Streaming Transmission over Block Fading Channels with Delay Constraint

Giuseppe Cocco†, Deniz Gündüz* and Christian Ibars†

†CTTC, Barcelona, Spain
* Imperial College London, London, UK
giuseppe.cocco@cttc.es, d.gunduz@imperial.ac.uk, christian.ibars@cttc.es

***Abstract*—Streaming transmission over a block fading channel is studied assuming that the transmitter receives a new message at each channel block at a constant rate, which is fixed by an underlying application. A common deadline is assumed for all the messages, at which point the receiver tries to decode as many messages as possible. Various achievable schemes are proposed and compared with an informed transmitter upper bound in terms of average throughput. It is shown that the adaptive joint encoding (aJE) scheme is asymptotically optimal; that is, it achieves the ergodic capacity as the transmission deadline goes to infinity; and it closely follows the upper bound in the case of a finite transmission deadline. On the other hand, in the presence of multiple receivers with different signal-to-noise ratios (SNR), memoryless transmission (MT), generalized time-sharing (gTS) and superposition transmission (ST) schemes are shown to be more robust than the joint encoding (JE) scheme as they have gradual performance degradation with the decreasing SNR.**

***Index Terms*—Block-fading; Delay-constrained transmission; Multimedia streaming; Multiple access channel; Ergodic capacity; Satellite broadcasting**

I. Introduction

In a *streaming transmitter* data becomes available at the transmitter over time rather than being available at the beginning of transmission. Consider, for example, digital TV satellite broadcasting. The satellite receives video packets from a gateway on Earth at a fixed data rate and has to forward the received packets to the users within a certain deadline. Hence, the transmission of the first packet starts before the following packets arrive at the transmitter. We consider streaming transmission over a block fading channel with channel state information (CSI) available only at the receiver. This assumption results from practical constraints when the receiver belongs to a large population of terminals, or when the transmission delay is significantly larger than the channel coherence time[1] [2]. The data that arrives at the transmitter over a channel block can be modeled as an independent message whose rate is fixed by the quality of the gateway-satellite link and the video encoding scheme used for recording the event. We assume that the transmitter cannot modify the contents of the packets to change the data rate, as the satellite transmitter is oblivious to the underlying video coding scheme adopted by the source, and considers the accumulated data over each channel block as a single data packet that can be either transmitted or dropped.

We further impose a delay constraint on the transmission such that the receiver buffers the received messages for M channel blocks before displaying the content, which is typical of multimedia streaming applications (see Fig. I). As the messages arrive at the transmitter gradually over M channel blocks, the last message sees only a single channel realization, while the first message can be transmitted over the whole span of M channel blocks. For a finite number M of messages and M channel blocks, it is not possible to average out the effect of fading in the absence of CSI at the transmitter, and there is always a non-zero outage probability [3]. Hence, the performance measure we study is the average throughput, that is, the average decoded data rate by the user.

Communication over fading channels has been extensively studied [4]. The capacity of a fading channel depends on the available information about the channel behavior [5]. When both the transmitter and the receiver have CSI, the capacity is achieved though waterfilling [6]. This is called the *ergodic capacity* as the capacity is averaged over the fading distribution. In the case of a fast fading channel without CSI at the transmitter ergodic capacity is achieved with constant power transmission [4]. However, when there is a delay requirement on the transmission as in our model, and the delay constraint is short compared to the channel coherence time, we have a slow fading channel. In a slow-fading channel, if only the receiver can track the channel realization, outage becomes unavoidable [4]. An alternative performance measure in this case is the ϵ-

This work was partially supported by the European Commission under project ICT-FP7-258512 (EXALTED), by the Spanish Government under projects TEC2010-17816 (JUNTOS) and TEC2010-21100 (SOFOCLES), and by the Government of Catalonia under grant 2009-SGR-940. Giuseppe Cocco is partially supported by the European Space Agency under the Networking/Partnering Initiative. Deniz Gündüz was supported by the European Commission's Marie Curie IRG Fellowship with reference number 256410 (COOPMEDIA) under the Seventh Framework Programme.

[1] Transmission rate can be adjusted to the channel state through adaptive coding and modulation (ACM) driven by a feedback channel. However, in real-time broadcast systems with large delays and many receivers, such as satellite systems, this is not practical. For instance, in real-time video transmission the ACM bit-rate control-loop may drive the source bit-rate (e.g., variable bit rate video encoder), but this may lead to a large delay (hundreds of milliseconds) in executing rate variation commands. In such cases the total control loop delay is too large to allow real time compensation of fading [1, Section 4.5.2.1].

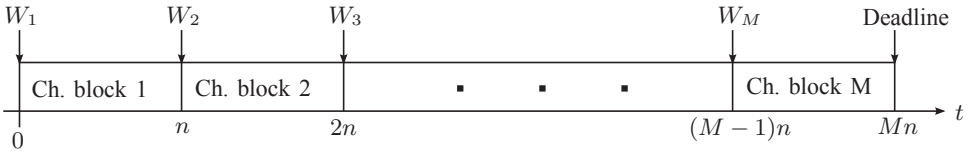


Fig. 1. The transmitter receives message W_i of rate R at the beginning of channel block i. All the M messages need to be transmitted to the receiver by the end of channel block M.

outage capacity [7]. In general it is hard to characterize the outage capacity exactly; hence, many works have focused on either the high signal-to-noise ratio (SNR) [8] or low SNR [9] asymptotic regimes. Another approach, which is also adopted in this work, is to study the average transmission rate, i.e., average throughput, as in [10] and [11]. Outages may occur even if the transmitter has access to CSI when the system is required to sustain a constant transmission rate at all channel states, called the delay-limited capacity [12], [13]. Due to the constant rate of the arriving messages at all channel blocks, our problem is similar to the delay-limited capacity concept. However, here we neither assume CSI at the transmitter nor require all arriving messages to be transmitted. Our work also differs from the average rate optimization in [10] since the transmitter in [10] can adapt the transmission rate based on the channel characteristics and the delay constraint, whereas in our model the message rate is fixed by the underlying application. The only degree-of-freedom the transmitter has in our setting is the multiple channel blocks it can use for transmitting the messages while being constrained by the causal arrival of the messages and the total delay constraint of M channel blocks.

Streaming transmission has received significant attention recently especially with the increasing demand for multimedia streaming applications [14]. Most of the work in this area focus on practical code construction [15], [16], [17]. The diversity-multiplexing tradeoff in a streaming transmission system with a maximum delay constraint for each message is studied in [18]. Unlike in [18], we assume that *the whole* set of messages has a common deadline; hence, in our setting the degree-of-freedom available to the first message is higher than the one available to the last.

In the present paper we extend our work in [19] by presenting analytical results and introducing more effective transmission schemes. We study joint encoding (JE), which encodes all the available messages into a single codeword at each channel block, and the more classical time-sharing (TS) and superposition (ST) coding schemes. The main contributions of the present work can be summarized as follows:

1) We introduce a channel model for streaming transmission over block fading channels with a common decoding deadline to study real-time multimedia streaming in networks with large delays, such as digital satellite broadcasting systems.
2) We introduce an informed transmitter upper bound on the performance assuming the availability of perfect CSI at the transmitter.
3) We find a simplified expression for the average throughput of the JE scheme, and use this expression to show that, in the limit of infinite channel blocks, the JE scheme has a threshold behavior that depends on the average channel quality.
4) We propose the *adaptive JE (aJE)* scheme, which drops certain packets depending on the average channel SNR, and show that it performs very close to the informed transmitter upper bound for a finite number of messages, and approaches the ergodic capacity as the number of channel blocks goes to infinity.
5) We propose a *generalized time-sharing* (gTS) scheme, in which each message is transmitted over a window of W channel blocks through time-sharing. We show that optimizing the window size W significantly improves the average throughput in the high SNR regime.
6) We show that the gTS and the ST schemes provide gradual performance improvement with increasing average SNR as opposed to the threshold behavior of the JE scheme. Focusing on the gTS scheme with equal time allocation and the ST scheme with equal power allocation, we show, through numerical simulations, that either scheme can outperform the other depending on the average SNR. Both schemes can be further improved by optimizing the time or power allocation, respectively.
7) We show that the aJE scheme is advantageous in a single receiver system whereas the simple gTS and ST schemes can be attractive when broadcasting to multiple users with a wide range of SNR values, or in a point-to-point system with inaccurate CSI.

The rest of the paper is organized as follows. In Section II we describe the system model. In Section III we provide an upper bound on the average throughput. In Section IV we describe the proposed transmission schemes in detail. Section V is devoted to the numerical results. Finally, Section VI contains the conclusions.

II. System Model

We consider streaming transmission over a block fading channel. The channel is constant for a block of n channel uses and changes in an independent and identically distributed (i.i.d.) manner from one block to the next. We assume that the transmitter accumulates the data that arrives at a fixed rate during a channel block, and considers the accumulated data as a single message to be transmitted during the following channel blocks. We consider streaming of M messages over M channel blocks, such that message W_t becomes available at the beginning of channel block t, for $t = 1, \dots, M$ (see Fig. I). Each message W_t has rate R bits per channel use (bpcu), i.e., W_t is chosen randomly with uniform distribution from the set $\mathcal{W}_t = \{1, \dots, 2^{nR}\}$, where n is the number of channel uses per channel block. Following a typical assumption in the

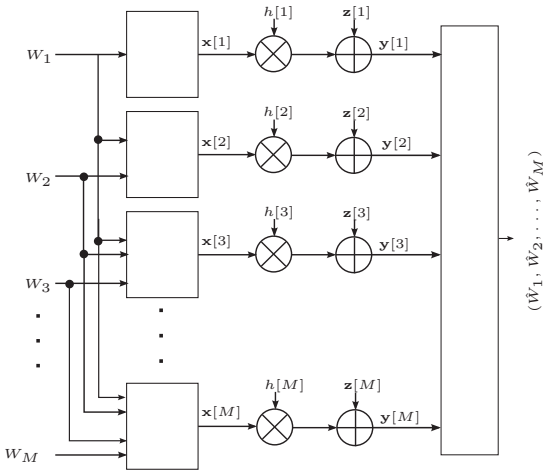


Fig. 2. Equivalent channel model for the sequential transmission of M messages over M channel blocks to a single receiver.

literature (see, e.g., [10]), we assume that n, though still large (as to give rise to the notion of reliable communication [20]), is much shorter than the dynamics of the slow fading process. The channel in block t is given by

$$\mathbf{y}[t] = h[t]\mathbf{x}[t] + \mathbf{z}[t], \tag{1}$$

where $h[t] \in \mathbb{C}$ is the channel state, $\mathbf{x}[t] \in \mathbb{C}^n$ is the channel input, $\mathbf{z}[t] \in \mathbb{C}^n$ is the i.i.d. unit-variance Gaussian noise, and $\mathbf{y}[t] \in \mathbb{C}^n$ is the channel output. The instantaneous channel gains are known at the receiver, while the transmitter only has knowledge of the statistics of the process $h[t]$. We have a short-term average power constraint of P, i.e., $E[\mathbf{x}[t]\mathbf{x}[t]^\dagger] \leq nP$ for $t = 1, \ldots, M$, where $\mathbf{x}[t]^\dagger$ represents the Hermitian transpose of $\mathbf{x}[t]$, and $E[x]$ is the mean value of x. As we assume a unitary noise power, in the following we will use interchangeably the quantities P and SNR.

This point to point channel can be seen as an orthogonal multiple access channel (MAC) with a special message hierarchy [21], in which the encoder at each channel block acts as a separate virtual transmitter (see Fig. 2). The receiver tries to decode as many of the messages as possible, and the performance measure is the average throughput, denoted by $\overline{R}$ and defined as the normalized average number of received messages multiplied by the transmission rate. We denote the instantaneous channel capacity over channel block t by $C_t \triangleq \log_2(1 + \phi[t]P)$, where $\phi[t]$ is a random variable distributed according to a generic probability density function (pdf) $f_\Phi(\phi)$. Note that C_t is also a random variable. We define $\overline{C} \triangleq E[\log_2(1 + \phi P)]$, where the expectation is taken over $f_\Phi(\phi)$. $\overline{C}$ is the ergodic capacity when there is no delay constraint on the transmission.

III. Informed Transmitter Upper Bound

We first provide an upper bound on the performance by assuming that the transmitter is informed about the exact channel realizations at the beginning of the transmission. This allows the transmitter to optimally allocate the resources among messages so that the average throughput $\overline{R}$ is maximized. Assume that $C_1, \ldots, C_M$ are known by the transmitter and the maximum number of messages that can be decoded is $m \leq M$. We can always have the first m messages to be the successfully decoded ones by reordering. When the channel state is known at the transmitter, the first m messages can be decoded successfully if and only if [21]:

$$iR \leq C_{m-i+1} + C_{m-i+2} + \cdots + C_M, \text{ for } i = 1, \ldots, m.$$

We can equivalently write these conditions as

$$R \leq \min_{i \in \{1,\ldots,m\}} \left[\frac{1}{m-i+1} \sum_{j=i}^{M} C_j \right]. \tag{2}$$

Then, for each channel realization $\{h[1], \ldots, h[M]\}$, the upper bound on the average throughput is given by $\frac{m^*}{M}R$, where m^* is the greatest m value that satisfies (2). This is an upper bound for each specific channel realization. An upper bound on $\overline{R}$ can be obtained by averaging this upper bound over the distribution of channel realizations.

Another upper bound on $\overline{R}$ can be found from the ergodic capacity assuming all messages are available at the encoder at the beginning, and letting M go to infinity. Finally, the bound $\overline{R} \leq R$ follows naturally from the data arrival rate. Thus, $\overline{R}$ is bounded above by $\min\left\{R, \overline{C}\right\}$.

IV. Transmission Schemes

The most straightforward transmission scheme consists in transmitting each message only within the channel block following its arrival, and discard it. This is called memoryless transmission (MT). Due to the i.i.d. nature of the channel, successful decoding probability is constant over messages. Denoting this probability by $p \triangleq Pr\{C_t \geq R\}$, the probability that exactly m messages are decoded is

$$\eta(m) \triangleq \binom{M}{m} p^m (1-p)^{M-m}. \tag{3}$$

The average throughput of the MT scheme $\overline{R}_{MT}$ is $\frac{R}{M}\sum_{m=1}^{M} m\eta(m) = Rp$. The MT scheme treats all messages equally. However, depending on the average channel conditions, it might be more beneficial to allocate more resources to some of the messages in order to increase the average throughput. In the following, we will consider three transmission schemes based on different types of resource allocation. We will find the average throughput for these schemes and compare them with the upper bound introduced in Section III.

A. Joint Encoding (JE) Transmission

In the *joint encoding (JE)* scheme we generate a single multiple-index codebook for each channel block. JE is also studied in [18] in terms of the diversity-multiplexing tradeoff (DMT) with a per-message delay constraint rather than the common deadline constraint we consider. Moreover, unlike in [18], here we study the JE scheme in terms of its average throughput for a fixed message rate R.

For channel block t, we generate a t dimensional codebook of size $s_1 \times \cdots \times s_t$, $s_i = 2^{nR}$, $\forall i \in \{1, \ldots, t\}$, with Gaussian distribution, and index the codewords as $\mathbf{x}_t(W_1, \ldots, W_t)$, where $W_i \in \mathcal{W} = \{1, \ldots, 2^{nR}\}$ for $i = 1, \ldots, t$. The receiver

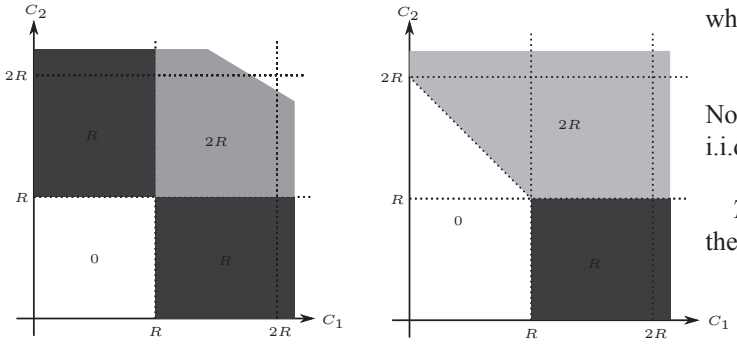


Fig. 3. Total decoded rate regions illustrated on the (C_1, C_2) plane with $M = 2$ messages for MT (on the left) and JE (on the right) schemes.

uses joint typicality decoder and tries to estimate as many messages as possible at the end of block M. The decoder will be able to decode the first m messages correctly if [21]:

$$(m-j+1)R \leq \sum_{t=j}^{m} C_t, \quad \forall \;\; j = 1, 2, \ldots, m. \tag{4}$$

As a comparison, we illustrate the achievable rate regions for MT and JE schemes for $M = 2$ in Fig. 3. In the case of MT, a total rate of $2R$ can be decoded successfully if both C_1 and C_2 are above R. We achieve a total rate of R if only C_1 or C_2 is above R. On the other hand, in the case of joint encoding, we tradeoff a part of the region of rate R for rate $2R$.

Using the conditions in Eqn. (4) we define functions $g^m(R)$, for $m = 0, 1, \ldots, M$, as follows:

$$g^m(R) = \begin{cases} 1, & \text{if } (m-j+1)R \leq \sum_{t=j}^{m} C_t, j = 1, \ldots, m, \\ 0, & \text{otherwise.} \end{cases}$$

Then the probability of decoding exactly m messages can be written as,

$$\eta(m) = Pr\left\{g^m(R) = 1 \text{ and } g^{m+1}(R) = 0\right\}. \tag{5}$$

After some manipulation, it is possible to prove that exactly m messages, $m = 0, 1, \ldots, M$, can be decoded if:

$$C_{m-i+1} + \cdots + C_m \geq iR, \quad i = 1, \ldots, m, \tag{6}$$

$$C_{m+1} + \cdots + C_{m+i} < iR, \quad i = 1, \ldots, M-m. \tag{7}$$

Then $\eta(m)$ can be calculated as in Eqn. (8) at the bottom of the next page, where we have defined $x^+ = \max\{0, x\}$, and $f_{C_1 \cdots C_m}(c_1, \ldots, c_m)$ as the joint pdf of $C_1, \ldots, C_m$, which is equal to the product of the marginal pdf's due to independence. The probability in Eqn. (8) cannot be easily evaluated for a generic M. However, we provide a much simpler way to calculate the average throughput $\overline{R}_{JE}$. This simplification is valid not only for i.i.d. but also for *conditionally i.i.d.* channels. Random variables $\{C_1, \cdots, C_M\}$ are said to be conditionally i.i.d. given a random variable U if the joint distribution is of the form

$$f_{C_1,\cdots,C_M,U}(c_1, \ldots, c_M, u) = f_{C_1|U}(c_1|u) \times \cdots \times f_{C_M|U}(c_M|u) f_U(u), \tag{9}$$

where

$$f_{C_i|U}(c_i|u) = f_{C_j|U}(c_j|u), \; \forall i, j \in \{1, \ldots M\}. \tag{10}$$

Note that i.i.d. channels is a particular case of conditionally i.i.d. channels where U is a constant.

Theorem 1: The average throughput for the JE scheme in the case of conditionally i.i.d. channel capacities is given by:

$$\overline{R}_{JE} = \frac{R}{M} \sum_{m=1}^{M} Pr\{C_1 + \cdots + C_m \geq mR\}. \tag{11}$$

Proof: See Appendix A.

In general it is still difficult to find an exact expression for $\overline{R}_{JE}$, but Theorem 1 simplifies the numerical analysis significantly. Moreover, it is possible to show that $\overline{R}_{JE}$ approaches R for large M if $\overline{C} > R$. To prove this, we rewrite Eqn. (11) as:

$$\overline{R}_{JE} = R - \frac{R}{M} \sum_{m=1}^{M} a_m, \tag{12}$$

where we have defined

$$a_m \triangleq Pr\left\{\frac{C_1 + \cdots + C_m}{m} < R\right\}. \tag{13}$$

It is sufficient to prove that, if $\overline{C} > R$, then $\lim_{M\to\infty} \sum_{m=1}^{M} a_m = c$, for some $0 < c < \infty$. We start by noting that $\lim_{m\to+\infty} a_m = 0$, since, by the law of large numbers, $\frac{C_1+\cdots+C_m}{m}$ converges to $\overline{C}$ in probability as m goes to infinity. To prove the convergence of the series sum we show that

$$\lim_{m\to+\infty} \frac{a_{m+1}}{a_m} = \lambda, \tag{14}$$

with $0 < \lambda < 1$.

We define:

$$l_m \triangleq \sqrt{m}\left(\overline{C} - \frac{C_1 + \cdots + C_m}{m}\right), \quad m = 1, 2, \ldots, M, \tag{15}$$

where each l_m is a random variable with zero mean and variance σ_c^2, which corresponds to the variance of the channel.

From the central limit theorem we can write:

$$\lim_{m\to+\infty} \frac{a_{m+1}}{a_m} = \lim_{m\to+\infty} \frac{Pr\left\{l_{m+1} > \frac{\overline{C}-r}{1/\sqrt{m+1}}\right\}}{Pr\left\{l_m > \frac{\overline{C}-r}{1/\sqrt{m}}\right\}} \tag{16}$$

$$= \lim_{m\to+\infty} \frac{Q\left(\frac{\overline{C}-R}{\sigma_c/\sqrt{m+1}}\right)}{Q\left(\frac{\overline{C}-R}{\sigma_c/\sqrt{m}}\right)} \tag{17}$$

$$\leq \lim_{m\to+\infty} \frac{\frac{\sigma_c/\sqrt{m+1}}{(\overline{C}-R)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\overline{C}-R}{\sigma_c/\sqrt{m+1}}\right)^2}}{\frac{\frac{\overline{C}-R}{\sigma_c/\sqrt{m}}}{1+\left(\frac{\overline{C}-R}{\sigma_c/\sqrt{m}}\right)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\overline{C}-R}{\sigma_c/\sqrt{m}}\right)^2}} \tag{18}$$

$$= \lim_{m\to+\infty} \frac{\sigma_c^2 + m(\overline{C}-R)^2}{\sqrt{m(m+1)}(\overline{C}-R)^2} e^{-\frac{(\overline{C}-R)^2}{2}\left[\frac{m+1}{\sigma_c^2} - \frac{m}{\sigma_c^2}\right]} \tag{19}$$

$$= e^{-\frac{(\overline{C}-R)^2}{2\sigma_c^2}} < 1, \tag{20}$$

where equality (17) follows from the fact that l_m converges in distribution to a Gaussian random variable with zero mean and variance σ_c^2, while inequality (18) follows from the bounds on the Q-function:

$$\frac{x}{(1+x^2)\sqrt{2\pi}}e^{-\frac{x^2}{2}} < Q(x) < \frac{1}{x\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \tag{21}$$

for $x > 0$. Similarly, we prove that if $\overline{C} < R$, the average rate tends to zero asymptotically with M. To see this, we consider the series in Eqn. (11) defining $b_m = Pr\{C_1 + \cdots + C_m \geq mR\}$. We want to prove that $\overline{R}_{JE} = M^{-1}\sum_{m=1}^{M} b_m$ converges to zero. It is sufficient to prove that $\sum_{m=1}^{M} b_m$ converges to a constant. We first notice that $\lim_{m\to+\infty} b_m = 0$ by the law of large numbers, as $M^{-1}\sum_{m=1}^{M} C_m$ converges in probability to $\overline{C}$, which is lower than R by hypothesis. One can also show that $\lim_{m\to+\infty} \frac{b_{m+1}}{b_m} = 0$; and hence, $\overline{R}_{JE}$ goes to zero with M. Overall we see that the average rate of the JE scheme shows a threshold behavior, i.e., we have:

$$\lim_{M\to\infty} \overline{R}_{JE} = \begin{cases} R, & \text{if } R < \overline{C} \\ 0, & \text{if } R > \overline{C}. \end{cases} \tag{22}$$

1) Adaptive Joint Encoding (aJE) Transmission: Eqn. (22) indicates a phase transition behavior such that $\overline{R}_{JE}$ is zero even for large M if $R > \overline{C}$. While the transmission rate R cannot be modified by the transmitter, it may choose to transmit only a fraction $\alpha = \frac{M'}{M} < 1$ of the messages, allocating the extra $M - M'$ channel blocks to the M' messages, effectively controlling the transmission rate. In other words, the M' messages are encoded and transmitted in M' channel blocks as described above, while each of the remaining $M - M'$ blocks is divided into M' equal parts, and the encoding process used for the first M' blocks is repeated, using independent codewords, across the M' parts of each block. For instance, if $M = 3$ and $M' = 2$, $\mathbf{x}_1(W_1)$ and $\mathbf{x}_2(W_1, W_2)$ are transmitted in the first and second channel blocks, respectively. The third channel block is divided into $M' = 2$ equal parts and the independent codewords $\mathbf{x}_{31}(W_1)$ and $\mathbf{x}_{32}(W_1, W_2)$ are transmitted in the first and in the second half of the block, respectively. We call this variant of the JE scheme *adaptive JE (aJE)* scheme. The conditions for decoding exactly m messages, $m = 0, 1, \ldots, M'$, in aJE can be obtained from those given in Eqn. (6) and Eqn. (7) by replacing C_i with $C_i^* = C_i + \frac{1}{M'}\sum_{j=M'+1}^{M} C_j$, $i \in \{1, \ldots, M'\}$. Note that the random variables C_i^*, $i \in \{1, \ldots, M'\}$, are conditionally i.i.d., i.e., they are i.i.d. once the variable $U = \frac{1}{M'}\sum_{j=M'+1}^{M} C_j$ is fixed. This implies that Theorem 1 holds.

In Appendix B we prove that, by choosing α such that $\alpha R \leq \overline{C}$, we can have:

$$\lim_{M\to\infty} \overline{R}_{aJE} = \min\{R, \overline{C}\}. \tag{23}$$

Eqn. (23) suggests that the average transmission rate can be adapted at the message level while keeping a fixed rate at the physical layer. Comparing the bound $\min\{R, \overline{C}\}$ and Eqn. (23) we see that the aJE scheme achieves the optimal average throughput in the limit of infinite M; hence, as the number of messages and the channel blocks go to infinity, the aJE scheme achieves the optimal performance. We will show in Section V through numerical analysis that near optimality of the aJE scheme is valid even for finite M. However, when there are multiple users or inaccuracy in the channel statistics information at the transmitter, aJE performs very poorly for users whose average received SNR is below the target value. In the following we propose alternative transmission schemes providing gradual performance degradation with decreasing SNR.

B. Time-Sharing (TS) Transmission

One of the resources that the encoder can allocate among different messages is the total number of channel uses within each channel block. While the whole first channel block has to be dedicated to message W_1 (the only available message), the second channel block can be divided among the messages W_1 and W_2, and so on so forth. Assume that the encoder divides the channel block t into t portions $\alpha_{1t}, \ldots, \alpha_{tt}$ such that $\alpha_{it} \geq 0$ and $\sum_{i=1}^{t} \alpha_{it} = 1$. In channel block t, $\alpha_{it} n$ channel uses are allocated to message W_i. A constant power P is used throughout the block. Then the total amount of received mutual information (MI) relative to message W_i is $I_i^{tot} \triangleq \sum_{t=i}^{M} \alpha_{it} C_t$. Letting $\alpha_{it} = 1$ if $t = i$ and $\alpha_{it} = 0$ otherwise, we obtain the MT scheme.

For simplicity, in the *time-sharing (TS)* scheme we assume equal time allocation among all the available messages; that is, for $i = 1, \ldots, M$, we have $\alpha_{it} = \frac{1}{t}$ for $t = i, i+1, \ldots, M$, and $\alpha_{it} = 0$ for $t = 1, \ldots, i$. The messages that arrive earlier are allocated more resources; and hence, are more likely to be decoded. We have $I_i^{tot} > I_j^{tot}$ for $1 \leq i < j \leq M$. Hence,

$$\overline{R}_{TS} = \frac{R}{M}\sum_{m=1}^{M} Pr\left\{\frac{C_m}{m} + \frac{C_{m+1}}{m+1} + \cdots + \frac{C_M}{M} \geq R\right\}. \tag{24}$$

1) Generalized Time-Sharing (gTS) Transmission: Note that, in TS transmission, message W_i is transmitted over $M - i + 1$ channel blocks, which allocates significantly more resources to the earlier messages. To balance the resource allocation between the messages, we consider transmitting each message over a limited window of channel blocks.

In *generalized time-sharing* transmission each message is encoded with equal time allocation over W consecutive blocks as long as the total deadline of M channel blocks is not met.

$$\eta(m) = \int_R^\infty \int_{(2R-x_m)^+}^\infty \cdots \int_{(mR-x_m-\cdots-x_2)^+}^\infty f_{C_1\cdots C_m}(x_1, \ldots, x_m)dx_1 \cdots dx_m \times \int_0^R \int_0^{2R-x_{m+1}} \cdots \int_0^{(M-m)R-x_{m+1}-\cdots-x_{M-1}} f_{C_{m+1}\cdots C_M}(x_{m+1}, \ldots, x_M)dx_{m+1} \cdots dx_M \tag{8}$$

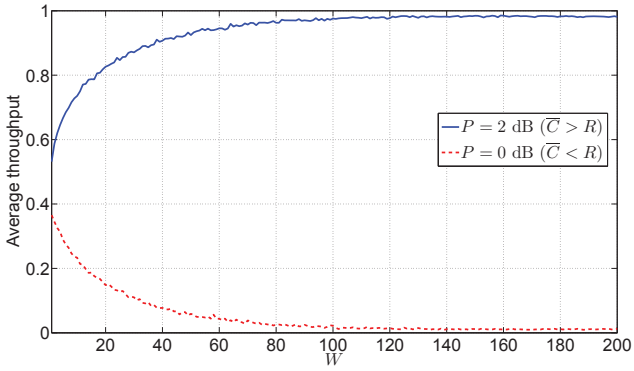


Fig. 4. Average throughput for the gTS scheme plotted against the window size W for $M = 10^4$ messages and $R = 1$ bpcu for two different average SNR values.

Messages from W_1 to W_{M-W+1} are encoded over a window of W blocks, while messages W_i, for $i \in \{M-W+2, M-W+3, \ldots, M\}$ are encoded over $M-i+1$ blocks. In particular we focus on the effect of variable W on the average throughput $\overline{R}_{gTS}$. In case $W \ll M$, most of the messages are transmitted over W slots together with $W-1$ other messages. In this case the MI accumulated for a generic message W_i is:

$$I_i^{tot} = \frac{1}{W} \sum_{t=i}^{i+W-1} C_t. \tag{25}$$

By the law of large numbers, Eqn. (25) converges in probability to the average channel capacity $\overline{C}$ as $W \to \infty$. Thus, we expect that, when the transmission rate R is above $\overline{C}$, the gTS scheme shows poor performance for large W (and hence, large M), while almost all messages are received successfully if $R < \overline{C}$. We confirm this by analyzing the effect of W on $\overline{R}_{gTS}$ numerically in Fig. 4 for $M = 10^4$ and $R = 1$ bpcu. For $P = 0$ dB we have $\overline{C} < R$, which leads to a decreasing $\overline{R}_{gTS}$ with increasing window size W. On the other hand, for $P = 2$ dB, we have $\overline{C} > R$, and accordingly $\overline{R}_{gTS}$ approaches 1 as W increases.

The same reasoning cannot be applied if the window size is on the order of the number of messages, as the number of initial messages which share the channel with less than $W-1$ other messages and the number of final messages which share the channel with more than $W-1$ messages are no longer negligible with respect to M. In Fig. 5(a), we plot $\overline{R}_{gTS}$ as a function of W for relatively small number of messages and $\overline{C} \geq R$. As seen in the figure, for a given M an optimal value of W can be chosen to maximize $\overline{R}_{gTS}$. Optimal W increases with M when $R < \bar{C}$. We plot $\overline{R}_{gTS}$ for $\overline{C} < R$ in Fig. 5(b). From the figure we see that $\overline{R}_{gTS}$ decreases monotonically with W up to a minimum, after which it increases almost linearly. The initial decrease in the average throughput is due to the averaging effect described above, which is relevant for $W \ll M$. The following increase in $\overline{R}_{gTS}$ is because the messages which are transmitted earlier (i.e., W_i with $i \ll W$) get an increasing amount of resources as W increases, and so the probability to be decoded increases. As a matter of fact, for each finite i, the average MI accumulated for message W_i grows indefinitely with W, i.e.:

$$\lim_{W\to\infty} E\left\{ \sum_{t=i}^{i+W-1} \frac{C_t}{\min\{t, W\}} \right\}$$
$$= \overline{C} \lim_{W\to\infty} \sum_{t=i}^{i+W-1} \frac{1}{\min\{t, W\}} = +\infty. \tag{26}$$

Thus, for a fixed i, letting W go to infinity leads to an infinite average MI, which translates into a higher $\overline{R}_{gTS}$. Note that this is valid only for relatively small i and large W, i.e., only messages transmitted earlier benefit from increasing W, while the rest of the messages are penalized. If W is small compared to M, as in the plot of Fig. 4 for $P = 0$ dB, the fraction of messages which benefit from the increasing W (i.e., messages $W_1, \ldots, W_L$ with $L \ll W$) remains small compared to M. For this reason $\overline{R}_{gTS}$ does not increase with W for the range of W considered in Fig. 4, while it does for the same range in Fig. 5(b).

Although the idea of encoding a message over a fraction of the available consecutive slots (e.g., $W < M$ for message W_1 in gTS) can be applied to all the schemes considered in this paper, the analysis becomes quite cumbersome. Hence, we restrict our analysis to the TS scheme as explained above.

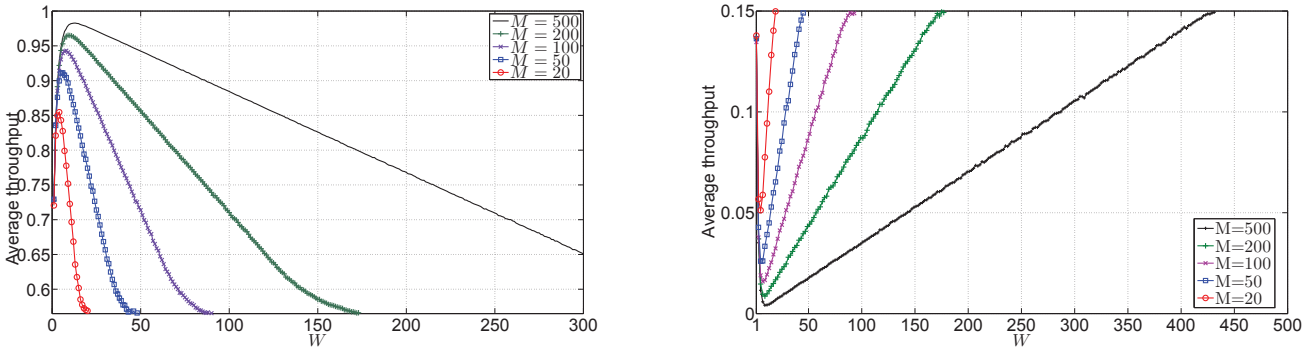
C. Superposition Transmission (ST)

In *superposition transmission (ST)* the superposition of t codewords, chosen from t independent Gaussian codebooks of size 2^{nR}, corresponding to the available messages $\{W_1, \ldots, W_t\}$ is transmitted in channel block t, $t \in \{1, \ldots, M\}$. The codewords are scaled such that the average total transmit power in each block is P. In the first block, only information about message W_1 is transmitted with average power $P_{11} = P$; in the second block we divide the total power P among the two messages, allocating P_{12} and P_{22} for W_1 and W_2, respectively. In general, over channel block t we allocate an average power P_{it} for W_i, while $\sum_{i=1}^{t} P_{it} = P$.

Let $\mathcal{S}$ be any subset of the set of messages $\mathcal{M} = \{1, \ldots, M\}$. We define $C(\mathcal{S})$ as follows:

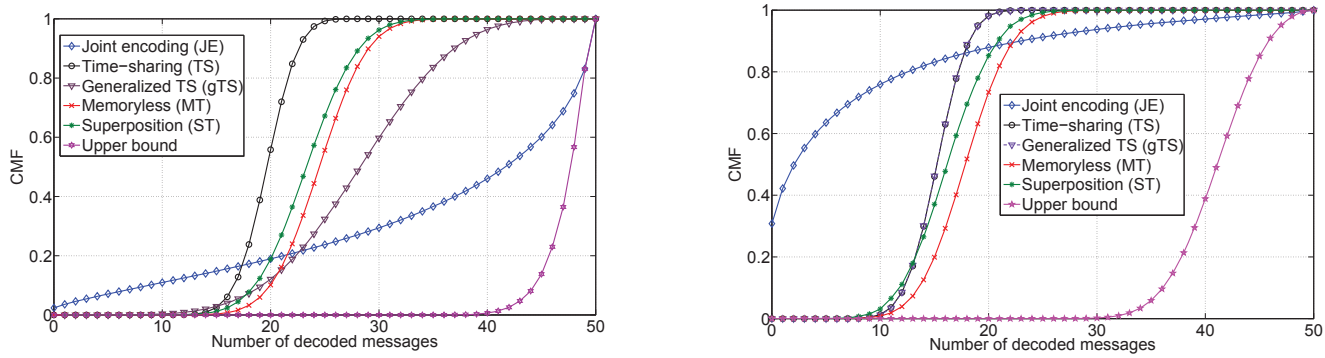
$$C(\mathcal{S}) \triangleq \sum_{t=1}^{M} \log_2 \left(1 + \frac{\phi[t] \sum_{s\in\mathcal{S}} P_{st}}{1 + \phi[t] \sum_{s\in\mathcal{M}\setminus\mathcal{S}} P_{st}} \right). \tag{27}$$

This provides an upper bound on the total rate of messages in set $\mathcal{S}$ that can be decoded jointly at the user considering the codewords corresponding to the remaining messages as noise. The receiver first checks if any of the messages can be decoded alone by considering the other transmissions as noise. If a message can be decoded, the corresponding signal is subtracted and the process is repeated over the remaining signal. If no message can be decoded alone, then the receiver considers joint decoding of message pairs, followed by triplets, and so on so forth. This algorithm gives us the maximum throughput. However, it is challenging in general to find a closed form expression for the average throughput, and optimize the power allocation. Hence, we focus here on the special case of equal power allocation, where we divide the total average power P among all the available messages at each channel block. The performance of the ST scheme will



(a) $R = 1$ bpcu, $P = 5$ dB ($\overline{C} > R$). (b) $R = 1$ bpcu, $P = -3$ dB ($\overline{C} < R$).

Fig. 5. Average throughput for the gTS scheme plotted against the window size W for different values of M.



(a) $P = 1.44$ dB ($\overline{C} > R$). (b) $P = 0$ dB ($\overline{C} < R$).

Fig. 6. The cumulative mass function (cmf) of the number of decoded messages for $R = 1$ bpcu and $M = 50$.

be studied in Section V numerically and compared with the other transmission schemes and the upper bound.

V. Numerical Results

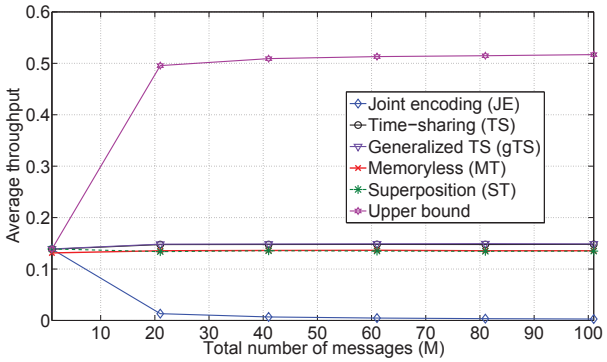
In this section we provide numerical results comparing the proposed transmission schemes. For the simulations we assume that the channel is Rayleigh fading, i.e., the channel gain $\phi(t)$ is exponentially distributed with parameter 1, i.e., $f_\Phi(\phi) = e^{-\phi}$ for $\phi > 0$, and zero otherwise.

In Fig. 6(a) the cumulative mass function (cmf) of the number of decoded messages is shown for the proposed transmission techniques for $R = 1$, $M = 50$ and $P = 1.44$ dB, which correspond to an average channel capacity $\overline{C} \simeq 1.07 > R$. We see that MT outperforms ST and TS schemes, as its cmf lays below the other two. The gTS scheme improves significantly compared to the ordinary TS scheme. On the other hand, the comparison with the JE scheme depends on the performance metric we choose. For instance, JE has the lowest probability to decode more than m messages, for $m \leq 15$, while it has the highest probability for $m \geq 22$.

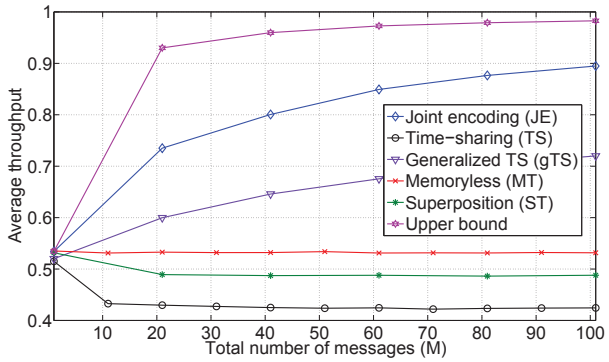
In Fig. 6(b) the cmf's for $P = 0$ dB are shown. In this case the average capacity is $\overline{C} \simeq 0.86$. Comparing Fig. 6(b) and Fig. 6(a), we see how the cmf of the JE scheme has different behaviors depending on whether $\overline{C}$ is above or below R. We see from Fig. 6(b) that for the JE scheme there is a probability of about 0.3 not to decode any message, while in all the other schemes such probability is zero. However, the JE scheme also has the highest probability to decode more than 23 messages. Furthermore, we note that the cmf of the gTS scheme converges to the cmf of the TS scheme at low SNR. This is because, as shown in Section IV-B1, when $\overline{C} < R$, the optimal window size W is equal to M, which is nothing but the TS scheme.

In Fig. 7(a) and Fig. 7(b) the average throughput is plotted against the delay constraint M for SNR values of -3 dB and 2 dB, respectively, and a message rate of $R = 1$ bpcu. While JE outperforms the other schemes at $SNR = 2$ dB, it has the poorest performance at $SNR = -3$ dB. This behavior is expected based on the threshold behavior of the JE scheme that we have outlined in Section IV-A. Note that the average capacity corresponding to $SNR = -3$ dB and 2 dB are $\overline{C} = 0.522 < R$ and $\overline{C} = 1.158 > R$, respectively.

When $\overline{C} > R$ the average throughput of the JE scheme gets close to the upper bound as the delay constraint M increases. It can be seen from Fig. 7(b) that, with a relatively short delay constraint of $M = 100$, the JE scheme achieves 90% of the average throughput of the upper bound. Also the gTS scheme improves its performance as the delay constraint gets larger,



(a) $P = -3$ dB ($\overline{C} < R$).



(b) $P = 2$ dB ($\overline{C} > R$).

Fig. 7. Average throughput vs delay constraint (total number of messages) M for $R = 1$ bpcu.

while the other schemes achieve an average throughput that is almost constant with M.

We observe that, in the low SNR regime, i.e., when $\overline{C} < R$, all the proposed schemes other than JE perform very close to each other, and significantly below the upper bound. The large gap to the informed transmitter upper bound is mostly due to the looseness of this bound in the low SNR regime.

Although none of the schemes dominates the others at all SNR values, it is interesting to note how, for a given SNR, the best scheme does not depend on the delay constraint. This may be useful in case of a practical implementation of the proposed schemes for the case of a single receiver or multiple receivers with similar average SNRs, as indicates that the SNR, rather than the delay, is the main design parameter. Further in this section we consider the case in which the receivers have different average SNR values.

In Fig. 8 the average throughput $\overline{R}$ is plotted against the transmission rate R for the case of $M = 100$ and $P = 20$ dB. The aJE scheme outperforms all the other schemes, performing very close to the upper bound, illustrating its rate adaptation capability. The number M' of messages transmitted in the aJE scheme is chosen so that $\frac{M'}{M} = 0.95\frac{\overline{C}}{R}$. In the figure we also show the upper bound obtained from the ergodic capacity $\min(R, \overline{C})$. It can be seen how it closely approximates the informed transmitter upper bound for $R < 6$. The JE scheme performs better than the others up to a certain transmission rate, beyond which rapidly becomes the worst one. This is due to the phase transition behavior exposed in Section IV-A in the case of asymptotically large delay, and observed here even for a relatively small M. Among the other schemes, MT achieves the highest average throughput in the region $R < 6.8$, while TS has the worst performance. The opposite is true in the region $R > 6.8$, where the curve of ST scheme is upper and lower bounded by the curves of the MT and TS schemes. We have repeated the simulations with different parameters (i.e., changing P and M) with similar results, that is, MT, TS, and ST schemes meet approximately at the same point, below which MT has the best performance of the three while above the intersection TS has the best performance. At the moment we have no analytical explanation for this observation, which would mean that there is always a scheme outperforming ST. We next study the performance of the considered schemes as a function of the distance from the transmitter.

We scale the average received power at the receiver with $d^{-\alpha_{pl}}$, where d is the distance from the transmitter to the receiver and α_{pl} is the path loss exponent. The results are shown in Fig. 9 for $P = 20$ dB, $M = 100$, $R = 1$ bpcu and a path loss exponent $\alpha_{pl} = 3$. The dependence of $\overline{R}$ on the distance is important, for instance, in the context of broadcast transmission in cellular networks, in which case the receiving terminals may have different distances from the transmitter. In such a scenario the range of the average channel SNR values at the receivers becomes important, and the transmitter should use a transmission scheme that performs well over this range. For instance, in a system in which all users have the same average SNR, such as a narrow-beam satellite system where the SNR within the beam footprint has variations of at most a few dB's on average [22], the transmission scheme should perform well around the average SNR of the beam. A similar situation may occur in a microcell, where the relatively small radius of the cell implies a limited variation in the average SNR. Instead, in the case of a macrocell, in which the received SNR may vary significantly from the proximity of the transmitter to the edge of the cell, the transmitter should adopt a scheme which performs well over a larger range of SNR values. In the range up to $d = 4$ the JE scheme achieves the highest average throughput while for $d \geq 6$ the TS scheme outperforms the others. The drop in the average throughput in the JE scheme when passing from $d = 4$ to $d = 5$ is similar to what we observe in Fig. 8 when the rate increases beyond $R = 6$ bpcu. In both cases the transition takes place as the transmission rate surpasses the average channel capacity. The aJE scheme, which selects the fraction of messages to transmit based on $\overline{C}$, outperforms all other schemes and gets relatively close to the informed transmitter upper bound and the ergodic capacity. It is interesting to observe that the behavior of the JE and the aJE schemes in case of finite delay constraint ($M = 100$) closely follows the results shown for the asymptotic case in Section IV-A. The aJE scheme adapts the average transmission rate at message

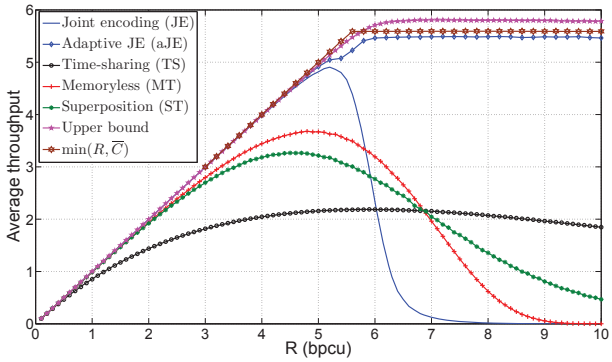


Fig. 8. Average throughput vs R for $P = 20$ dB and $M = 100$ messages. The upper bound $\min(R, \overline{C})$ is also shown.

level to the average channel capacity. We recall that, in the aJE scheme, the transmitter only has a statistical knowledge of the channel, and yet gets pretty close to the performance of a genie-aided transmitter even for a reasonably low number of channel blocks. We further notice how the adaptive JE scheme closely approaches the ergodic capacity, even though data arrives gradually at the transmitter during the transmission, instead of being available at the beginning, which is generally assumed for the achievability of the ergodic capacity [6]. We should note that in Fig. 9 the average transmission rate is optimized for each given distance for the aJE scheme, while such optimization is not done for the other schemes. Thus, when two (or more) terminals have different distances from the transmitter, the performance for the receivers might not be optimized simultaneously by the same scheme, and a tradeoff between the average throughputs of the two nodes is required. The performance can be improved by considering a combination of the aJE scheme with the TS or ST schemes. The plots in Fig. 9 show how TS, MT and ST schemes are more robust compared to the JE scheme, as their average throughput decreases smoothly with the distance, unlike the JE scheme, which has a sudden drop. This provides robustness in the case of multiple receivers with different average SNRs or when the channel statistics information at the transmitter is not accurate.

VI. Conclusions

We have considered a transmitter streaming data to a receiver over a block fading channel, such that the transmitter is provided with an independent message at a fixed rate at the beginning of each channel block. We have used the average throughput as our performance metric. We have proposed several new transmission schemes based on joint encoding, time-division and superposition encoding. A general upper bound on the average throughput has also been introduced assuming the availability of CSI at the transmitter.

We have shown analytically that the joint encoding (JE) scheme has a threshold behavior. It performs well when the target rate is below the average channel capacity $\overline{C}$, while its performance drops sharply when the target rate surpasses $\overline{C}$. To adapt to an average channel capacity that is below the fixed message rate R, the adaptive joint encoding (aJE) scheme transmits only some of the messages. We have proved analytically that the aJE scheme is asymptotically optimal as the number of channel blocks goes to infinity, even though data arrives gradually over time at a fixed rate, rather than being available initially. We have also shown numerically that, even for a finite number of messages, the aJE scheme outperforms other schemes in all the considered settings and performs close to the upper bound.

The JE and the aJE schemes create an M-block long concatenated code, and imposes a certain structure on the way messages can be decoded. For example, message W_k can be decoded only if message W_{k-1} can also be decoded. This is useful when the underlying application has a minimum rate requirement that needs to be satisfied over M channel blocks, or when the average SNRs of the users vary over a limited range of SNR values. Independent encoding used in time-sharing based schemes (TS, gTS, MT), instead, makes messages less dependent on the decoding of the other messages. The ST scheme, based on message superposition, collocates itself between JE- and TS-based schemes, as messages are encoded independently, but the probability of correctly decoding each one is affected by decoding of others.

We conclude that the aJE scheme is advantageous in systems with a single receiver or with multiple receivers having similar average SNR values, as the performance of the user with the highest average SNR is limited by the user with the lowest average SNR. On contrast, the gTS and ST schemes can be attractive when broadcasting to multiple users with a wide range of SNR values, or in a point-to-point system with inaccurate CSI, as the average throughput decreases gradually with decreasing SNR in these transmission schemes.

Appendix A

Proof of Theorem 1

Let B_k denote the event "the first k messages can be decoded at the end of channel block k", while B_k^c denotes the complementary event. The event B_k holds if and only if

$$C_{k-i+1} + C_{k-i+2} + \cdots + C_k \geq iR \tag{28}$$

is satisfied for all $i = 1, \ldots, k$. Let $E_{k,j}$ denote the event "the j-th inequality needed to decode the first k messages in k channel blocks is satisfied", that is:

$$E_{k,j} \triangleq \{C_{k-j+1} + \cdots + C_k \geq jR\}, \tag{29}$$

for $j = 1, \ldots, k$, while $E_{k,j}^c$ denotes the complementary event.

Note that in the JE scheme if m messages are decoded these are the first m messages. Let n_d denote the number of decoded messages at the end of channel block M. Then the average throughput is

$$\begin{aligned}\overline{R}_{JE} = R[&Pr\{n_d \geq 1\} + Pr\{n_d \geq 2\} + \\ &\cdots + \Pr\{n_d \geq M-1\} + \Pr\{n_d \geq M\}].\end{aligned} \tag{30}$$

The k-th term in the sum of Eqn. (30) is the probability of decoding *at least* k (i.e. k or more) messages. Each term in Eqn. (30) can be expressed as the sum of two terms as:

$$Pr\{n_d \geq k\} = Pr\{B_k, n_d \geq k\} + Pr\{B_k^c, n_d \geq k\} \tag{31}$$

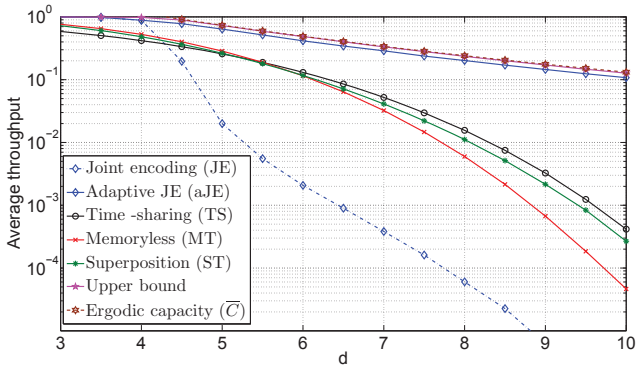


Fig. 9. Average throughput $\overline{R}$ vs distance from the transmitter for $R = 1$ bpcu, $M = 100$, $P = 20$ dB and $\alpha_{pl} = 3$.

The first term in Eqn. (31) is the probability of "decoding k messages at the end of channel block k *and* decoding at least k messages at the end of M channel blocks". If B_k holds, the event "decode at least k messages at the end of channel block M" is satisfied; hence, we have:

$$Pr\{B_k, n_d \geq k\} = Pr\{B_k\} = Pr\{E_{k,1}, \ldots, E_{k,k}\}. \quad (32)$$

As for the second term in Eqn. (31), it is the probability of decoding at least k messages but *not* k at the end of channel block k. It can be further decomposed into the sum of two terms:

$$\begin{aligned} Pr\{B_k^c, n_d \geq k\} &= Pr\{B_k^c, B_{k+1}, n_d \geq k\} \\ &+ Pr\{B_k^c, B_{k+1}^c, n_d \geq k\}. \end{aligned} \quad (33)$$

The event $n_d \geq k$ holds if the condition B_{k+1} is satisfied (i.e., if $k+1$ messages are decoded at the end of block $k+1$, then more than k messages are decoded at the end of channel block M); hence, we have:

$$Pr\{B_k^c, B_{k+1}, n_d \geq k\} = Pr\{B_k^c, B_{k+1}\}.$$

Plugging these into Eqn. (31), we obtain

$$\begin{aligned} Pr\{n_d \geq k\} &= Pr\{B_k\} + Pr\{B_k^c, B_{k+1}\} \\ &+ Pr\{B_k^c, B_{k+1}^c, n_d \geq k\}. \end{aligned} \quad (34)$$

We can continue in a similar fashion, so that, the event "at least k messages are decoded" can be written as the union of the disjoint events ("k messages are decoded in k slots") $\bigcup$ ("k messages are not decoded in k slots but $k+1$ messages are decoded in $k+1$ slots") $\bigcup \cdots \bigcup$ ("no message can be decoded before slot M but M messages are decoded in slot M"). Hence, by the law of total probability, we have:

$$Pr\{n_d \geq k\} = \sum_{j=k}^{M} Pr\{B_k^c, B_{k+1}^c, \ldots, B_{j-1}^c, B_j\}. \quad (35)$$

Note that each term of the sum in Eqn. (35) says nothing about what happens to messages beyond the j-th, which can either be decoded or not. Plugging Eqn. (35) in Eqn. (30) we find:

$$\begin{aligned} E[m] &= \sum_{k=1}^{M} \sum_{j=k}^{M} Pr\{B_k^c, B_{k+1}^c, \ldots, B_{j-1}^c, B_j\} \\ &= \sum_{j=1}^{M} \sum_{k=1}^{j} Pr\{B_k^c, B_{k+1}^c, \ldots, B_{j-1}^c, B_j\}. \end{aligned} \quad (36)$$

We can rewrite each of these events as the intersection of events of the kind $E_{k,i}$ and $E_{k,i}^c$. Each term of the sum in Eqn. (36) can be written as:

$$\begin{aligned} &Pr\{B_k^c, B_{k+1}^c, \ldots, B_{j-1}^c, B_j\} \\ &= Pr\{E_{k,1}, B_k^c, B_{k+1}, \ldots, B_{j-1}^c, B_j\} \\ &+ Pr\{E_{k,1}^c, B_k^c, B_{k+1}^c, \ldots, B_{j-1}^c, B_j\}. \end{aligned} \quad (37)$$

As the event $E_{k,1}^c$ implies the event B_k^c, this can be removed from the second term on the right hand side of Eqn. (37). Note that, in general, the event $E_{k,i}^c$, $i \in \{1, \cdots, k\}$ implies the event B_k^c. In order to remove the event B_k^c from the first term as well, we write it as the sum of probabilities of two disjoint events: one intersecting with $E_{k,2}$ and the other with $E_{k,2}^c$. Then we get:

$$\begin{aligned} &Pr\{B_k^c, B_{k+1}^c, \ldots, B_{j-1}^c, B_j\} \\ &= Pr\{E_{k,1}, E_{k,2}, B_k^c, \ldots, B_{j-1}^c, B_j\} \\ &+ Pr\{E_{k,1}, E_{k,2}^c, B_k^c, \ldots, B_{j-1}^c, B_j\} \\ &+ Pr\{E_{k,1}^c, B_{k+1}^c, \ldots, B_{j-1}^c, B_j\}. \end{aligned} \quad (38)$$

Now B_k^c can be removed from the second term of the sum thanks to the presence of $E_{k,2}^c$. Each of the terms in the right hand side of Eqn. (38) can be further written as the sum of the probabilities of two disjoint events, and so on so forth. The process is iterated until all the B_d^c, $d < j$ events are eliminated and we are left with the intersections of events only of the type $E_{p,q}$ and $E_{p,q}^c$, for some $p, q \in \{k, k+1, \ldots, M\}$ and B_j. The iteration is done as follows:

For each term of the summation, we take the B_l^c event with the lowest index. If any $E_{l,j}^c$ event is present, then B_l^c can be eliminated. If not, we write the term as the sum of the two probabilities corresponding to the events which are the intersections of the B_l^c event with $E_{l,d+1}$ and $E_{l,d+1}^c$, respectively, where d is the highest index j among the events in which $E_{l,j}$ is already present. The iteration process stops when $l = j$.

At the end of the process all the probabilities involving events $B_k^c, \ldots, B_{j-1}^c$ will be removed and replaced by sequences of the kind:

$$\begin{aligned} \{&E_{k,1}, E_{k,2}, \ldots, E_{k,i_k}^c, E_{k+1,i_k+1}, \ldots, \\ &E_{k+1,i_{k+1}}^c, \ldots, E_{j-1,i_{j-2}+1}, E_{j-1,i_{j-1}}^c, B_j\}, \end{aligned} \quad (39)$$

where $i_{j-1} \in \{j-1-k, \ldots, j-1\}$ is the index corresponding to the last inequality required to decode $j-1$ messages, which is not satisfied. Note that exactly one $E_{l,r}^c$ event for each B_l^c is present after the iteration.

For B_j to hold, all the events $E_{j,1}, \ldots, E_{j,j}$ must hold. It is easy to show that, after the iterative process used to remove the

B_l^c's, the event $E_{j,i_{j-1}+1}$ ensures that all the events required for B_j with indices lower than or equal to i_{j-1} automatically hold. Thus, we can add the events $\{E_{j,i_{j-1}+1}, \dots, E_{j,j}\}$ to guarantee that B_j holds, and remove it from the list. It is important to notice that the term $E_{j,j}$ is always present. At this point we are left with the sum of probabilities of events, which we call *E-events*, each of which is the intersection of events of the form $E_{i,j}$ and $E_{i,j}^c$. Thus, an E-event S_k^j has the following form:

$$S_k^j \triangleq \{E_{k,1}, E_{k,2}, \dots, E_{k,i_k}^c, E_{k+1,i_k+1}, \dots, E_{k+1,i_{k+1}}^c, \dots, E_{j-1,i_{j-2}+1}, E_{j-1,i_{j-1}}^c, E_{j,i_{j-1}+1}, \dots, E_{j,j}\}. \quad (40)$$

By construction, the number of E-events for the generic term j of the sum in Eqn. (36) is equal to the number of possible dispositions of $j-k$ E^c's over $j-1$ positions. As the number of events of type E^c is different for the E-events of different terms in Eqn. (36), the E-events relative to two different terms of Eqn. (36) are different. We define $\mathcal{S}_j$ as the set of all E-events which contain the event $E_{j,j}$. The elements of $\mathcal{S}_j$ correspond to all the possible ways in which j messages can be decoded at the end of block number j. The cardinality of $\mathcal{S}_j$ is equal to:

$$|\mathcal{S}_j| = \sum_{k=1}^{j} \frac{(j-1)!}{(k-1)!(j-k)!} = 2^{j-1}. \quad (41)$$

Now we want to prove that

$$\sum_{S_k^j \in \mathcal{S}_j} Pr\{S_k^j\} = Pr\{E_{j,j}\}. \quad (42)$$

Note that each $E_{k,l}$ corresponds to a different event if the index k is different, even for the same index l; thus, the law of total probability can not be directly applied to prove Eqn. (42). However, we will prove in the following preposition that $Pr\{E_{k_1,l}\} = Pr\{E_{k_2,l}\}$, $\forall k_1, k_2$.

Proposition 1: Let us consider a set of random variables $C_1, \cdots, C_j$ that are conditionally i.i.d. given U. Given any two ordering vectors $\mathbf{i} = i_1, i_2, \cdots, i_j$ and $\mathbf{l} = l_1, l_2, \cdots, l_j$, we have

$$\begin{aligned} &Pr\{C_{i_1} \gtrless R, \dots, C_{i_1} + \cdots + C_{i_j} \gtrless jR\} \\ &= Pr\{C_{l_1} \gtrless R, \dots, C_{l_1} + \cdots + C_{l_j} \gtrless jR\}. \end{aligned} \quad (43)$$

Proof: The left hand side of Eqn. (43) can be rewritten as:

$$\begin{aligned} &Pr\{C_{i_1} \gtrless R, \dots, C_{i_1} + \cdots + C_{i_j} \gtrless jR\} \\ &= \int_{-\infty}^{+\infty} du \int_{\theta_1^{low}}^{\theta_1^{up}} dc_{i_1} \dots \int_{\theta_j^{low}}^{\theta_j^{up}} dc_{i_j} f_{\mathbf{C_i}|U}(\mathbf{c_i}|u) f_U(u), \end{aligned} \quad (44)$$

where $\mathbf{C_i} = C_{i_1}, \dots, C_{i_j}$ and $\mathbf{c_i} = c_{i_1}, \dots, c_{i_j}$, while θ_h^{low} and θ_h^{up} are the lower and upper extremes of the integration interval. θ_h^{low} is either equal to $-\infty$ or to $hR - c_{i_1} - \cdots - c_{i_{h-1}}$, $\forall h \in \{1, \dots, j\}$, depending on whether there is a $<$ or a $\geq$ in the h-th inequality within brackets in Eqn. (44), respectively, while θ_h^{up} is either equal to $hR - c_{i_1} - \cdots - c_{i_{h-1}}$ or to $+\infty$ depending on whether there is a $<$ or a $\geq$ in the h-th inequality of Eqn. (44), respectively. By using Eqn. (9) and Eqn. (10) we can write:

$$\begin{aligned} &Pr\{C_{i_1} \gtrless R, \dots, C_{i_1} + \cdots + C_{i_j} \gtrless jR\} \\ &= \int_{-\infty}^{+\infty} du f_U(u) \int_{\theta_1^{low}}^{\theta_1^{up}} dc_{i_1} \dots \int_{\theta_j^{low}}^{\theta_j^{up}} dc_{i_j} f_{\mathbf{C}_i|U} \\ &= \int_{-\infty}^{+\infty} du f_U(u) \int_{\theta_1^{low}}^{\theta_1^{up}} dc_{l_1} \dots \int_{\theta_j^{low}}^{\theta_j^{up}} dc_{l_j} f_{\mathbf{C}_l|U} \\ &= Pr\{C_{l_1} \gtrless R, \dots, C_{l_1} + \cdots + C_{l_j} \gtrless jR\}, \end{aligned} \quad (45)$$

where we defined $f_{\mathbf{C}_i|U} \triangleq f_{C_{i_1}|U}(c_{i_1}|u) \cdots f_{C_{i_j}|U}(c_{i_j}|u)$ and $f_{\mathbf{C}_l|U} \triangleq f_{C_{l_1}|U}(c_{l_1}|u) \times \cdots \times f_{C_{l_j}|U}(c_{l_j}|u)$ The proposition above guarantees that, although these events do not partition the whole probability space of $E_{j,j}$, their probabilities add up to that of $E_{j,j}$, i.e.:

$$\begin{aligned} \sum_{k=1}^{2^{j-1}} Pr\{S_k^j\} &= Pr\{E_{j,j}\} \\ &= Pr\{C_1 + \cdots + C_j \geq jR\}. \end{aligned} \quad (46)$$

Finally, plugging Eqn. (46) into Eqn. (36) we can write:

$$E[m] = \sum_{j=1}^{M} \sum_{k=1}^{j} Pr\{B_k^c, B_{k+1}^c, \dots, B_{j-1}^c, B_j\} \quad (47)$$

$$= \sum_{j=1}^{M} \sum_{S_k^j \in \mathcal{S}_j} Pr\{S_k^j\}$$

$$= \sum_{j=1}^{M} Pr\{C_1 + \cdots + C_j \geq jR\}. \quad (48)$$

APPENDIX B

In the following we prove that the average throughput of the aJE scheme $\overline{R}_{aJE}$ approaches αR for large M if $\overline{C} > \alpha R$. Similarly to the JE scheme, it is sufficient to prove that, if $\overline{C} > \alpha R$,

$$\lim_{M \to \infty} \sum_{m=1}^{M\alpha} a_m^* = c, \quad (49)$$

for some $0 < c < \infty$, where $a_m^* \triangleq Pr\left\{\frac{C_1^* + \cdots + C_m^*}{m} < R\right\}$. We can rewrite a_m^* as:

$$a_m^* = Pr\left\{ l_m > \frac{\overline{C}/\alpha - R}{\sigma_c \sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{M\alpha^2}\right)}} \right\}, \quad (50)$$

where

$$l_m \triangleq \frac{\overline{C}/\alpha - \frac{C_1 + \cdots + C_m}{m} - \frac{(1-\alpha)}{\alpha} \frac{1}{M(1-\alpha)} \sum_{j=M\alpha+1}^{M} C_j}{\sigma_c \sqrt{\frac{1}{m} + \frac{1-\alpha}{M\alpha^2}}} \quad (51)$$

is a random variable with zero mean and unit variance. From the law of large numbers applied to Eqn. (50), we have $\lim_{m \to +\infty} a_m^* = 0$. First we show that

$$\lim_{m \to +\infty} \left(\frac{a_m^*}{d_m} \right) = c', \quad (52)$$

for some $0 < c' < +\infty$ where we have defined:

$$d_m \triangleq Pr\left\{ l'_m > \frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}\right)}} \right\}, \tag{53}$$

and

$$l'_m \triangleq \frac{\overline{C}/\alpha - \frac{C_1+\cdots+C_m}{m} - \frac{(1-\alpha)}{\alpha}\frac{1}{m(1-\alpha)}\sum_{j=M\alpha+1}^{M\alpha+m} C_j}{\sigma_c\sqrt{\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}}} \tag{54}$$

such that l'_m is a random variable with zero mean and unit variance. We have

$$\begin{aligned}
\lim_{m\to+\infty}\left(\frac{a^*_m}{d_m}\right) &= \lim_{m\to+\infty} \frac{Pr\left\{ l_m > \frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{M\alpha^2}\right)}} \right\}}{Pr\left\{ l'_m > \frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}\right)}} \right\}} \\
&= \lim_{m\to+\infty} \frac{Q\left(\frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{M\alpha^2}\right)}}\right)}{Q\left(\frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}\right)}}\right)} \qquad (55) \\
&\leq 1, \qquad (56)
\end{aligned}$$

where inequality (56) follows from the fact that $m < M$ and $Q(x)$ is monotonically decreasing in x. Then we show that

$$\lim_{M\to\infty}\sum_{m=1}^{M\alpha} d_m = c'', \tag{57}$$

for some $0 < c'' < +\infty$. To prove the convergence of the series sum we show that $\lim_{m\to+\infty}\frac{d_{m+1}}{d_m} = \lambda'$, for some $0 < \lambda' < 1$. From the central limit theorem we can write:

$$\begin{aligned}
\lim_{m\to+\infty}\frac{d_{m+1}}{d_m} &= \lim_{m\to+\infty} \frac{Pr\left\{ l_{m+1} > \frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m+1} + \frac{1-\alpha}{(m+1)\alpha^2}\right)}} \right\}}{Pr\left\{ l_m > \frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}\right)}} \right\}} \\
&= \lim_{m\to+\infty} \frac{Q\left(\frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m+1} + \frac{1-\alpha}{(m+1)\alpha^2}\right)}}\right)}{Q\left(\frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}\right)}}\right)} \\
&\leq \lim_{m\to+\infty} \frac{\frac{\sigma_c\sqrt{\left(\frac{1}{m+1} + \frac{1-\alpha}{(m+1)\alpha^2}\right)}}{(\overline{C}/\alpha - R)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m+1} + \frac{1-\alpha}{(m+1)\alpha^2}\right)}}\right)^2}}{\frac{\frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}\right)}}}{1+\left(\frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}\right)}}\right)^2}\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\overline{C}/\alpha - R}{\sigma_c\sqrt{\left(\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}\right)}}\right)^2}} \qquad (58) \\
&= \lim_{m\to+\infty} \frac{\sqrt{\left(\frac{1}{m+1} + \frac{1-\alpha}{(m+1)\alpha^2}\right)}}{(\overline{C}/\alpha - R)^2} \\
&\quad \cdot \frac{\left[\sigma_c^2\left(\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}\right) + (\overline{C}/\alpha - R)^2\right]}{\sqrt{\frac{1}{m} + \frac{1-\alpha}{m\alpha^2}}} e^{-\frac{(\overline{C}/\alpha - R)^2}{2\sigma_c^2}\left(\frac{\alpha^2}{\alpha^2-\alpha+1}\right)} \qquad (59) \\
&= e^{-\frac{(\overline{C}/\alpha - R)^2}{2\sigma_c^2}\left(\frac{\alpha^2}{\alpha^2-\alpha+1}\right)} < 1, \qquad (60)
\end{aligned}$$

where inequality (58) follows from Eqn. (21).

From Eqn. (60) it follows that $\lim_{M\to\infty}\overline{R}_{aJE} = R$ if $\alpha R < \overline{C}$. Similarly, it can be easily shown that $\lim_{M\to\infty}\overline{R}_{aJE} = 0$ if $\alpha R > \overline{C}$.

REFERENCES

[1] European Telecommunications Standards Institute, "ETSI TR 102 376 V1.1.1 (2005-02) Digital Video Broadcasting (DVB) user guidelines for the second generation system for broadcasting, interactive services, news gathering and other broadband satellite applications (DVB-S2)," Feb. 2005.

[2] D. Tse, R. Yates, and L. Zang, "Fading broadcast channels with state information at the receivers," in *IEEE Allerton Conf. on Comm., Control, and Computing*, Monticello, IL, U.S.A., Sep. 2008.

[3] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. on Info. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[4] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[5] A. Goldsmith, *Wireless Communication*. Cambridge University Press, 2005.

[6] A. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. on Info. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[7] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Comm.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[8] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. on Info. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[9] A. S. Avestimehr and D. N. C. Tse, "Outage capacity of the fading relay channel in the low snr regime," *IEEE Trans. on Info. Theory*, vol. 53, no. 4, pp. 1401–1415, Apr. 2007.

[10] S. Shamai, "A broadcast strategy for the Gaussian slowly fading channel," in *IEEE Int'l Symp. on Info. Theory*, Ulm, Germany, June-July 1997.

[11] S. Shamai and A. Steiner, "A broadcast approach for a single user slowly fading mimo channel," *IEEE Trans. on Info. Theory*, vol. 49, no. 10, pp. 2617–2635, Oct. 2003.

[12] D. Gündüz and E. Erkip, "Opportunistic cooperation by dynamic resource allocation," *IEEE Trans. on Wireless Comm.*, vol. 6, no. 4, pp. 1446–1454, Apr. 2007.

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

[14] O. Oyman and S. Singh, "Quality of experience for HTTP adaptive streaming services," *IEEE Comm. Magazine*, vol. 50, no. 4, pp. 20–27, Apr. 2012.

[15] M. C. O. Bogino, P. Cataldi, M. Grangetto, E. Magli, and G. Olmo, "Sliding-window digital fountain codes for streaming multimedia contents," in *Int'l Symp. on Circuits Systems*, New Orleans, LA, U.S.A., May 2007.

[16] A. Badr, A. Khisti, and E. Martinian, "Diversity embedded streaming erasure codes (DE-SCo): Constructions and optimalty," in *IEEE Global Telecomm. Conf.*, Miami, FL, U.S.A., Dec. 2010.

[17] D. Leong and T. Ho, "Erasure coding for real-time streaming," in *IEEE Int'l Symp. Info. Theory (ISIT)*, Boston, MA, U.S.A., July 2012.

[18] A. Khisti and S. C. Draper, "Streaming data over fading wireless channels: The diversity-multiplexing tradeoff," in *IEEE Int'l Symp. on Info. Theory*, St. Petersburg, Russia, Aug. 2011.

[19] G. Cocco, D. Gunduz, and C. Ibars, "Real-time broadcasting over block-fading channels," in *Int'l Symp. on Wireless Comm. Systems*, Aachen, Germany, Nov. 2011.
[20] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, May 1994.
[21] V. V. Prelov, "Transmission over a multiple access channel with a special source hierarchy," *Problemy Peredachi Informatsii*, vol. 20, no. 4, pp. 3–10, Oct.-Dec. 1984.
[22] G. Maral and M. Bousquet, *Satellite Communications Systems - systems, techniques and technologies*. John Wiley & Sons Ltd, 2009.



Giuseppe Cocco (Italy) received his M.S. degree in Telecommunications Engineering (*Cum Laude*) from the University of Pisa in Pisa, Italy, in 2008 and his PhD in Telecommunications Engineering from the Polytechnic University of Catalonia in Barcelona, Spain, in 2013. Currently he is a post-doctoral researcher at CTTC. He took part in several research projects funded by the European Commission, the European Space Agency and the Spanish Government such as ESA's NGW, ICT-Exalted, ICT-Newcom++ and IST-WIP.

In 2009 he was awarded the Networking/Partnering Initiative (NPI) grant by the European Space Agency (ESA) for his research proposal on gap-filler-less mobile satellite networks. From July 2010 to July 2011 he stayed at ESA's European Space Research and Technology Center (ESTEC) in the Netherlands as a visiting PhD student. His research interests lie in the fields of satellite communications, network coding and information theory.



Deniz Gündüz received the B.S. degree in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey in 2002, and the M.S. and Ph.D. degrees in electrical engineering from Polytechnic Institute of New York University, Brooklyn, NY in 2004 and 2007, respectively.

Currently he is a lecturer in the Electrical and Electronic Engineering Department of Imperial College London, London, UK. He was a research associate at CTTC in Barcelona, Spain from November 2009 until September 2012. He also held a visiting researcher position at Princeton University from November 2009 until November 2011. Previously he was a consulting assistant professor at the Department of Electrical Engineering, Stanford University, and a postdoctoral Research Associate at the Department of Electrical Engineering, Princeton University.

He is the recipient of a Marie Curie Reintegration Grant funded by the European Commission, the 2008 Alexander Hessel Award of Polytechnic Institute of New York University given to the best PhD Dissertation, and a recipient of the Best Student Paper Award at the 2007 IEEE International Symposium on Information Theory (ISIT).

He is an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS, and served as a guest editor of the EURASIP Journal on Wireless Communications and Networking, Special Issue on Recent Advances in Optimization Techniques in Wireless Communication Networks. He is a co-chair of the Network Theory Symposium at the 2013 IEEE Global Conference on Signal and Information Processing (GlobalSIP), and also served as a co-chair of the 2012 IEEE European School of Information Theory (ESIT). His research interests lie in the areas of communication theory and information theory with special emphasis on joint source-channel coding, multi-user networks, energy efficient communications and security.



Christian Ibars received degrees in electrical engineering from Universitat Politecnica de Catalunya, Barcelona, Spain, and Politecnico di Torino, Torino, Italy, in 1999, and a Ph. D. degree in electrical engineering from the New Jersey Institute of Technology, Newark, NJ, in 2003.

During 2000, he was a visiting student at Stanford University. Since 2003 he has been with the Centre Tecnologic de Telecomunicacions de Catalunya, Castelldefels, Spain, where he is now head of the Communications Systems Division.

He has also held external lecturer appointments at Universitat Pompeu Fabra during 2006-2008. During 2010, he was a research scholar at the Cisco Research Center, and during 2010-2011 he was a Marie Curie Fellow at the University of Padova. In 2012 he served in a Specialist Task Force on interference mitigation in ETSI. He participated in several projects funded by the National Science Foundation, the European Commission, the European Space Agency and the Spanish Government, such as Free Bits, IST-Winner, IST-WIP, IST-Coopcom, ICT-Exalted, ICT-Emphatic, and MIMOSat. He has published over 80 international journal and conference papers. His current research interests include wireless multiuser communications, cooperative systems, interference coordination, and radio resource management.