On the Transport Capacity of Gaussian Multiple Access and Broadcast Channels

G. A. Gupta, S. Toumpis, J. Sayir, R. R. Müller

***Abstract*— We study the transport capacity of a Gaussian multiple access channel (MAC), which consists of multiple transmitters and a single receiver, and a Gaussian broadcast channel (BC), which consists of a single transmitter and multiple receivers. The transport capacity is defined as the sum, over all transmitters (for the MAC) or receivers (for the BC), of the product of the data rate with a reward $r(x)$ which is a function of the distance x that the data travels.**

In the case of the MAC, assuming that the sum of the transmitter powers is upper bounded, we calculate in closed form the optimal power allocation among the transmitters, that maximizes the transport capacity, using Karush-Kuhn-Tucker (KKT) conditions. We also derive asymptotic expressions for the optimal power allocation, that hold as the number of transmitters approaches infinity, using the most-rapid-approach method of the calculus of variations. In the case of the BC, we calculate in closed form the optimal allocation of the transmitter power among the signals to the different receivers, both for a finite number of receivers and for the case of asymptotically many receivers, using our results for the MAC together with duality arguments.

Our results can be used to gain intuition and develop good design principles in a variety of settings. For example, they apply to the uplink and downlink channel of cellular networks, and also to sensor networks which consist of multiple sensors that communicate with a single central station.

I. Introduction

A. Transport Capacity

Consider a wireless multihop network in which a particular node T scheduled to transmit has two options: either transmit to a destination node D_1 with rate R_1, or transmit to a destination node D_2 with rate $R_2 < R_1$. Assume that both transmissions will require the same amount of bandwidth and power, and will convey information of equal importance. In this setting, which destination should node T prefer? Traditional thinking suggests that T should transmit to the node to which it can send data with the highest rate, i.e., D_1. However, in a *multihop* wireless network, in which every packet will have to be transmitted multiple times to reach its final destination, it is not only important that a node transmits *with high rate*, but also that the signal travels *a large distance*. Indeed, the smaller the distance that a transmission covers, the higher is the number of transmissions of a similar type that are needed, before the transmitted packet reaches its final destination.

Taking this argument a step further, we can argue that a natural figure of merit for the usefulness of a transmission is neither its rate nor the distance covered, but rather the *product* of the two, measured in $\mathrm{bps} \times \mathrm{m}$. Indeed, if two transmissions have the same rate-distance product, using either of the two repeatedly to transmit a given volume of data to a distant destination would consume the same power and bandwidth, even if their rates and covered distances differ significantly. It follows that the summation of the rate-distance products over all transmissions that are active at a given time instant in a wireless network, termed the **transport capacity**, is a natural figure of merit about how efficiently the network operates at that particular instant.

B. Related Work

The importance of transmitting over large distances was recognized in the early 1980's [4], [5]. However, transport capacity was defined more recently in [6]. There, the authors consider a wireless network of n nodes, placed in a bounded two-dimensional region. It is assumed that the power of transmitted signals decays with distance according to a power law, and that a signal is successfully received if the Signal to Interference and Noise Ratio (SINR) at the receiver is above a fixed threshold. All transmissions are with a fixed global rate W. It is shown that the transport capacity under *any* placement of nodes will have to be smaller than $k_1\sqrt{n}$, where k_1 is a constant independent of n. The bound comes from the fact that any transmission invariably creates interference to other transmissions near by. On the other hand, the authors give examples of network topologies that can sustain a transport capacity greater than $k_2\sqrt{n}$, where $k_2 < k_1$ is another constant, also independent of n.

In [7], the usefulness of a link is described in terms of the product of the communication rate with a *reward function* $r(x)$, where x is the distance between the transmitter and the receiver. In the special case where $r(x) = x$, we get the standard rate-distance product. The authors study the Gaussian Broadcast Channel (BC) of Fig. 1, which consists of a single transmitter T and multiple receivers $V_1,\ V_2, \ldots, V_n$, placed at increasing distances from the transmitter. The capacity region of this channel, i.e., the set of simultaneously achievable rates of communication from the transmitter to each of the receivers, is known [8]. The authors build on this knowledge to calculate

Work supported by **K***plus* funding for the ftw. project I0 "Signal and Information Processing."

G. A. Gupta is with the Mathematics Department of the Indian Institute of Technology (IIT), New Delhi, India, 110016 (email: gupta@ftw.at).

S. Toumpis and J. Sayir are with the Telecommunications Research Center Vienna (ftw.), Donau-City-Strasse 1/3, A-1220, Austria (email: {toumpis,sayir}@ftw.at).

R. R. Müller is with Norwegian University of Science and Technology, 7491 Trondheim, Norway (email: mueller@iet.ntnu.no).

Parts of this work have appeared, in preliminary form, in [1], [2], [3].

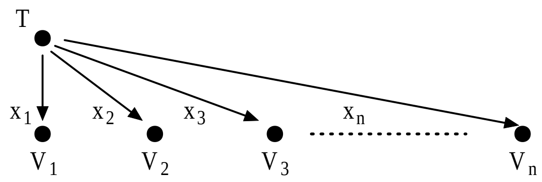

Fig. 1. The Gaussian broadcast channel, which consists of a single transmitter T and n receivers $V_1, V_2, \ldots, V_n$, placed at increasing distances $x_1 < x_2 < \ldots < x_n$ from the transmitter.

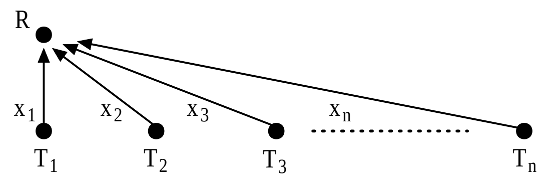

Fig. 2. The Gaussian multiple access channel, which consists of a single receiver R and n transmitters $T_1, T_2, \ldots, T_n$, placed at increasing distances $x_1 < x_2 < \ldots < x_n$ from the receiver.

the point in the capacity region that maximizes the transport capacity of the channel, defined as the summation, over all receivers, of their respective rate-reward products.

C. Contributions

In this work, we study the transport capacity of the Gaussian multiple access and broadcast channels, along the information theoretic tangent initiated in [7], and using the notions of reward and transport capacity defined there.

We start in Section II by considering the Gaussian Multiple Access Channel (MAC) of Fig. 2, that consists of a single receiver and n transmitters $T_1,\ T_2, \ldots,\ T_n$, placed at increasing distances from the receiver. Each transmitter T_i can transmit with a maximum power P_i. The capacity region of this channel is known [8]. Building on this knowledge we calculate the point in the capacity region that maximizes the transport capacity, defined for this channel as the summation, over all transmitters, of their respective rate-reward products.

In Section III, we relax the individual constraints on the powers of the transmitters, and instead assume that the sum of their powers is upper bounded. Under this sum-power constraint, we derive a closed form solution for the optimal allocation of the total power, and the resulting transport capacity. The calculation is done in a very straightforward manner, using the Karush-Kuhn-Tucker (KKT) conditions. We also derive asymptotic expressions for the optimal power allocation and the transport capacity it induces, that hold as the number of transmitters approaches infinity. The expressions are derived using the most-rapid-approach method from the calculus of variations. In all derivations, we adopt a reasonable assumption that essentially specifies that rewards increase with distance slower that the rate at which the signal strengths decay with distance.

In Section IV, we examine the Gaussian broadcast channel. We calculate, in closed form, the allocation of the transmitter power among the signals to the different receivers, that maximizes the transport capacity, both in the case of a finite number of receivers, and for the asymptotic case where the number of receivers goes to infinity. Although similar results have been derived in [7], our derivations are much shorter because they are based on the duality between the broadcast channel and the multiple access channel under the sum-power constraint [9], and on the results of the previous section (which are also much shorter than the derivation of [7]). Also, the expressions we arrive at are simpler.

We conclude in Section V with a discussion of our results and their implications in the design of practical wireless systems. We also discuss the connection of our results with related works on the calculation of arbitrary points of the capacity region.

II. The Multiple Access Channel

As shown in Fig. 2, the MAC consists of a single receiver R and n transmitters $T_1, T_2, \ldots, T_n$ placed at increasing distances $x_1 < x_2 < \ldots < x_n$ from the receiver (but not necessarily along a straight line, or even on the same plane). The receiver is subject to additive white Gaussian noise with spectral density η. Transmitter T_i can transmit with a maximum power[1] P_i, and the total bandwidth available for communication is equal to B.

When T_i transmits with power p, R will receive the signal with power $h(x_i) \times p$, where the **gain function** $h(\cdot)$ captures the dependence of the signal power on distance. For compactness, we use the notation $h_i = h(x_i)$. A gain function of particular interest is the **monomial gain function**, defined by $h(x) = K_h x^{-\gamma}$, where $\gamma > 0$ is the **gain exponent** and $K_h > 0$ is a constant.

The capacity region $\mathcal{C}^{\mathrm{MAC}}$ of the multiple access channel is defined as the set of all the combinations of rates $\mathbf{R} = (R_1, R_2, \ldots, R_n)$ with which each of the n transmitters can simultaneously send data to the receiver. The capacity region is a closed, convex polyhedron, given by [8]:

$$\mathcal{C}^{\mathrm{MAC}} = \Bigg\{ \mathbf{R} : \sum_{i \in \mathcal{I}} R_i \leq B \log_2 \Big(1 + \frac{\sum_{i \in \mathcal{I}} h_i P_i}{\eta B}\Big) \ \ \forall \, \mathcal{I} \subseteq \{1, 2, \ldots, n\} \Bigg\}.$$

It can be shown that the number of vertices of $\mathcal{C}^{\mathrm{MAC}}$ whose coordinates are *all* positive is exactly $n!$. Each of these vertices can be achieved by a successive decoding scheme, in which the signals from the n transmitters are decoded by the receiver one by one. When decoding the signal of transmitter T_i, those signals that have already been decoded do not affect the decoding. On the other hand, those signals that have not been decoded yet appear as additive white Gaussian noise. In particular, consider a successive decoding scheme in which the signal from $T_{\pi(j)}$ is decoded j-th, and $\pi(\cdot)$ is a permutation of the set $\{1, 2, \ldots, n\}$. (Consequently, $\pi^{-1}(i)$ is the rank with which the signal of T_i is decoded.) The components of the

[1] More formally, T_i can transmit with any power any time it uses the channel, as long as the average power over time converges with probability 1 to a value equal or smaller than P_i. Alternative constraints have been considered in a similar setting in [10], [11].

vertex $\mathbf{R}_\pi \triangleq (R_{1,\pi}, \ldots, R_{n,\pi})$, achieved by the permutation $\pi(\cdot)$, are given by:

$$R_{i,\pi} = B \log_2(1 + \frac{h_i P_i}{\eta B + \sum_{k\,:\,\pi^{-1}(k) > \pi^{-1}(i)} h_k P_k}).$$

We now associate the transmission of a bit of information across a distance x with a reward $r(x)$, where the **reward function** $r(\cdot)$ is strictly increasing, with $r(0) = 0$. For compactness, we use the notation $r_i \triangleq r(x_i)$. A reward function of particular interest is the **monomial reward function** $r(x) = K_r x^\rho$ where $\rho > 0$ is the **reward exponent** and $K_r > 0$ is a constant. The **transport capacity** associated with the point $\mathbf{R} = (R_1, R_2, \ldots, R_n) \in \mathcal{C}^{\mathrm{MAC}}$ is defined as

$$C_T^{\mathrm{MAC}}(\mathbf{R}) \triangleq \sum_{i=1}^{n} r_i R_i.$$

In the special case of the monomial reward function with $\rho = 1$, $K_r = 1$, this definition coincides with the original definition of transport capacity given in [6].

The transport capacity is a linear function of the rates R_i, who in turn must belong in the convex polyhedron $\mathcal{C}^{\mathrm{MAC}}$. Therefore, its maximization is a linear program [12], and the supremum is actually achieved at one of the $n!$ vertices $\mathbf{R}_\pi$. One could expect the complexity of the problem to increase factorially with the number of nodes. However, because of the special structure of the problem, the solution is remarkably simple, as the next theorem shows[2]:

Theorem 1: The maximum transport capacity is only achieved by the successive decoding scheme under which the signal of transmitter T_j is decoded j-th. In other words, the unique optimal permutation $\pi(\cdot)$ is the identity permutation $\pi(i) = i$. Therefore, the maximum transport capacity is given by:

$$B \sum_{i=1}^{n} r_i \log_2(1 + \frac{h_i P_i}{\eta B + \sum_{k=i+1}^{n} h_k P_k}).$$

Proof: Let us assume that the transport capacity is achieved by using a permutation $\pi_1(\cdot)$ other than the identity permutation. Therefore, there is a j_0 such that $k \triangleq \pi_1(j_0) > \pi_1(j_0+1) \triangleq l$. We define a new decoding order, specified by the permutation $\pi_2(\cdot)$:

$$\pi_2(j) \triangleq \begin{cases} \pi_1(j+1) & \text{if } j = j_0, \\ \pi_1(j-1) & \text{if } j = j_0 + 1, \\ \pi_i(j) & \text{otherwise.} \end{cases}$$

Both orders of decoding achieve exactly the same transmission rate for all transmitters other than T_k and T_l. Any difference in the transport capacities $C_T^{\mathrm{MAC}}(\mathbf{R}_{\pi_2}) - C_T^{\mathrm{MAC}}(\mathbf{R}_{\pi_1})$ will be due to the different rates achieved by T_k and T_l. Let I be the combined noise and interference power that the receiver sees when decoding the signal coming from T_l, under the original decoding order π_1. Also let $p_k = h_k P_k$ and $p_l = h_l P_l$.

[2]This theorem appeared first, in a different setting, in [13].

Then:

$$\begin{aligned} &\frac{1}{B}\left[C_T^{\mathrm{MAC}}(\mathbf{R}_{\pi_2}) - C_T^{\mathrm{MAC}}(\mathbf{R}_{\pi_1})\right] \\ &= r_k \log_2(1 + \frac{p_k}{I}) + r_l \log_2(1 + \frac{p_l}{I + p_k}) \\ &\quad - r_k \log_2(1 + \frac{p_k}{I + p_l}) - r_l \log_2(1 + \frac{p_l}{I}) \\ &= (r_k - r_l)\left[\log_2(I + p_k) + \log_2(I + p_l)\right] \\ &\quad + (r_l - r_k)\left[\log_2(I) + \log_2(I + p_l + p_k)\right] \\ &= (r_k - r_l) \log_2\left[\frac{I^2 + I p_k + I p_l + p_k p_l}{I^2 + I p_k + I p_l}\right] \\ &> 0. \end{aligned}$$

Therefore, the transport capacity strictly increases if we exchange the decoding orders of nodes k and l. Repeating the process, we can create a finite sequence of permutations $\pi_1, \pi_2, \ldots, \pi_m$ of strictly increasing transport capacity, with π_m being the identity permutation. The result follows. □

III. The MAC under a Sum-Power Constraint

A. Problem Formulation

In the previous section, it was assumed that each transmitter T_i has a maximum power P_i with which it can transmit. A natural extension of our investigation is to assume that transmitters no longer have individual constraints, but rather that the sum of powers must be smaller than or equal to some global constant P_0. For each distribution of powers whose sum does not exceed P_0, Theorem 1 applies. Therefore, to maximize the transport capacity in this setting, we need to solve the following optimization problem:

$$\begin{aligned} \text{maximize:} \quad & B \textstyle\sum_{i=1}^{n} r_i \log_2(1 + \frac{h_i P_i}{\eta B + \sum_{k=i+1}^{n} h_k P_k}), \\ \text{subject to:} \quad & \textstyle\sum_{k=1}^{n} P_k \le P_0, \quad P_i \ge 0, \; i = 1, \ldots, n. \end{aligned} \quad (1)$$

This problem is pertinent in a number of settings. For example, in the deployment phase of a sensor network consisting of a single central node and many sensors that must forward information to the central node, if our power sources are limited, we would like to know what is their optimal distribution over the sensors that will lead to the most efficient operation of the network, where we quantify efficiency with the notion of transport capacity. As another example, consider a cellular network, in which many mobile stations in a cell want to access the base station, and the network has placed an upper bound on the total transmitted power coming from that cell, in order to bound the interference experienced in neighboring cells that share the same frequency band. In such a setting, we would like to determine the maximum possible transport capacity, because this information will suggest how large the cell can be made. Finally, as we show in Section IV, knowing the transport capacity of the MAC with a sum-power constraint will allow us to determine the transport capacity of the BC.

B. Basic Properties

We rewrite the optimization problem (1) as:

$$\begin{array}{ll}\text{minimize:} & f_0(P_1,\dots,P_n) = Br_n\log_2(\eta B) \\ & -B\sum_{i=1}^n (r_i - r_{i-1})\log_2(\eta B + \sum_{k=i}^n h_k P_k) \\ \text{subject to:} & \sum_{k=1}^n P_k \le P_0, \quad P_i \ge 0,\ i = 1,\dots,n.\end{array} \tag{2}$$

We note that the objective function $f_0(P_1,\dots,P_n)$ is convex. Indeed, it can be written as the composite function $u(a_1(P_1,\dots,P_n),\dots,a_n(P_1,\dots,P_n))$, where the function $u : (\mathbf{R}^+)^n \to \mathbf{R}$ is defined by $u(a_1,\dots,a_n) = Br_n\log_2(\eta B) - B\sum_{i=1}^n (r_i - r_{i-1})\log_2 a_i$, and the functions $a_i : (\mathbf{R}^+)^n \to \mathbf{R}^+$ are defined by:

$$a_i(P_1,\dots,P_n) \triangleq \eta B + \sum_{k=i}^n h_k P_k, \quad i = 1,\dots,n. \tag{3}$$

The functions $a_i(P_1,\dots,P_n)$ are linear, and hence concave. Noting that the sequence $\{r_i\}$ is strictly increasing, we can easily show that the Hessian of the function $u(\cdot)$ is positive definite, therefore $u(\cdot)$ is (strictly) convex. In addition, $u(\cdot)$ is non-increasing in each argument. It follows that the composition $u(a_1(P_1,\dots,P_n),\dots,a_n(P_1,\dots,P_n))$ is convex [12]. As the $n+1$ inequality constraints of (2) are linear, it follows that (2) is a convex optimization problem.

The optimization function f_0 is continuous, and the domain of the problem, i.e., the set of power vectors where the constraints are satisfied, is compact. Therefore, the infimum of f_0 is actually achieved, and it makes sense to discuss about a minimum. Furthermore, only one power distribution achieves this minimum. To see why, let us assume that there are actually two power distributions, $\{P_i^1\}$ and $\{P_i^2\}$ achieving it. Because the mapping from the space of $\{P_i\}$ to the space of $\{a_i\}$ is one-on-one, there are two distinct points $\{a_i^1\}$ and $\{a_i^2\}$ where the function $u(\cdot)$ achieves its minimum. However, the Hessian of $u(\cdot)$ is positive definite in $(\mathbf{R}^+)^n$, therefore $u(\cdot)$ is strictly convex and must have a unique minimum. Therefore, we arrive at a contradiction.

C. A Basic Assumption

Assumption 1: The function $l(x) \triangleq \frac{r'(x)}{(\frac{1}{h(x)})'}$ *is decreasing:*

$$0 < x < y \Rightarrow \frac{r'(x)}{(\frac{1}{h})'(x)} \ge \frac{r'(y)}{(\frac{1}{h})'(y)}. \tag{4}$$

Roughly speaking, the assumption states that rates increase with distance not as fast as the channel decays with distance. It is not as restricting as it would first seem, as it is satisfied in most cases of practical interest. For example, it is clearly satisfied for the case of the monomial gain and reward functions with $\gamma \ge \rho$. For all realistic propagation environment models, $\gamma > 2$. In addition, as discussed in the introduction, the most interesting case of the monomial reward function is when $\rho = 1$. Therefore, the assumption is very reasonable assuming monomial reward and gain functions.

Assumption 1 allows us to determine the optimal power allocation in a straightforward manner. The optimal power allocation without using it can also be found, but the derivations are significantly more complicated, as we discuss in Section V.

Assumption 1 implies a number of facts that we will need later on, and we collect here in the form of a lemma. Their proofs appear in Appendix I.

Lemma 1:

(i) *Assumption 1 is equivalent to each of the following:*

$$\frac{r_i - r_{i-1}}{\frac{1}{h_i} - \frac{1}{h_{i-1}}} \ge \frac{r_{i+1} - r_i}{\frac{1}{h_{i+1}} - \frac{1}{h_i}},\ \forall i = 1,\dots,n-1, \tag{5}$$

$$\frac{\frac{r_{i-1}}{h_i} - \frac{r_i}{h_{i-1}}}{r_i - r_{i-1}} \le \frac{\frac{r_i}{h_{i+1}} - \frac{r_{i+1}}{h_i}}{r_{i+1} - r_i},\ \forall i = 1,\dots,n-1, \tag{6}$$

for any placement of transmitters $\{x_i\}$, *and where* $r_0 \triangleq 0$ *and* $h_0 \triangleq \infty$.

(ii) *Assumption 1 implies that for any placement of transmitters* $\{x_i\}$, *and* $\forall i = 1,\dots,n-1,\ \forall k = 1,\dots,n-i,$

$$\frac{r_i - r_{i-1}}{\frac{1}{h_i} - \frac{1}{h_{i-1}}} \ge \frac{r_{i+k} - r_{i-1}}{\frac{1}{h_{i+k}} - \frac{1}{h_{i-1}}}. \tag{7}$$

(iii) *Assumption 1 implies that, for* $x > 0$, $l(x) \le r(x)h(x)$.

(iv) *Assumption 1 implies that the function*

$$g(x) \triangleq \frac{r(x)(\frac{1}{h})'(x)}{r'(x)} - \frac{1}{h(x)} = -\frac{(r(x)h(x))'}{h^2(x)r'(x)}$$

is increasing.

D. The Closed Form Solution

It follows by the convexity of the problem [12], that, to prove the optimality of a power allocation $\{P_i\}$, it suffices to show that it satisfies the Karush-Kuhn-Tucker (KKT) conditions, which in our problem become:

$$\sum_{k=1}^n P_k = P_0,\ P_i \ge 0,\ \lambda_i \ge 0,\ \frac{\partial f_0}{\partial P_i} - \lambda_i + \nu = 0,\ \lambda_i \cdot P_i = 0,$$

for $i = 1,\dots,n$ and for some $\nu, \lambda_i \in \mathbf{R}$. This set of equations can easily be shown to be equivalent to the following:

$$\sum_{k=1}^n P_k = P_0,\ P_i \ge 0,\ \frac{\partial f_0}{\partial P_i} + \nu \ge 0,\ P_i \cdot \left[\frac{\partial f_0}{\partial P_i} + \nu\right] = 0,$$

for $i = 1,\dots,n$, and for some $\nu \in \mathbf{R}$. By substituting the partial derivatives of f_0, and setting $\lambda \triangleq \frac{\log 2}{B}\nu$, this set of equations becomes

$$\sum_{k=1}^n P_k = P_0,\ P_i \ge 0,\ h_i \sum_{k=1}^i \frac{r_k - r_{k-1}}{a_k} - \lambda \le 0,$$
$$P_i \cdot \left[h_i \sum_{k=1}^i \frac{r_k - r_{k-1}}{a_k} - \lambda\right] = 0, \tag{8}$$

for $i = 1,\dots,n$ and some $\lambda \in \mathbf{R}$. Note that we have defined the a_i's in (3). We also let $a_{n+1} \triangleq \eta B$, so that

$$P_i = \frac{a_i - a_{i+1}}{h_i}, \quad i = 1,\dots,n.$$

We now make the claim that the optimal power distribution $\{P_i\}$ satisfies the following set of equations:

$$\sum_{k=1}^{n} P_k = P_0, \tag{9}$$
$$P_i \geq 0, \quad i=1,\ldots,L, \tag{10}$$
$$h_i \sum_{k=1}^{i} \frac{r_k - r_{k-1}}{a_k} - \lambda = 0, \quad i=1,\ldots,L, \tag{11}$$
$$P_i = 0, \quad i=L+1,\ldots,n, \tag{12}$$
$$h_i \sum_{k=1}^{i} \frac{r_k - r_{k-1}}{a_k} - \lambda < 0, \quad i=L+1,\ldots,n, \tag{13}$$

for some $\lambda \in \mathbf{R}$ and an index L, which we call the **cutoff index**. If (9)-(13) hold, then (8) will also hold, and the $\{P_i\}$ form the unique optimal power allocation. Our strategy therefore will be to find a power distribution $\{P_i\}$ that satisfies (9)-(13).

Equations (10), (11), and (12) will be satisfied if we set:

$$a_i = \begin{cases} \frac{1}{\lambda}\frac{r_i - r_{i-1}}{\frac{1}{h_i}-\frac{1}{h_{i-1}}} & i=1,\ldots,L, \\ \eta B & i=L+1,\ldots,n, \end{cases} \tag{14}$$

provided that this sequence is decreasing. By (5), it suffices to show that $a_L \geq \eta B$. But this will depend on the values of λ and L, which we calculate next.

We first note that (14) implies that:

$$\sum_{k=1}^{i} P_k = \frac{1}{\lambda}\left[\frac{r_i h_i - r_{i+1}h_{i+1}}{h_i - h_{i+1}}\right], \; i=1,\ldots,L-1. \tag{15}$$

This can be shown by straightforward induction. In addition,

$$P_L = \frac{a_L - a_{L+1}}{h_L} = \frac{1}{h_L}\left[\frac{1}{\lambda}\frac{r_L - r_{L-1}}{\frac{1}{h_L}-\frac{1}{h_{L-1}}} - \eta B\right]. \tag{16}$$

Combining (16) with (15) for $i = L-1$, we derive the value of λ that satisfies the sum-power constraint (9):

$$\lambda = \frac{r_L h_L}{P_0 h_L + \eta B},$$

therefore we have that $a_L = \frac{P_0 h_L + \eta B}{r_L h_L} \times \frac{r_L - r_{L-1}}{\frac{1}{h_L}-\frac{1}{h_{L-1}}}$. We now simply define the cutoff index L to be the largest i for which the following inequality holds:

$$\frac{P_0 h_i + \eta B}{r_i h_i} \times \frac{r_i - r_{i-1}}{\frac{1}{h_i}-\frac{1}{h_{i-1}}} \geq \eta B \Leftrightarrow \frac{P_0}{\eta B} \geq \frac{\frac{r_{i-1}}{h_i} - \frac{r_i}{h_{i-1}}}{r_i - r_{i-1}}. \tag{17}$$

Note that the expression on the far right side of (17) is increasing, as follows from (6). Therefore, (17) will hold for a contiguous range of indices i, and the largest is selected as L. Also, note that the set of indices is never empty, as for $i = 1$ we arrive at the trivial identity $\frac{P_0}{\eta B} \geq 0$. With this selection of L, $a_L \geq \eta B$ and the sequence $\{a_i\}$ defined in (14) is decreasing. Therefore, all the constraints (9)-(12) are satisfied.

It remains to show that (13) is also satisfied. By using (14), we readily have that $\sum_{k=1}^{L} \frac{r_k - r_{k-1}}{a_k} = \frac{\lambda}{h_L}$. Therefore, it easily follows that the (13) are equivalent to

$$\eta B > \frac{1}{\lambda}\frac{r_i - r_L}{\frac{1}{h_i}-\frac{1}{h_L}}, \quad i=L+1,\ldots,n. \tag{18}$$

Because of (7), it suffices to show that (18) holds for $i = L+1$. In addition, straightforward algebra shows that the following equivalence also holds:

$$\eta B > \frac{1}{\lambda}\frac{r_{L+1} - r_L}{\frac{1}{h_{L+1}}-\frac{1}{h_L}} \Leftrightarrow \frac{P_0}{\eta B} < \frac{\frac{r_L}{h_{L+1}} - \frac{r_{L+1}}{h_L}}{r_{L+1} - r_L}.$$

However, the second relation is satisfied, by the way we have defined L. Therefore, the inequalities (13) are also satisfied. This concludes our proof, as we have found a power distribution that satisfies all the requirements (9)-(13). We now state our result in the form of a theorem:

Theorem 2: Let L be the largest index $i \in \{1,\ldots,n\}$ that satisfies the inequality

$$\frac{P_0}{\eta B} \geq \frac{\frac{r_{i-1}}{h_i} - \frac{r_i}{h_{i-1}}}{r_i - r_{i-1}},$$

where the sequence on the right hand side is increasing, by Assumption 1. Also let

$$a_i = \begin{cases} \frac{1}{\lambda}\cdot\frac{r_i - r_{i-1}}{\frac{1}{h_i}-\frac{1}{h_{i-1}}} & i=1,\ldots,L, \\ \eta B & i=L+1, \end{cases}$$

where $\lambda = \frac{r_L h_L}{P_0 h_L + \eta B}$. The maximum transport capacity is

$$C_{T,\max}^{\mathrm{MAC}} = B\sum_{i=1}^{L}(r_i - r_{i-1})\log_2(a_i) - B r_L \log_2(\eta B),$$

and the unique power distribution that achieves it is given by:

$$P_i = \begin{cases} \frac{a_i - a_{i+1}}{h_i} & i=1,\ldots,L, \\ 0 & i=L+1,\ldots,n. \end{cases} \tag{19}$$

Our result is conceptually straightforward: The available power should be distributed among the first L transmitters, and, the greater the available power, the larger L becomes. Therefore, the solution resembles a water filling that starts from near the receiver and moves outwards. The precise allocation of power among the first L users will depend on the exact shape of the reward and gain functions.

As a numerical example, let us consider a multiple access channel that consists of a single receiver, placed at the origin, and 200 transmitters, placed uniformly along the x-axis with a separation of 25 m from each other, starting at 25 m from the receiver. The total power available is $P_0 = 400$ W, the available bandwidth $B = 10$ MHz, the noise spectral density is $\eta = 10^{-16}\ \frac{\mathrm{W}}{\mathrm{Hz}}$, and a monomial power gain function with $\gamma = 3$ and $K_h = 0.1\ \mathrm{m}^3$ is assumed. In Fig. 3, we plot the optimal distribution of powers, assuming a monomial reward function with $K_r = 1$, and for the cases $\rho = 1$, $\rho = 1.5$, and $\rho = 2$. In Fig. 4, we plot the corresponding distributions of the rate-reward products of the individual transmitters. Finally, in Fig. 5 we compare the optimal power allocation for the case $\rho = 1$, with the allocation induced if all the nodes outside the intervals $[200\ \mathrm{m}, 900\ \mathrm{m}]$ and $[1300\ \mathrm{m}, 1600\ \mathrm{m}]$ are removed.

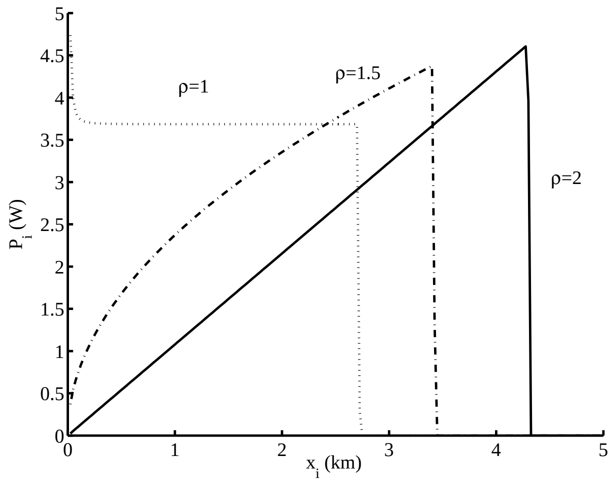

Fig. 3. The optimal power allocation in a multiple access channel consisting of a receiver placed in the origin, and 200 transmitters placed uniformly along the x-axis with a separation of $25\ \mathrm{m}$ from each other, and for various values of the reward exponent ρ.

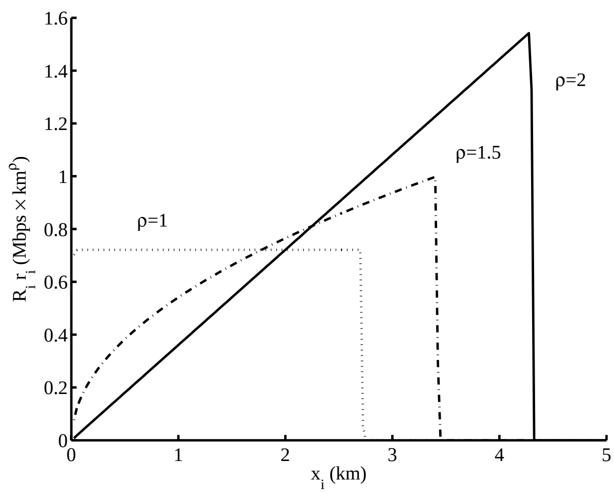

Fig. 4. The distribution of rate-reward products induced by the power distribution of Fig. 3.

As can be seen from the figure, the nodes that lie directly on the borders of the 'forbidden regions' take for themselves most of the power that was allocated to the nodes that were removed. The powers allocated to the rest of the nodes also change, and in fact in the same proportion, through the change in the value of λ.

E. Large Number of Transmitters

Let us now consider the case where the number of transmitters is very large, ideally approaching infinity. Our aim is to suppress the effects of the particular node placements, and draw better intuition about the inherent capabilities of the multiple access channel. Formally, we assume that a large number n of transmitters are placed in the interval $[a, b]$, where $a \geq 0$ and $b < \infty$, with $n \to \infty$. We also require that the distance between any point $x \in [a, b]$ and its nearest transmitter goes to zero.

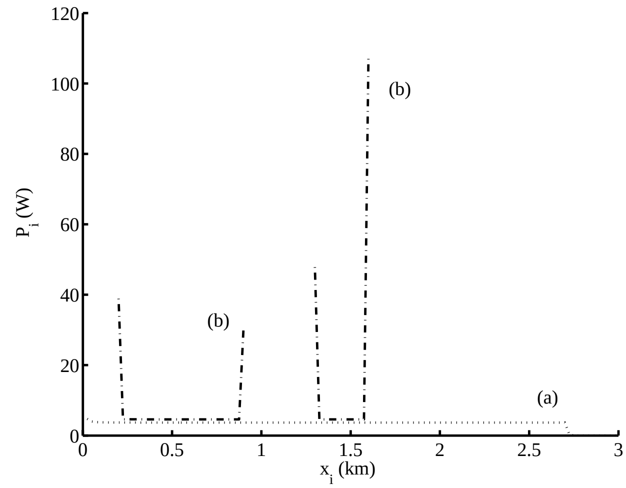

Fig. 5. **(a)** The optimal power distribution of the channel of Fig. 3, for the case $\rho = 1$. **(b)** The optimal power allocation in the channel of Fig. 3, for the case $\rho = 1$, and if all the nodes outside the intervals $[200\ \mathrm{m}, 900\ \mathrm{m}]$ and $[1300\ \mathrm{m}, 1600\ \mathrm{m}]$ are removed.

Clearly, it no longer makes sense to discuss in terms of the powers allocated to individual nodes, since the power allocated to almost all of them will have to converge to 0, but rather in terms of the **power density function** $p(x)$, with $x \in [a, b]$, defined such that the power allocated to the transmitters lying in the set $A \subset [a, b]$ converges to $\int_A p(x)\, dx$. Our sum-power constraint becomes $\int_a^b p(x)\, dx = P_0$.

As the distance between any point $x \in [a, b]$ and its nearest transmitter goes to zero, by the standard theory of Riemann integrals, we have that the transport capacity, in the form of the finite summation of (2), will converge to the following Riemann integral:

$$B \int_a^b r'(x) \log_2[\eta B + \int_{x^-}^b h(t)p(t)\, dt]\, dx - Br(b) \log_2(\eta B).$$

Theorem 3: The shape of the optimal power allocation $p^{\mathrm{MAC}}(x)$, $a \leq x \leq b$, will depend on the ratio $\frac{P_0}{\eta B}$. In particular[3]*:*

(i) *If $\frac{P_0}{\eta B} < g(a)$, then $p^{\mathrm{MAC}}(x) = P_0 \delta(x - a)$.*
(ii) *If $g(a) \leq \frac{P_0}{\eta B} \leq g(b)$, then*

$$p^{\mathrm{MAC}}(x) = \left[\frac{P_0 + \frac{\eta B}{h(x_c)}}{r(x_c)}\right] \times \left[l(a)g(a)\delta(x-a) - \boldsymbol{I}_{[a,x_c]} \frac{l'(x)}{h(x)}\right], \tag{20}$$

*where the **cutoff point** x_c satisfies the equation:*

$$\frac{P_0}{\eta B} = g(x_c). \tag{21}$$

[3]Note that function $g(\cdot)$ was defined in Lemma 1(iv).

(iii) *If* $g(b) < \frac{P_0}{\eta B}$, *then*

$$p^{\rm MAC}(x) = \left[l(a)g(a)\frac{P_0 + \frac{\eta B}{h(b)}}{r(b)} \right] \delta(x-a) + \frac{l(b)}{r(b)h(b)} \left[P_0 - g(b)\eta B\right] \delta(x-b) - \left[\frac{P_0 + \frac{\eta B}{h(b)}}{r(b)} \right] \frac{l'(x)}{h(x)}. \quad (22)$$

As expected, the water filling structure of the discrete case is maintained. Theorem 3 can be proved in a straightforward manner by starting from Theorem 2 and taking the appropriate limits. However, in Appendix III, we prove it starting from scratch, using calculus of variations. There are two reasons for this approach. Firstly, the calculus of variations proof is very short and gives additional intuition which is obscured by the heavily algebraic nature of the proof of the finite transmitter case. Secondly, it uses the most-rapid-approach method, which is very powerful and might be applicable in other problems of the same nature, and therefore might be of independent interest.

As an illustrative application of the theorem, let us consider the case of the monomial reward and gain functions, with $a = 0$, and b large enough so that $\frac{P_0}{\eta B} < g(b)$. After straightforward substitutions we have:

$$p^{\rm MAC}(x) = \left\{ \left[\eta B(\gamma-\rho)\right]^{\frac{\rho}{\gamma}} (\rho P_0)^{\frac{\gamma-\rho}{\gamma}} K_h^{-\frac{\rho}{\gamma}} \right\} x^{\rho-1} \mathbf{I}_{[0,x_c]},$$

$$x_c = \left[\left(\frac{K_h P_0}{\eta B} \right) \left(\frac{\rho}{\gamma-\rho} \right) \right]^{\frac{1}{\gamma}}, \quad (23)$$

$$C^{\rm MAC}_{T,\max} = \frac{K_r B}{\log 2} \left[\frac{K_h P_0}{\eta B} \right]^{\frac{\rho}{\gamma}} (\frac{\gamma}{\rho} - 1)^{1-\frac{\rho}{\gamma}}. \quad (24)$$

It is interesting to compare the transport capacity with $C^{1-1}_{T,\max}$, the rate-reward product of a single transmitter-receiver pair, separated by the distance $x_{\rm opt}$ that maximizes it. $C^{1-1}_{T,\max}$ is calculated in Appendix II, and is given by (36). Combining that equation with (24) shows that the quotient

$$\frac{C^{\rm MAC}_{T,\max}}{C^{1-1}_{T,\max}} = \frac{(\frac{\gamma}{\rho} - 1)^{1-\frac{\rho}{\gamma}}}{\left[e^{g(z_0)} - 1\right]^{-\frac{\rho}{\gamma}} g(z_0)} \triangleq H(\frac{\rho}{\gamma})$$

is only a function of $\frac{\rho}{\gamma}$. As shown in Fig. 6, $H(\frac{\rho}{\gamma})$ is a strictly decreasing, convex function with $\lim_{\frac{\rho}{\gamma}\to 0^+} H(\frac{\rho}{\gamma}) = e$ and $\lim_{\frac{\rho}{\gamma}\to 1^-} H(\frac{\rho}{\gamma}) = 1$. It is interesting to note that the gains of receiving from multiple transmitters at the same time, versus receiving from a single transmitter, albeit placed at the optimal distance, are rather limited, for example only around 25% when $\rho/\gamma = 0.5$.

IV. The Broadcast Channel

We now turn our attention to the Gaussian broadcast channel of Fig. 1, that consists of a transmitter T with total power P_0, and n receivers $V_1, V_2, \ldots, V_n$, placed at increasing distances $0 < x_1 < x_2 < \ldots < x_n$ from the transmitter (but not necessarily along a straight line, or even on the same plane). As with the multiple access channel of Fig. 2, we assume that if the transmitter sends a signal with power p, the signal

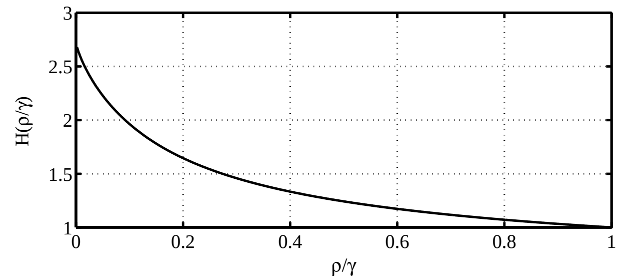

Fig. 6. The function $H(\frac{\rho}{\gamma})$, that represents the gains of using multiple transmitters over using a single transmitter, placed at the optimal distance $x_{\rm opt}$, in the case of monomial gain and reward functions.

will arrive at receiver V_i with power $h(x_i) \times p$, where $h(\cdot)$ is the gain function. A receiver placed at distance x from the transmitter is susceptible to additive white Gaussian noise of spectral density $\eta(x)$, and for simplicity we set $\eta_i \triangleq \eta(x_i)$. The total bandwidth available for communication is B.

The capacity region $\mathcal{C}^{\rm BC}$, i.e., the set of all combinations of rates $\mathbf{R} = (R_1, R_2, \ldots, R_n)$ with which the transmitter can simultaneously send data to the receivers, is known [8]. In particular, let $\pi(\cdot)$ be a permutation function that gives an ordering of the receivers in terms of increasing channel qualities, i.e., $\frac{h_{\pi(1)}}{\eta_{\pi(1)}} \le \frac{h_{\pi(2)}}{\eta_{\pi(2)}} \le \ldots \le \frac{h_{\pi(n)}}{\eta_{\pi(n)}}$. Clearly, if the $\frac{h_i}{\eta_i}$ are distinct there is only one such ordering. Then:

$$\mathcal{C}^{\rm BC} = \Bigg\{ \mathbf{R} : R_i \le B\log_2(1 + \frac{h_i P_i}{\eta_i B + h_i \sum_{j:\pi^{-1}(j)>\pi^{-1}(i)} P_j},\ i = 1,\ldots,n, \sum_{j=1}^{n} P_j = P_0,\ P_i \ge 0,\ i = 1,\ldots,n \Bigg\}. \quad (25)$$

To achieve a point $\mathbf{R} = (R_1, R_2, \ldots, R_n)$ in the capacity region, the transmitter encodes with rate R_i the message intended for receiver V_i independently of the others, and transmits it with power P_i, simultaneously with the signals intended for all other receivers. Each receiver will start to successively decode the signals for each of the receivers, in the order specified by $\pi(\cdot)$, (i.e., in the order of increasing channel quality) stopping after it decodes its own signal. When a receiver decodes a signal, other signals that have been already decoded do no create any interference, but the rest of the signals appear as thermal noise. Although alternative decoding orders are also acceptable, they will not in general attain points on the boundary of the capacity region.

Let the **normalized gain function** $h^n(x)$ be defined as

$$h^n(x) \triangleq \frac{\eta}{\eta(x)} h(x),$$

where η is arbitrary, and let $h^n_i \triangleq h^n(x_i)$. By inspecting (25), it is clear that the capacity region is identical to the capacity region of the broadcast channel with the normalized gain function, and in which all receivers are susceptible to thermal noise of a common spectral density η:

$$\mathcal{C}^{\rm BC} = \Bigg\{ \mathbf{R} : R_i \leq$$

$$B \log_2(1 + \frac{(\frac{h_i \eta}{\eta_i}) P_i}{\eta B + (\frac{h_i \eta}{\eta_i}) \sum_{j:\pi^{-1}(j) > \pi^{-1}(i)} P_j},\ i = 1, \ldots, n,$$

$$\sum_{j=1}^{n} P_j = P_0,\ P_i \geq 0,\ i = 1, \ldots, n \Bigg\}. \quad (26)$$

Similar to the MAC, we define the transport capacity of the BC, associated with a point in the capacity region $\mathbf{R} = (R_1, R_2, \ldots, R_n)$ as:

$$C_T^{\rm BC}(\mathbf{R}) \triangleq \sum_{i=1}^{n} r_i R_i, \quad (27)$$

where $r_i \triangleq r(x_i)$, and $r(\cdot)$ is the reward function. To calculate the point $\mathbf{R}$ in the capacity region that maximizes the transport capacity, in principle we could start from scratch, for example using the KKT conditions, as in Section III. However, the following theorem (Theorem 1 of [9]) allows us to use the results of the previous section on the MAC:

Theorem 4: Consider the following dual *channels:*

1) *A broadcast channel in which the power gains between the transmitter and the n receivers are $\boldsymbol{h} = (h_1, h_2, \ldots, h_n)$, and the receivers are susceptible to thermal noise with a common spectral power η. Let $\mathcal{C}^{\rm BC}(P_0; \boldsymbol{h})$ be its capacity region when the power available to the transmitter is P_0.*
2) *A multiple access channel in which the power gains between the receiver and the n transmitters are also $\boldsymbol{h} = (h_1, h_2, \ldots, h_n)$, and the receiver is susceptible to thermal noise, also of spectral power η. Let $\mathcal{C}^{\rm MAC}(\boldsymbol{P}; \boldsymbol{h})$ be its capacity region assuming that the powers available to the transmitters are $\boldsymbol{P} = (P_1, P_2, \ldots, P_n)$.*

The capacity region of the BC is equal to the union of the capacity regions of the dual MAC over all power distributions $(P_1, \ldots, P_n)$ such that $\boldsymbol{1} \cdot \boldsymbol{P} \triangleq \sum_{i=1}^{n} P_i = P_0$:

$$\mathcal{C}^{\rm BC}(P_0; \boldsymbol{h}) = \bigcup_{\{\boldsymbol{P}: \boldsymbol{1} \cdot \boldsymbol{P} = P_0\}} \mathcal{C}^{\rm MAC}(\boldsymbol{P}; \boldsymbol{h}).$$

Furthermore, let $\boldsymbol{R}$ be a point in the capacity region of the MAC achieved if the receiver decodes the incoming signals in the decoding order $\pi(1), \pi(2), \ldots$ and the powers available to the transmitters are $(P_1^{\rm MAC}, P_2^{\rm MAC}, \ldots, P_n^{\rm MAC})$. The dual BC will achieve the same point in the capacity region if each receiver decodes the incoming signals in the inverse *order $\pi(n), \pi(n-1), \ldots$, stopping after it decodes its own signal, and the powers $(P_1^{\rm BC}, P_2^{\rm BC}, \ldots, P_n^{\rm BC})$ allocated to the individual signals are given by:*

$$P_{\pi(i)}^{\rm BC} = P_{\pi(i)}^{\rm MAC} \frac{\eta B + h_{\pi(i)} \sum_{j=1}^{i-1} P_{\pi(j)}^{\rm BC}}{\eta B + \sum_{j=i+1}^{n} h_{\pi(j)} P_{\pi(j)}^{\rm MAC}},\ i = 1, \ldots, n. \quad (28)$$

Using this result, the following theorem easily follows:

Theorem 5: Assume that the function $l^n(x) \triangleq \frac{r'(x)}{(\frac{1}{h^n(x)})'}$ is decreasing. Let L be the largest index $i \in \{1, \ldots, n\}$ that satisfies the inequality

$$\frac{P_0}{\eta B} \geq \frac{\frac{r_{i-1}}{h_i^n} - \frac{r_i}{h_{i-1}^n}}{r_i - r_{i-1}},$$

and let

$$\beta_i = \begin{cases} \eta B \frac{\frac{r_i}{h_{i+1}^n} - \frac{r_{i+1}}{h_i^n}}{r_{i+1} - r_i} & i = 0, \ldots, L-1, \\ P_0, & i = L, \ldots, n. \end{cases}$$

The maximum transport capacity of the BC is:

$$C_{T,\max}^{\rm BC} = B \sum_{i=1}^{L} r_i \log_2 \left(\frac{\frac{\eta B}{h_i^n} + \beta_i}{\frac{\eta B}{h_i^n} + \beta_{i-1}} \right). \quad (29)$$

The optimal power allocation that achieves it is:

$$P_i = \begin{cases} \beta_i - \beta_{i-1}, & i = 1, \ldots, L, \\ 0, & i = L+1, \ldots, n. \end{cases} \quad (30)$$

Furthermore, each receiver V_i should decode first the signal intended for V_n, then the signal intended for V_{n-1}, etc., eventually decoding its own signal.

Proof: As discussed, the capacity region of the original BC, given by (25), is identical to the capacity region of the modified BC with the normalized gains $h_i^n \triangleq \frac{\eta}{\eta_i} h_i$ and the common spectral power density η for all receivers. That capacity region is given by (26). In addition, by Theorem 4 the points in this capacity region are exactly the points that can be achieved by its dual MAC under a sum power constraint. Therefore, maximizing the transport capacity of the BC is the same as maximizing the transport capacity of the dual MAC under a sum power constraint. For this problem, we can use Theorem 2.

To apply Theorem 2, we need to ensure that Assumption 1 holds. This translates to the requirement that the function $l^n(x)$ is decreasing.

By Theorems 1 and 2, in the dual MAC, the set of rates that maximizes the transport capacity is achieved by the decoding order in which the signal from transmitter T_i is decoded i-th, and the power allocation given by (19). By Theorem 4, in the BC, the same set of rates is achieved by the inverse decoding order, (i.e., receivers decode the signal intended for V_n, then the signal intended for V_{n-1}, etc.) and for a power allocation that is uniquely specified by (28) with $\pi(i) = i$. To prove that the optimal power allocation is given by (30), we simply substitute (30) and (19) in (28) (with $\pi(i) = i$) and we arrive at an identity, for all $i = 1, \ldots, n$.

To prove (29), we note that, since receiver V_i decodes signals intended for receivers further away first, and then its own, only the power of signals intended for receivers $V_1, \ldots, V_{i-1}$ will affect the decoding of its own signal. Therefore, the rate R_i will be

$$R_i = B \log_2(1 + \frac{(\frac{h_i \eta}{\eta_i}) P_i}{\eta B + (\frac{h_i \eta}{\eta_i}) \sum_{j=1}^{i-1} P_j}).$$

Combining these with (27) and (30), (29) follows. □

The structure of the solution is straightforward: The power of the transmitter P_0 is divided among the first L receivers, and the larger P_0 is, the larger L will be. Therefore, as in the case of the MAC, the solution resembles a water filling from left to right. In addition, the optimal distribution of rate-reward products is exactly the same as that of the dual MAC. On the other hand, as the following numerical example shows, the optimal power allocations of the dual BC and MAC will in general be different.

As a numerical example, let us consider the dual BC of the MAC of Section III: it consists of a single transmitter, placed at the origin, and 200 receivers, placed uniformly along the x-axis with a separation of $25\ \mathrm{m}$ from each other, starting at $25\ \mathrm{m}$ from the transmitter. The total power available at the transmitter is $P_0 = 400\ \mathrm{W}$, the available bandwidth $B = 10\ \mathrm{MHz}$, the noise spectral density of all receivers is $\eta = 10^{-16}\ \frac{\mathrm{W}}{\mathrm{Hz}}$, and a monomial power gain function with $\gamma = 3$ and $K_h = 0.1\ \mathrm{m}^3$ is assumed. In Fig. 7 we plot the optimal distribution of powers, assuming a monomial reward function with $K_r = 1$, and for the cases $\rho = 1$, $\rho = 1.5$, and $\rho = 2$. Note that the power allocations have a different shape from those of the dual MAC. On the other hand, by duality, the rate-reward products of the BC will be in all cases the same of the rate-reward products of the dual MAC, which are plotted in Fig. 4. Finally, in Fig. 8 we compare the optimal power allocation for the case $\rho = 1$, with the allocation induced if the nodes lying outside the intervals $(600\ \mathrm{m}, 900\ \mathrm{m})$ and $(1300\ \mathrm{m}, 2650\ \mathrm{m})$ are removed. As can be seen from Fig. 8, and also from from Theorem 5, the nodes that lie directly on the borders of the 'forbidden regions' take for themselves *all* the power that was allocated to the nodes that were removed. Contrary to the multiple access case, the powers allocated to the rest of the nodes do not change at all.

Similarly to the MAC, we would like to consider the case where a large number n of receivers is placed in the interval $[a, b]$. The following theorem is proved in Appendix III:

Theorem 6: Assume that the function $l^n(x) \triangleq \frac{r'(x)}{(\frac{1}{h^n(x)})'}$ *is decreasing, in which case by Lemma 1(iv) the function*

$$g^n(x) \triangleq \frac{(\frac{1}{h^n(x)})'r(x) - (\frac{1}{h^n(x)})r'(x)}{r'(x)}$$

is increasing. The shape of the optimal power allocation $p^{\mathrm{BC}}(x)$ *will depend on the ratio* $\frac{P_0}{\eta B}$*. In particular:*

(i) *If* $\frac{P_0}{\eta B} < g^n(a)$*, then* $p^{\mathrm{BC}}(x) = P_0\delta(x - a)$*.*

(ii) *If* $g^n(a) \leq \frac{P_0}{\eta B} \leq g^n(b)$*, then*

$$p^{\mathrm{BC}}(x) = g^n(a)\delta(x - a) + \boldsymbol{I}_{[a,x_c]}(g^n)'(x), \qquad (31)$$

where the ***cutoff point*** x_c *satisfies the equation:*

$$\frac{P_0}{\eta B} = g^n(x_c). \qquad (32)$$

(i) *If* $g^n(b) < \frac{P_0}{\eta B}$*, then*

$$p^{\mathrm{BC}}(x) = g^n(a)\delta(x-a) + [g^n(b) - P_0]\delta(x-b) + (g^n)'(x). \qquad (33)$$

Furthermore, each receiver should perform successive decoding of the received signals, starting from the signal intended for the far-most receiver, and moving inwards, until it decodes its own signal.

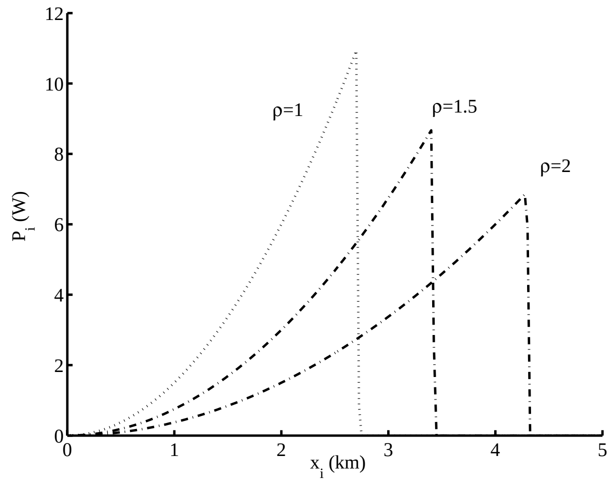

Fig. 7. The optimal power allocation in a broadcast channel consisting of a transmitter placed in the origin, and 200 receivers placed uniformly along the x-axis with a separation of $25\ \mathrm{m}$ from each other, and for various values of the reward exponent ρ.

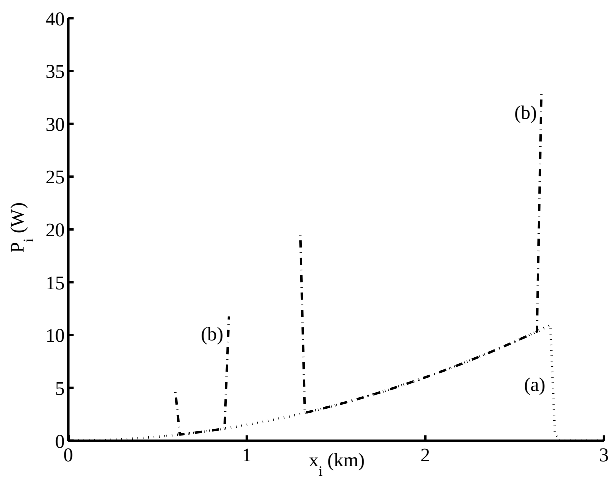

Fig. 8. **(a)** The optimal power distribution of the channel of Fig. 7, for the case $\rho = 1$. **(b)** The optimal power allocation in the channel of Fig. 7, for the case $\rho = 1$, and if all the nodes outside the intervals $[300\ \mathrm{m}, 900\ \mathrm{m}]$ and $[1300\ \mathrm{m}, 2650\ \mathrm{m}]$ are removed.

Let us consider the case of a BC with monomial reward and gain functions, where $\eta(x) \triangleq \eta$, $a = 0$, and b is large enough so that $\frac{P_0}{\eta B} < g^n(b)$. By duality, it follows that x_c will be given by (23), as in the MAC, and $C^{\mathrm{BC}}_{T,\max} = C^{\mathrm{MAC}}_{T,\max}$, with $C^{\mathrm{MAC}}_{T,\max}$ given by (24). The optimal power allocation, however, will be different. Applying Theorem 6, we have that

$$p^{\mathrm{BC}}(x) = \left[\frac{\eta B(\gamma - \rho)\gamma}{\rho K_h}\right] x^{\gamma-1}\mathbf{I}_{[0,x_c]}.$$

Therefore, the optimal power allocation changes with the distance x as $x^{\rho-1}$ for the MAC, but as $x^{\gamma-1}$ for the BC.

V. Discussion and Conclusions

In this work we study the transport capacity of a Gaussian Multiple Access Channel (MAC), which consists of a single receiver and multiple transmitters, and a Gaussian Broadcast Channel (BC), which consists of a single transmitter and multiple receivers. The transport capacity is defined as the summation, over all simultaneous transmissions, of the product of the data rate with a reward $r(x)$ which is a function of the distance x that the data covers. In the special case where $r(x) = x$, the transport capacity is measured in $\mathrm{bps}{\times}\mathrm{m}$, and is a natural figure of merit of the efficiency with which a network is operating at a given time instant. However, the general form of $r(\cdot)$ allows for alternative notions of usefulness of a transmission.

In the case of the MAC, we first show that the receiver should decode the incoming signals starting from the signal of the nearest transmitter and moving outwards. Under a sum-power constraint, we determine in closed form the optimal allocation of transmitter powers that maximizes the transport capacity. Our proof is conceptually straightforward and short in length, and is based on KKT conditions.

In the case of the BC, we calculate in closed form the optimal distribution of the transmitter power to the signals of the different receivers. The proof is very simple, and is based on our results for the MAC, and duality arguments that became available only recently [9].

We also present asymptotic results, that only hold as the number of transmitters (for the MAC) or receivers (for the BC) go to infinity. The results are derived using the most-rapid-descent method of the calculus of variations.

We must emphasize that a closed form solution for the optimal power allocation of the BC, under an assumption very similar to our assumption that the function $h^n(x)$ is decreasing, and compatible to our result, appeared first in [7]. There, however, it was shown that the transmitter may allocate power to a contiguous group of receivers, and perhaps one or two extra, outlying receivers. Therefore, there may be inactive receivers (i.e. receivers that receive no power) separating the contiguous group from the outlying receivers. Our work shows that in fact only a much less general scenario can occur, i.e., the first L receivers will receive all the power, for some L, and the rest will receive no power at all. In addition, our derivations are much shorter, because we are using duality arguments and our results for the MAC (which are also shorter than the derivations in [7]).

Our work brings forward a number of aspects of the MAC and the BC that designers of practical systems may want to take into consideration:

First of all, in both the MAC and the BC the optimal power allocation resembles a water filling: Only the signals coming from the nearest L transmitters (in the case of the MAC) or going to the nearest L receivers (in the case of the BC) are allocated positive power, and the larger the available power is, the larger L will be.

Secondly, if the MAC and the BC are dual, then the transport capacities and the rate-reward distributions are identical, but not the power allocations. For example, in a cellular system in which the uplink MAC and the downlink BC are duals, there is a very nice coupling in the sense that the maximization of the transport capacities of the uplink and the downlink also ensures that each user will be transmitting in the uplink with the same rate with which it will be receiving in the downlink, but the two power allocations that must be used will be different.

Thirdly, as Theorems 2 and 5 show, in order to maximize the transport capacity of the MAC (BC), a lot of power must be allocated to nodes (transmitters or receivers) that sparsely populate the same distance range. For example, in the special case where the nodes are placed along a straight line, as Figs. 5 and 8 show, a lot of power must be allocated to nodes that are neighboring areas where no other nodes are placed.

Finally, the transport capacity of both the MAC and the BC may only be marginally better than the transport capacity of a simple transmitter-receiver pair, provided the distance between the two can be optimized. For example, in the case of monomial reward and gain functions with $\frac{\rho}{\gamma} = \frac{1}{2}$, the gains by using successive decoding, in terms of transport capacity, are around 25%. Given the complexity of receivers that employ successive decoding, it is clear that in certain situations using successive decoding may not be worth the investment.

It should be stressed that all our results crucially depend on Assumption 1, which essentially requires that, as the distance x between a transmitter and a receiver increases, the reward $r(x)$ is not increasing as fast as the channel gain $h(x)$ is decreasing. As discussed, this is a reasonable assumption for most cases of interest.

Note that, from a purely mathematical perspective, the transport capacity is simply a weighted sum, over all simultaneous transmissions, of the achieved data rates. The maximization of such a weighted sum is an interesting problem even outside the context of transport capacity. Indeed, the weights $\{r_i\}$ can be thought of as specifying a particular direction in the n-dimensional Euclidean space, therefore the maximization of the weighted sum corresponds to finding the boundary of the capacity region along this particular direction. This is actually an old problem that has been studied by a number of different researchers. (See, for example, [14], [10], [15], [16] and the references therein.) As all these works predate the duality results of [9], in all cases the authors either examine the BC or the MAC. To the best of our knowledge, in all cases the optimal power allocation is not given in closed form, but rather an algorithm is provided that can be used to determine it. The novelty of our work is that we adopt Assumption 1, which is very reasonable if the weighted sum is interpreted as the transport capacity. This leads to very short derivations, and the determination of the power allocation in closed form. An additional novelty of our work is that we use duality arguments to derive the results for the BC using the results of the MAC.

Acknowledgment

The authors would like to thank Prof. S. Vishwanath, of the University of Texas, for pointing to [13], which contains the idea of Theorem 1 in a different context, and Prof. N. Jindal, of the University of Minnesota, for suggesting the use of duality

arguments for the derivation of the optimal transport capacity of the BC.

Appendix I
Proof of Lemma 1

(i) To prove the equivalence of (5) and (6), we cross-multiply both of them and we arrive at identical inequalities.

Next, we show that (5) implies (4). If (5) holds for all transmitter placements, it will then hold for the case of four transmitters placed at locations $x-\epsilon, x, y-\epsilon, y$, where $\epsilon < y-x$. By applying (5) twice, we have:

$$\frac{r(x)-r(x-\epsilon)}{\frac{1}{h(x)}-\frac{1}{h(x-\epsilon)}} \geq \frac{r(y-\epsilon)-r(x)}{\frac{1}{h(y-\epsilon)}-\frac{1}{h(x)}} \geq \frac{r(y)-r(y-\epsilon)}{\frac{1}{h(y)}-\frac{1}{h(y-\epsilon)}}.$$

By dividing the numerators and denominators of the left and right hand side by ϵ, and taking $\epsilon \to 0$, we find that $\frac{r'(x)}{(\frac{1}{h})'(x)} \geq \frac{r'(y)}{(\frac{1}{h})'(y)}$.

The proof that (4) implies (5) follows similarly to Lemma 1 of [7], and so is omitted.

(ii) We use induction. In particular, we prove (7) first for $k=1$ and for all $i=1,\dots,n-1$. For this, we note that, from (5), we have that

$$\frac{r_{i+1}-r_i}{r_i-r_{i-1}} \leq \frac{\frac{1}{h_{i+1}}-\frac{1}{h_i}}{\frac{1}{h_i}-\frac{1}{h_{i-1}}}, \quad i=1,\dots,n-1,$$

Adding 1 to each size, simplifying and rearranging terms, we arrive at (7) for $k=1$ and for all $i=1,\dots,n-1$. We now make the inductive hypothesis that (7) holds for some $k \in 1,\dots,n-2$, and for all $i=1,\dots,n-k$. We will show that it then holds for $k+1$, and for all $i=1,\dots,n-k-1$. For this, we note that

$$\frac{r_i-r_{i-1}}{\frac{1}{h_i}-\frac{1}{h_{i-1}}} \geq \frac{r_{i+1}-r_i}{\frac{1}{h_{i+1}}-\frac{1}{h_i}} \geq \frac{r_{i+k+1}-r_i}{\frac{1}{h_{i+k+1}}-\frac{1}{h_i}}, \; i=1,\dots,n-k-1.$$

The first inequality comes from (5), and the second from the induction hypothesis. Therefore,

$$\frac{r_{i+k+1}-r_i}{r_i-r_{i-1}} \leq \frac{\frac{1}{h_{i+k+1}}-\frac{1}{h_i}}{\frac{1}{h_i}-\frac{1}{h_{i-1}}}, \quad i=1,\dots,n-k-1.$$

By adding 1 to each side, simplifying and rearranging terms, we arrive at (7) for $k+1$, and for all $i=1,\dots,n-k-1$.

(iii) To prove the inequality, we apply (5) for $x_{i-1}=0$, $x_i=x$, $x_{i+1}=x+\epsilon$, and then take $\epsilon \to 0$.

(iv) If Assumption 1 holds, then (6) follows for all transmitter placements, and in particular for the case of four transmitters placed at locations $x-\epsilon, x, y-\epsilon, y$, where $\epsilon < y-x$:

$$\frac{\frac{r(x-\epsilon)}{h(x)}-\frac{r(x)}{h(x-\epsilon)}}{r(x)-r(x-\epsilon)} \leq \frac{\frac{r(x)}{h(y-\epsilon)}-\frac{r(y-\epsilon)}{h(x)}}{r(y-\epsilon)-r(x)} \leq \frac{\frac{r(y-\epsilon)}{h(y)}-\frac{r(y)}{h(y-\epsilon)}}{r(y)-r(y-\epsilon)}.$$

Dropping the middle part, it follows that

$$\frac{r(x-\epsilon)h(x-\epsilon)-r(x)h(x)}{[r(x)-r(x-\epsilon)]h(x)h(x-\epsilon)} \leq \frac{r(y-\epsilon)h(y-\epsilon)-r(y)h(y)}{[r(y)-r(y-\epsilon)]h(y)h(y-\epsilon)}.$$

Taking the limit $\epsilon \to 0$, we arrive at $g(x) \leq g(y)$.

Appendix II
Optimal separation of a transmitter-receiver pair

In this appendix we calculate the optimal separation between a single transmitter and a single receiver, that maximizes the reward-distance product, assuming monomial reward and gain functions. This optimization problem was first considered in [7]. There, however, only an asymptotic analysis as the gain exponent $\gamma \to \infty$ was offered.

Let a transmitter T and a receiver R be separated by a distance x which is allowed to vary. The transmitter power is P_0, the bandwidth available for the communication is B, and the receiver is susceptible to additive white Gaussian noise of density η. We assume that the signal power changes with distance according to the gain function $h(x)$, and the transmission of a bit of information over a distance x is rewarded by a reward $r(x)$. Finally, we assume that the channel between the transmitter and receiver operates at the Shannon capacity $C = B\log_2(1+\frac{P_0 h(x)}{\eta B})$. We define the transport capacity $C_T^{1-1}(x)$ of this setting as the reward-distance product:

$$C_T^{1-1}(x) \triangleq r(x) B \log_2(1+\frac{P_0 h(x)}{\eta B}).$$

We are interested in determining the maximum possible value for $C_T^{1-1}(x)$, $C_{T,\max}^{1-1} \triangleq \sup_{0<x<\infty} C_T(x)$. Clearly, unless specific cases for $r(\cdot)$ and $h(\cdot)$ are considered, we can not go much further. So let us limit the discussion to the monomial reward and gain functions. In this case,

$$C_{T,\max}^{1-1} = \sup_{0<x<\infty} BK_r x^\rho \log_2(1+\frac{K_h P_0}{\eta B x^\gamma}).$$

Let $A \triangleq \frac{K_h P_0}{\eta B}$ and $f(x) \triangleq x^\rho \log_2(1+\frac{A}{x^\gamma})$, so that

$$C_{T,,\max}^{1-1} = BK_r \sup_{0<x<\infty} f(x).$$

When $\gamma < \rho$, clearly $\lim_{x\to\infty} f(x) = \infty$, so that $C_{T,\max}^{1-1} = \infty$. When $\gamma = \rho$, $f(x)$ is monotonically increasing so that its supremum is approached as $x \to \infty$, and $C_{T,\max}^{1-1} = ABK_r \log_2(e)$. However, as in the main text, we are mostly interested in the case $\gamma > \rho$. In this case, $f(x)$ achieves a single maximum for an optimum value of x, x_{opt}.

To find x_{opt}, we set the derivative of $f(x)$ equal to 0:

$$\frac{\rho}{\gamma}\log(1+\frac{A}{x^\gamma}) - \frac{A}{A+x^\gamma} = 0.$$

We make the substitution

$$y = \log(1+\frac{A}{x^\gamma}) - \frac{\gamma}{\rho} \tag{34}$$

and after simplifying we arrive at:

$$ye^y = (-\frac{\gamma}{\rho})e^{(-\frac{\gamma}{\rho})} \triangleq z_0. \tag{35}$$

This equation is of the form $ye^y = z_0$ where z_0 is given and we must find y. In other words, solving (35) is equivalent to calculating the inverse of the function $y \mapsto z = ye^y$. This is a very old problem, actually predating Euler, who has himself worked on it [17]. The inverse is known in the literature as Lambert's W function, and it appears often in a variety of

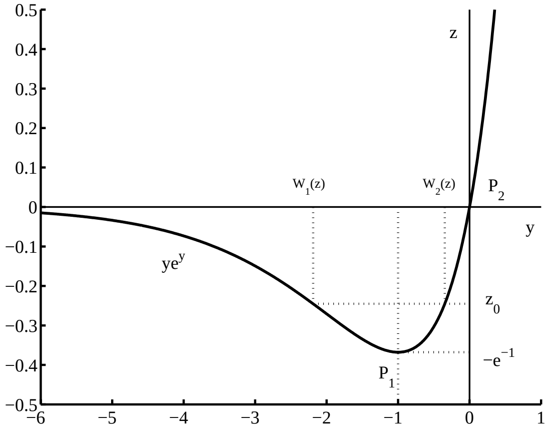

Fig. 9. The plot of the function ye^y.

situations from the enumeration of trees in graph theory to the calculation of wave heights in physics [18].

In general, $W(z)$ is defined for complex z and is complex and multivalued. In our context, however, both z and $W(z)$ are real, and the situation is relatively simple. In Fig. 9 we plot $z = ye^y$ for real y. From the figure it is clear that for $z \geq 0$ $W(\cdot)$ has a single branch, the curve on the right of point P_2. For $z < -e^{-1}$ it has no branches, and for $-e^{-1} \leq z < 0$ it has two branches, $W_1(\cdot)$ and $W_2(\cdot)$. The branch $W_1(\cdot)$ is the curve that lies on the left of point P_1, and the branch $W_2(\cdot)$ is the curve that lies between the points P_1 and P_2. Unfortunately, no closed-form expressions are known for the two branches. We define

$$g(z) \triangleq W_2(z) - W_1(z) \qquad (z \in [-e^{-1}, 0)).$$

The value z_0 of (35) lies in the interval $(-e^{-1}, 0)$, so (35) has two solutions, one for each branch. $W_1(z_0)$ is clearly equal to $-\frac{\gamma}{\rho}$ and is not acceptable, since plugging this to (34) implies that $x = +\infty$. However $W_2(z_0)$ is acceptable, and using (34) it leads to the following solution:

$$x_{\text{opt}} = \left[\frac{A}{e^{W_2(z_0)+\frac{\gamma}{\rho}} - 1} \right]^{\frac{1}{\gamma}}.$$

Noting that $W_2(z_0) + \frac{\gamma}{\rho} = g(z_0)$ and that $A \triangleq \frac{K_h P_0}{\eta B}$, we have that:

$$x_{\text{opt}} = \left[\frac{K_h P_0}{\eta B} \right]^{\frac{1}{\gamma}} \left[\frac{1}{e^{g(z_0)} - 1} \right]^{\frac{1}{\gamma}}.$$

The optimal transport capacity becomes

$$C_{T,\max}^{1-1} = \frac{K_r B}{\log 2} \left[\frac{K_h P_0}{\eta B} \right]^{\frac{\rho}{\gamma}} \left[e^{g(z_0)} - 1 \right]^{-\frac{\rho}{\gamma}} g(z_0). \qquad (36)$$

A few interesting observations can be made from (36). For example, ρ and γ affect the transport capacity only through their quotient $\frac{\rho}{\gamma}$. Therefore, changing both their values while leaving their quotient fixed does not change the value of the transport capacity. In addition, in contrast to Shannon capacity, the dependence of the transport capacity on the available bandwidth and transmitter power is monomial.

APPENDIX III
PROOFS OF THEOREMS 3 AND 6

The fist three subsections of this appendix present a proof of Theorem 3, using calculus of variations. The last subsection sketches a proof for Theorem 6.

A. Problem Formulation

We make the technical assumption $p(x) \leq T$, where T can be arbitrarily large, but not a function of the number of nodes n. This ensures that all terms in the summation of the objective function of (1) are very small, and the objective function can be approximated by a Riemann integral. The optimization problem (1) becomes:

$$\begin{aligned} &\text{maximize:} && \tfrac{B}{\log 2} \int_a^b \tfrac{r(x)h(x)p(x)}{\eta B + \int_x^b p(t)h(t)\,dt}\, dx, \\ &\text{subject to:} && 0 \leq p(x) \leq T, \quad \int_a^b p(x)\,dx = P_0. \end{aligned} \qquad (37)$$

To derive the above, we have used the fact that $\log_2(1+x) \simeq \frac{x}{\log 2}$ for $x \to 0$.

To bring (37) to a more standard form, we set:

$$y(x) \triangleq \eta B + \int_x^b h(t)p(t)\,dt \Rightarrow p(x) = -\frac{y'(x)}{h(x)}.$$

It follows that $y(b) = \eta B$. The problem now becomes (ignoring the factor $\frac{B}{\log 2}$):

$$\begin{aligned} &\text{minimize:} && \int_a^b \tfrac{r(x)y'(x)}{y(x)}\, dx, \\ &\text{subject to:} && \begin{cases} -Th(x) \leq y'(x) \leq 0, \;\; y(b) = \eta B, \\ -\int_a^b \frac{y'(x)}{h(x)}\, dx = P_0. \end{cases} \end{aligned} \qquad (38)$$

To remove the second equality constraint, we modify the objective by subtracting the left hand side of the constraint, multiplied by a *Lagrange multiplier* λ. After the optimization is performed, λ will be chosen to satisfy the equality constraint, but until then it will be treated as yet another parameter of the problem[4]. The problem now becomes:

$$\begin{aligned} &\text{minimize:} && \int_a^b \left[\tfrac{r(x)}{y(x)} + \tfrac{\lambda}{h(x)} \right] y'(x)\, dx, \\ &\text{subject to:} && -Th(x) \leq y'(x) \leq 0, \quad y(b) = \eta B. \end{aligned} \qquad (39)$$

We will call the integrand of the objective function the *Lagrangian*, and we will denote it by $F(x, y(x), y'(x))$.

The Lagrange multiplier λ has a very simple interpretation. As shown in Theorem **[X.3]** of [19], it equals the rate of change of the minimum of (38) with respect to P_0. Therefore, although we still need to find λ, we know that it must be strictly negative. Indeed, it is easy to show that if more power is available to the transmitters, the optimal transport capacity will strictly increase, and so the minimum of the objective of (38) will strictly decrease.

Note that we have a constraint on the value of $y(x)$ on the right hand side of the interval, i.e., $y(b) = \eta B$. However, no constraint is placed on $y(a)$. Standard theory requires that the

[4]For a justification of this procedure, see any text on calculus of variations, for example [19].

following *natural*, or *Euler* boundary condition $\frac{\partial F}{\partial y'}(a) = 0$ be adopted [19]:

$$\frac{\partial F}{\partial y'}(a) = 0 \Rightarrow y(a) = -\frac{r(a)h(a)}{\lambda}.$$

The next task is to formulate the Euler Differential Equation (DE) that describes the shape of any smooth (i.e. with continuous first order derivative) $y(x)$ that minimizes the objective. The Euler DE for this problem is given by [19]:

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = \frac{\partial F}{\partial y} \Rightarrow y_s(x) = -\frac{r'(x)}{\lambda(\frac{1}{h})'(x)} = -\frac{l(x)}{\lambda}. \quad (40)$$

B. The Most-Rapid-Approach Method

In most problems in the calculus of variations, the Euler DE turns out to be a second order differential equation, that can be uniquely determined by requiring that the optimal function $y(x)$ satisfies the two boundary conditions. In our case however, the Euler DE is actually a plain algebraic equation, which contains no derivatives, and so can not be made, in general, to satisfy our boundary conditions! It follows that the solution can not be smooth everywhere.

This problem is actually an instance of a more general class of problems, in which the Lagrangian is linear with $y'(x)$ (as can readily be seen by (39)). This case is called the *singular case*, and a solution to such a problem is called a *singular solution*. Such cases are rarely discussed in introductory courses in calculus of variations [19]. For this class of problems, Theorem **[XII.5]** of [19] applies:

Theorem 7: (Most-Rapid-Approach Method) Consider the minimization of the integral $\int_a^b F(x, y(x), y'(x))\,dx$ with respect to $y(x)$, where the Lagrangian $F(x, y(x), y'(x)) = F_0(x, y(x)) + F_1(x, y(x))y'(x)$, and under the constraints $y(a) = A$, $y(b) = B$, and

$$L(x, y) \leq y'(x) \leq U(x, y) \quad (a \leq x \leq b). \quad (41)$$

Assume that the following conditions hold:

1) *The Euler DE has a unique singular solution $y_s(x)$.*
2) $\frac{\partial F_0}{\partial y} - \frac{\partial F_1}{\partial x} > 0\ (< 0)$ *if* $y(x) - y_s(x) > 0\ (< 0)$.
3) *$y_s(x)$ satisfies the inequality constraints (41).*

Then the global minimum is achieved by the composite function

$$y(x) \triangleq \begin{cases} y_a(x), & a \leq x \leq x_1, \\ y_s(x), & x_1 \leq x \leq x_2, \\ y_b(x), & x_2 \leq x \leq b. \end{cases}$$

If $y_s(a) \leq A$, then $y_a(x)$ is the uniquely defined function that starts at the point $y(a) = A$, and descends *as fast as possible toward the singular solution, satisfying at all times the equality $y'_a(x) = L(x, y_a(x))$. If, on the other hand, $y_s(a) \geq A$, then $y_a(x)$ is the uniquely defined function that starts at the point $y(a) = A$ and* ascends *as fast as possible toward the singular solution, satisfying the equality $y'_a(x) = U(x, y_a(x))$. The point x_1 is where $y_a(x)$ and $y_s(x)$ cross. The function $y_b(x)$ and the point x_2 are defined in a similar manner.*

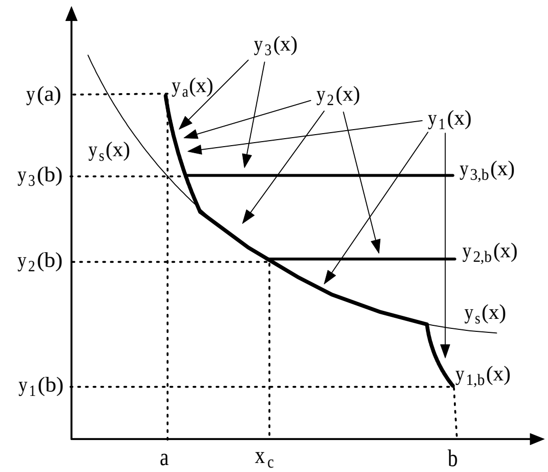

Fig. 10. The singular solution $y_s(x)$ (in thin line) and three possible forms for the composite extremal function $y(x)$ (denoted by thick line).

If, while moving toward the singular solution, functions $y_a(x)$ and $y_b(x)$ meet each other first, at some point x_3, then the optimal solution $y(x)$ is given by

$$y(x) \triangleq \begin{cases} y_a(x), & a \leq x \leq x_3, \\ y_b(x), & x_3 \leq x \leq b. \end{cases}$$

In all other cases (for example when $x_1 > x_2$, or when $y_a(x)$ or $y_b(x)$ never intersect $y_s(x)$ or each other), the problem has no solution.

The singular solution of our problem, given by (40), is unique, therefore the first condition of Theorem 7 is satisfied. It is also straightforward to show that the second condition is satisfied, by simply writing down the partial derivatives of F_0 and F_1. It remains to be shown that the third condition is also satisfied, i.e., that the singular solution satisfies the inequality constraints (41), i.e., $-Th(x) \leq [-\frac{r'(x)}{\lambda(\frac{1}{h})'(x)}]' \leq 0$. The left hand side inequality will be satisfied if we take T to be large enough. The right hand side is also satisfied, by Assumption 1 and the fact that $\lambda < 0$.

C. The Optimal Power Allocation

All the conditions of Theorem 7 are thus satisfied, and we are ready to apply it in our case. By Lemma 1(iii), it follows that $r(a)h(a) \geq \frac{r'(a)}{(\frac{1}{h})'(a)} \Rightarrow y_s(a) \leq y(a)$, i.e., the singular solution $y_s(x) = -\frac{r'(x)}{\lambda(\frac{1}{h})'(x)}$ will always pass below the left boundary condition. Regarding whether the singular solution passes below or above the right boundary condition $y(b) = \eta B$, a complication arises from the fact that we do not know the actual value of λ. Therefore, we can not tell beforehand the resulting form of the composite solution $y(x)$. Three different cases, all corresponding to different functional forms for $y(x)$, appear in Fig. 10. Note that the horizontal potions $y_{2,b}(x)$ and $y_{3,b}(x)$ of the composite solution appear in intervals where no power is allocated, i.e., $p(x) = 0$. Similarly, the steeply descending portions $y_a(x)$ and $y_{1,b}(x)$ of the composite solution appear in intervals where the power allocated is the maximum possible, i.e., $p(x) = T$. As T increases, these portions become steeper. Intuitively, which case is active will depend on how much power P_0 we have available.

To simplify things, let us assume that $T \to \infty$. This is equivalent to allowing $y_a(x)$ and $y_{1,b}(x)$ to be vertical, thus causing delta functions to appear in the power distribution. Note that we have already required that T be finite. Therefore, our solution will not be the power distribution for $T = \infty$, but rather the limit of the power distribution as $T \to \infty$.

Next, we assume that we are in each of the three cases, and then develop conditions on the total available power P_0:

Case 1 ($y(x)$ has the form of $y_1(x)$): The power of the delta function at $x = a$ will be

$$y(a) - y_s(a) = -\frac{1}{\lambda}\left[r(a) - \frac{r'(a)}{h(a)(\frac{1}{h})'(a)}\right] = -\frac{1}{\lambda}l(a)g(a),$$

and the power of the delta function at $x = b$ will be

$$y_s(b) - y(b) = \frac{1}{h(b)}\left[-\frac{1}{\lambda}\frac{r'(b)}{(\frac{1}{h})'(b)} - \eta B\right].$$

In the interval (a, b), the power distribution will be given by

$$p(x) = -\frac{y_s'(x)}{h(x)} = \frac{1}{\lambda}\frac{l'(x)}{h(x)}.$$

Requiring that the total power is P_0 readily gives that $\lambda = -\frac{r(b)}{P_0 + \frac{\eta B}{h(b)}}$. The power allocated to the delta function at $x = b$ must be positive, and this readily gives that $\frac{P_0}{\eta B} > g(b)$. Combining our findings, we arrive at (22).

Case 2 ($y(x)$ has the form of $y_2(x)$): In this case, the distribution of power consists of a delta function at $x = a$ with amplitude $-\frac{1}{\lambda}l(a)g(a)$, and a finite density $p(x) = \frac{1}{\lambda}\frac{l'(x)}{h(x)}$ in the interval (a, x_c), where the **cutoff point** x_c is given by the equation

$$y_s(x_c) = y(b) \Rightarrow -\frac{1}{\lambda}l(x_c) = \eta B. \tag{42}$$

Requiring the total power to be P_0 gives $\lambda = -\frac{r(x_c)}{P_0 + \frac{\eta B}{h(x_c)}}$ and, combining this with (42), we have that x_c is given by (32). As $g(x)$ is increasing (by Lemma 1(iv)), and we need to have $a \le x_c \le b$, we must have $g(a) \le \frac{P_0}{\eta B} \le g(b)$. Combining our findings, we arrive at (20).

Case 3 ($y(x)$ has the form of $y_3(x)$): Working as in the first two cases, we readily arrive at $p^{\mathrm{MAC}}(x) = P_0\delta(x - a)$.

D. Sketch of Proof for Theorem 6

The proof is very similar to the proof of Theorem 5, and uses duality, i.e., Theorem 4, but together with the *continuous* version of the MAC results, i.e., Theorem 3, instead of the *discrete* version, i.e., Theorem 2. The continuous version of (28), with $\pi(i) = i$, can be easily seen to be:

$$p^{\mathrm{BC}}(x) = p^{\mathrm{MAC}}(x)\frac{\eta B + h(x)\int_a^{x^-} p^{\mathrm{BC}}(t)\,dt}{\eta B + \int_{x^+}^b h(t)p^{\mathrm{MAC}}(t)\,dt}. \tag{43}$$

To verify equations (31), (32), and (33), we plug each of them in (43) together with, respectively, (20), (21), (22) and in all cases arrive at an identity.

References

[1] S. Toumpis, R. Müller, and J. Sayir, "On the transport capacity of a multiple access gaussian channel," in *Proc. IEEE Inter. Workshop on Wireless Ad Hoc Networks*, Oulu, Finland, May-June 2004.

[2] G. A. Gupta, S. Toumpis, J. Sayir, and R. R. Müller, "On the transport capacity of gaussian multiple access and broadcast channels," in *Proc. IEEE WiOpt*, Riva del Garda, Trentino, Italy, Apr. 2005.

[3] ——, "Transport capacity of gaussian multiple access and broadcast channels with a large number of nodes," in *Proc. IEEE ISIT*, Adelaide, Australia, Sep. 2005.

[4] H. Takagi and L. Kleinrock, "Optimal transmission ranges for randomly distributed packet radio networks," *IEEE Trans. Commun.*, vol. 32, no. 3, pp. 246–257, March 1984.

[5] R. Nelson and L. Kleinrock, "The spatial capacity of a slotted ALOHA multihop packet radio network with capture," *IEEE Trans. Commun.*, vol. 32, no. 6, pp. 684–694, June 1984.

[6] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[7] A. Reznik and S. Verdú, "On the transport capacity of a broadcast Gaussian channel," *Communications in Information and Systems*, vol. 2, no. 2, pp. 183–216, Dec. 2002.

[8] T. Cover and J. Thomas, *Elements of Information Theory*, 1st ed. New York: Wiley Interscience, Aug. 1991.

[9] N. Jindal, S. Vishwanath, and A. J. Goldsmith, "On the duality of gaussian multiple-access and broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 5, pp. 768–783, May 2004.

[10] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels — Part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.

[11] S. V. Hanly and D. N. C. Tse, "Multiaccess fading channels — Part II: Delay-limited capacities," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2816–2831, Nov. 1998.

[12] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. Cambridge University Press, 2004.

[13] S. Vishwanath, S. A. Jafar, and A. J. Goldsmith, "Optimum power and rate allocation strategies for multiple access fading channels," in *Proc. Spring IEEE VTC*, vol. 4, Rhodes, Greece, May 2001, pp. 2888–2892.

[14] D. Hughes-Hartogs, "The capacity of a degraded spectral Gaussian broadcast channel," Ph.D. dissertation, Stanford University, July 1975.

[15] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels — Part I: Ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, p. 1083, March 2001.

[16] R. Knopp and P. A. Humblet, "Information capacity and power control in single cell multiuser communications," in *Proc. IEEE ICC*, Seattle, WA, June 1995.

[17] L. Euler, "De serie Lambertina plurimisque eius insignibus proprieatibus," *Opera Mathematica 1.*, vol. 6, pp. 350–369, 1921 (original data 1779).

[18] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, vol. 5, pp. 329–359, 1996.

[19] F. Y. M. Wan, *Introduction to the Calculus of Variations and its Applications*. Chapman & Hall Mathematics, 1995.