On the Capacity of Multiple Access and Broadcast Fading Channels with Full Channel State Information at Low SNR

Zouheir Rezki Senior Member, IEEE, and Mohamed-Slim Alouini, Fellow, IEEE,

Abstract—We study the throughput capacity region of the Gaussian multi-access (MAC) fading channel with perfect channel state information (CSI) at the receiver and at the transmitters, at low power regime. We show that it has a multidimensional rectangle structure and thus is simply characterized by single user capacity points. More specifically, we show that at low power regime, the boundary surface of the capacity region shrinks to a single point corresponding to the sum rate maximizer and that the coordinates of this point coincide with single user capacity bounds. Inspired by this result, we propose an on-off scheme, compute its achievable rate, and show that this scheme achieves single user capacity bounds of the MAC channel for a wide class of fading channels at asymptotically low power regime. We argue that this class of fading encompasses all known wireless channels for which the capacity region of the MAC channel has even a simpler expression in terms of users’ average power constraints only. Using the duality of Gaussian MAC and broadcast channels (BC), we deduce a simple characterization of the BC capacity region at low power regime and show that for a class of fading channels (including Rayleigh fading), time-sharing is asymptotically optimal.

Index Terms—Multi-access, broadcast, ergodic capacity, capacity region, low-SNR, low power, fading channel, on-off signaling.

I. INTRODUCTION

It is now widely accepted that energy efficiency is a key parameter in designing wireless communication systems. This has catalyzed interests of many researchers inside the information/communication theory communities in order to better understand performance limits of wireless communication in the low power regime, and develop new techniques to achieve/approach these limits, e.g., [1]–[5]. For instance, in wide band communications, although the signal strength is generally very low, one can capitalize on the huge bandwidth and still achieve a high capacity [2], [6], [7]. The low-SNR framework is also useful to model cellular networks in some specific cases [4], [8], sensor networks where power saving is detrimental [9], [10] and more generally any communication scenario where the bandwidth and the power are fixed, but the system degree of freedom is large enough such that the power per degree of freedom is very low [1], [11]. Mainly, there are two trends in the literature of communications at low-power regime. The first one focuses on studying fundamental limits in terms of channel capacity, error probability, error exponent, etc; e.g., [1]–[5], [7], [12]. The second one deals more with signaling design and practical schemes that achieve these performance limits asymptotically, e.g., [1], [13]–[16]. A comprehensive list of references regarding communications at low-power regime and energy efficiency can be found in [11], [17], [18].

In this paper, we aim at studying the throughput capacity region of fading multi-access channels (MAC) and fading broadcast channels (BC) with Gaussian noise, where perfect channel state information (CSI) at the receiver(s) (CSI-R) and at the transmitter(s) (CSI-T) is assumed, at low power regime. The throughput capacity of the Gaussian MAC fading channel has been derived in [19]. Therein, it has been shown that each point on the boundary surface of the capacity region can be obtained by successive decoding and that the optimal rate and power allocations can be seen as the generalization of the single-user water-filling construction to MAC channels. The boundary surface is defined as the set of users’ rates such that no component can be increased with other components remaining fixed, while staying in the capacity region [19].
For instance, the capacity region of a 2-user symmetric MAC channel is depicted in Fig. 1. The boundary surface is the curved line joining the points B and C. Differently from the nonfading Gaussian MAC channel where the boundary surface of the capacity region is a line with slope -1, there is no linear part on the boundary surface of the Gaussian MAC fading channels capacity region. Furthermore, since the horizontal line joining A and B and the vertical line joining C and D are completely defined by single user capacity bounds, then determining the boundary surface B-C is enough to fully characterize the MAC capacity region. However, to obtain each point on the boundary surface, a parameterized optimization problem must be solved.

On the other hand, the BC is generally a good model for downlink communications in cellular networks where a base station is sending common and/or independent messages to different users. The single-input single-output (SISO) Gaussian BC is by nature a degraded channel for which the capacity is known [20]. Even with fading along with perfect CSI at both the transmitter and the receivers (CSI-TR), the capacity region has been obtained in [21] and in the context of parallel Gaussian channels in [22]. Here again, although the optimal power profile structure is essentially a water filling, an explicit characterization of the boundary capacity region seems difficult to obtain.

We focus in this paper on the low power regime (formally defined later) and analyze the capacity region of the Gaussian MAC fading channel. We show that interestingly, at low power regime, the MAC capacity region has a multidimensional rectangle structure, i.e. the boundary surface reduces to a single point. This point corresponds to the sum rate maximizer, the components of which are single-user capacity bounds. Then, we propose an on-off scheme and provide a necessary condition on the fading channels under which this scheme achieves the boundary point so that each user achieves single-user performance as if others were turned off. Using the duality of Gaussian MAC and BC, we provide a simple characterization of the BC capacity region at low power regime and show that for a class of fading channels (including Rayleigh fading), time-sharing is asymptotically optimal.

We note that the conference versions of this paper [23], [24] contain part of the results presented herein. However, we believe that the actual version is a substantial extension of the conference versions. We highlight below the main extra contributions beyond the conference versions:

- We present an explicit and simplified power profile (c.f. (29) in Section V) that achieves any point on the boundary of the Gaussian BC. Its simplicity is coming from the fact that the transmitter needs not solve multi-variable equations to set the optimal power. Instead, likewise in point-to-point communications, the optimal water-filling is derived by solving single-variable equations.
- We provide a necessary and sufficient condition under which time-sharing is asymptotically optimal (c.f. (35) in Section V), thereby fully characterizing the class of fading channels for which time-sharing is asymptotically capacity-achieving. Illustrative examples (Example 2 and Example 3 in Section V) are given to support this result.
- We include the proof of Lemma 1 (Appendix A) which is instrumental to the proof of our main result (Theorem 1).
- We insert the proof of Theorem 2 (Appendix B) which we believe is an important result that characterizes the capacity of point-to-point communications at low power regime. While this result is instrumental to the optimality of the on-off power control for the MAC channel (c.f. (23) in Section IV), it could be also of independent interest.
- We reinforce our numerical results and include in all figures the capacity regions of the additive white Gaussian noise (AWGN) channel and the corresponding fading with channel state information at the receiver (CSI-R) only as benchmarks. We also include Fig. 12 as an evidence that time-sharing might not be optimal at low power regime for the log-logistic broadcast fading channel as discussed in Example 3.

The remainder of this paper is organized as follows. Section II introduces our system model. Section III contains the main result of this paper along with its proof. An off-on scheme is proposed in Section IV where the MAC capacity region is explicitly characterized for a class of fading channels. The capacity of BC is deduced in Section V. In Section VI, selected numerical results are provided. Finally, Section VII concludes the paper.

II. SYSTEM MODEL

We consider an uplink scenario where $K$ users are communicating with a base station. All terminals have a single antenna each, i.e., a SISO MAC channel. We focus on a discrete-time Gaussian MAC in which the received signal at the base station at time instant $n$, $n = 1, \ldots, \infty$, is given by

$$y(n) = \sum_{k=1}^{K} h_k(n) x_k(n) + v(n),$$

where $h_k(n), x_k(n)$ and $v(n)$ are complex random variables (r.v.) that represent the channel gain and the transmitted signal of user $k$ and the additive noise, respectively. For convenience, we denote a circularly symmetric complex Gaussian r.v., say $r$, with mean zero, and variance $\sigma^2$, as $r \sim \mathcal{CN}(0, \sigma^2)$. We assume without loss of generality a normalized Gaussian noise so that $v \sim \mathcal{CN}(0, 1)$. Each user is constrained by an average transmit power $P_k$.

We also consider the $K$-user discrete-time Gaussian BC in which the received signal at user $k$, $k = 1, \ldots, K$, is given by

$$y_k(n) = h_k(n) x(n) + v_k(n),$$

where $x(n)$ and $v_k(n)$ are complex r.v. that represent the transmitted signal and the additive noise at user $k$. We also assume without loss of generality a normalized Gaussian noise so that $v_k \sim \mathcal{CN}(0, 1)$. The transmitter is constrained by an average transmit power $\bar{P}$.

We focus on fading processes with continuous probability density function (pdf) and with infinite support (although the later assumption is not mandatory). Furthermore, we assume
that the channel gains of all users are independent, but not necessarily identically distributed. In addition, in both the MAC and the BC described by (1) and (2), respectively, perfect CSI-TR is assumed, implying that at time instant \( n \), each terminal knows perfectly all channel gains \( h(n, k) \), \( k = 1, \ldots, K \). For convenience, we let \( y_k = |h_k|^2 \). We focus on asymptotically low power regime meaning that \( \bar{P}_k \to 0 \) for the MAC and that \( \bar{P} \to 0 \) for the BC; and we say that \( f(x) \approx g(x) \) if and only if \( \lim_{x \to \infty} \frac{f(x)}{g(x)} = 1 \). Inequalities \( \preceq \) and \( \succeq \) are defined analogously. The last definition extends to functions of severable variables where the definition of limits is standard. When it is clear from the context, we omit \( a \) in \( \approx \) for convenience. Unless the base is specified, all the log(\( \cdot \)) functions in this paper represent the natural logarithm functions. Finally, bold face letters indicate vectors of dimension \( K \), i.e., \( \mathbf{R} = (R_1, \ldots, R_K) \).

### III. MAC Capacity Region At Low Power Regime

In this section, we first present our main result in Theorem 1 followed by the proof.

**Theorem 1:** Let \( \mathcal{R}(\bar{P}) \) be the multidimensional rectangle defined by the single user capacities, i.e., \( \mathcal{R}(\bar{P}) = \left\{ \mathbf{R} : R_k \leq C_k(\bar{P}_k), \ k = 1, \ldots, K \right\} \), where \( C_k(\bar{P}_k) \) is the single user capacity of user \( k \), with average transmit power \( \bar{P}_k \), and with perfect CSI-TR. For the SISO MAC described by (1), the capacity region \( C_{\text{MAC}}(\bar{P}) \) coincides with \( \mathcal{R}(\bar{P}) \) at asymptotically low power regime. That is, for any point \( \mathbf{R} \in \) on the boundary of the capacity region \( C_{\text{MAC}}(\bar{P}) \), we have:

\[
\lim_{\bar{P} \to 0} \frac{R_k(\bar{P})}{C_k(\bar{P})} = 1, \tag{3}
\]

for all \( k = 1, \ldots, K \).

Before proving Theorem 1, we note first that there is no loss of rigor by considering \( C_k(\bar{P}) \) in (3) although \( C_k(\cdot) \) depends only on \( \bar{P}_k \). Then, since our focus is on asymptotically low power regime, both \( \mathcal{R}(\bar{P}) \) and \( C_{\text{MAC}}(\bar{P}) \) collapse to the zero point. However, the characterization of the capacity region adopted in Theorem 1 is in the sense that the limit of the ratio between the achievable rate and the single user capacity is equal to 1 as \( \bar{P} \to 0 \) [25].

**Proof:** We want to show that \( \forall \ \mu = (\mu_1, \ldots, \mu_K) \in \mathbb{N}_k^K \), the point on boundary surface of the capacity region, \( \mathbf{R}^*(\mu) \) is independent of \( \mu \), i.e., all surfaces parameterizing the boundary surface of the capacity region intersect in exactly one point. This direct approach seems a bit complicated since the expression of \( R_k^*(\mu) \) in [19, Theorem 3.16] is somehow complicated. Instead, we adopt an information-theoretical approach to prove Theorem 1. Clearly, the region on the right hand side (RHS) of (3) is an upper bound on the MAC capacity region since the former is only constrained by single user capacities of all users. To show that all points in this set are achievable, it suffices to sow that the point \( C(\bar{P}) = (C_1(\bar{P}_1), \ldots, C_K(\bar{P}_K)) \) is asymptotically achievable. For this purpose, let us compute \( \mathbf{R}^*(\mu) \) for \( \mu_1 = \mu_2 = \ldots = \mu_K = 1 \), i.e., the point on the boundary of the capacity region that maximizes the sum rate. The \( k \)th components, \( k = 1, \ldots, K \), of this point is given by [19]:

\[
R_k^* = \int_{\lambda_k}^{\infty} \log \left( \frac{h}{\lambda_k} \right) \prod_{i \neq k} F_i \left( \frac{\lambda_i}{\lambda_k} h \right) f_i(h) dh, \tag{4}
\]

where \( f_k(\cdot) \) and \( F_k(\cdot) \) denote the probability density function (pdf) of the channel gain power \( y_k \) and its cumulative distribution function (cdf), respectively; and where the constants \( \lambda_k \)'s satisfy

\[
\bar{P}_k = \int_{\lambda_k}^{\infty} \left( \frac{1}{\lambda_k} - \frac{1}{h} \right) \prod_{i \neq k} F_i \left( \frac{x_i}{x_k} h \right) f_i(h) dh. \tag{5}
\]

Let us define a function, say \( G_k(x_1, \ldots, x_K) \), as the RHS of (5), i.e.,

\[
G_k(x_1, \ldots, x_K) = \int_{x_k}^{\infty} \left( \frac{1}{x_k} - \frac{1}{h} \right) \prod_{i \neq k} F_i \left( \frac{x_i}{x_k} h \right) f_i(h) dh, \tag{6}
\]

for positive \( x_1, \ldots, x_K \). Note that \( G_k(\lambda_1, \ldots, \lambda_K) = \bar{P}_k, k = 1, \ldots, K \). Because each average power constraint depends on all \( \lambda_k \)'s, it is natural to consider that each \( \lambda_k \) is a function of all \( \bar{P} \)'s, i.e., \( \lambda_k = \lambda_k(\bar{P}_1, \ldots, \bar{P}_K) \).

**Lemma 1:** For the \( \lambda_k \)'s that satisfy the average power constraint (5), it holds that:

\[
\lim_{\bar{P} \to 0} \lambda_k(\bar{P}_1, \ldots, \bar{P}_K) = \infty \tag{7}
\]

for all \( k = 1, \ldots, K \). Where \( \bar{P} \to 0 \) stands for \( (\bar{P}_1, \ldots, \bar{P}_K) \to (0, \ldots, 0) \).

**Proof:** For convenience, the proof is presented in Appendix A.

The intuition behind (7) is that at low power regime, \( \lambda_k \) which has an interpretation of the power cost, converges toward infinity since at low power regime, the power becomes more expensive. This is true because the fading channels considered have infinite support.

Now, since \( \forall h \in [\lambda_k, \infty) \), we have \( F_i(\lambda_i \gamma) \leq F_i \left( \frac{h}{\lambda_k} \right) \leq 1 \), the following inequalities hold true for all \( \lambda_k > 0 \):

\[
\prod_{i \neq k} F_i(\lambda_i \gamma) \int_{\lambda_k}^{\infty} \left( \frac{1}{\lambda_k} - \frac{1}{h} \right) f_i(h) dh \leq 
\int_{\lambda_k}^{\infty} \left( \frac{1}{\lambda_k} - \frac{1}{h} \right) \prod_{i \neq k} F_i \left( \frac{\lambda_i}{\lambda_k} h \right) f_i(h) dh \leq 
\int_{\lambda_k}^{\infty} \left( \frac{1}{\lambda_k} - \frac{1}{h} \right) f_i(h) dh, \tag{8}
\]

or equivalently,

\[
\prod_{i \neq k} F_i(\lambda_i \gamma) \leq \frac{1}{\lambda_k} \left( \frac{1}{\lambda_k} - \frac{1}{\gamma} \right) \prod_{i \neq k} F_i \left( \frac{\lambda_i}{\lambda_k} \gamma \right) f_i(h) dh \leq \frac{1}{\lambda_k} \int_{\lambda_k}^{\infty} \left( \frac{1}{\lambda_k} - \frac{1}{h} \right) f_i(h) dh \leq 1. \tag{9}
\]

1 Although rigorously speaking, in the context of the MAC, for any \( k = 1, \ldots, K \), \( \lambda_k (\cdot) \) is a multivariate function of all \( \bar{P}_i, i = 1, \ldots, K \), i.e., \( \lambda_k = \lambda_k(\bar{P}_1, \ldots, \bar{P}_K) \), we will find it convenient to denote it simply as \( \lambda_k \) when there is no ambiguity.
Taking the limits as $\hat{P} \to 0$ on both sides of (9), we establish that
\begin{equation}
\int_{s_k}^\infty \left(1 - \frac{1}{h} \right) \prod_{i \neq k} F_i \left(\frac{h}{\lambda_k} \right) f_k(h) \, dh \approx \int_{s_k}^\infty \left(1 - \frac{1}{h} \right) f_k(h) \, dh. \tag{10}
\end{equation}
Along similar lines, one can also prove that
\begin{equation}
\int_{s_k}^\infty \log \left(\frac{h}{\lambda_k} \right) \prod_{i \neq k} F_i \left(\frac{h}{\lambda_k} \right) f_k(h) \, dh \approx \int_{s_k}^\infty \log \left(\frac{h}{\lambda_k} \right) f_k(h) \, dh. \tag{11}
\end{equation}
Note that the RHS of (11) may be written as $E \left[ \log \left(1 + \gamma_k \left(\frac{h}{\lambda_k} - \frac{1}{\lambda_k} \right) \right) \right]$, which takes clearly the form of an ergodic capacity expression due to the averaging over $\gamma_k$ and to the water-filling of the power structure inside the $\log(\cdot)$ function. Therefore, the RHS of (11) describes the capacity of fading channel $k$ in function of $\lambda_k$ and not in function of $P_k$ as we would have wished. But since $\lambda_k$ on the RHS of (11) also satisfies (10), then we have
\begin{equation}
\hat{P}_k \approx \int_{s_k}^\infty \left(1 - \frac{1}{\lambda_k} \right) f_k(h) \, dh \tag{12}
\end{equation}
due to (5). Hence, we conclude that at asymptotically low power regime, $R_k^* (\hat{P}_k) \approx C_k (\hat{P}_k)$, for all $k = 1, \ldots, K$. Finally, we highlight the fact that while $\lambda_k = \lambda_k \left(\hat{P}_1, \ldots, \hat{P}_K\right)$ for an arbitrary $\hat{P}_k$ values, (12) stipulates that at asymptotically low power regime, the dependence on $\hat{P}_i$’s, $i \neq k$, breaks down so that $\lambda_k = \lambda_k (\hat{P}_k)$ only.

It is worthwhile to mention that in order for our result to hold, the fading of different users need to be independent, not necessarily identically distributed. We also note that Theorem 1 emphasizes on how communication at low power regime can be energy efficient: at this regime, each user benefits from a single user performance if others do not exist. The strategy to achieve this is the same as the one that maximizes the sum rate, a time-division multiple-access strategy where at most one user is allowed to transmit at any given state.

We note that since in the MAC scenario we perform water-filling over the maximum of users’ channel gains $\max_{k=1,\ldots,K} \gamma_k$ to achieve the maximum sum-rate, whereas in the single-user case, we perform water-filling on $\gamma_k$, for each individual user, then these rates seem to be different a priori. However, the two strategies perform similarly at asymptotically low power regime as stated by Theorem 1. At low power regime, since the power cost $\lambda_k \to \infty$, for all $k = 1, \ldots, K$, if each user transmits only when its channel gain is “extremely good” ($\gamma_i \geq \lambda_k$), then this channel gain is the strongest with high probability. That is, the probability that the channel gain $\gamma_j$ be bigger than all $\gamma_j, j \neq k$, given that $\gamma_k \geq \lambda_k$ converges to 1 as $\lambda_k \to \infty$, or equivalently as $\hat{P}_k \to 0$. The proof of Theorem 1 is somehow too technical and to get the feel of the insight behind Theorem 1, we consider the following simple example.

**Example 1:** Let us consider a symmetric MAC channel where all users have the same fading statistics and are constrained by the same average power constraint, i.e., $\hat{P} = \hat{P}$, for $k = 1, \ldots, K$. As discussed above, TDMA achieves the maximum sum-rate. This yields a sum-rate equal to $R_s = E \left[ \log (1 + \gamma_{\max} P (\gamma_{\max})) \right]$, where $\gamma_{\max} = \max_{k=1,\ldots,K} \gamma_k$ and the function $P(\cdot)$ is defined on $\left(0, \infty\right)$ by: $P(x) = \left[\frac{1}{x} - \frac{1}{\lambda}\right]$, with $\lambda$ chosen such that the average power constraint is satisfied with equality, i.e., $E \left[ P (\gamma_{\max}) \right] = K \hat{P}$. Note that the factor $K$ in the later equation is due to the fact that each user has probability $1/K$ (symmetric fading) of being the one that has maximum channel gain. By symmetry, each user gets the rate $R_k = \frac{1}{K} \gamma_k E \left[ \log (1 + \gamma_{\max} P (\gamma_{\max})) \right]$ which can be lower-bounded as follows:
\begin{equation}
R_k = \frac{1}{K} E \left[ \log (1 + \gamma_k P (\gamma_k)) \right] \text{Prob} [\gamma_j \leq \gamma_k, j \neq k] \tag{13}
\end{equation}
\begin{equation}
\geq \frac{1}{K} E \left[ \log (1 + \gamma_k P (\gamma_k)) \right] \text{Prob} [\gamma_j \leq \lambda, j \neq k] \tag{14}
\end{equation}
\begin{equation}
= \text{Prob} [\gamma_j \leq \lambda, j \neq k] E \left[ \log (1 + \gamma_k P (\gamma_k)) \right] \tag{15}
\end{equation}
where (14) follows because only $\gamma_k \geq \lambda$ matters in (13). Since $\lambda \to \infty$ as $\hat{P} \to 0$, then using (15), we have:
\begin{equation}
\lim_{\hat{P} \to 0} \frac{R_k}{E \left[ \log (1 + \gamma_k P (\gamma_k)) \right]} \geq 1. \tag{16}
\end{equation}
From (13), we also have:
\begin{equation}
\frac{R_k}{E \left[ \log (1 + \gamma_k P (\gamma_k)) \right]} \leq 1. \tag{17}
\end{equation}
Combining (16) and (17), we obtain:
\begin{equation}
\lim_{\hat{P} \to 0} \frac{R_k}{E \left[ \log (1 + \gamma_k P (\gamma_k)) \right]} = 1. \tag{18}
\end{equation}
Along similar steps, one can also show that:
\begin{equation}
\lim_{\hat{P} \to 0} \frac{\hat{P}}{E \left[ P (\gamma_k) \right]} = 1. \tag{19}
\end{equation}

Note that (19) asserts that $\lambda$ satisfies the single user power constraint asymptotically at low power regime, i.e., $E \left[ P (\gamma_k) \right] \approx \hat{P}$; hence $C_k (\hat{P}) \approx E \left[ \log (1 + \gamma_k P (\gamma_k)) \right]$, which combined with (18) yields $R_k \approx C_k (\hat{P})$ as predicted by Theorem 1.

**IV. ON-OFF POWER CONTROL ACHIEVABLE RATE**

It is well known that for the MAC channel, successive decoding is optimal [19]. Here, we investigate the rate achieved by an on-off signaling. We show that at asymptotically low power regime, a simple on-off policy is optimal for a class of fading channels, in the sense that each user achieves the single user capacity bound.

In our scheme, each user communicates solely with the base station while other users are turned off so that the receiver can perform any optimal single-user decoding without the need of successive decoding. The power policy of each user is an appropriate on-off power control.

Before introducing the on-off signaling scheme, let us first introduce the class of fading channels we are interested in. Indeed, we focus on fading channels such that $\lim_{t\to\infty} t^{-\beta} f_t(t)$ exists (including infinity) and is strictly positive. For convenience, let $\bar{h}_k$ be that limit, i.e., $\bar{h}_k = \lim_{t\to\infty} t^{-\beta} f_t(t)$. Then, our focus is on
fading channels such that \( l_k > 0, k = 1, \ldots, K \). In reliability theory, the function \( \zeta_k(t) = \frac{f_k(t)}{P_k(t)} \) is known as Generalized Failure Rate (GFR) and the probability distributions with increasing GFR (also known as IGFR) satisfy \( \lim_{t \to \infty} \zeta_k(t) > n \) if and only if \( E(y^n_k) \), for \( n > 0 \), is finite [26]. That is, for probability distributions with IGFR, having a finite mean is enough to show that \( l_k > 1 \). We do not assume that the fading channels considered herein have an IGFR; however, we focus on fading channels with \( l_k > 0 \) as the later condition is less restrictive than the IGFR property.

First, let us digress from the MAC setting and focus on a SISO point-to-point communication with perfect CSI-TR and claim the result in Theorem 2 that characterizes the capacity of this class of fading channels at low power regime.

**Theorem 2:** Let \( l_k = \lim_{t \to \infty} \frac{f_k(t)}{P_k(t)} \). For fading channels with \( l_k > 0 \) (including \( \infty \)), the capacity of the discrete-time memoryless channel described by \( y_k(n) = h_k(n) x_k(n) + v_k(n) \), \( n = 1, \ldots, \infty \), with perfect CSI-TR and under an average transmit power constraint \( \tilde{P}_k \), is given by

\[
C_k(\tilde{P}_k) \approx \left(1 + \frac{1}{l_k}\right) \lambda_k(\tilde{P}_k) \tilde{P}_k, \tag{20}
\]

where \( \lambda_k(\tilde{P}_k) \) is the water level corresponding to the capacity-achieving power profile. Furthermore, an on-off signaling defined by

\[
P_k(\gamma_k) = \begin{cases} \frac{\tilde{P}_k}{P_k(\gamma_k)} & \text{if } \gamma_k \geq \tau_k \\ 0 & \text{otherwise}, \end{cases} \tag{21}
\]

where \( \tau_k = \left(1 + \frac{1}{l_k}\right) \lambda_k(\tilde{P}_k) \), is asymptotically capacity-achieving.

**Proof:** The proof is presented in Appendix B.

Theorem 2 provides an asymptotic expression of the capacity over a wide class of fading channels. Should a specific fading be chosen, the Lagrange multiplier \( \lambda_k(\cdot) \) can be explicitly expressed in function of the average power \( \bar{P}_k \). For instance, considering Rayleigh or the more general Nakagami-m fading, one can easily verify that \( l_k \to \infty \), for both fading. Hence, they belong to the class of fading we are focusing in. Moreover, it has been shown in [27] that \( \lambda_k(\bar{P}_k) \approx \log \left(\frac{\bar{P}_k}{\tau_k}\right) \). Hence, using Theorem 2, it can be seen immediately that the capacity scales essentially as \( C_k(\bar{P}_k) \approx \bar{P}_k \log \left(\frac{\bar{P}_k}{\tau_k}\right) \) at low power regime. Note that this result is in full agreement with the ones established via different techniques in [28] and [27]. Needless to recall that at asymptotically low power regime, the Lagrange multiplier \( \lambda(\bar{P}_k) \) converges toward infinity. The proof is technical and can be found in Appendix B. However, the intuition behind this fact can be easily understood by giving to \( \lambda(\bar{P}_k) \) its economic interpretation as the price rate. Since at low power regime, the power becomes more expensive, then it is natural that \( \lambda(\bar{P}_k) \) increases indefinitely as \( \bar{P}_k \to 0 \). Note that for the class of fading channels in Theorem 2, it has been shown in [1] that the ratio of the capacity and \( \bar{P}_k \) goes to infinity. Theorem 2 describes this behavior more precisely stating that the capacity per unit power goes to infinity as \( 1 + \frac{1}{l_k} \lambda_k(\bar{P}_k) \). This is in fact the capacity gain provided by perfect CSI-T, but the second statement of Theorem 2 clearly emphasizes that 1-bit feedback is actually enough for this to be true and proposes a general capacity-achieving on-off scheme that yields this gain. Finally, we note that in order for Theorem 2 to apply, we only require the \( l_k \) to be positive. We do not claim that this condition is necessary, but we have verified that for all typical wireless channels encountered in the literature, this condition holds true. For instance, for Rayleigh, Rician and Nakagami fading channels, it can be easily verified that \( l_k \to \infty \) and that the ergodic capacity is essentially equal to \( K_0 \bar{P}_k \log \left(\frac{1}{\bar{P}_k}\right) \), where \( K_0 \) is a constant that depends on the fading statistics [27–29].

We now return back to the MAC channel and propose a communication strategy that achieves the sum-rate capacity. For user \( k \), let us consider an on-off power control scheme that transmits whenever \( \gamma_k \geq \gamma_i, i \neq k \), and \( \gamma_k \geq \tau_k \), where \( \tau_k = \left(1 + \frac{1}{l_k}\right) \lambda_k(\bar{P}_k) \) and \( \lambda_k(\bar{P}_k) \) satisfies (12); with a fixed power equal to \( \bar{Q}_k \), and remains silent otherwise. To fulfill the average power constraint of each user, \( \bar{Q}_k \) is chosen such that:

\[
\bar{Q}_k = \frac{\bar{P}_k}{\prod_{i \neq k} F_i(\gamma_k) f_k(\gamma_k) d\gamma} \tag{22}
\]

That is, the instantaneous transmit power \( P_k \) of user \( k \) is defined by

\[
P_k(\gamma_1, \ldots, \gamma_K) = \begin{cases} \bar{Q}_k & \text{if } \gamma_k \geq \gamma_i \text{ and } \gamma_k \geq \tau_k, \\ 0 & \text{otherwise}. \end{cases} \tag{23}
\]

Before computing the achievable rate by this scheme, we note that because \( \tau_k \to \infty \) as \( \bar{P}_k \to 0 \). We have:

\[
\bar{Q}_k \leq \frac{\bar{P}_k}{\prod_{i \neq k} F_i(\tau_k) \int_{\tau_k}^{\infty} f_k(\gamma_k) d\gamma} \tag{24}
\]

\[
\approx \frac{\bar{P}_k}{1 - F_k(\tau_k)}. \tag{25}
\]

Since the RHS of (25) converges to 0 as \( \bar{P}_k \to 0 \) as shown in Appendix B, so does \( \bar{Q}_k \). We are now ready to compute the rate achieved by user \( k \) as follows:

\[
R_k = \int_0^{\infty} \int_0^{\infty} \sum_{i \neq k} \log (1 + \gamma_i P_k(\gamma_1, \ldots, \gamma_K)) \prod_{i=1}^{K} f_i(\gamma_i) d\gamma_1 \ldots d\gamma_K
\]

\[
= \int_{\tau_k}^{\infty} \log (1 + \gamma_k Q_k) f_k(\gamma_k) \prod_{i \neq k} F_i(\gamma_k) d\gamma_k
\]

\[
\geq \int_{\tau_k}^{\infty} (1 - \epsilon) \gamma_k Q_k f_k(\gamma_k) \prod_{i \neq k} F_i(\gamma_k) d\gamma_k \tag{26}
\]

\[
\geq (1 - \epsilon) \tau_k \bar{P}_k, \tag{27}
\]

where (26) follows because \( \lim_{\bar{P}_k \to 0} Q_k = 0 \), then \( \forall \gamma_k \in [\tau_k, \infty) \), we have \( \log (1 + \gamma_k Q_k) \approx \gamma_k Q_k \) and thus \( \frac{\log (1 + \gamma_k Q_k)}{\gamma_k Q_k} - 1 \leq \epsilon \), for all \( \epsilon > 0 \), at sufficiently low \( \bar{P}_k \). By taking the limits on both sides of (27) as \( \epsilon \to 0 \), we establish that a rate equal to \( \left(1 + \frac{1}{l_k}\right) \lambda_k(\bar{P}_k) \bar{P}_k \) is asymptotically achievable. This rate corresponds to the asymptotic capacity described by Theorem 2 and hence is the best rate we can achieve.

We finally note that in order to set the above achievability
scheme, only \( \log_2(K+1) \) feedback bits are required at each fading realization. These bits identify the user \( k, k=1, \ldots, K \), if any, that is allowed to communicate at that fading realization. Furthermore, while at asymptotically low power regime, both TDMA and the proposed on-off scheme are optimal, as far as the optimality criterion is in the limit sense; we argue that the proposed on-off is easier to implement than TDMA. This is because in TDMA, the active user requires perfect CSI of its channel gain to perform a water-filling, whereas this is not necessary for the proposed scheme as a constant power is used instead.

V. Capacity Region of the BC At Low Power Regime

In order to characterize the BC capacity region at low power regime, we utilize the established duality between the Gaussian MAC and BC so that we can deduce the BC capacity region from that of the MAC given in Theorem 1 [30]. Recall that the duality implies that the capacity region of the BC with power \( \bar{P} \), is exactly equal to the capacity of the dual MAC, which has the same channel gains, and a sum power constraint of \( \bar{P} \) across all \( K \) transmitters. That is \( C_{BC}(\bar{P}) = \bigcup_{\alpha \in R} C_{MAC}(P) \), where \( \alpha \cdot P = \sum_{k=1}^{K} \alpha_k P_k \) [30]. Our result is formalized in Theorem 3.

**Theorem 3:** Let \( \mathcal{R}(\bar{P}) \) be the region defined by:

\[
\mathcal{R}(\bar{P}) = \left\{ R : R_k \leq C_k(\alpha_k \bar{P}), \ k = 1, \ldots, K, \sum_{k=1}^{K} \alpha_k = 1 \right\},
\]

where \( C_k(\alpha_k \bar{P}) \) is the single user capacity of user \( k \), with average transmit power \( \alpha_k \bar{P} \), and with perfect CSI-TR; and where \( \alpha_k \)'s are arbitrary positive coefficients. For the SISO BC described by (2), the capacity region \( C_{BC}(\bar{P}) \) coincides with \( \mathcal{R}(\bar{P}) \) at asymptotically low power regime. That is, all points in \( \mathcal{R}(\bar{P}) \) are achievable and vice versa any point in \( C_{BC}(\bar{P}) \) is necessarily in \( \mathcal{R}(\bar{P}) \) too.\(^2\)

**Proof:** To show that any point in \( \mathcal{R}(\bar{P}) \) is achievable, we need only to focus on rates on the boundary of \( \mathcal{R}(\bar{P}) \), i.e., the rates such that \( R_k = C_k(\alpha_k \bar{P}) \), for all \( k = 1, \ldots, K \). Then, we know from Theorem 1 that the point with coordinates \( (C_1(\alpha_1 \bar{P}), \ldots, C_K(\alpha_K \bar{P})) \) belongs to \( C_{MAC}(\bar{P}, \alpha) \). Since \( \sum_{k=1}^{K} \alpha_k \bar{P} = \bar{P} \), then the point \( (C_1(\alpha_1 \bar{P}), \ldots, C_K(\alpha_K \bar{P})) \) also belongs to \( C_{BC}(\bar{P}) \) by the duality between the Gaussian MAC and BC. This completes the achievability part.

To prove the converse, we only need to show that the points on the boundary of the BC capacity region necessarily belong to \( \mathcal{R}(\bar{P}) \) too. To that end, let \( R \) be a point on the boundary of the BC capacity region. Then again, by the duality between the MAC and the BC, there exists a power policy \( P \) such \( P = \sum_{k=1}^{K} P_k \) and \( R \in C_{MAC}(P) \). For convenience, we let \( P_k = \alpha_k \bar{P} \), for some positive \( \alpha_k \)'s, so that we have \( \sum_{k=1}^{K} \alpha_k = 1 \). Furthermore, since \( R \) is on the boundary of the BC capacity region, then it is necessarily on the boundary of the \( C_{MAC}(P) \) too, otherwise this would contradict the definition of the boundary curve. But, from Theorem 1 we know that the boundary of the MAC capacity region shrinks to a single point at low power regime, corresponding to the sum-rate maximizer and that \( R_k(\bar{P}_k) \approx C_k(\bar{P}_k) \), or equivalently \( R_k(\alpha_k \bar{P}) \approx C_k(\alpha_k \bar{P}) \). Therefore, the converse is also true and Theorem 3 is thus proved.

Again, in order for Theorem 3 to hold, the fading of different users need not be identically distributed, only the independence is instrumental. We also note that from the proof above, since each point on the BC boundary region corresponds to a sum rate maximizer of a certain dual MAC, then to achieve a given point on the boundary of the BC capacity region, an optimal strategy is to use the MAC-BC transformation in each fading state to find a BC power policy that achieves the same average rate as the corresponding dual MAC [30]. That is, at any given state, the transmitter allocates all the power to at most one user as follows:

\[
P(\gamma) = \begin{cases} \frac{1}{\lambda_j} - \frac{1}{\gamma_j} & \text{if } \gamma_j > \frac{1}{\lambda_j} \gamma, \forall j \\ 0 & \text{else} \end{cases}
\]

where \( \lambda_j \)'s are solution of \( G_k(\lambda) = \alpha_k \bar{P} \), for \( k = 1, \ldots, K \). Since at low power regime, \( G_k(\lambda) \approx \int_{\lambda_k}^{\infty} \left( \frac{1}{\lambda} - \frac{1}{\gamma} \right) f_j(h) \, dh \) due to (10), then \( \lambda_j \)'s can be obtained by solving a one-variable equation, i.e., \( \int_{\lambda_k}^{\infty} \left( \frac{1}{\lambda} - \frac{1}{\gamma} \right) f_j(h) \, dh = \alpha_k \bar{P} \). Note that \( E[P(\gamma)|\gamma] = \sum_{k=1}^{K} \alpha_k P_k = \bar{P} \) confirming that the power policy given by (29) satisfies the power constraint. It is then easy to check that rate achieved by each user using this strategy is asymptotically equal to \( C_k(\alpha_k \bar{P}) \). Finally, we emphasize the fact that due to the concavity of \( C_k(\cdot) \), we have \( C_k(\alpha_k \bar{P}) \geq \alpha_k C_k(\bar{P}) \), confirming also that the BC capacity region \( \mathcal{R}(\bar{P}) \) contains the one corresponding to a time-sharing strategy where the available resource is exclusively allocated to user \( k \) during a fraction of time \( \alpha_k \). Consequently, time-sharing is asymptotically optimal if and only if \( C_k(\alpha_k \bar{P}) \approx \alpha_k C_k(\bar{P}) \), for all \( k = 1, \ldots, K \). Clearly, the later condition characterizes a class of fading where the power price \( \lambda_k \) is asymptotically invariant to a scaling, i.e., \( \lambda_k(\alpha \bar{P}) \approx \lambda_k(\bar{P}) \), for all \( \alpha \in [0,1] \). To see this, let us first recall the expression of the single user capacity that depends only on the power constraint and not on the Lagrange multiplier \( \lambda \) as established in Appendix B:

\[
C(\bar{P}) = \int_{0}^{\bar{P}} G^{-1}(t) \, dt,
\]

where \( G^{-1}(\cdot) \) is the inverse function of \( G(\cdot) \) and where the function \( G(\cdot) \) is defined on \((0, \infty)\) by \( G(\lambda) = E \left[ \left( \frac{1}{\lambda} - \frac{1}{\gamma} \right)^+ \right] \). We note that although not explicit, (30) is interesting in the sense that it gives a simple expression of the capacity that does not depend on \( \lambda \) and that is valid for an arbitrary \( \bar{P} \) value.

Now, time-sharing is asymptotically optimal if and only if:

\[
C_k(\alpha_k \bar{P}) \approx \alpha_k C_k(\bar{P}) \iff \lim_{P \to 0} \frac{C_k(\alpha_k \bar{P})}{\alpha_k C_k(\bar{P})} = 1
\]

\(^2\)Here again, as mentioned previously at the beginning of the proof of Theorem 1, the achievability and the converse are in the sense that the limit of the ratio is equal to 1 as \( P \to 0 \).
where (33) follows by l’Hospital rule and (35) is true by definition of the function $G_k(\cdot)$. Below we present an example of fading where this is actually the case.

Example 2 (BC with independent Rayleigh fading channels). Consider a BC where the power channel gain for each user follows an exponential distribution $f_k(x) = \frac{1}{\lambda_k} \exp(-\frac{x}{\lambda_k})$. For these channels, it is shown that the single user capacity scales essentially as $C_k(\bar{P}) \approx \lambda_k \bar{P}_k \log(\frac{\bar{P}_k}{\lambda_k})$ and that the Lagrange multiplier $\lambda_k(\bar{P}_k)$ scales essentially as $\lambda_k(\bar{P}_k) \approx \log(\frac{\bar{P}_k}{\lambda_k}) - 2 \log(\log(\frac{\bar{P}_k}{\lambda_k}))$ [28]. We note that for these channels, $\lambda_k(\bar{P}_k) \approx \log(\frac{\alpha}{\lambda_k}) - 2 \log(\log(\frac{\alpha}{\lambda_k})) \approx \log(\frac{1}{\bar{P}_k}) - 2 \log(\log(\frac{1}{\bar{P}_k})) \approx \lambda_k(\bar{P}_k)$. Hence, for these popular fading channels, time-sharing among users achieves the boundary of the capacity region $R(\bar{P})$.

However, property (35) does not hold in general as evidenced by the following example.

Example 3 (BC with i.i.d. log-logistic fading channels): Consider a BC where the power channel gain for each user follows a log-logistic distribution $f_k(x) = \frac{1}{(1+x^{-\gamma})}$. For these channels, using Theorem 2, it can be easily verified that the single user capacity scales essentially as $C_k(\bar{P}_k) \approx 2 \lambda_k(\bar{P}_k) \bar{P}_k$. On the other hand, it can be shown that the Lagrange multiplier $\lambda_k(\bar{P}_k) \approx \frac{1}{\sqrt{2\lambda_k}}$. Therefore, property (35) does not hold in this case since $\lambda_k(\bar{P}_k) \approx \frac{1}{\sqrt{2\lambda_k}} \approx \frac{1}{\sqrt{\gamma}} \lambda_k(\bar{P}_k)$. Hence, for these fading channels, time-sharing among users is strictly suboptimal.

VI. Numerical Results and Discussion

We present numerical results for a 2-user MAC channel where both users undergo independent Rayleigh fading channels. In all figures, we depict the actual throughput capacity region obtained numerically using the characterization in [19, Theorem 3.16], TDMA’s achievable rate-region, the throughput achievable region by the proposed on-off scheme, the capacity region of an additive white Gaussian noise (AWGN) MAC channel and that of a Rayleigh fading MAC channel with CSI-R only. We note first that CSI-T provides a tremendous capacity gain over CSI-R only, especially at low power regime as can be seen in Fig. 2, Fig. 3, Fig. 4 and Fig. 6. Furthermore, progressively from Fig. 2 to Fig. 4, the throughput capacity region converges to a square, since the transmit powers of the 2 users are equal. In the previous figures, both TDMA’s rate-region and the rate-region of the proposed on-off scheme converge toward the capacity region at low power regime. However, the convergence of TDMA’s region toward the capacity region is faster than that of the proposed scheme. Note that the ratio between the maximum on-off achievable rate and the single user capacity bound of the MAC capacity region increases from 0.9 for $P_1 = P_2 = 0$ dB in Fig. 2, to 0.95 for $P_1 = P_2 = -30$ dB in Fig. 4. Decreasing $P_1$ and $P_2$ further, results in the convergence of the ratio to 1. For a better illustration as to how the proposed on-off scheme approaches the MAC capacity region at low-power regime, we provide in Fig. 5 a plot that describes the ratio $\eta$, between the rate achieved by the proposed on-off scheme and the single user capacity, versus $P_1$. Clearly, since the rate region achieved by the on-off scheme is certainly a multi-dimensional rectangle, then $\eta$ describes how close is the on-off achievable rate region to the MAC capacity region. As can be seen in Fig. 5, below -10 dB, the on-off scheme has already achieved more than 94% of the single user capacity bound. Similar observations can be noted from Fig. 6 for $P_1 = -30$ dB and $P_2 = -40$ dB. It is also worthwhile to observe that the ratio between the achievable rate and the average power increases with the power budget. This implies that a better spectral efficiency is achieved at low power regime. Since this ratio has dimension of spectral efficiency per unit power (bits per second per Hertz over Joules per second), we found it more convenient to call it spectral efficiency per unit power (SEUP), the unit of which is bits per Hertz per Joule (b/Hz/J). Furthermore, one may define a SEUP region for the MAC that describes the maximum spectral efficiency per unit power users can achieve. The SEUP region follows in a straightforward manner from the capacity region of the MAC channel. Indeed, in a $K$-user MAC channel, given a transmit power budget $(\bar{P}_1,\ldots,\bar{P}_K)$ of the users, each point of the capacity region defined by $(R_1,\ldots,R_K)$, specifies a point of the SEUP region, say $(SEUP_1,\ldots,SEUP_K)$, such that $SEUP(x) = \frac{R_1}{\bar{P}_1}$. Analogously to single-user communications, the low-power regime provides the best (highest) SEUP, or equivalently the best (lowest) minimum energy per bit as defined in [1]. Figure 7 depicts the SEUP region for a 2-user MAC Rayleigh fading channel versus for different $\bar{P}$ values, where $\bar{P}_1 = \bar{P}_2 = \bar{P}$. Also shown in Fig. 7 is the SEUP region achieved by the proposed on-off scheme. As can be seen in Fig. 7, the SEUP region gets larger as $\bar{P}$ decreases. Albeit, we have presented results where both users’ channels undergo Rayleigh fading channel, we emphasize that the same trend holds as long as both channels are independent, but not necessarily identically distributed.

For the BC, Fig. 8, Fig. 9, Fig. 10 and Fig. 11 display the capacity region obtained numerically, the one given by Theorem 3 along with time-sharing achievable rate-region, for $\bar{P} = -10$ dB, $\bar{P} = -30$ dB, $\bar{P} = -30$ dB with unequal power channel gain where the difference of the power gains of the two users is equal to 3 dB, and $\bar{P} = -70$ dB, respectively. Recall that while the capacity region of the fading BC with CSI-R only is generally not know, it is actually known in the symmetric fading case considered in Fig. 8, Fig. 9, and Fig. 11. Here again, beyond the huge capacity gain provided by CSI-T over CSI-R only, we note in all figures the accuracy of the characterization in Theorem 3. Although
time-sharing achievable rate region is strictly suboptimal at $-10$ dB, it converges slowly to the capacity region as shown in Fig. 11. However, this is not true in general as established by (35). Indeed, Fig. 12 depicts the BC capacity region, the characterization in Theorem 3 and time-sharing performance for log-logistic fading as described by Example 3. Expectedly, the characterization in Theorem 3 is still very accurate as it matches perfectly well the actual BC capacity region, but time-sharing is strictly sub-optimal in this case even at $P = -70$ dB.

VII. Conclusion

We have analyzed the throughput capacity region of the MAC fading channel with perfect CSI at the transmitters and at the receiver at low power regime. While the capacity region has a polymatroid structure at arbitrary power regime, We have shown that it is simply a multidimensional rectangle at asymptotically low power regime and that each user can achieve a single user performance as if others do not exist. We have also proposed a simple on-off scheme, computed its achievable rate and characterized a class of fading channels for which the proposed scheme is asymptotically optimal. We have deduced the BC capacity region using the duality and provided a simple asymptotic characterization of the BC capacity region as well. A class of fading channels for which time-sharing is asymptotically optimal has been identified.

Appendix A

Proof of Lemma 1

We recall that $\lambda_k$’s, $k = 1, \ldots, K$, are solution of (5) and thus we have $G_k(\lambda_1, \ldots, \lambda_K) = \bar{P}_k$. We also note that $\forall k, \lambda_k > 0$, otherwise if there exists a certain $k_0$ such that $\lambda_{k_0} = 0$, then $\bar{P}_{k_0} \to \infty$ according to (5) which contradicts the fact that user’s average powers are finite. Let $\lambda_m = \min_{k=1,\ldots,K} \lambda_k$, then the following inequalities hold true:

$$
\bar{P}_m = \int_{\lambda_m}^{\infty} \left( \frac{1}{\lambda_m} - \frac{1}{\gamma} \right) \prod_{i \neq m} F_i \left( \frac{\lambda_i}{\lambda_m} \gamma \right) f_m(\gamma) \, d\gamma
$$

$$
\geq \int_{\lambda_m}^{\infty} \left( \frac{1}{\lambda_m} - \frac{1}{\gamma} \right) \prod_{i \neq m} F_i (\lambda_i) f_m(\gamma) \, d\gamma
$$

Figure 2. Capacity region and on-off achievable rate region of a 2-user MAC Rayleigh fading channel, with $P_1 = P_2 = 0$ dB.

Figure 3. Capacity region and on-off achievable rate region of a 2-user MAC Rayleigh fading channel, with $P_1 = P_2 = -10$ dB.

Figure 4. Capacity region and on-off achievable rate region of a 2-user MAC Rayleigh fading channel, with $P_1 = P_2 = -30$ dB.

Figure 5. Ratio $\eta$, between the rate achieved by the proposed on-off scheme and the single user capacity versus $P_1$, for a 2-user MAC Rayleigh fading channel.
Figure 6. Capacity region and on-off achievable rate region of a 2-user MAC Rayleigh fading channel, with $P_1 = -40$ dB and $P_2 = -30$ dB.

Figure 7. Spectral efficiency per unit power (SEPUP) region for a 2-user MAC Rayleigh fading channel for different $P$ values.

Figure 8. Capacity region of a 2-user BC Rayleigh fading channel, with $P = -10$ dB.

Figure 9. Capacity region of a 2-user BC Rayleigh fading channel, with $P = -30$ dB.

Figure 10. Capacity region of a 2-user BC Rayleigh fading channel, with $P = -30$ dB and the difference of the power gains of the two users equal to 3 dB.

Figure 11. Capacity region of a 2-user BC Rayleigh fading channel, with $P = -70$ dB.
average transmit power constraint \( \bar{P} \). For the channel described by \( \gamma \), the Lagrange multiplier obtained by satisfying the average transmit power constraint is nothing but the function defined by \( G_m(\cdot) \) in Lemma 2.

\[
\int_{\lambda_m}^{\infty} \left( \frac{1}{\lambda_m} - \frac{1}{\gamma} \right) f_m(\gamma) \, d\gamma.
\]

The RHS of (36) is positive and \( \lim_{P \to 0} \tilde{P}_m = 0 \), then by the Sandwich Theorem, we have

\[
\lim_{P \to 0} \left( \prod_{i \neq m} F_i(\lambda_m) \int_{\lambda_m}^{\infty} \left( \frac{1}{\lambda_m} - \frac{1}{\gamma} \right) f_m(\gamma) \, d\gamma \right) = 0.
\]

Since \( \lambda_m > 0 \), then (37) implies necessarily that \( \lim_{P \to 0} \int_{\lambda_m}^{\infty} \left( \frac{1}{\lambda_m} - \frac{1}{\gamma} \right) f_m(\gamma) \, d\gamma = 0 \). The target in the later equation is nothing but the \( G_m(\cdot) \) function defined in [27] which is continuous and monotonically decreasing with limit \( G_m(\infty) = 0 \). Hence, \( \lim_{P \to 0} \lambda_m = \infty \). In summary, we have shown that the \( \lambda_m = \min_k \lambda_k \to \infty \) as \( \tilde{P} \to 0 \), we conclude that for all \( k = 1, \ldots, K \), we have \( \lim_{P \to 0} \lambda_k(\tilde{P}_1, \ldots, \tilde{P}_k) = \infty \) and Lemma 1 is thus proved.

\section*{Appendix B}

\textbf{Proof of Theorem 2}

We first prove the asymptotic capacity expression given by (20). For the channel described by \( \gamma_k(n) = h_k(n) x_k(n) + v_k(n), n = 1, \ldots, \infty \), with perfect CSI-TR and under an average transmit power constraint \( \tilde{P}_k \), it is well-known that the instantaneous optimal power, \( P_k(\gamma) \), is a water-filling policy given by [31]: \( P_k(\gamma_k) = \left[ \frac{1}{\lambda_k(\tilde{P}_k)} - \frac{1}{\gamma_k} \right] \) where \( \lambda_k(\tilde{P}_k) \) is the Lagrange multiplier obtained by satisfying the average power constraint with equality. That is, \( G_k(\lambda_k(\tilde{P}_k)) \triangleq \int_{\lambda_k(\tilde{P}_k)}^{\infty} \left( \frac{1}{\lambda_k(\tilde{P}_k)} - \frac{1}{\gamma} \right) f_k(\gamma) \, d\gamma = \tilde{P}_k \). The capacity is then obtained by averaging log \((1 + P_k(\gamma_k) \gamma_k)\) and is given by

\[
C_k(\lambda_k(\tilde{P}_k)) = \int_{\lambda_k(\tilde{P}_k)}^{\infty} \log \left( 1 + \frac{t}{\lambda_k(\tilde{P}_k)} \right) f_k(t) \, dt.
\]

Now, let us summarize few properties of the function \( G_k(\cdot) \) in Lemma 2.

\textbf{Lemma 2}: The function \( G_k(\cdot) \) defined by \( G_k(t) = \int_{\lambda_k(\tilde{P}_k)}^{\infty} \left( \frac{1}{\lambda_k(\tilde{P}_k)} - \frac{1}{\gamma} \right) f_k(\gamma) \, d\gamma \) is i) continuous and positive definite; ii) strictly monotonically decreasing; iii) invertible on its domain.

\[
l_{k}(t) = \frac{1}{G_k^{-1}(t)},
\]

von (20) \( \bar{P} \) for the channel described by \( \gamma \), the Lagrange multiplier obtained by satisfying the average transmit power constraint is nothing but the function defined by \( G_m(\cdot) \) in Lemma 2.

\[
The RHS of (36) is positive and \( \lim_{P \to 0} \tilde{P}_m = 0 \), then by the Sandwich Theorem, we have

\[
\lim_{P \to 0} \left( \prod_{i \neq m} F_i(\lambda_m) \int_{\lambda_m}^{\infty} \left( \frac{1}{\lambda_m} - \frac{1}{\gamma} \right) f_m(\gamma) \, d\gamma \right) = 0.
\]

Since \( \lambda_m > 0 \), then (37) implies necessarily that \( \lim_{P \to 0} \int_{\lambda_m}^{\infty} \left( \frac{1}{\lambda_m} - \frac{1}{\gamma} \right) f_m(\gamma) \, d\gamma = 0 \). The target in the later equation is nothing but the \( G_m(\cdot) \) function defined in [27] which is continuous and monotonically decreasing with limit \( G_m(\infty) = 0 \). Hence, \( \lim_{P \to 0} \lambda_m = \infty \). In summary, we have shown that the \( \lambda_m = \min_k \lambda_k \to \infty \) as \( \tilde{P} \to 0 \), we conclude that for all \( k = 1, \ldots, K \), we have \( \lim_{P \to 0} \lambda_k(\tilde{P}_1, \ldots, \tilde{P}_k) = \infty \) and Lemma 1 is thus proved.

\section*{Appendix B}

\textbf{Proof of Theorem 2}

We first prove the asymptotic capacity expression given by (20). For the channel described by \( \gamma_k(n) = h_k(n) x_k(n) + v_k(n), n = 1, \ldots, \infty \), with perfect CSI-TR and under an average transmit power constraint \( \tilde{P}_k \), it is well-known that the instantaneous optimal power, \( P_k(\gamma) \), is a water-filling policy given by [31]: \( P_k(\gamma_k) = \left[ \frac{1}{\lambda_k(\tilde{P}_k)} - \frac{1}{\gamma_k} \right] \) where \( \lambda_k(\tilde{P}_k) \) is the Lagrange multiplier obtained by satisfying the average power constraint with equality. That is, \( G_k(\lambda_k(\tilde{P}_k)) \triangleq \int_{\lambda_k(\tilde{P}_k)}^{\infty} \left( \frac{1}{\lambda_k(\tilde{P}_k)} - \frac{1}{\gamma} \right) f_k(\gamma) \, d\gamma = \tilde{P}_k \). The capacity is then obtained by averaging log \((1 + P_k(\gamma_k) \gamma_k)\) and is given by

\[
C_k(\lambda_k(\tilde{P}_k)) = \int_{\lambda_k(\tilde{P}_k)}^{\infty} \log \left( 1 + \frac{t}{\lambda_k(\tilde{P}_k)} \right) f_k(t) \, dt.
\]

Now, let us summarize few properties of the function \( G_k(\cdot) \) in Lemma 2.

\textbf{Lemma 2}: The function \( G_k(\cdot) \) defined by \( G_k(t) = \int_{\lambda_k(\tilde{P}_k)}^{\infty} \left( \frac{1}{\lambda_k(\tilde{P}_k)} - \frac{1}{\gamma} \right) f_k(\gamma) \, d\gamma \) is i) continuous and positive definite; ii) strictly monotonically decreasing; iii) invertible on its domain.

\[
l_{k}(t) = \frac{1}{G_k^{-1}(t)},
\]

where \( l_{k}(\cdot) \) is the function defined on \((0, \infty)\) by \( l_{k}(t) = \int_{0}^{t} \log \left( \frac{1}{u} \right) f_{k}(u) \, du \). For this particular optimization problem, it can be easily verified that

\[
\frac{dl_{k}}{dt} = \frac{dG_{k}}{dt},
\]

t for all \( t > 0 \). Now, for any \( t > 0 \), the following equalities hold true:

\[
l_{k}(t) = \int_{0}^{t} \frac{dI_{k}}{du} \, du = \int_{0}^{t} \frac{dG_{k}(u)}{du} \, du = \int_{0}^{t} G_{k}^{-1}(u) \, du,
\]

where \( G_{k}^{-1}(\cdot) \) is the inverse function of \( G_{k}(\cdot) \). Equality (44) is true because \( \lim_{t \to 0} l_{k}(t) = 0 \), whereas (45) follows from (43); and (46) is obtained by change of variables. Applying (46) to \( t = \lambda_k(\tilde{P}_k) \) and using the fact that \( I_{k}(\lambda_k(\tilde{P}_k)) = C_k(\tilde{P}_k) \), we obtain

\[
C_k(\tilde{P}_k) = \int_{0}^{G_{k}^{-1}(1)(\tilde{P}_k)} G_{k}^{-1}(t) \, dt.
\]
To prove (20) in Theorem 2, it suffices to use (47) in order to verify that \( \lim_{\hat{P}_k \to 0} G_{k}(\hat{P}_k) = 1 + \frac{1}{\lambda} \). This can be shown as follows:

\[
\lim_{\hat{P}_k \to 0} G_{k}(\hat{P}_k) = \lim_{\hat{P}_k \to 0} \frac{\int_{0}^{\hat{P}_k} G_{k}^{-1}(t) \, dt}{\int_{0}^{\hat{P}_k} G_{k}^{-1}(t) \, dt} = \frac{G_{k}^{-1}(\hat{P}_k)}{1 - F_{k}(\hat{P}_k)} (48)
\]

\[
= \frac{1}{1 - \frac{1}{\lambda}}. (49)
\]

where (48) follows by the l’Hôpital rule and where (49) follows from vi) in Lemma 2. This establishes (20) in Theorem 2. Next, we prove that the on-off power given by (21) is capacity-achieving. We treat the cases \( l_k \to \infty \) and \( 0 < l_k < \infty \) separately.

For the class of fading we are interested in, note that if \( l_k \to \infty \), then \( \tau_k \) in (21) is equal to \( \lambda_k \left( \frac{P_k}{1 - F_{k}(\tau_k)} \right) \) and we have:

\[
\frac{P_k}{1 - F_{k}(\tau_k)} = \frac{P_k}{1 - \lambda_k \left( \frac{P_k}{1 - F_{k}(\tau_k)} \right)} = \frac{G_{k}^{-1}(\hat{P}_k)}{1 - F_{k}(\hat{P}_k)} (50)
\]

where (50) follows from the proof of iv) in Lemma 2. Since \( \lambda_k \left( \frac{P_k}{1 - F_{k}(\tau_k)} \right) \to \infty \) as \( \hat{P}_k \to 0 \), then \( \frac{P_k}{1 - F_{k}(\tau_k)} \to 0 \) as \( \hat{P}_k \to 0 \). The rate achieved by the on-off power policy can be computed as follows:

\[
R_k = \int_{\tau_k}^{\infty} \log \left( 1 + \frac{P_k}{1 - F_{k}(\tau_k)} \right) f_k(\gamma_k) \, d\gamma_k
\]

\[
= \int_{\lambda_k(\hat{P}_k)}^{\infty} \log \left( 1 + \frac{P_k}{1 - F_{k}(\lambda_k(\hat{P}_k))} \right) f_k(\gamma_k) \, d\gamma_k
\]

\[
\geq \int_{\lambda_k(\hat{P}_k)}^{\infty} (1 - \epsilon) \gamma_k \frac{P_k}{1 - F_{k}(\lambda_k(\hat{P}_k))} f_k(\gamma_k) \, d\gamma_k
\]

\[
\geq (1 - \epsilon) \lambda_k \left( \frac{P_k}{1 - F_{k}(\tau_k)} \right) \hat{P}_k
\]

(52)

where (51) follows because \( \lim_{\hat{P}_k \to 0} \frac{P_k}{1 - F_{k}(\tau_k)} = 0 \) and thus \( \forall \gamma_k \in [\tau_k, \infty) \), we have \( \log \left( 1 + \frac{P_k}{1 - F_{k}(\tau_k)} \right) \approx \gamma_k \frac{P_k}{1 - F_{k}(\tau_k)} \), or equivalently \( \frac{\log(1 + \gamma_k \frac{P_k}{1 - F_{k}(\tau_k)})}{\gamma_k \frac{P_k}{1 - F_{k}(\tau_k)}} \to 1 \), \( \forall \epsilon > 0 \), at sufficiently low \( \hat{P}_k \). By taking the limits on both sides of (52) as \( \epsilon \to 0 \), we establish that a rate equal to \( \lambda_k \left( \frac{P_k}{1 - F_{k}(\tau_k)} \right) \hat{P}_k \) is asymptotically achievable. This rate corresponds to the asymptotic capacity described by Theorem 2 when \( l_k \to \infty \) and hence is the best rate one can achieve.

On the other hand, if \( 0 < l_k < \infty \), then the following statements hold true \( \forall t \in [t_0, \infty) \), for some \( t_0 \) sufficiently large:

\[
\lim_{t \to \infty} \frac{f_k(t)}{1 - F_k(t)} = l_k \Leftrightarrow \frac{f_k(t)}{1 - F_k(t)} \approx \frac{l_k}{t}
\]

\[
\Rightarrow - \frac{d}{dt} (\log(1 - F_k(t))) \approx l_k \frac{d}{dt} (\log(t))
\]

\[
\Rightarrow - \log \left( \frac{1 - F_k(t)}{1 - F_k(t_0)} \right) \approx l_k \log \left( \frac{t}{t_0} \right) (53)
\]

\[
\Rightarrow - \log(1 - F_k(t)) \approx l_k \log(t) (54)
\]

\[
\Rightarrow \frac{1}{1 - F_k(t)} \approx A \frac{l_k}{t} (55)
\]

where (53) follows because if \( f(t) \approx g(t) \), then \( \forall \epsilon > 0 \), there exists a certain \( t_0 \) such that \( \forall t \geq t_0 \), we have

\[
(1 - \epsilon) \, g(t) < f(t) < (1 + \epsilon) \, g(t). (56)
\]

Integrating (56) between \( t_0 \) and \( t > t_0 \), we establish that

\[
\int_{t_0}^{t} g(u) \, du \approx \int_{t_0}^{t} f(u) \, du. (57)
\]

The equivalence (54) is due to the fact that \( t_0 \) is a constant that vanishes as \( t \to \infty \) whereas \( A \) in (55) is some constant. Therefore, the fixed power in the on-off power policy in this case can be computed as follows:

\[
\frac{\hat{P}_k}{1 - F_k(\tau_k)} = \frac{\hat{P}_k}{1 - F_k(1 + \frac{1}{\lambda_k} \lambda_k(\hat{P}_k))}
\]

\[
\approx \left( 1 + \frac{1}{\lambda_k} \right) \left( 1 + \frac{P_k}{1 - F_{k}(\lambda_k(\hat{P}_k))} \right) \hat{P}_k
\]

(58)

where (58) follows from (55). Since \( \frac{P_k}{1 - F_{k}(\lambda_k(\hat{P}_k))} \to 0 \) as \( \hat{P}_k \to 0 \) due to (50), then so does \( \frac{\hat{P}_k}{\lambda_k(\hat{P}_k)} \). Following similar lines as in the case where \( l_k \to \infty \), it can be shown that a rate equal to

\[
R_k \geq (1 - \epsilon) \quad \tau_k \quad \hat{P}_k
\]

\[
\approx (1 - \epsilon) \left( 1 + \frac{1}{\lambda_k} \right) \lambda_k \left( \frac{P_k}{1 - F_{k}(\tau_k)} \right) \hat{P}_k
\]

(59)

is achievable. The RHS of (59) coincides with the asymptotic capacity in Theorem 2 as an arbitrary small \( \epsilon \) can be chosen. This completes the proof of Theorem 2.

**Acknowledgment**

The authors would like to thank the editor Dr. Rui Zhang for volunteering his time to handle this paper and the anonymous reviewers for their valuable comments that have enhanced the technical quality and the lucidity of this paper.

**References**

[1] S. Verdu, “Spectral efficiency in the wideband regime,” IEEE Trans. Inform. Theory, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[2] L. Zheng, N. C. D. Tse, and M. Médard, “Channel coherence in the low-SNR regime,” IEEE Trans. Inform. Theory, vol. 53, no. 3, pp. 976–997, Mar. 2007.

[3] X. Li, S. Jin, M. R. McKay, X. Gao, and K. K. Wong, “Capacity of MIMO-MAC with transmit channel knowledge in the low SNR regime,” IEEE Trans. Wireless Commun., vol. 9, no. 3, pp. 926–931, Mar. 2010.

[4] A. Lozano, A. Tulino, and S. Verdu, “Multiple-antenna capacity in the low-power regime,” IEEE Trans. Inform. Theory, vol. 49, no. 10, pp. 2527–2544, Oct. 2003.

[5] D. Qiao, M. Gsouris, and S. Velipasalar, “Energy efficiency in the low-SNR regime under queueing constraints and channel uncertainty,” IEEE Trans. on Comm., vol. 59, no. 7, pp. 2006–2017, Jul. 2011.
V. Raghavan, G. Harilhan, and A. Sayeed, “Capacity of sparse multipath channels in the ultra-wideband regime,” IEEE J. Select. Areas Signal Processing, vol. 1, no. 3, pp. 357–371, Oct. 2007.

S. Ray, M. Medard, and L. Zheng, “On noncoherent MIMO channels in the wideband regime: Capacity and reliability,” IEEE Trans. Inform. Theory, vol. 53, no. 6, pp. 1983–2009, Jun. 2007.

M. Beko, J. Xavier, and V. Barroso, “Further results on the capacity and error probability analysis of noncoherent MIMO systems in the low SNR regime,” IEEE Trans. Signal Processing, vol. 56, no. 7, pp. 2915–2930, Jul. 2008.

J.-F. Chamberland and V. Veeravalli, “Asymptotic results for decentralized detection in power constrained wireless sensor networks,” IEEE J. Select. Areas Commun., vol. 22, no. 6, pp. 1007–1015, Aug. 2004.

X. Wu and R. Srikan, “MIMO channels in the low-SNR regime: Capacity and reliability of energy-constrained wireless sensor networks,” IEEE Trans. Wireless Commun., vol. 5, no. 5, pp. 984–989, May 2006.

D. Tse and S. Hanly, “Multiaccess fading channels. Part I: Polymatroid structure, optimal resource allocation and throughput capacities,” IEEE Trans. Inform. Theory, vol. 47, no. 6, pp. 2207–2221, 2005.

Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, “Linear precoding for finite-alphabet signaling over MIMO/M wireless channels,” IEEE Trans. Vehicular Technology, vol. 61, no. 6, pp. 2599–2612, Jul. 2012.

M. Gursoy, “On the low-SNR capacity of phase-shift keying with hard-decision detection,” in Int. Symp. Inform. Theory ISIT’2007, Nice, France, Jun. 2007, pp. 166–170.

M. Gursoy, H. Poor, and S. Verdu, “Noncoherent Rician fading channel-part II: spectral efficiency in the low-power regime,” IEEE Trans. Wireless Communications, vol. 4, no. 5, pp. 2207–2221, 2005.

L. Yang and G. Giannakis, “Ultra-wideband communications: an idea whose time has come,” IEEE Signal Processing Magazine, vol. 21, no. 6, pp. 26–54, 2004.

D. Tse and S. 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.

Mohamed-Slim Alouini (S’94, M’98, SM’03, F’09) was born in Tunis, Tunisia. He received the Ph.D. degree in Electrical Engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 1998. He served as a faculty member in the University of Minnesota, Minneapolis, MN, USA, then in the Texas A&M University at Qatar, Education City, Doha, Qatar before joining King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia. His research interests include: performance limits of communication systems, cognitive and sensor networks, physical-layer security, and low-complexity detection algorithms.

His current research interests include the modeling, design, and performance analysis of wireless communication systems.