints, the capacity of each channel

$$C_k = \log(1 + g_k p_k) \text{ [bit/symbol]}, \tag{2}$$

where p_k is the power of channel k, and $p_k g_k = \mathrm{SNR}_k$ is its Rx SNR (since the noise variance is 1). When there is the total (sum) power constraint (TPC) at the transmitter (Tx), the total (sum) capacity can be expressed as

$$\max_{p_k \geq 0} \sum_{k=1}^{n} C_k(p_k) \text{ s.t. } \sum_{k=1}^{n} p_k \leq P \tag{3}$$

where P is the maximum Tx power. The solution of this problem (i.e. optimal power allocation) is given by the well-known and widely-used water-filling (WF) procedure [8]-[11]:

$$p_{k(WF)} = (\lambda_{WF}^{-1} - g_k^{-1})_+ \tag{4}$$

where λ_{WF} is the dual variable ("water level"), which is found from the total power constraint

$$\sum_k p_{k(WF)}(\lambda_{WF}) = P \tag{5}$$

by e.g. bisection method, and $(x)_+ = \max\{0, x\}$. The WF procedure is widely used in many areas of information and communication theory as well as in signal processing and control.

A. Modulation-constrained channels

The complexity of the problem in (3) increases dramatically when modulation constraints are present (when the respective mutual information should be used instead of the unconstrained capacity C_k); no closed-form solution is known in this case. Indeed, when modulation is constrained for each channel so that its constellation has given (fixed) cardinality M_k (due to e.g. complexity, implementation issues etc.), the relevant performance metric is modulation-constrained achievable information rate, which is given by the mutual information (MI) between input and output of the equivalent channel (including the modulator as its part) [10][14][15] and which is used extensively as a performance metric of modern optical fiber transmission systems [7][12][13]. While its analytical expression is available [10][14][15], it includes integrals over infinite intervals for which no closed-form solutions are available and which are time-consuming to evaluate numerically, especially in iterative numerical optimization.

To overcome these difficulties, the MMSE-based formulation was used in [16][17]. Based on it, the following optimal power allocation $p_{k(MWF)}$, termed "mercury/waterfilling" (MWF), for the problem in (3) under the added modulation constraints was obtained:

$$p_{k(MWF)} = g_k^{-1} \cdot \mathrm{MMSE}_k^{-1}\{\min(1, \eta g_k^{-1})\} \tag{6}$$

where the "water level" η is determined from the TPC $\sum_k p_{k(MWF)} = P$, and $\mathrm{MMSE}_k^{-1}\{\cdot\}$ is the inverse MMSE function for k-th channel, whose constellation cardinality is M_k. However, the key difficulty here is that inverse MMSE functions are not available in closed-form (except for Gaussian inputs) making this solution analytically-intractable in general. One has to resort to their numerical evaluation, which can be time-consuming and numerically-unstable (especially at high SNR and for high-order modulations) [18]. This limits significantly the available insights as well as related design and optimization procedures.

To overcome this drawback, we use here an approximation R_k of the modulation-constrained MI obtained in [20] via the sphere-packing method extended to modulation-constrained inputs,

$$R_k = \log(1 + g_k p_k) - \log(1 + g_k p_k / M_k) \tag{7}$$

where M_k is the constellation cardinality of k-th channel (e.g. M_k-QAM), and 2nd term represents the rate loss due to using a constellation of finite cardinality. Albeit its approximate nature, this expression has a number of advantages: it is in an explicit closed-form and yet sufficiently accurate over the whole SNR range and for constellations of various cardinalities; it compares favorably to the rates achieved in recent state-of-the-art fiber-optics transmission systems using coded modulation (being in fact more close to those rates than the exact MI); it possesses a number of useful analytical properties (important for optimization) and allows one to obtain a closed-form solution to the related optimization problem. For completeness, its derivation from "first principles" (extending the standard sphere-packing method to modulation-constrained inputs) is given in Appendix I. For its accuracy, see Fig. 4 in [20] and Fig. 2 in this paper.

III. Optimal Power Allocation

Under modulation constraints and using the above approximation, the sum rate maximization problem becomes

$$\text{(P1)} \quad \max_{p_k \geq 0} \sum_{k=1}^{n} R_k(p_k) \text{ s.t. } \sum_{k=1}^{n} p_k \leq P \tag{8}$$

where $R_k(p_k)$ is as in (7). Unlike the MWF in (6), the above problem admits a closed-form solution in the general case as follows.

Theorem 1. *The optimal power allocation for the problem in* (8) *is unique and is given by*

$$p_k = \frac{1}{2g_k}\left(\sqrt{(M_k-1)^2 + \frac{4g_k}{\lambda}(M_k-1)} - M_k - 1\right)_+ \tag{9}$$

where the dual variable $\lambda > 0$ is found as a unique solution of the following (nonlinear) equation (the TPC):

$$\sum_k p_k(\lambda) = P. \tag{10}$$

Proof. see Appendix II. □

Since the left-hand side of (10) is a monotonically-decreasing function of dual variable λ, it can be solved efficiently using the bisection method. In fact, in some special cases, this equation can also be solved analytically.

The OPA in (9) possesses the following properties:

1. Channel k is active, i.e. $p_k > 0$, if and only if its channel gain (or SNR) is sufficiently large, $g_k(1 - M_k^{-1}) > \lambda$.

2. Since λ is a monotonically-decreasing function of P, the number of active channel increases with P: only the channel with largest $g_k(1-M_k^{-1})$ is active at low SNR (small P) while all channels are active at large SNR (large P).

3. If all channel gains g_k are the same, then channels with larger constellations (larger M_k) get more power.

4. For identical M_k, weaker channels get more power at high SNR while the opposite is true at low SNR and only the strongest channel is active at sufficiently low SNR,

$$P(1+g_1P/M_1) \leq g_2^{-1} - g_1^{-1} \tag{11}$$

where $M_k \gg 1$, and g_k are in descending order.

Note that while some of these properties are qualitatively similar to those of the standard WF, others are dramatically different (e.g. weaker channels get more power at high SNR).

To see further analogy of the above OPA with the standard WF in (4), we re-write (9) as follows:

$$p_k = \left(\frac{\alpha_k}{\lambda} - \frac{1}{g_k}\right)_+,\ \alpha_k = \frac{2}{\sqrt{1+\frac{4g_k}{\lambda(M_k-1)}}+1} \tag{12}$$

Note that α_k represents a correction to the standard WF in (4) due to the constrained cardinality M_k and can be interpreted as variable ”water level” for each channel depending on its gain and cardinality: higher $g_k/(M_k-1)$ call for lower water levels. This is in stark contrast to the standard WF in (4), where the ”water level” is the same for all channels. The variable ”water-level” above is consistent with the ”mercury-waterfilling” interpretation in [16][17] (for which no explicit expressions are available for variable water levels).

It follows from (12) that $\alpha_k \to 1$, $p_k \to p_{k(WF)}$ as $M_k \to \infty$,

$$p_k = \left(\lambda^{-1} - g_k^{-1}\right)_+ + o(1) \approx p_{k(WF)} \tag{13}$$

where the approximation holds if $M_k \gg 4g_k/\lambda + 1$. Further note that, under a more relaxed (SNR-independent) condition $M_k \gg 1$, the channel activity condition $g_k(1-M_k^{-1}) > \lambda$ converges to that of the standard WF $g_k > \lambda_{WF}$. However, we caution the reader that while these two conditions have similar appearance, the respective ”water levels” are different since they are found from different conditions in (10) and (5). In fact, $\lambda \leq \lambda_{WF}$, since $p_k(\lambda) \leq p_{k(WF)}(\lambda)$.

To gain further insights into the OPA and to obtain closed-form solutions for ”water level” λ, we study some special cases below.

IV. High SNR regime

In this section, we will assume that all $g_k > 0$ and study the high SNR regime. This is motivated by the fact that high SNR is needed to achieve high spectral efficiency and thus overall high transmission rate, which is an important practical objective for many applications [1][2][6][7]. We use the tools of asymptotic analysis [24] to obtain the following approximation.

Proposition 1. *In the high-SNR regime $P \to \infty$, all channels are active and the OPA in* (9) *can be expressed as follows:*

$$p_k = \theta_k P(1+o(1)) \approx \theta_k P,\ \theta_k = \frac{\beta_k}{\sum_{i=1}^n \beta_i} \tag{14}$$

where $\beta_k = \sqrt{(M_k-1)/g_k}$ and the approximation holds for

$$P \gg \frac{1}{2}\sum_{i=1}^n \beta_i \max_k \beta_k \tag{15}$$

Proof. Since $P \to \infty$ implies $\lambda \to 0$, it follows from (9) that

$$p_k = \sqrt{(M_k-1)/\lambda g_k}(1+o(1)) \tag{16}$$

where $o(1)$ collects all asymptotically-negligible terms [24], $o(1) \to 0$ as $P \to \infty$. Using this in the TPC $\sum_k p_k = P$, one obtains

$$\frac{1}{\sqrt{\lambda}} = \frac{P}{\sum_k \sqrt{(M_k-1)/g_k}}(1+o(1)) \tag{17}$$

Combining the last two equalities, one obtains (14). The approximation follows by omitting $o(1)$. □

Note from (14) that, at high SNR, p_k is proportional to $\sqrt{(M_k-1)/g_k}$, i.e. weaker channels get more power. This is in stark contrast to the standard WF in (4), where all channels get the same power at high SNR, $p_{k(WF)} \approx 1/\lambda = P/n$, and, in general (any SNR), weaker channels get less power.

Further note from (14) that channels with larger M_k get more power, as expected intuitively (the above OPA quantifies this intuition).

V. Optimal Cardinality Allocation

Quite often in practice multiple channels are allocated the same Tx power while they have different gains so that they also have different SNRs at the Rx end. A question arises as to how to chose constellation cardinalities to fit these conditions. An intuitively-appealing choice is to use large M_k for channels with larger SNR. However, even if one follows this intuition, it remains unclear what are the specific values of M_k to be used. Usually, they are selected in ad-hoc way in practice (following the qualitative intuition above) [21]. To address this issue, we consider here an optimal cardinality allocation (OCA), based on different channel gains with uniform Tx power allocation (UPA), i.e. the same p_k for all k: $p_k = P/n$.

To this end, we turn the OPA problem around and ask the question: under what selection of M_k, the OPA in (9) is uniform?

To simplify the analysis, we consider here the practically-important large SNR regime, when per-channel SNRs are large: $\text{SNR}_k = p_k g_k = Pg_k/n \gg 1$ (this is necessary for the spectral efficiency to be high). Using Proposition 1, it follows that the optimal power allocation becomes uniform, i.e. $p_k = P/n$ for all k, if all θ_k are the same, i.e. the cardinalities satisfy:

$$M_k = ag_k + 1 \approx ag_k \tag{18}$$

for some $a > 0$, where the approximation holds for moderately-large cardinalities $M_k \gg 1$, which is the case for modern optical fiber systems [6][7]; setting $a = 4P/n$ allows one to approach closely the modulation-unconstrained rates without using unnecessarily large constellations. It follows from (18) that *proportional cardinality allocation* $M_k \sim g_k$ is optimal under the uniform power allocation. This rule makes precise the intuitive and well-known observation that stronger channels can support higher-order modulation formats. Since M_k is integer (and, in practice, is often a power of 2 or 4), the value in (18) should be rounded off to the nearest available one.

VI. Dual Problem

In this section, we consider a problem dual of (P1), namely, minimizing the total Tx power P subject to the rate constraint:

$$\text{(P2)} \quad \min_{p_k \geq 0,\ P} P \quad \text{s.t.} \sum_{k=1}^n p_k \leq P,\ \sum_{k=1}^n R_k(p_k) \geq C_2 \tag{19}$$

where C_2 is a given target sum rate, and the problem's variables are $\{p_k\}$ and P. The next Proposition shows that the OPA of (P1) in (9) can also be used to solve (P2). To this end, let $p_k^{(1)}$ and $p_k^{(2)}$ be the OPA of (P1) and (P2), respectively, and $P_i = \sum_k p_k^{(i)}$, $i = 1, 2$, i.e. P_1 is the total Tx power of (P1), and let $C_1 = \sum_k R_k(p_k^{(1)})$ be the optimal value (rate) of (P1), and likewise for (P2).

Proposition 2. *The optimal power allocations of (P1) and (P2) are the same, i.e.* $p_k^{(1)} = p_k^{(2)}$, *if* $C_2 = C_1$. *Hence, the OPA of (P1) not only maximizes the sum rate but also minimizes the total Tx power needed to achieve this sum rate.*

Proof. First, observe that $P_2 \leq P_1$, since $p_k^{(1)}$ is also feasible for (P2) under $C_2 = C_1$. Next, we show by contradiction that $P_2 < P_1$ is not possible and hence $P_1 = P_2$. Indeed, if

$$P_1 > P_2 = \sum_k p_k^{(2)} \tag{20}$$

then $p_k^{(2)}$ is also feasible for (P1) and hence

$$\sum_k R_k(p_k^{(2)}) \leq C_1 \tag{21}$$

On the other hand, from (19),

$$\sum_{k=1}^{n} R_k(p_k^{(2)}) \geq C_2 = C_1 \tag{22}$$

so that

$$\sum_{k=1}^{n} R_k(p_k^{(2)}) = C_1 \tag{23}$$

i.e. $p_k^{(2)}$ is also optimal for (P1). This, however, is impossible since (i) $p_k^{(2)} \neq p_k^{(1)}$ (from $P_1 > P_2$) and (ii) the OPA of (P1) is unique. Thus, $P_1 = P_2$ and (23) follows from (21) and (22). Hence, $p_k^{(2)}$ also solves (P1), $p_k^{(2)} = p_k^{(1)}$, and is unique. □

VII. An Example

To validate and illustrate the analytical results above, we consider the following (representative) example: $n = 2$, $g_1 = 100, g_2 = 1$ and use 16-QAM for both channels, $M_k = 16$. For convenience of presentation over a wide range of P, we use normalized power allocations $p'_k = p_k n/P$ (so that $\sum_k p'_k = n$ regardless on P). Fig. 1 shows normalized power allocations: the OPA in (9), the MWF in (6) and the WF in (4). Note that all 3 allocations coincide at low SNR $P < -9$ dB, where all power goes to the strongest channel. The OPA in (9) and the MWF in (6) are close to each other for whole SNR range while being significantly different from the WF for $P > -8$ dB. The strongest channel gets more (or all) power if $P < -6$ dB, while the weakest channel gets more power if $P > -4$ dB and almost all power if $P > 10$ dB. This is in sharp contrast to the standard WF where stronger channel always gets more power and its power allocations approaches uniform one at high SNR.

While there is some discrepancy between the OPA in (9) and the MWF in (6), this has negligible impact on the true sum rate, as Fig. 2 shows. In fact, the OPA in (9) and the MWF in (6) deliver almost the same sum rate at any SNR while the standard WF is noticeably below at the transition range -5 dB < P < 15 dB. The similarity between the OPA in (9) and the MWF in (6) can be explained via sensitivity analysis, which shows that the sum rate is not very sensitive to power allocation, especially at high SNR.

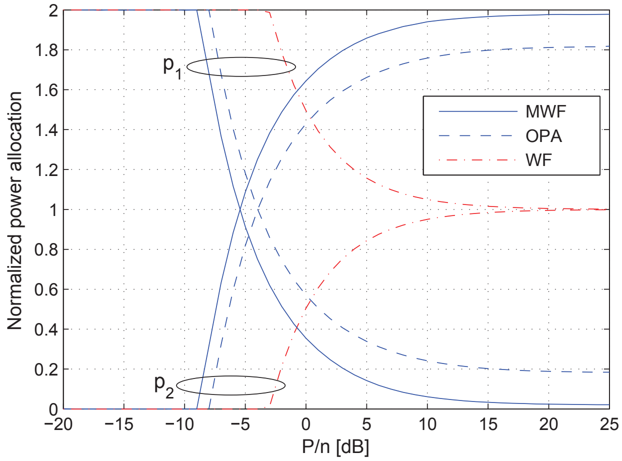

Fig. 1. Normalized power allocations: the OPA in (9), the MWF in (6) and the WF in (4); $n = 2$, $g_1 = 100, g_2 = 1$, 16-QAM for both channels.

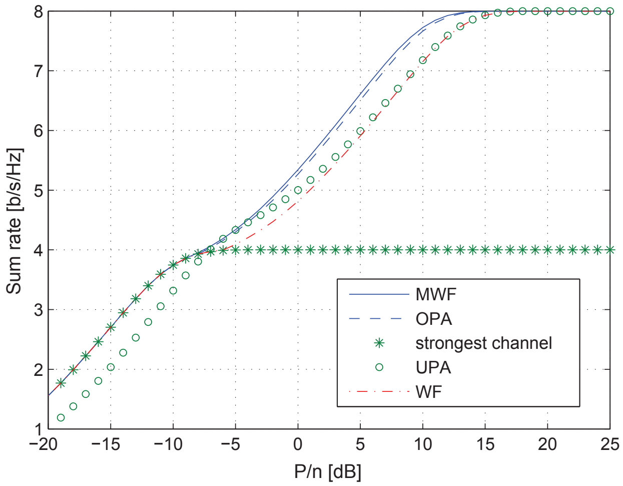

Fig. 2. The true sum rate (MI) attained by the power allocations in Fig. 1 along with the UPA and strongest channel only rates. While the MWF and OPA deliver almost the same sum rate at any SNR, the other strategies are sub-optimal over some intervals.

Since $g_1 \gg g_2$ in this example, it may be argued that allocating all power to the strongest channel is a good strategy. This, however, is not the case at any but very low $P < -7$ dB, as Fig. 2 shows: since modulation-constrained rate saturates at high SNR, the excess power (beyond the saturation point) should be allocated to weaker channel to boost the sum rate. However, if SNR is sufficiently low, $P < -7$ dB, allocating all power to the strongest channel is optimal.

VIII. Appendix I: Sphere Packing Method

The sphere-packing method was originally used by Shannon to provide an intuitive and insightful (albeit approximate) derivation of the AWGN channel capacity [8]-[11]. We exploit this approach here and extend it to modulation-constrained inputs.

Consider first a (single) real-valued AWGN channel with codewords of blocklength N, $y_i = x_i + \xi_i$, $i = 1...N$, where ξ_i is i.i.d. Gaussian noise of variance $N\sigma_0^2$. As N increases,

$\sum_i \xi_i^2$ approaches $N\sigma_0^2$ with high probability (known as "sphere hardening"), so that the received sequence belongs to a noise sphere centered on the transmitted codeword (codeword region) with high probability. As long as the noise spheres corresponding to different codewords do not overlap, probability of error can be made as small as desired. Without modulation constraints, the maximum possible number of codewords is given by the ratio of volumes of the received signal sphere and noise sphere [8]-[11]. However, for a fixed constellation of M points, the number of codewords of length N can be at most M^N. To evaluate the impact of this constraint on the rate, we present the per-symbol MI in the following form:

$$C_M = C - \Delta C, \tag{24}$$

where C is the (unconstrained) channel capacity (as above) and $\Delta C \geq 0$ is the rate loss due to a fixed modulation of order M. To estimate ΔC, consider a hypothetical channel with noise power σ_1^2 such that the number of distinct codewords (noise spheres) is exactly M^N:

$$M^N = \frac{V}{V_1} = \frac{(N\sigma_x^2 + N\sigma_1^2)^{N/2}}{(N\sigma_1^2)^{N/2}}, \tag{25}$$

where $V_1 = \alpha(N\sigma_1^2)^{N/2}$ is the volume of the hypothetical noise sphere, which is also the volume of a codeword region when there are exactly M^N codewords, and V is the volume of received signal sphere (to which the received signal belongs with high probability for any coderword). For this channel, there is no loss in capacity due to a fixed constellation (within the sphere packing approximation) since the noise power is "right" (i.e. noise spheres are the same as respective codeword regions so that no more codewords can fit without increasing the error probability):

$$\sigma_1^2 = \frac{\sigma_x^2}{M^2 - 1} \approx \frac{\sigma_x^2}{M^2}, \tag{26}$$

where the last approximation holds when M is reasonably large, $M^2 \gg 1$. However, if the true noise power is less than the hypothetical one, $\sigma_0^2 < \sigma_1^2$, more than one noise sphere can fit within the hypothetical noise sphere (codeword region), as shown in Fig. 3 (the central sphere). The resulting ΔC can be interpreted as the capacity of the fictitious channel with signal power σ_1^2 and the noise power σ_0^2, which can be estimated via the ratio of volumes again:

$$\Delta C \approx \frac{1}{N} \log \frac{(N\sigma_1^2 + N\sigma_0^2)^{N/2}}{(N\sigma_0^2)^{N/2}} \approx \frac{1}{2} \log\left(1 + \frac{\sigma_x^2}{M^2\sigma_0^2}\right). \tag{27}$$

Substituting this in (24), one finally obtains an approximation C_a for the per-symbol MI of 1-D constellation (e.g. M-PAM):

$$C_M \approx C_a = \frac{1}{2} \log \frac{1+\gamma}{1+\gamma/M^2}. \tag{28}$$

where $\gamma = \sigma_x^2/\sigma_0^2$ is the SNR.

Note that as M increases, both C_M and C_a approach C: $C_a \approx C_M \approx C$ if $M \gg \sqrt{\gamma}$, from which one can estimate minimum M required to approach closely the channel capacity without using unnecessarily large constellations:

$$M_{min} \approx 2\max\{1, \sqrt{\gamma}\}. \tag{29}$$

This demonstrates that the upper bound in [22], which can be put in the form $M_{min} \leq 2\sqrt{1+\gamma}$, is actually tight. Note also that the upper bound was obtained in [22] via an elaborate information-theoretic analysis (which does not yield a rate approximation) while our approximation to M_{min} follows directly from (28).

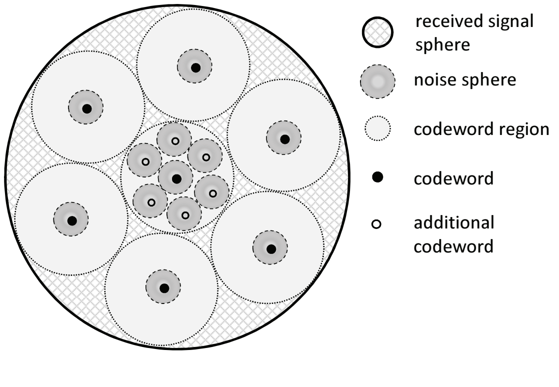

Fig. 3. The impact of the limited number of codewords: more noise spheres representing additional codewords could be packed into existing codeword regions when noise is small.

Since complex-valued channel with 2-D constellations (e.g. M-QAM) can be considered as two real-valued channels with 1-D constellations ($\sqrt{M}$-PAM), its per-symbol MI can be expressed as follows:

$$C_{M-QAM} = 2C_{\sqrt{M}-PAM} \approx \log \frac{1+\gamma}{1+\gamma/M}, \tag{30}$$

and the minimum constellation cardinality to approach closely the channel capacity is $M_{min} \approx 4\max\{1, \gamma\}$.

IX. Appendix II: Proof of Theorem 1

Since all $R_k(p_k)$ are concave and differentiable (to any order) functions, (P1) in (8) is convex and hence its KKT conditions are sufficient for global optimality [23]. The Lagrangian for this problem is

$$L = -\sum_{k=1}^{n} \ln \frac{1+p_k g_k}{1+p_k g_k/M_k} + \lambda\left(\sum_k p_k - P\right) - \sum_k \mu_k p_k$$

where λ is the Lagrange multiplier (dual variable) responsible for the TPC, and μ_k are dual variables responsible for $p_k \geq 0$. The respective KKT conditions are

$$\frac{\partial L}{\partial p_k} = -\frac{(M_k - 1)g_k}{(1+p_k g_k)(M_k + p_k g_k)} + \lambda - \mu_k = 0 \tag{31}$$

$$\lambda\left(\sum_k p_k - P\right) = 0, \quad \mu_k p_k = 0 \tag{32}$$

$$\sum_k p_k \leq P, \quad p_k \geq 0 \quad \lambda \geq 0,\ \mu_k \geq 0 \tag{33}$$

where (31) are stationarity conditions, (32) are complementary slackness conditions, (33) are primal and dual feasibility conditions. If $p_k > 0$ (active channel), then $\mu_k = 0$ and (31) reduces to

$$g_k^2 p_k^2 + (M_k + 1)p_k g_k + M_k - \frac{M_k - 1}{\lambda} g_k = 0 \tag{34}$$

where $p_k > 0$ implies $\lambda < (1 - 1/M_k)g_k$. Solving (34), one obtains (9). If $\lambda \geq (1 - 1/M_k)g_k$, then $p_k = 0$. It can be further seen that $\lambda > 0$ so that, using complementary slackness in (33), (10) follows (i.e. the TPC is always active, unless all $g_k = 0$ - a trivial case not considered here).

Since the objective in (8) is strictly concave (unless all $g_k = 0$), the above solution is unique.

References

[1] D. Tse, P. Viswanath, Fundamentals of Wireless Communications, Cambridge University Press, 2005.

[2] R.W. Heath, A. Lozano, Foundations of MIMO Communications, Cambridge University Press, 2019.

[3] B.S. Tsybakov, “The capacity of a memoryless Gaussian vector channel,” Probl. Inf. Transm., vol. 1, no. 1, pp. 18–29, 1965.

[4] I. E. Telatar, Capacity of Multi-Antenna Gaussian Channels, AT&T Bell Labs, Internal Tech. Memo, June 1995, (European Trans. Telecom., v.10, no. 6, Dec. 1999).

[5] B.S. Tsybakov, Capacity of a Discrete-Time Gaussian Channel with a Filter,” Probl. Inf. Transm., vol. 6, pp. 253–256, Jul./Sep. 1970.

[6] E. Agrell et al, Roadmap of Optical Communications, Journal of Optics, v. 18, no. 6, pp. 1–40, May 2016.

[7] P.J. Winzer, D.T. Neilson, A.R. Chraplyvy, Fiber-optic transmission and networking: the previous 20 and the next 20 years, Optics Express, v. 26, no. 18, pp. 24190–24239, Sep. 2018.

[8] C.E. Shannon, Communication in the Presence of Noise, Proc. IRE, vol. 37, no. 1, pp. 10–21, Jan. 1949.

[9] R. G. Gallager, Information Theory and Reliable Communication. New York: Wiley, 1968.

[10] R.E. Blahut, Principles and Practice of Information Theory, Addison-Wesley, 1987.

[11] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, 2006.

[12] M. Karlsson, E. Agrell, Multidimensional Modulation and Coding in Optical Transport, IEEE J. Lightwave Tech., vol. 35, no. 5, pp. 876–884, Feb. 2017.

[13] J. Cho, P.J. Winzer, Probabilistic Constellation Shaping for Optical Fiber Communications, IEEE J. Lightwave Tech., vol. 37, no. 6, pp. 1590–1607, Mar. 2019.

[14] G. Ungerboeck, Channel Coding with Multilevel/Phase Signals, IEEE Trans. Info. Theory, v. 28, No. 1, pp. 55–67, Jan. 1982.

[15] G.D. Forney and G. Ungerboeck, Modulation and Coding for Linear Gaussian Channels, IEEE Trans. Inf. Theory, vol. 44, no. 6, pp. 2384—2415, Oct. 1998.

[16] A. Lozano, A.M. Tulino, S. Verdu, Optimum Power Allocation for Parallel Gaussian Channels With Arbitrary Input Distributions, IEEE Trans. Info. Theory, v. 52, No. 7, pp. 3033–3051, July 2006.

[17] A. Lozano, A. Tulino, and S. Verdu, Optimum Power Allocation for Multiuser OFDM with Arbitrary Signal Constellations, IEEE Trans. Commun., vol. 56, no. 5, pp. 828-837, May 2008.

[18] Y. Han, Optimization of Modulation Constrained Digital Transmission Systems, MS Thesis, University of Ottawa, 2017.

[19] M. Secondini, E. Forestieri, and G. Prati, “Achievable information rate in nonlinear WDM fiber-optic systems with arbitrary modulation formats and dispersion maps,” IEEE J. Lightwave Tech., vol. 31, no. 23, pp. 3839–3852, Dec. 2013.

[20] M. Urlea, S. Loyka, Simple Closed-Form Approximations for Achievable Information Rates of Coded Modulation Systems, IEEE J. Lightwave Tech., vol. 39, no. 5, pp. 1306-1311, Mar. 2021.

[21] S. Okamoto et al, A study on the effect of ultra-wide band WDM on optical transmission systems, IEEE J. Lightwave Tech., v. 38, no. 5, pp. 1061–1070, Mar. 2020.

[22] L.H. Ozarow, A.D. Wyner, On the Capacity of the Gaussian Channel with a Finite Number of Input Levels, IEEE Trans. Info. Theory, v. 36, No. 6, pp. 1426–1428, July 1990.

[23] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[24] F.W.J. Olver, Asymptotics and Special Functions, Moscow: Nauka, 1990.