



CHALMERS

Chalmers Publication Library

Feedback Subsampling in Temporally-Correlated Slowly-Fading Channels using Quantized CSI

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

IEEE Transactions on Communications (ISSN: 0090-6778)

Citation for the published paper:

Makki, B. ; Eriksson, T. (2013) "Feedback Subsampling in Temporally-Correlated Slowly-Fading Channels using Quantized CSI". IEEE Transactions on Communications(99), pp. 1-13.

<http://dx.doi.org/10.1109/TCOMM.2013.042313.120167>

Downloaded from: <http://publications.lib.chalmers.se/publication/178646>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

Feedback Subsampling in Temporally-Correlated Slowly-Fading Channels using Quantized CSI

Behrooz Makki and Thomas Eriksson

Abstract

This paper studies the problem of feedback subsampling in temporally-correlated wireless networks utilizing quantized channel state information (CSI). Under both peak and average power constraints, the system data transmission efficiency is studied in two scenarios. First, we focus on the case where the codewords span one fading block. In the second scenario, the throughput is determined for *piecewise* slowly-fading channels where the codewords are so long that a finite number of correlated gain realizations are experienced during each codeword transmission. Considering different temporal correlation conditions in both scenarios, substantial throughput increment is observed with feedback rates well below 1 bit per slot.

Behrooz Makki and Thomas Eriksson are with Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden,
Email: {behrooz.makki, thomase}@chalmers.se

I. INTRODUCTION

Background: Adaptive resource allocation techniques are useful tools to increase the data transmission efficiency and reliability of wireless networks [1]–[7]. The main idea behind these schemes is to adapt some transmission parameters, such as the rate and power, based on the channel quality. Therefore, it is essential to have good estimates of the channel quality at both end-points of the communication link. Channel state estimation at the receiver is relatively simple and incurs negligible loss in the transmission rate, particularly when the channel experiences slow variations. On the other hand, due to the signaling load caused by reporting the channel information, assuming perfect channel knowledge at the transmitter is an overly optimistic assumption, which does not match with reality¹. This is the main motivation for the present limited channel state information (CSI) feedback systems, e.g., [1]–[10].

Among different techniques proposed for feedback load reduction in wireless networks, CSI quantization schemes [11]–[17] and feedback subsampling [1]–[4] have received considerable attention recently. Implementing quantized CSI, the transmitter is provided with a rough measure of the channel quality obtained by the channel gain quantization. Subsampling techniques, on the other hand, deal with the time-domain correlations between the adjacent blocks². Here, the CSI is fed back in specific time slots and in the other blocks it is estimated via, e.g., minimum mean square error (MMSE) estimators [1]–[4].

In theoretical investigations, the communication channels are normally modeled as slowly-fading channels. In this channel model, the fading coefficients are assumed to be random variables that remain constant over time intervals of fixed duration, determined by the channel coherence time, and vary across successive blocks in an independent and identically distributed (iid) manner [11]–[19]. This is a useful model particularly for analytical performance analysis, e.g., in computation of error probability. For feedback compression, however, the slowly-fading is not an appropriate model, because the channel temporal correlations can not be exploited. Also, the basic slowly-fading channel is not always a realistic model [20]–[30]. For instance, the indoor ultra wideband (UWB) channels vary slowly and smoothly in successive fading blocks. Although there may be occasional sharp transitions due to sudden changes in the surrounding environment (e.g., when a person walks into the room [28], [29]), these changes occur infrequently, and the adjacent fading blocks are typically temporally-

¹For instance, in the 3G cellular systems the available feedback rate is limited to only 1.5 kbits per slot [8].

²Subsampling could also be applied in the frequency domain, but transform coding approaches have proven quite successful for exploiting frequency correlation, e.g., [1]–[4].

correlated. Moreover, for the fast-moving users the channel may change during a codeword transmission period. For these reasons, it is interesting to extend the basic slowly-fading models to the case where there is temporal correlation between the successive channel realizations.

Related works: There are many scientific reports tackling the partial CSI problem from both practical and theoretical aspects³ [1], [2]. For instance, different practical feedback reduction schemes have been proposed by [3]–[7], [31]–[34].

From an information-theoretic perspective, there are many results dealing with quantized CSI in slowly-fading channels; Kim and Skoglund [12] found the channel average rate in the presence of quantized CSI and investigated different optimality conditions for the quantization boundaries. Further, Ekbatani et. al, [13], [14] studied the effect of feedback channel noise on the performance of the slowly-fading channels. As another kind of partial CSI feedback, [15]–[19] implemented hybrid automatic repeat request (HARQ) in slowly-fading channels and [16], [17] compared the results with the ones obtained by quantized CSI. References [12]–[19] are all based on the basic uncorrelated slowly-fading channel assumption.

Temporal correlation in limited-feedback schemes has been considered by a number of information theorists. For example, [20], [21] investigated the performance of temporally-correlated multiple-input-multiple-output (MIMO) channels in the presence of feedback delay or other cells interference. A linear precoding technique was proposed in [22] for temporally-correlated MIMO channels utilizing quantized CSI. Implementation of HARQ in temporally-correlated channels has been considered by, e.g., [23]–[26]. Finally, we studied the effect of dynamic quantizers on the throughput and feedback load of correlated channels [27].

Reviewing the literature, there are some points that are the main motivators for this paper:

- It has been previously shown that feedback compression via exploiting the temporal correlations is crucial for practical implementation of many communication systems⁴. However, feedback subsampling, which is one of the most powerful methods for feedback compression, has not been well analyzed yet. (Meanwhile there are a number of simulation-based works dealing with feedback subsampling [3], [4].)
- To the best of our knowledge, all related slowly-fading channel-based papers assume the codewords to

³Due to extremely high number of papers dealing with partial CSI, it is not possible to mention all related works here. We apologize to the authors whose papers we have not included in our list and refer the readers to, e.g., [1], [2] for deeper review of the partial CSI schemes.

⁴For instance, the amount of CSI required for proper implementation of orthogonal frequency-division multiplexing (OFDM) and MIMO broadcast channels is not practically affordable if temporal and frequency correlations are not exploited for feedback compression [4], [35].

span a single fading block. On the other hand, modern codes often use very long codewords, which may exceed the channel coherence time [36]. Moreover, as stated before, the fading parameters of the fast-moving users or the UWB communication setups may change during a codeword transmission. Therefore, it is interesting to evaluate the system performance in the cases where the channel takes a finite number of values during a codeword transmission period.

Contributions: This paper studies the problem of feedback subsampling in temporally-correlated channels. The problem is cast in form of maximizing the throughput subject to different peak and average power constraints, when the transmitter is provided with subsampled quantized CSI. The results are obtained for different fading models where the channel remains constant during a codeword transmission or takes a finite number of correlated values within a codeword transmission period, which models the time-varying channels.

As opposed to [3], [4], we present analytical results for the feedback subsampling problem. Also, our problem setup is different from the ones studied in [12]–[27], because 1) feedback subsampling, on which we focus, was not investigated in those papers and 2) we consider the fading variations during a codeword transmission.

The results show that, depending on the temporal correlation conditions, substantial throughput increment can be achieved with feedback rates well below 1 bit per slot (bps). Also, there are average power thresholds below which an on-off approach is the best scheme maximizing the throughput of subsampling-based approaches. We show that the network data transmission efficiency is overestimated if the fading variations during a codeword transmission are approximated by their average value. Finally, the arguments of the paper can be used to extend many schemes previously proposed for slowly-fading networks to the ones in fast-fading channels.

An overview of the paper: The results of the paper are obtained in two scenarios, as stated in Table 1. First, it is assumed that, while there is correlation between successive gain realizations, each codeword spans a single fading block (Fig. 1). In this case, called *short-length coding* scheme, the transmitter utilizes the quantized CSI received in 1 out of K blocks to estimate the optimal transmission parameters of the next codewords (feedback subsampling). Here, the throughput is obtained for delay-free and delayed feedback conditions and is compared with the throughput in HARQ protocols.

In the second scenario, the codewords are assumed to span a number of fading blocks (Fig. 1). In other words, the basic slowly-fading model assumption is relaxed to a piecewise slowly-fading model where the channel gain can have a finite number of correlated values in each codeword transmission time slot. We denote this model as *long-length coding* approach. Here, the system throughput is obtained for different power constraints

and the effect of repetition codes on the system performance is evaluated.

II. SYSTEM MODEL

Channel model: We consider a correlated slowly-fading channel. That is, the fading random variable is assumed to be constant in fading blocks of length L channel uses, normally determined based on the channel coherence time. In this way, utilizing codewords of length L , the channel output at, e.g., the k -th fading block can be represented as

$$Y_k[i] = H_k X_k[i] + Z_k[i], \quad i = 1, \dots, L \quad (1)$$

where the input message X_k multiplied by the fading random variable H_k is summed with an iid complex Gaussian noise $Z_k \sim \mathcal{CN}(0, N_0)$ resulting in the output Y_k . With no loss of generality, we set $N_0 = 1$ and define $G_k \doteq |H_k|^2$ which is denoted *channel gain* in the sequel. Also, the results are obtained in an information theoretic perspective where the blocks are supposed to be (asymptotically) long [11]–[19]⁵.

We assume the successive fading realizations to be dependent such that given $G_k = g_k$, $k = 0, \dots, K - 1$, the fading realization in time slot K is obtained according to the conditional probability density function (pdf) $f_{G_K|G_{K-1}\dots G_0}(g_K|g_{K-1}\dots g_0)$. The analytical results are valid for a fairly general case where the gain distribution can be any continuous pdf taking positive values over $(0, \infty)$. For the simulation results, however, we focus on Rayleigh fading channels where the successive fading channel realizations are related to each other according to the first order Gauss-Markov process

$$H_{k+1} = \alpha H_k + \sqrt{1 - \alpha^2} \varepsilon, \quad \varepsilon \sim \mathcal{CN}(0, \mu), \quad G_k = |H_k|^2. \quad (2)$$

Here, ε is a Gaussian noise $\mathcal{CN}(0, \mu)$ uncorrelated with H_k . Also, α is a known correlation factor which, defining $\mathbf{E}\{\cdot\}$ as the expectation operator and \check{x} as the conjugate of x , demonstrates two successive gain realizations dependencies by $\alpha = \frac{\mathbf{E}\{H_{k+1}\check{H}_k\}}{\mathbf{E}\{|H_k|^2\}}$, i.e., the one-lag normalized autocorrelation⁶. Under this model, the gain joint and marginal pdfs are found as

$$f_{G_k, G_{k-1}}(x, y) = \frac{1}{(1 - \alpha^2)\mu^2} e^{-\frac{x+y}{(1-\alpha^2)\mu}} \Psi_0\left(\frac{2\alpha\sqrt{xy}}{(1 - \alpha^2)\mu}\right) \quad (3)$$

and

⁵As discussed in, e.g., [37], the information theoretical results of slowly-fading channels match the results of actual codes for practical code lengths, e.g., $L \simeq 100$ channel uses, and provide appropriate performance bounds for systems with smaller code lengths.

⁶As an extension, one can consider the Rician fading models where H_k and ε in (2) are redefined as $H_k = H_k^r + jH_k^i$ and $\varepsilon = \varepsilon^r + j\varepsilon^i$ with $j = \sqrt{-1}$, $H_k^r \sim \mathcal{N}(a^r, \frac{\mu}{\sqrt{2}})$, $H_k^i \sim \mathcal{N}(a^i, \frac{\mu}{\sqrt{2}})$, $\varepsilon^r \sim \mathcal{N}(\frac{1-\alpha}{\sqrt{1-\alpha^2}}a^r, \frac{\mu}{\sqrt{2}})$, $\varepsilon^i \sim \mathcal{N}(\frac{1-\alpha}{\sqrt{1-\alpha^2}}a^i, \frac{\mu}{\sqrt{2}})$. Here, a^r and a^i are the corresponding means of the Gaussian variables H_k^r and H_k^i .

$$f_{G_k}(g) = \frac{1}{\mu} e^{-\frac{g}{\mu}}, g \geq 0, \quad (4)$$

respectively, where $\Psi_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind [38]. Also, one can extend (2) to

$$H_{k+n} = \alpha^{n-l} H_{k+l} + \sqrt{1 - \alpha^{2(n-l)}} \varepsilon, n > l \quad (5)$$

which, as seen in the following, can be used in cases with delayed feedback or arbitrary subsampling rates (More details about the fading pdf are given in Section V.). Finally, all results are presented in natural logarithm basis, and in all simulations the throughput is presented in nats-per-channel-use (npcu).

CSI feedback scheme: It is assumed that the receiver has perfect instantaneous knowledge about the channel gain, which is an acceptable assumption in slowly-fading conditions [11]–[19]. On the other hand, the transmitter is provided with subsampled quantized CSI which is fed back by the receiver every K -th block; in time slots mK , $m = \dots, -1, 0, 1, \dots$, a deterministic mapping function (quantizer)

$$\Omega(g_{mK}) = n \text{ if } g_{mK} \in S_n = [\tilde{g}_{n-1}, \tilde{g}_n), \tilde{g}_0 = 0, \tilde{g}_N = \infty \quad (6)$$

is implemented by the receiver which partitions the nonnegative real line into N nonoverlapping quantization regions with quantization boundaries $\tilde{g}_n, n = 0, \dots, N$. Then, if the channel realization falls into the n -th quantization region, i.e., $g_{mK} \in S_n = [\tilde{g}_{n-1}, \tilde{g}_n)$, the quantization index n is sent back to the transmitter, where the transmission rate and power are selected based on the received CSI. Note that the transmitter has no CSI except the region in which the channel gain falls. Also, considering the feedback subsampling, the feedback rate is given by $\frac{1}{K} \log_2 N$ bps. Both delay-free and delayed feedback conditions are studied in the following. Finally, we define $p_n = \Pr\{g_{mK} \in S_n\}$ as the probability of the gain belonging to S_n .

Forward channel data transmission approaches: We focus on the *bursting* communication model where, while K successive fading blocks are used in each transmission interval, there is a long idle period between the transmission of different packets⁷. Therefore, we can ignore the gain dependencies between different packet transmissions. Two different transmission schemes are considered for every K successive fading blocks, as illustrated in Fig. 1. In both scenarios, the transmitter is only informed about the quantized CSI of the first block out of the K blocks of a packet. The only difference between two methods returns back to the length of the codewords. In the short-length coding scenario, each codeword spans only one fading block. In the

⁷In the following, the data transmitted in K successive fading blocks is denoted a packet (See Fig. 1).

long-length coding scenario, on the other hand, each codeword is supposed to span over the entire packet, consisting of K fading blocks.

Evaluation yardstick: Our figure of merit is the long-term throughput (in npcu) defined as [12]–[17], [39]

$$\hat{\eta} = \lim_{I \rightarrow \infty} \frac{\sum_{i=1}^I Q(i)}{\sum_{i=1}^I \tau(i)} \stackrel{(a)}{=} \frac{\mathbf{E}\{Q\}_{k=0}^{K-1}}{\mathbf{E}\{\tau\}_{k=0}^{K-1}}. \quad (7)$$

Here, $Q(i)$ and $\tau(i)$ are the number of successfully decoded information nats and the number of channel uses in the i -th time slot, respectively. Then, (a) follows from the law of large numbers where $\mathbf{E}\{Q\}_{k=0}^{K-1}$ is the expected number of successfully decoded information nats at the end of each packet transmission and $\mathbf{E}\{\tau\}_{k=0}^{K-1}$ is the expected number of channel uses in a packet transmission period.

In both quantized CSI-based transmission scenarios studied in this paper the *whole* K blocks are used for transmission of fixed-length codes. That is, $\tau(i) = KL, \forall i$, i.e., the number of channel uses in each packet is the same regardless of the channel conditions. Thus, the denominator of (7) is constant in both quantized CSI-based scenarios and, using law of large numbers, the throughput simplifies to

$$\hat{\eta} = \lim_{I \rightarrow \infty} \frac{\sum_{i=1}^I Q(i)}{IKL} = \lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I \frac{Q(i)}{KL} = \lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I R(i) = \mathbf{E}_{G_0, \dots, G_{K-1}} \{R(g_0, \dots, g_{K-1})\} = \hat{R}, \quad (8)$$

where $R(i)$ is the achievable rate in the i -th packet period. In other words, as stated in (8), in both quantized CSI-based scenarios the throughput degenerates to the average rate $\hat{R} = \mathbf{E}_{G_0, \dots, G_{K-1}} \{R(g_0, \dots, g_{K-1})\}$, i.e., the expectation of the successful transmission rates when the channel is in states $G_k = g_k, k = 0, \dots, K-1$ [12]–[17].

For the HARQ protocols with a bursting communication model, the length of the packets, i.e., τ_i 's, is a random variable and not constant. This is because, depending on the channel condition, we may need to retransmit the data for a number of times. Thus, as also stated in [15], [17]–[19], [39], the packets have different weights on the long-term throughput and channel empirical pdf does not match the true channel distribution. As a result, the long-term throughput should be directly calculated based on (7), as explained in the Appendix.

III. THROUGHPUT ANALYSIS; SHORT-LENGTH CODING SCENARIO

Given that the gain realization at time slot $k = 0$ falls in the n -th quantization region S_n , defined in (6), the data is transmitted by K codewords having rates $R_{n,k} = \log(1 + g_{n,k}^* T_{n,k}), k = 0, \dots, K-1$, if short-length coding is considered. Here, $g_{n,k}^*, k = 0, \dots, K-1$, are auxiliary variables, one-to-one related with rates $R_{n,k}$, which simplify the equations (Thus, we may use $R_{n,k}$ and $g_{n,k}^*$ interchangeably in the following discussions).

These parameters can be interpreted as fixed values estimated by the transmitter if $g_0 \in S_n$. Also, $T_{n,k}$ is the power considered for the k -th block such that, for a given quantization region S_n , we have

$$\frac{1}{L} \sum_{l=1}^L |X_k[l]|^2 \leq T_{n,k}. \quad (9)$$

Remark 1: As the gain realization in time slot $k = 0$ is in the region $S_n = [\tilde{g}_{n-1}, \tilde{g}_n)$, the optimal value of the auxiliary variable $g_{n,0}^*$ must be within this region as well⁸. However, this argument is not valid for the other considered values $g_{n,k}^*$, $k > 0$, (and consequently, the rates $R_{n,k}$) which can be any positive value.

The codewords are decoded at the receiver separately. Therefore, if the gain realization at time slot k supports the considered rate in that block, i.e., $g_k \geq g_{n,k}^* = \frac{e^{R_{n,k}} - 1}{T_{n,k}}$, the transmitted data is successfully decoded, otherwise outage occurs. Consequently, for the n -th quantization region, the expected received information in time slot k is found as

$$\begin{aligned} \hat{Q}_{n,k} &= L \times R_{n,k} \times \Pr\{\text{successful decoding} | g_0 \in S_n\} = L \times \log(1 + g_{n,k}^* T_{n,k}) \Pr\{g_k \geq g_{n,k}^* | g_0 \in S_n\} \\ &= \frac{L}{p_n} \log(1 + g_{n,k}^* T_{n,k}) \Pr\{g_k \geq g_{n,k}^* \& g_0 \in S_n\} = \frac{\beta_{n,k} L}{p_n} \log(1 + g_{n,k}^* T_{n,k}), \\ \beta_{n,k} &= \begin{cases} \int_{g_{n,0}^*}^{\tilde{g}_n} \int_0^\infty \cdots \int_0^\infty f_{G_0, \dots, G_{K-1}}(x_0, \dots, x_{K-1}) dx_0 \cdots dx_{K-1} = F_{G_0}(\tilde{g}_n) - F_{G_0}(g_{n,0}^*), & k = 0 \\ \int_{\tilde{g}_{n-1}}^{\tilde{g}_n} \int_0^\infty \cdots \int_{g_{n,k}^*}^\infty \cdots \int_0^\infty f_{G_0, \dots, G_{K-1}}(x_0, \dots, x_{K-1}) dx_0 \cdots dx_{K-1} \\ & = \int_{\tilde{g}_{n-1}}^{\tilde{g}_n} \int_{g_{n,k}^*}^\infty f_{G_0, G_k}(x_0, x_k) dx_0 dx_k, & k \neq 0 \end{cases} \quad (10) \end{aligned}$$

where $f_{G_0, \dots, G_{K-1}}$ is the gain joint pdf in K successive realizations and F_{G_0} denotes the gain marginal cumulative distribution function (cdf). Thus, $\beta_{n,k}$ is the probability of successful message decoding in time slot k when the transmission rate is set to $R_{n,k} = \log(1 + g_{n,k}^* T_{n,k})$ and the quantization index n has been received in the slot $k = 0$. In this way, using (10), the throughput is obtained by

$$\hat{\eta} = \frac{\mathbf{E}\{Q\}_{k=0}^{K-1}}{\mathbf{E}\{\tau\}_{k=0}^{K-1}} = \frac{\sum_{n=1}^N p_n \sum_{k=0}^{K-1} \hat{Q}_{n,k}}{KL} = \frac{1}{K} \sum_{n=1}^N \sum_{k=0}^{K-1} \beta_{n,k} \log(1 + g_{n,k}^* T_{n,k}). \quad (11)$$

Also, the average transmission power, defined in [11]–[18], [37], [39], is simply found as⁹

$$\hat{T} = \sum_{n=1}^N p_n \left(\frac{T_{n,0}L + \cdots + T_{n,K}L}{KL} \right) = \frac{1}{K} \sum_{n=1}^N \sum_{k=0}^{K-1} p_n T_{n,k}. \quad (12)$$

⁸In fact, with the same procedure as in [12] which studies unsampled CSI in uncorrelated slowly-fading channels, we can show that in the optimal case we have $g_{n,0}^* = \tilde{g}_{n-1}$ if $n \neq 1$.

⁹In general, the average power is defined as $\hat{T} = \lim_{I \rightarrow \infty} \frac{\sum_{i=1}^I \xi(i)}{\sum_{i=1}^I \tau(i)}$ where $\xi(i)$ is the consumed energy in the i -th time slot. With the same procedure as in (8), the average power is rephrased as (12) for the quantized CSI schemes using packets of fixed length.

In this perspective, considering T as the transmission power constraint, the general power-limited throughput optimization problem can be stated as

$$\begin{aligned} \hat{\eta}_{\max} = & \max_{\forall g_{n,k}^*, \hat{g}_n, T_{n,k}} \frac{1}{K} \sum_{n=1}^N \sum_{k=0}^{K-1} \beta_{n,k} \log(1 + g_{n,k}^* T_{n,k}) \\ & \text{subject to } \frac{1}{K} \sum_{n=1}^N \sum_{k=0}^{K-1} p_n T_{n,k} \leq T \end{aligned} \quad (13)$$

which, based on the power allocation strategy and the fading distributions, can be solved numerically or analytically. Finally, setting $K = 1$ the results are simplified to the ones presented in [12] with no CSI subsampling and no correlation between the channel realizations.

Transmission power constraints: Based on the transmitter power adaptation capabilities, there may be different power constraints; due to, e.g., hardware or complexity limitations, there are cases where, independently of the feedback index, the power allocated to each codeword can not exceed a maximum value T . In this case, as the transmission rate of AWGN channels is an increasing function of the signal-to-noise ratio (SNR) [11]–[19], the optimal powers maximizing the throughput are obtained by $T_{n,k} = T, \forall n, k$. This constraint is normally called *short-term* power allocation [11]–[19].

Under the more relaxed *long-term* (battery-limited) power constraint, the transmitter can adapt the power based on the channel conditions such that $\hat{T} \leq T$. In this way, the optimal powers maximizing the throughput can be found based on (11), (12) and a Lagrange multiplier function $\Upsilon = \hat{\eta} - \lambda \hat{T}$ leading to the following water-filling [40, chapter 9.4] equation

$$\frac{\partial \Upsilon}{\partial T_{n,k}} = 0 \Rightarrow T_{n,k} = \left[\frac{\beta_{n,k}}{\lambda p_n} - \frac{1}{g_{n,k}^*} \right]^+. \quad (14)$$

Here, λ is the Lagrange multiplier satisfying $\hat{T} \leq T$ constraint and $[x]^+ \doteq \max(0, x)$.

Remark 2: It has been previously shown that to maximize the throughput more resources should be allocated to the *better* channels, e.g., [12], [16], [17]. Thus, it can be easily shown that, maximizing the throughput, we have $T_{n,k} \geq T_{n,l}$ if $\Pr\{g_k \geq x | g_0 \in S_n\} \geq \Pr\{g_l \geq x | g_0 \in S_n\}, \forall x$. That is, for a given quantization index n , higher powers are allocated to the blocks in which the channel gains are estimated to be higher¹⁰.

Remark 3: With $\Pr\{g_k \geq x | g_0 \in S_n\} \geq \Pr\{g_l \geq x | g_0 \in S_n\}, \exists k, l, \forall x$, there exist average power thresholds below which an on-off approach is the optimal scheme maximizing the throughput in the short-length coding scheme. This point follows from Remark 2; the property $T_{n,k} \geq T_{n,l}$ in Remark 2 simply means that, for every

¹⁰Note that the two power terms $T_{n,k}$ and $T_{n,l}$ are interchangeable in the average power (12).

given quantization index n , no power is allocated to the blocks estimated to be weak, until the strong blocks of the packet have received their required powers. That is, for every given average power threshold T , there is a set of block indices $\tilde{l}_n \subset \{0, \dots, K-1\}$ where $T_{n,l} = 0$ if $l \in \tilde{l}_n$ and $T_{n,l} > 0$ if $l \notin \tilde{l}_n$. Hence, the maximum throughput is achieved when, depending on the average power constraint, the transmitter turns off in weak time slots and allocates the power in the time slots where the channel is estimated to be strong.

As an example, if the current channel gain is in the first quantization region, it is expected that the channel gain increases in the next slot. Hence, we should have $T_{1,1} \geq T_{1,0}$ (For simulation results, see Fig. 3). Thus, with a low total power T , no power may be assigned to $T_{1,0}$, i.e., $T_{1,0} = 0$, while we may have $T_{1,1} > 0$.

Remark 4: With the same arguments as in Remark 2, we have $g_{n,k}^* \geq g_{n,l}^*$ (or equivalently $R_{n,k} \geq R_{n,l}$) if, compared to the l -th slot, better channel condition is expected in the k -th slot. Also, in harmony with intuitions, the simulation results show that when the subsampling rate increases or the channel correlation decreases the auxiliary variables $g_{n,k}^*$ converge to $g_{N_0}^*$ which is the fixed value considered for data transmission with no CSI at the transmitter.

Finally, the paper has focused on the single-layer transmission (SLT) scheme. However, the same procedure as in, e.g., [12], can be used to extend the results to the case with a multi-layer transmission (MLT) approach, which improves the system performance, particularly for low feedback rates.

Iterative throughput optimization algorithm: Equation (13) is a nonconvex optimization problem in terms of quantization parameters. Hence, to solve the problem, we propose an iterative algorithm stated in Algorithm I. Here, we should mention that, as the problem is nonconvex, there is no guarantee that the globally optimal parameters are determined by any algorithm, except the algorithms with an exhaustive search-based fashion which are run infinitely many times. Step VII of Algorithm I implies that a number of randomly generated answers are checked in each iteration. Thus, it can be easily proved that the algorithm can asymptotically reach the global optimum. However, as the algorithm is practically run for a finite number of iterations, the simulation results of Section V can not be claimed to be globally optimal. For this reason, 1) we have checked the proposed algorithm for many different initial settings, to reduce the effect of local minima. For almost all of the initial settings the algorithm reaches the same result with high accuracy, which is an indication of a good result. Moreover, 2) we have double-checked the results of short-length coding scheme under short-term power constraint by modifying the gradient-based algorithm of [12], which has been used for uncorrelated channels. The simulations show that in some cases Algorithm I works better than the algorithm in [12], which might be

because the algorithm of [12] needs to numerically differentiate the Marcum-Q function that shows up in the Rayleigh-fading model (see (24)). Finally, note that, while Algorithm I is much faster than exhaustive search, the gradient-based algorithm of [12] is faster than the proposed algorithm. However, as the parameters of, e.g., (13), are determined off-line, the complexity is not as important as in online applications.

Algorithm 1 Throughput optimization

- I. For a given power T and the fading pdf $f_{G_0, \dots, G_{K-1}}$, consider J , e.g. $J = 20$, randomly generated vectors $\Lambda^j = [\tilde{g}_0^j \dots \tilde{g}_N^j g_{1,0}^{j*} \dots g_{N,K-1}^{j*}]$, $0 = \tilde{g}_0^j \leq \tilde{g}_1^j \leq \dots \leq \tilde{g}_N^j$, $j = 1, \dots, J$.
 - II. For each vector, do the following procedures
 - Find the probability coefficients $\beta_{n,k}$'s and p_n 's based on the joint pdf.
 - Find the optimal transmission powers according to (12) and (14).
 - Determine the throughput $\hat{\eta}^j$ based on the transmission powers and (11).
 - III. Determine the vector which results in highest throughput, i.e., Λ^i where $\hat{\eta}^j \leq \hat{\eta}^i$, $\forall j = 1, \dots, J$.
 - IV. $\Lambda^1 \leftarrow \Lambda^i$.
 - V. Generate $b \ll J$, e.g., $b = 5$, vectors $\Lambda^{j,\text{new}}$, $j = 1, \dots, b$ around Λ^1 . These vectors should also satisfy the constraints introduced in I.
 - VI. $\Lambda^{j+1} \leftarrow \Lambda^{j,\text{new}}$, $j = 1, \dots, b$.
 - VII. Regenerate the remaining vectors Λ^j , $j = b + 2, \dots, J$ randomly such that $\Lambda^j = [\tilde{g}_0^j \dots \tilde{g}_N^j g_{1,0}^{j*} \dots g_{N,K-1}^{j*}]$, $0 = \tilde{g}_0^j \leq \tilde{g}_1^j \leq \dots \leq \tilde{g}_N^j$, $j = b + 2, \dots, J$.
 - VIII. Go to II and continue until convergence.
-

A. Discussions

In the following, we extend the results of the short-length coding approach to the cases with delayed feedback. Also, some discussions about the continuous communication model, practical implementation of feedback subsampling and HARQ protocols are presented. Note that, although the results are given for the short-length coding scenario, the same arguments can also be used in the long-length coding scenario.

1) *Delayed feedback*: The results of, e.g. (11), can be extended to the case where there is $D \geq 1$ steps delay in the received feedback, as the delay can be modeled as utilizing the received quantized CSI of the $k = 0$ fading block for parameter setting in blocks $k = D, \dots, D + K - 1$. In this way, with the same procedure as

before, the throughput is found as

$$\begin{aligned}\hat{\eta}_{\text{delayed}}(D) &= \frac{1}{K} \sum_{n=1}^N \sum_{k=0}^{K-1} \beta_{n,k+D} \log(1 + g_{n,k+D}^* T_{n,k+D}) \\ \beta_{n,k+D} &= \int_{\hat{g}_{n-1}}^{\hat{g}_n} \int_0^\infty \cdots \int_{g_{n,k+D}^*}^\infty \cdots \int_0^\infty f_{g_0, \dots, g_{D+K-1}}(x_0, \dots, x_{D+K-1}) dx_0 \cdots dx_{D+K-1}.\end{aligned}\quad (15)$$

Also, the average transmission power is changed to $\hat{T} = \frac{1}{K} \sum_{n=1}^N \sum_{k=0}^{K-1} p_n T_{n,k+D}$ which, along with (15), change the throughput optimization problem (For simulation results, see Fig. 10).

2) *On the continuous communication model:* The same approach as in (8) can be used to determine the throughput of the quantized CSI-based schemes if, instead of the bursting, the *continuous* communication model is considered. However, the main difference between these two models is that, in contrast to the bursting model, the subsampled quantized CSI of the previous packets provide considerable information for the current packet, if continuous communication is considered. Thus, for the continuous model, all feedback bits received during the previous packet periods should be taken into account, which makes the fading pdf a function of (infinitely) many quantization indices, and not mathematically tractable. This is the reason for considering the bursting model. Note that 1) the bursting model is an appropriate model for many practical setups, as the users are not always active [15]–[17], [19], [41]. Also, 2) with quantized CSI, the results of the bursting model, e.g., (11), are valid also for the continuous model if, in each packet transmission period, the quantized CSI of the previous packet periods is ignored. Finally, 3) as discussed in, e.g., [15], [17], [19], there are fundamental differences between the performance of HARQ protocols in bursting and continuous communication models. This point, which is because of the variant length of HARQ packets, is further studied in the Appendix.

3) *On practical implementation of quantized feedback subsampling:* In practice, variable-rate coding is achieved by adaptive modulation and coding (AMC) where the received SNR is compared with a set of thresholds, listed in a number of tables, and an appropriate modulation and coding index is chosen based on the comparisons [1]–[4]. Also, utilizing subsampled CSI, MMSE estimators are normally used to estimate the SNR of the missing blocks as in, e.g., [4]. The thresholds can be mapped to the transmission powers $T_{n,k}$ and rates $R_{n,k}$. Moreover, the successful message decoding probabilities $\beta_{n,k}$ are obtained by the empirical pdf of the successfully decoded codewords for each SNR threshold in the tables.

Practically, feedback subsampling is of particular interest when the number of users or the number of chunks in OFDM setups increases. Here, the feedback resources can be reutilized by different users/chunks in a time division multiple access (TDMA) fashion [1]–[4]. Moreover, frequency correlations between, e.g., neighbor chunks of the OFDM channels can be exploited with the same methods as the ones we developed for temporal

correlation (There are also a number of transform coding approaches such as discrete cosine transform (DCT) which have been shown to be efficient for feedback compression in the frequency domain [1]–[4], [6], [7], [31]). Finally, special attentions are required when designing the schedulers in the presence of feedback subsampling. This is because the scheduler should take the accuracy of the channel estimations into account, and decide whether to exploit the channels with more updated CSI, or to explore the channels with an outdated CSI [26]. Thus, the scheduler should be designed based on the interplay between user scheduling, channel memory evolution and queue status.

4) *Data transmission with HARQ feedback*: HARQ is one of the most well-known schemes to increase the data transmission efficiency of wireless networks via limited feedback resources [15]–[19], [23]–[25], [39]. Therefore, it is interesting to compare the results of the quantized CSI approach with the throughput obtained by the HARQ feedback.

Among different HARQ protocols, repetition time diversity (RTD) is a good and simple fixed-length coding HARQ technique where the same codeword is retransmitted in each retransmission and the receiver performs maximum ratio combining of all received signals [17]–[19], [24], [25], [39]. Hence, its complexity is comparable with the ones in the quantized CSI schemes. This is particularly because not only the code lengths are the same in all retransmissions but also no new parity symbols needs to be generated in the retransmission rounds. In the Appendix, we determine the throughput of the RTD protocol for different channel conditions. Comparisons between the quantized CSI- and HARQ-based approaches are given in Section V which demonstrate the superiority of the quantized CSI scheme over the RTD protocol in the bursting communication model.

IV. THROUGHPUT ANALYSIS; LONG-LENGTH CODING SCENARIO

Utilizing long-length codes of length KL (in channel uses), Q_n information nats is encoded into a single codeword $\{\chi[i], i = 1, \dots, KL\}$ of rate $R_n = \frac{Q_n}{KL}$, if the n -th quantization index is received at the beginning of the packet (Fig. 1). The optimal coding scheme is considered where each codeword is constructed by concatenation of a standard *Gaussian* code and a power controller which allocates the power $T_{n,k}$, $k = 0, \dots, K - 1$, to the symbols $\{\chi[1 + kL] \dots \chi[(k + 1)L]\}$ of the code [42, chapter 7]. That is, the code is constructed based on $\mathcal{CN}(0, 1)$ and then the values of each block are rescaled by $\sqrt{T_{n,k}}$. Using this code construction scheme, the results of, e.g., [37], [40, chapter 15], [42, chapter 7], can be used to show that, with the set of powers $T_{n,k}$, the maximum number of decodable information nats using codewords of length KL is limited to

$$Q_{T_{n,0}, \dots, T_{n,K-1}}^{\max} = L \sum_{k=0}^{K-1} \log(1 + g_k T_{n,k}). \quad (16)$$

Therefore, for each quantization index n , the expected achievable information in this scenario is found as

$$\begin{aligned} \hat{Q}_n &= Q_n \times \Pr \left\{ Q_n \leq L \sum_{k=0}^{K-1} \log(1 + g_k T_{n,k}) \middle| g_0 \in S_n \right\} \\ &= \frac{1}{p_n} R_n \times KL \times \Pr \left\{ R_n \leq \frac{1}{K} \sum_{k=0}^{K-1} \log(1 + g_k T_{n,k}) \& g_0 \in S_n \right\} \end{aligned} \quad (17)$$

and the throughput is

$$\begin{aligned} \hat{\eta} &= \frac{\mathbf{E}\{Q\}_{k=0}^{K-1}}{KL} = \frac{\sum_{n=1}^N p_n (R_n \times KL \times \Pr\{R_n \leq \frac{1}{K} \sum_{k=0}^{K-1} \log(1 + g_k T_{n,k}) | g_0 \in S_n\})}{KL} \\ &= \sum_{n=1}^N R_n \Pr\{R_n \leq \frac{1}{K} \sum_{k=0}^{K-1} \log(1 + g_k T_{n,k}) \& g_0 \in S_n\}. \end{aligned} \quad (18)$$

Consequently, as the average transmission power is obtained by (12), the power-limited throughput optimization problem can be stated as

$$\begin{aligned} \hat{\eta}_{\max} &= \max_{\hat{g}_n, T_{n,k}, R_n} \sum_{n=1}^N R_n \Pr\{R_n \leq \frac{1}{K} \sum_{k=0}^{K-1} \log(1 + g_k T_{n,k}) \& g_0 \in S_n\} \\ &\text{subject to } \frac{1}{K} \sum_{n=1}^N \sum_{k=0}^{K-1} p_n T_{n,k} \leq T, \end{aligned} \quad (19)$$

which can be solved with the same procedure as in Algorithm I.

Remark 5: Using (18) and the same argument as in Remark 2, it follows that, to maximize the throughput, more powers should be allocated to the subcodewords in which the channel gain is estimated to be higher, for a given $g_0 \in S_n$.

Remark 6: Assume short-term power allocation and no CSI feedback. Then, (I): the throughput in a given piecewise slowly-fading channel is less than the throughput in another slowly-fading channel whose fading pdf is obtained by averaging the behavior of the piecewise slowly-fading channel in a codeword transmission period. However, (II): the system throughput in these cases converges at low SNRs. These points can be proved as follows.

With no CSI feedback, the piecewise slowly-fading channel throughput, i.e., (18), is rephrased as

$$\hat{\eta} = R \Pr\{R \leq \frac{1}{K} \sum_{k=0}^{K-1} \log(1 + g_k T)\}. \quad (20)$$

Moreover, the throughput in a slowly-fading channel following the pdf $f_{U^{(K)}}$, $U^{(K)} = \frac{1}{K} \sum_{k=0}^{K-1} g_k$ is

$$\hat{\eta} = R \Pr\{R \leq \log(1 + U^{(K)} T)\} = R \Pr\{R \leq \log(1 + \frac{1}{K} \sum_{k=0}^{K-1} g_k T)\}. \quad (21)$$

Now, comparing (20) and (21), part (I) is easily proved based on the concavity of the function $f(x) = \log(1+x)$, which leads to $\frac{1}{K} \sum_{k=0}^{K-1} \log(1 + g_k T) \leq \log(1 + \frac{1}{K} \sum_{k=0}^{K-1} g_k T)$. Finally, using $\log(1+x) \rightarrow x$ for low SNRs, the throughput in (20) and (21) converge to $\hat{\eta} = R \Pr\{R \leq \frac{1}{K} \sum_{k=0}^{K-1} g_k T\}$, as stated in part (II).

The remark is interesting when we remember that in practice the channel does not remain constant, even at low speeds, although it is approximated to be fixed. The remark shows that the practical data transmission efficiency of the communication systems is worse than what is theoretically obtained by such approximations (please see [43, chapter 4] as well). Also, although the remark is proved for the case of no CSI, it seems to be intuitively valid when the transmitter is provided with quantized CSI too, since the channel with quantized CSI can be modeled as a channel with no CSI and a modified fading pdf [16].

A. Discussions

With the same procedure as in Section III, the results of the long-length coding scenario can be extended to cases with delayed feedback. Here, we investigate the throughput in the presence of repetition codes. The results are interesting because implementation of repetition codes leads to substantial code design and decoding complexity reduction particularly in long codes. Also, in some high-performing coding schemes, it is easy to construct codes with rates close to one, while the low rates, e.g., rate $\frac{1}{2}$, are difficult to design. However, the low rates can be easily constructed by repetition of codes having rates close to one.

Using repetition codes, the codeword $\{\chi[i], i = 1, \dots, KL\}$ is constructed by K times repetition of a subcodeword $\{\tilde{\chi}[i], i = 1, \dots, L\}$ in the K blocks. The only difference between the subcodewords is their powers where, for a quantization index n , the k -th subcodeword is scaled to have power $T_{n,k}$. The scheme works as follows. For a quantization region S_n , construct a length- L Gaussian subcode $\{\tilde{\chi}[i], i = 1, \dots, L\}$ with \tilde{Q}_n information nats, power $\tilde{T}_{n,0}$ and rate $\tilde{R}_n = \frac{\tilde{Q}_n}{L}$. Then, repeat the subcode for K times and rescale the powers by $T_{n,k}$, $k = 0, \dots, K-1$, to make a single codeword $\{\chi[i], i = 1, \dots, KL\}$. At the receiver, maximum ratio combining of the subcodewords is performed. In this way, as the codeword has \tilde{Q}_n information nats and the data is repeated, it is successfully decoded if $\tilde{Q}_n \leq L \log(1 + \sum_{k=0}^{K-1} g_k T_{n,k})$. Hence, the throughput is found as

$$\begin{aligned} \hat{\eta}_{\text{rep}} &= \frac{\mathbf{E}\{Q\}_{k=0}^{K-1}}{\mathbf{E}\{\tau\}_{k=0}^{K-1}} = \frac{\sum_{n=1}^N p_n \tilde{Q}_n \Pr\{\tilde{Q}_n \leq L \log(1 + \sum_{k=0}^{K-1} T_{n,k} g_k) | g_0 \in S_n\}}{KL} \\ &= \sum_{n=1}^N p_n \frac{1}{K} \tilde{R}_n \Pr\{\tilde{R}_n \leq \log(1 + \sum_{k=0}^{K-1} T_{n,k} g_k) | g_0 \in S_n\} \stackrel{(a)}{\simeq} \sum_{n=1}^N p_n \frac{1}{K} \tilde{R}_n \Pr\{\tilde{R}_n \leq \sum_{k=0}^{K-1} T_{n,k} g_k | g_0 \in S_n\} \quad (22) \end{aligned}$$

where (a) is obtained by $\log(1+x) \simeq x$ for low SNRs.

On the other hand, considering (18), the maximum throughput of the long-length coding scheme at low SNRs is obtained by

$$\hat{\eta} \simeq \sum_{n=1}^N p_n R_n \Pr\{R_n \leq \frac{1}{K} \sum_{k=0}^{K-1} T_{n,k} g_k | g_0 \in S_n\}. \quad (23)$$

Thus, comparing (22) and (23) while setting $\tilde{R}_n = KR_n$, it is found that concatenation of repetition codes and power controllers along with maximum ratio combining at the receiver is optimal, in terms of throughput, at low SNRs (This point is in harmony with the literature [44]). In this case, the rate of the subcodewords considered in the repetition coding scheme should be K times of the original codewords rates.

V. SIMULATION RESULTS; RAYLEIGH FADING CORRELATED CHANNELS

The simulations focus on Rayleigh fading channels where the successive fading channel realizations are related to each other according to (2)-(5). Moreover, the probability terms $\Pr\{G_k \in [u, v] \& G_{k-1} \in [w, z]\}$ in, e.g., (10), are determined based on the following integration procedure

$$\begin{aligned} \Pr\{G_k \in [u, v] \& G_{k-1} \in [w, z]\} &= \int_u^v \int_w^z f_{G_k, G_{k-1}}(x, y) dx dy \stackrel{(a)}{=} \int_u^v \frac{1}{\mu} e^{-\frac{x}{\mu}} \left(\int_{\sqrt{\frac{2w}{r}}}^{\sqrt{\frac{2z}{r}}} \theta e^{-\frac{\theta^2}{2}} \Psi_0(s\sqrt{x}\theta) d\theta \right) dx \\ &\stackrel{(b)}{=} \int_u^v \frac{1}{\mu} e^{-\frac{x}{\mu}} \left\{ \xi(s\sqrt{x}, \sqrt{\frac{2w}{r}}) - \xi(s\sqrt{x}, \sqrt{\frac{2z}{r}}) \right\} dx \stackrel{(c)}{=} (1 - \alpha^2) e^{-\frac{u}{\mu}} \left\{ \xi\left(\sqrt{\frac{2w}{r}}\alpha, \sqrt{\frac{2u}{r}}\right) - \xi\left(\sqrt{\frac{2w}{r}}\alpha, \sqrt{\frac{2v}{r}}\right) \right\} \\ &\quad - (1 - \alpha^2) e^{-\frac{z}{\mu}} \left\{ \xi\left(\sqrt{\frac{2z}{r}}\alpha, \sqrt{\frac{2u}{r}}\right) - \xi\left(\sqrt{\frac{2z}{r}}\alpha, \sqrt{\frac{2v}{r}}\right) \right\} + \frac{1}{\mu} \int_u^v e^{-\frac{x}{\mu}} \left\{ \xi\left(\sqrt{\frac{2z}{r}}, s\sqrt{x}\right) - \xi\left(\sqrt{\frac{2w}{r}}, s\sqrt{x}\right) \right\} dx \\ &\stackrel{(d)}{=} e^{-\frac{w}{\mu}} \left\{ \xi\left(\sqrt{\frac{2w}{r}}\alpha, \sqrt{\frac{2u}{r}}\right) - \xi\left(\sqrt{\frac{2w}{r}}\alpha, \sqrt{\frac{2v}{r}}\right) \right\} - e^{-\frac{z}{\mu}} \left\{ \xi\left(\sqrt{\frac{2z}{r}}\alpha, \sqrt{\frac{2u}{r}}\right) - \xi\left(\sqrt{\frac{2z}{r}}\alpha, \sqrt{\frac{2v}{r}}\right) \right\} \\ &\quad + e^{-\frac{v}{\mu}} \xi\left(\sqrt{\frac{2w}{r}}, \sqrt{\frac{2v}{r}}\alpha\right) - e^{-\frac{u}{\mu}} \xi\left(\sqrt{\frac{2w}{r}}, \sqrt{\frac{2u}{r}}\alpha\right) - e^{-\frac{v}{\mu}} \xi\left(\sqrt{\frac{2z}{r}}, \sqrt{\frac{2v}{r}}\alpha\right) + e^{-\frac{u}{\mu}} \xi\left(\sqrt{\frac{2z}{r}}, \sqrt{\frac{2u}{r}}\alpha\right). \end{aligned} \quad (24)$$

Here, (a) is obtained by defining $r \doteq (1 - \alpha^2)\mu$, $s \doteq \sqrt{2/r}\alpha$ and using variable transform $\theta = \sqrt{2y/r}$. Then, (b) is directly obtained from the definition of the Marcum Q-function $\xi(x, y) = \int_y^\infty te^{-\frac{t^2+x^2}{2}} \Psi_0(xt) dt$. Finally, (c) is based on the fact that $\xi(x, y) = 1 + e^{-(x^2+y^2)/2} \Psi_0(xy) - \xi(y, x)$ and (d) is derived by using variable transform $t = \sqrt{x}$, partial integration and some calculations.

The simulation results are illustrated as follows¹¹.

On the effect of power allocation: Figures 2-4 study the effect of power allocation on the performance of the short- and long-length coding schemes¹². Here, the following points are deduced from the figures:

¹¹In all figures, we set $N = 2$, $K = 2$ and $\mu = 1$, unless otherwise stated.

¹²For adaptive power allocation in Fig. 4, the results are obtained for the case where, while the powers are optimally allocated to the subcodewords, the total transmission power in each codeword is constant. That is, $T_{n,0} + T_{n,1} = 2T$, $n = 1, 2$.

- With adaptive power allocation there are thresholds below which the on-off approach (not transmitting at all in the time slots where the gain is estimated to be low) is the optimal data transmission approach (Figs. 3 and 4b, Remark 3). Also, the simulation results are in harmony with Remarks 2 and 5; with adaptive power allocation more power is allocated to the slots in which the channel gain is estimated to be higher. For instance, we have $T_{2,0} \geq T_{1,0}$, that is, at $k = 0$ time slot more power is allocated to the second region. Also, given that the channel is in the first (the second) quantization region, it is expected that the channel gain increases (decreases) in the next slot. Hence, we have $T_{1,1} \geq T_{1,0}$ and $T_{2,0} \geq T_{2,1}$, as also seen in Figs. 3 and 4b.
- With more correlation, i.e., higher α , the power terms in a packet period, for instance $T_{1,0}$ and $T_{1,1}$ (or $T_{2,0}$ and $T_{2,1}$) get closer together (Fig. 3). The intuition behind this is that with more temporal correlation the probability that a good channel remains good (and a bad channel remains bad) increases.
- Also, in harmony with the literature, e.g., [12], [16], [17], adaptive power allocation leads to considerable increment in the throughput of the two scenarios at low SNRs (Fig. 2 and 4a).

Comparison between the short- and long-length coding scenarios: Considering short-term power constraint, Fig. 5 compares the throughput in the short- and long-length coding scenarios. The results show that, for Rayleigh fading channels, higher (but not much higher) throughput is achieved with long-length coding, compared to short-length coding. Particularly, the throughput of the two schemes are different at low temporal correlation conditions. However, increasing the temporal correlation, the throughput of the two scenarios get closer, converging to the one obtained by full correlation. Note that, while the long-length coding outperforms the short-length coding approach in terms of throughput, it leads to harder message encoding and decoding and channel estimation at the receiver. This is because longer codes are utilized in the long-length coding model and the channel changes during a codeword transmission period¹³.

On the efficiency of feedback subsampling: Considering equal-probability quantization¹⁴, Fig. 6 studies the throughput for different feedback rates. Compared to the cases with only CSI quantization and no feedback subsampling, the results show that higher throughput can be achieved with less or equal feedback rate when feedback subsampling is utilized (Using Fig. 6, this point can, for example, be verified by comparing the points

¹³In general, more time diversity can be exploited when the channel variations increases in time but, on the other hand, the channel estimation becomes harder for time-varying channels.

¹⁴With equal-probability quantization the channel marginal pdf is divided into N regions having probability $\frac{1}{N}$.

C or E with point D). Thus, depending on the correlation condition, the feedback subsampling can be an efficient feedback compression scheme for the communication systems.

Comparison between the quantized CSI and HARQ protocols: With a bursting communication model, Fig. 7 shows the RTD-based system throughput and compares the results with the ones obtained by the short-length coding based quantized CSI approach. Here, the results are obtained for a maximum of $M = 1$ RTD-based retransmission round and short-term power constraint. As demonstrated in the figure, the quantized CSI scheme outperforms the HARQ approach, in terms of throughput, and the difference between the throughput of the two schemes increases with power/temporal correlation. In contrast to the quantized CSI scheme, the HARQ-based throughput decreases with temporal correlation. Intuitively, this is because with more correlation less time diversity is exploited by the HARQ. The readers are referred to [16], [17] for comparison between the HARQ and quantized CSI schemes in the continuous communication model and uncorrelated slowly-fading channels.

On the effect of temporal correlation in the long-length coding scenario: Figure 8 verifies the validity of Remark 6 for different correlation conditions. Also, Fig. 9 compares the performance of optimal and repetition coding schemes in the long-length coding scenario. Considering Figs. 5, 8 and 9 we observe that the channel temporal correlation may play different roles in the throughput of the long-length coding scenario; As the correlation increases, less time diversity is exploited in the long-length coding scenario. On the other hand, we gain more from the subsampled CSI when the correlation increases, because the subsampled CSI provides considerable information for the slots with no CSI feedback. In this way, there is a tradeoff and depending on the fading pdf the limited-feedback throughput may increase or decrease with the correlation in the long-length coding scenario. For instance, the throughput decreases with the correlation if no CSI is provided at the transmitter (Fig. 8) (As another example for the case where the throughput decreases with correlation, please see the results of the repetition codes with $N = 2$ in Fig. 9). On the other hand, Figs. 5 and 9 demonstrate cases where the throughput increases with correlation, as the gain of feedback subsampling becomes higher than the loss in the time diversity when the correlation increases¹⁵.

Other results: According to Fig. 9, the repetition codes lead to considerable performance loss at medium/high SNRs. Also, the effect of channel temporal dependency on the throughput decreases when the suboptimal codes are utilized. Finally, Fig. 10 investigates the effect of different correlation conditions and $D = 1$ slot feedback

¹⁵Note that the subsampled CSI is always useful, in the sense that the throughput increases with the number of quantization regions.

delay on the throughput of the system utilizing short-length coding scheme. Here, the results show that with feedback delay, partial CSI is beneficial only when there is considerable temporal correlation between successive blocks (The same point is valid for the long-length coding scenario, although not seen in the figure.).

VI. CONCLUSION

This paper presented different schemes for exploiting time diversity in temporally-correlated channels. Using feedback subsampling, the results were obtained under both delayed and delay-free feedback conditions and with different transmission power constraints. The results show that with temporal correlation on-off approach is the optimal scheme maximizing the throughput of the feedback subsampling-based schemes. Approximating the channel variations by their average value leads to overestimating the performance of the wireless networks, unless if low-SNR regime is considered. Finally, depending on the channel condition, the feedback subsampling can be an efficient scheme which, compared to the traditional CSI quantization models, leads to higher throughput with less or equal feedback load.

APPENDIX A

PERFORMANCE ANALYSIS FOR THE RTD PROTOCOL

In contrast to the considered quantized CSI schemes, the length of the HARQ packets is not constant, but is a random variable which depends on the channel quality. Thus, each HARQ packet has a different weight on the system data transmission efficiency, and the long-term throughput does not degenerate to the average rate in the bursting data communication model [15], [17], [19]. More specifically, the empirical channel pdf does not match the true channel distribution, if the HARQ packets are sent in a bursting fashion. The reason is that if the channel is *good*, the packet transmission ends at the first transmission round. However, many channel uses are utilized for sending a packet when the channel is *bad*. Hence, a large portion of the data transmission is carried out when the channel experiences low quality, while the transmitter is mostly off when the channel is good.

With a continuous communication model, on the other hand, the channel gains are observed proportional to their realization probabilities, because the transmitter is always transmitting the codewords. Hence, the empirical channel pdf matches the true one and, with the same arguments as in [15], [17], [19], long-term throughput of the HARQ protocols degenerates to the average rate in the continuous communication model¹⁶.

¹⁶The same discussions are valid for the average power definition as well [15], [17], [19]

As stated before, we concentrate on the bursting communication model. Thus, the long-term throughput is directly calculated based on (7), as follows. For analysis of HARQ protocols in the continuous data communication models, the readers are referred to [16], [17].

We consider a maximum of M data retransmission rounds, i.e., each codeword is (re)transmitted a maximum of $M + 1$ rounds. In this case, Q information nats is encoded into a codeword of length L and rate $R = \frac{Q}{L}$ which is retransmitted a maximum of M times. In the m -th, $m = 1, \dots, M + 1$, (re)transmission round the codeword is rescaled to have power T_m . The (re)transmission continues until the codeword is correctly decoded by the receiver or the maximum permitted retransmission rounds is reached.

Let us define V_m as the event that the data is successfully decoded at the end of the m -th (re)transmission and not before, and \bar{V}_m as its complement. Provided that the receiver can decode the data, Q nats is received. On the other hand, mL channel uses are spent if the data is decoded at the m -th (re)transmission round. Also, independent of the message decoding status, there will be $(M + 1)L$ channel uses if all possible retransmission rounds are used. In this way, the throughput is found as

$$\begin{aligned} \hat{\eta}_{\text{RTD}} &= \frac{\mathbf{E}\{Q\}_{m=1}^{M+1}}{\mathbf{E}\{\tau\}_{m=1}^{M+1}} = \frac{Q(1 - \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\})}{L \sum_{m=1}^{M+1} m \Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\} + (M + 1)L \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}} \\ &= R \frac{1 - \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}}{\sum_{m=1}^{M+1} m \Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\} + (M + 1) \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}}. \end{aligned} \quad (25)$$

Also, with the same procedure and as the energy consumed at the first m (re)transmission rounds is $\xi^{(m)} = L \sum_{n=1}^m T_n$, the average transmission power [15]–[18], [37], [39] is found as

$$\begin{aligned} \hat{T}_{\text{RTD}} &= \frac{\mathbf{E}\{\xi\}_{m=1}^{M+1}}{\mathbf{E}\{\tau\}_{m=1}^{M+1}} = \frac{\sum_{m=1}^{M+1} \xi^{(m)} \Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\} + \xi^{(M+1)} \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}}{L \sum_{m=1}^{M+1} m \Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\} + (M + 1)L \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}} \\ &= \frac{\sum_{m=1}^{M+1} (\sum_{n=1}^m T_n) \Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\} + (\sum_{n=1}^{M+1} T_n) \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}}{\sum_{m=1}^{M+1} m \Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\} + (M + 1) \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}}. \end{aligned} \quad (26)$$

At the end of the m -th (re)transmission, the receiver performs maximum ratio combining of the m received signals. This process effectively increases the received SNR to $\sum_{n=1}^m g_n T_n$ and reduces the data rate to $\frac{R}{m}$ which is decoded if $R \leq \log(1 + \sum_{n=1}^m g_n T_n)$. This is based on the fact that with an equivalent SNR x the maximum decodable transmission rate is $\frac{1}{m} \log(1 + x)$ if a codeword is repeated m times [17]–[19], [39], [40].

In this way, the probability terms of (25) and (26) are found by

$$\begin{aligned} \Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\} &= \Pr\{\log(1 + \sum_{j=1}^{m-1} g_j T_j) < R \leq \log(1 + \sum_{j=1}^m g_j T_j)\} \\ \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\} &= \Pr\{\log(1 + \sum_{j=1}^{M+1} g_j T_j) < R\} \end{aligned} \quad (27)$$

which, as illustrated in Section V, can be simplified according to the fading joint distribution.

Finally, the following theorem provides a connection between the average rate and the long-term throughput of the HARQ protocols in the bursting communication model.

Theorem 1: The throughput of the RTD HARQ scheme is upper bounded by

$$\hat{\eta}_{\text{RTD}} \leq R \sum_{m=1}^{M+1} \frac{\Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\}}{m(1 - \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\})}. \quad (28)$$

Proof: The upper bound is found based on the following inequalities

$$\hat{\eta}_{\text{RTD}} \stackrel{(a)}{\leq} R \frac{1 - \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}}{\sum_{m=1}^{M+1} m \Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\}} \stackrel{(b)}{\leq} R \sum_{m=1}^{M+1} \frac{\Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\}}{m(1 - \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\})} = \frac{\hat{R}}{1 - \Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}}$$

where (a) is based on (25) and (b) follows from Jensen's inequality, convexity of the function $f(x) = \frac{1}{x}$ and $\sum_{m=1}^{M+1} \Pr\{\bar{V}_1, \bar{V}_2, \dots, \bar{V}_{m-1}, V_m\} = 1 - \Pr\{\bar{V}_1, \bar{V}_2, \dots, \bar{V}_{M+1}\}$. Also, the last equality comes from the definition of average rate, i.e., $\hat{R} = R \sum_{m=1}^{M+1} \frac{\Pr\{\bar{V}_1, \dots, \bar{V}_{m-1}, V_m\}}{m}$. This is particularly interesting when the number of retransmission rounds increases, where $\Pr\{\bar{V}_1, \dots, \bar{V}_{M+1}\}$ diminishes. ■

REFERENCES

- [1] T. Eriksson and T. Ottosson, "Compression of feedback for adaptive transmission and scheduling," *Proc. IEEE*, vol. 95, no. 12, pp. 2314–2321, Dec. 2007.
- [2] D. J. Love, et. al, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.
- [3] B. Makki and T. Eriksson, "Efficient channel quality feedback signaling using transform coding and bit allocation," in *VTC*, 2010, pp. 1–5.
- [4] T. Eriksson and T. Ottosson, "Compression of feedback in adaptive OFDM-based systems using scheduling," *IEEE Commun. Lett.*, vol. 11, no. 11, pp. 859–861, Nov. 2007.
- [5] S. Sanayei and A. Nosratinia, "Exploiting multiuser diversity with only 1-bit feedback," in *WCNC*, vol. 2, March 2005, pp. 978–983.
- [6] R1-070368, "System level comparison of best-M and DCT-based CQI compression schemes," Huawei, RAN1 meeting 47bis, Sorrento, Italy, Jan. 2007, [Online]. Available at www.quintillion.co.jp/3GPP/TSG_RAN/TSG_RAN2007/TSG_RAN_WG1_RL1_1.html.
- [7] R1-060018, "Performance evaluation of EUTRA downlink CQI feedback schemes," Motorola, RAN1 meeting 44, Helsinki, Finland, Jan. 2006, [Online]. Available at www.quintillion.co.jp/3GPP/TSG_RAN/TSG_RAN2007/TSG_RAN_WG1_RL1_1.html.
- [8] RP-010344, "Presentation of specification to TSG or WG," 3GPP TSG RAN Meeting 12, Stockholm, Sweden, June 2001, [Online]. Available at www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_12/Docs/PDFs/RP-010344.pdf.
- [9] "Technical specification group radio access network, physical layer procedures," 3rd Generation Partnership Project, Tech. Spec. 3GPP ts 25.214 V3.10.0, 2002-2003.
- [10] "Physical and medium access control layers for combined fixed and mobile operation in licensed bands," IEEE Std. Std 802.16e-2005, 2005.
- [11] Y. He and S. Dey, "Power allocation in spectrum sharing cognitive radio networks with quantized channel information," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1644–1656, June 2011.

- [12] T. T. Kim and M. Skoglund, "On the expected rate of slowly fading channels with quantized side information," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 820–829, April 2007.
- [13] S. Ekbatani, F. Etemadi, and H. Jafarkhani, "Throughput maximization over slowly fading channels using quantized and erroneous feedback," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2528–2533, Sept. 2009.
- [14] —, "Outage behavior of slow fading channels with power control using partial and erroneous CSIT," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6097–6102, Dec. 2010.
- [15] B. Makki, et. al, "Power allocation in repetition time diversity hybrid automatic repeat request feedback," in *WCNC*, April 2012, pp. 2329–2334.
- [16] B. Makki and T. Eriksson, "On hybrid ARQ and quantized CSI feedback schemes in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 986–997, April 2012.
- [17] C. Shen, T. Liu, and M. P. Fitz, "On the average rate performance of hybrid-ARQ in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3339–3352, Nov. 2009.
- [18] D. Tuninetti, "On the benefits of partial channel state information for repetition protocols in block fading channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5036–5053, Aug. 2011.
- [19] B. Makki, et. al, "HARQ feedback in spectrum sharing networks," *IEEE Commun. Lett.*, vol. 16, no. 9, pp. 1337–1340, Sept. 2012.
- [20] K. Huang, R. W. Heath, and J. G. Andrews, "Limited feedback beamforming over temporally-correlated channels," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1959–1975, May 2009.
- [21] S. Akoum and R. W. Heath, "Limited feedback for temporally correlated MIMO channels with other cell interference," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5219–5232, 2010.
- [22] T. Kim, et. al, "Differential rotation feedback MIMO system for temporally correlated channels," in *GLOBECOM*, Dec. 2008, pp. 1–5.
- [23] J. S. Harsini, F. Lahouti, M. Levorato, and M. Zorzi, "Analysis of non-cooperative and cooperative type II hybrid ARQ protocols with AMC over correlated fading channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 877–889, March 2011.
- [24] S. M. Kim, w. Choi, T. W. Ban, and D. K. Sung, "Optimal rate adaptation for hybrid ARQ in time-correlated rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 968–979, March 2011.
- [25] C. Li and X. Wang, "Throughput analysis for parallel ARQ over correlated MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1322–1332, Sept. 2007.
- [26] W. Ouyang, A. Eryilmaz, and N. B. Shroff, "Low-complexity optimal scheduling over correlated fading channels with ARQ feedback," in *WiOpt*, May 2012, pp. 270–277.
- [27] B. Makki and T. Eriksson, "CSI feedback in correlated slow-fading channels," *IEEE Commun. Lett.*, vol. 15, no. 12, pp. 1294–1297, Dec. 2011.
- [28] J. Tsao, D. Porrat, and D. Tse, "Prediction and modeling for the time-evolving ultra-wideband channel," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 3, pp. 340–356, 2007.
- [29] P. Pagani and P. Pajusco, "Characterization and modeling of temporal variations on an ultrawideband radio link," *IEEE Trans. Antennas and Propag.*, vol. 54, no. 11, pp. 3198–3206, 2006.
- [30] A. J. Goldsmith and M. Medard, "Capacity of time-varying channels with causal channel side information," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 881–899, March 2007.
- [31] A. Haghighat, et. al, "Haar compression for efficient CQI feedback signaling in 3GPP LTE systems," in *WCNC*, April 2008, pp. 819–823.
- [32] H. Gao, et. al, "Compression of CQI feedback with compressive sensing in adaptive OFDM systems," in *WCSP*, 2010, pp. 1–4.
- [33] K. Chakraborty, et. al, "Compression schemes for the feedback channel in multiband-OFDM UWB systems," in *ICUWB*, 2009, pp. 443–448.

- [34] V. Hassel, et. al, "Exploiting multiuser diversity using multiple feedback thresholds," in *VTC*, vol. 2, May 2005, pp. 1302–1306.
- [35] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.
- [36] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, June 2006.
- [37] G. Caire, et. al, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, 1999.
- [38] C. Tellambura and A. D. S. Jayalath, "Generation of bivariate rayleigh and nakagami-M fading envelopes," *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 170–172, May 2000.
- [39] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1971–1988, 2001.
- [40] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley Interscience, 2006.
- [41] F. Berggren, et. al, "Dynamic spectrum access, phase 1: scenarios and research challenges," Sept. 2004, available: <http://www.qeseth.se/olav/pubs/DSAReportPhase1.pdf>.
- [42] A. El Gamal and Y.-H. Kim, "Lecture notes on network information theory," 2010, [Online]. Available at <http://arxiv.org/abs/1001.3404v4/>.
- [43] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [44] Y. H. Kim, L. B. Milstein, and I. Song, "Performance of a turbo coded multicarrier DS/CDMA system with nonuniform repetition coding," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1764–1774, Sept. 2001.

LIST OF FIGURES

1	An example of a packet and different codewords lengths considered in the paper.	25
2	Throughput versus the average transmission power T . Short-length coding scenario. Adaptive power allocation is more useful when temporal correlation increases.	26
3	Optimal powers $T_{n,k}$ versus the average transmission power T . The power $T_{n,k}$ is allocated to the k -th time slot if the quantized index n has been fed back at time slot $k = 0$. Long-term power constraint, short-length coding scenario. At low transmission powers the on-off approach (not transmitting at all in the time slots where the gain is estimated to be low) is the best data transmission scheme, maximizing the throughput (Remark 3). Also, low powers are allocated to the slots in which the channel is estimated to be weak (Remark 2).	27
4	(a): Throughput and (b): the optimal power terms versus the average transmission power T . Long-length coding scenario, $\alpha = 0.1$. The results are in harmony with Remark 5.	28
5	Comparison between the short- and long-length coding scenarios. Short-term power constraint. For low correlation conditions, higher throughput is achieved by long-length coding scheme, compared to short-length coding model. However, the results converge when the temporal correlation increases.	29
6	Throughput vs the feedback rate $\frac{1}{K} \log_2 N$, equal-probability quantization, short-term power constraint. The points A, B, C, D and E correspond to the cases $(K = 2, N = 2)$, $(K = 4, N = 4)$, $(K = 3, N = 4)$ $(K = 1, N = 2)$ and $(K = 2, N = 4)$, respectively. Moreover, the feedback rates $\frac{1}{4}$, $\frac{1}{3}$ and 2 are obtained by setting $(K = 4, N = 2)$, $(K = 3, N = 2)$ and $(K = 1, N = 4)$, respectively.	30
7	Comparison between the short-length coding scheme and RTD HARQ feedback approach in the bursting communication model. Short-term power constraint. The quantized CSI approach outperforms the HARQ scheme, in terms of throughput.	31
8	Comparison between the piecewise slowly-fading and the equivalent averaged channel model in Remark 6, $K = 2$. Less time diversity is exploited when the temporal correlation increases. Therefore, with no CSI at the transmitter, the throughput decreases with the correlation parameter α	32
9	Comparison between the optimal and repetition codes in the long-length coding scenario. The relative efficiency of the repetition codes, compared to the optimal coding approach, decreases with the transmission power.	33
10	Throughput versus the transmission power T and different feedback delay conditions. Short-term power constraint, short-length coding scenario. With small temporal correlation, feedback delay deteriorates the performance of the partial CSI scheme substantially.	34

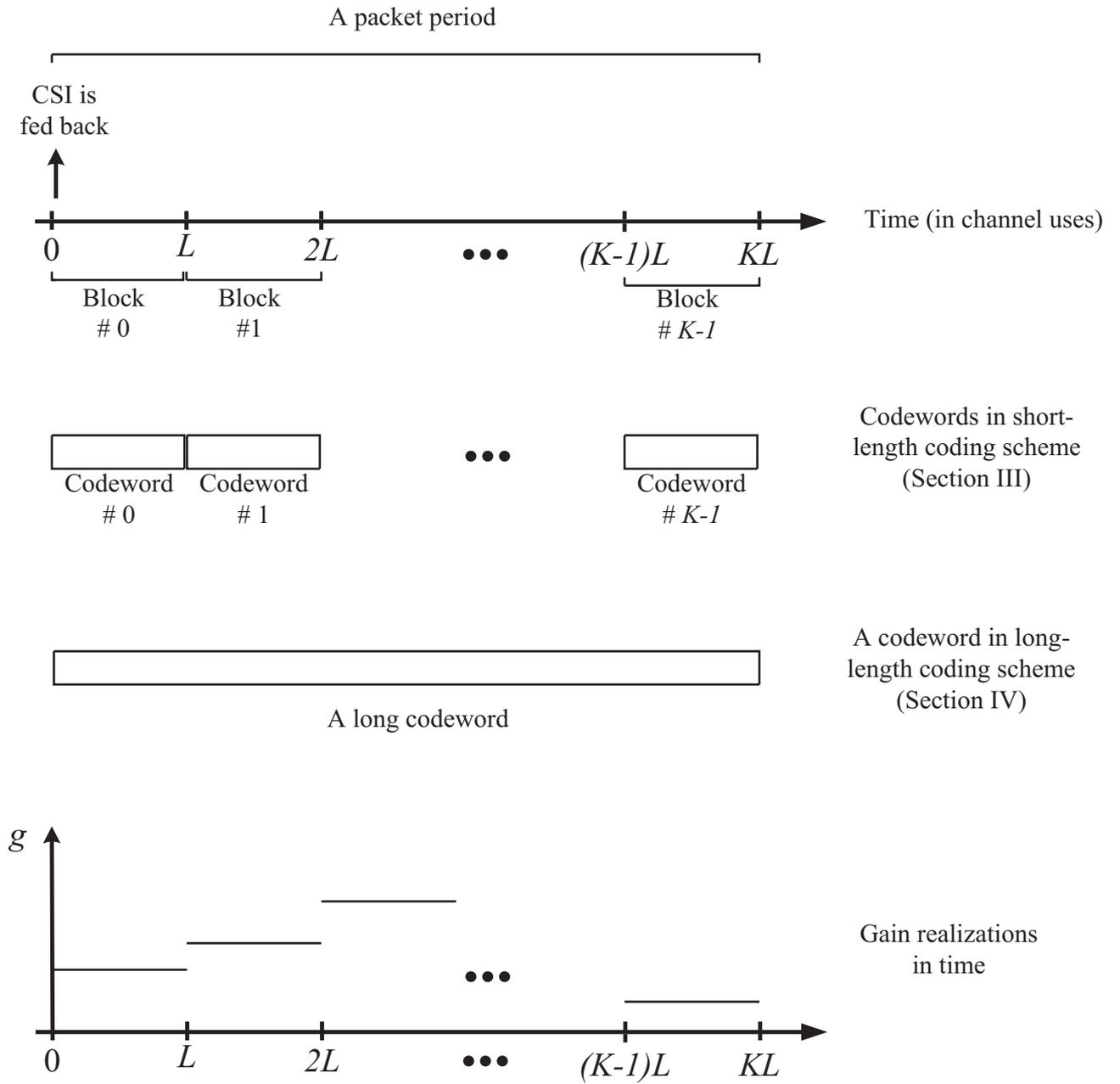


Fig. 1. An example of a packet and different codewords lengths considered in the paper.

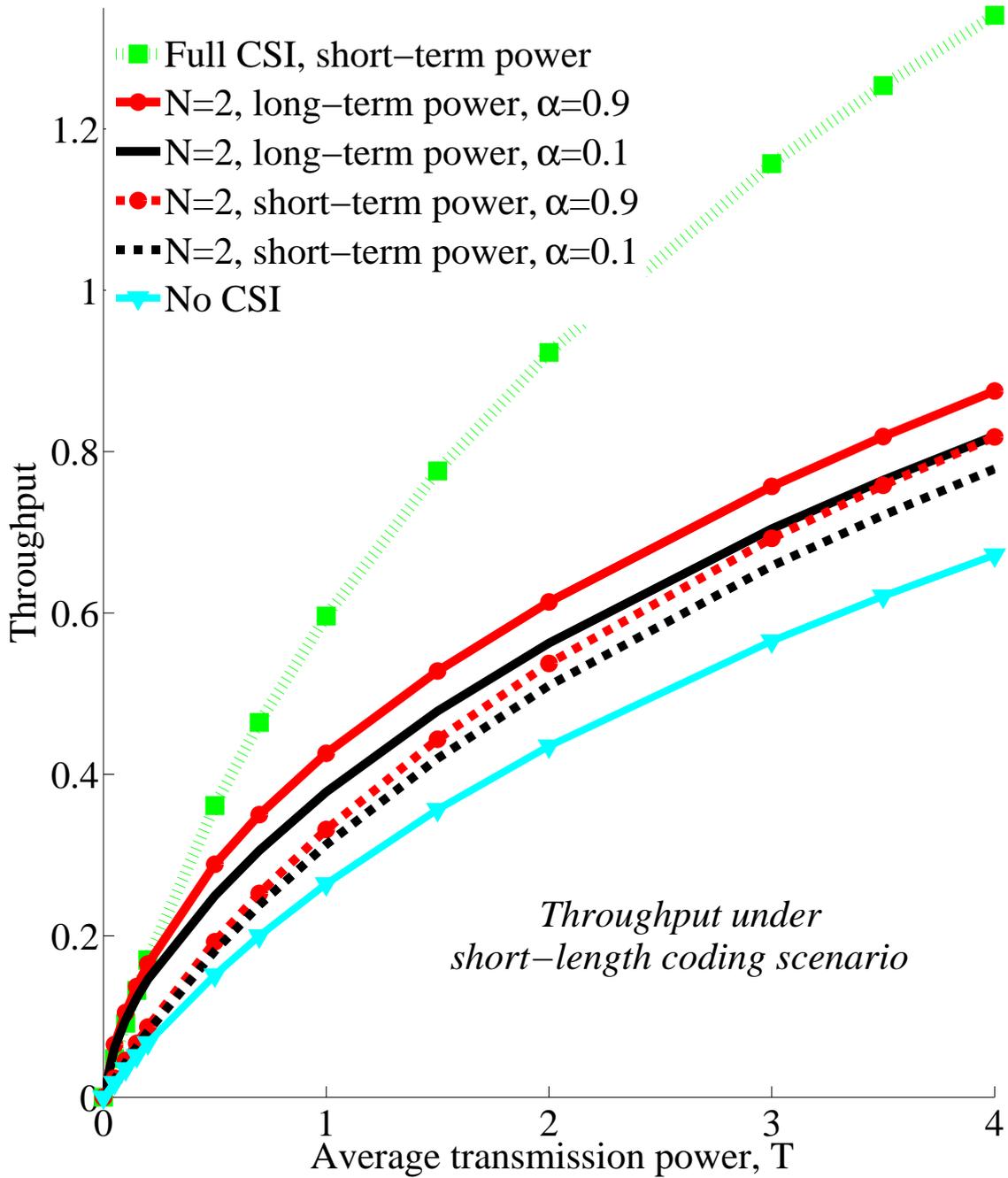


Fig. 2. Throughput versus the average transmission power T . Short-length coding scenario. Adaptive power allocation is more useful when temporal correlation increases.

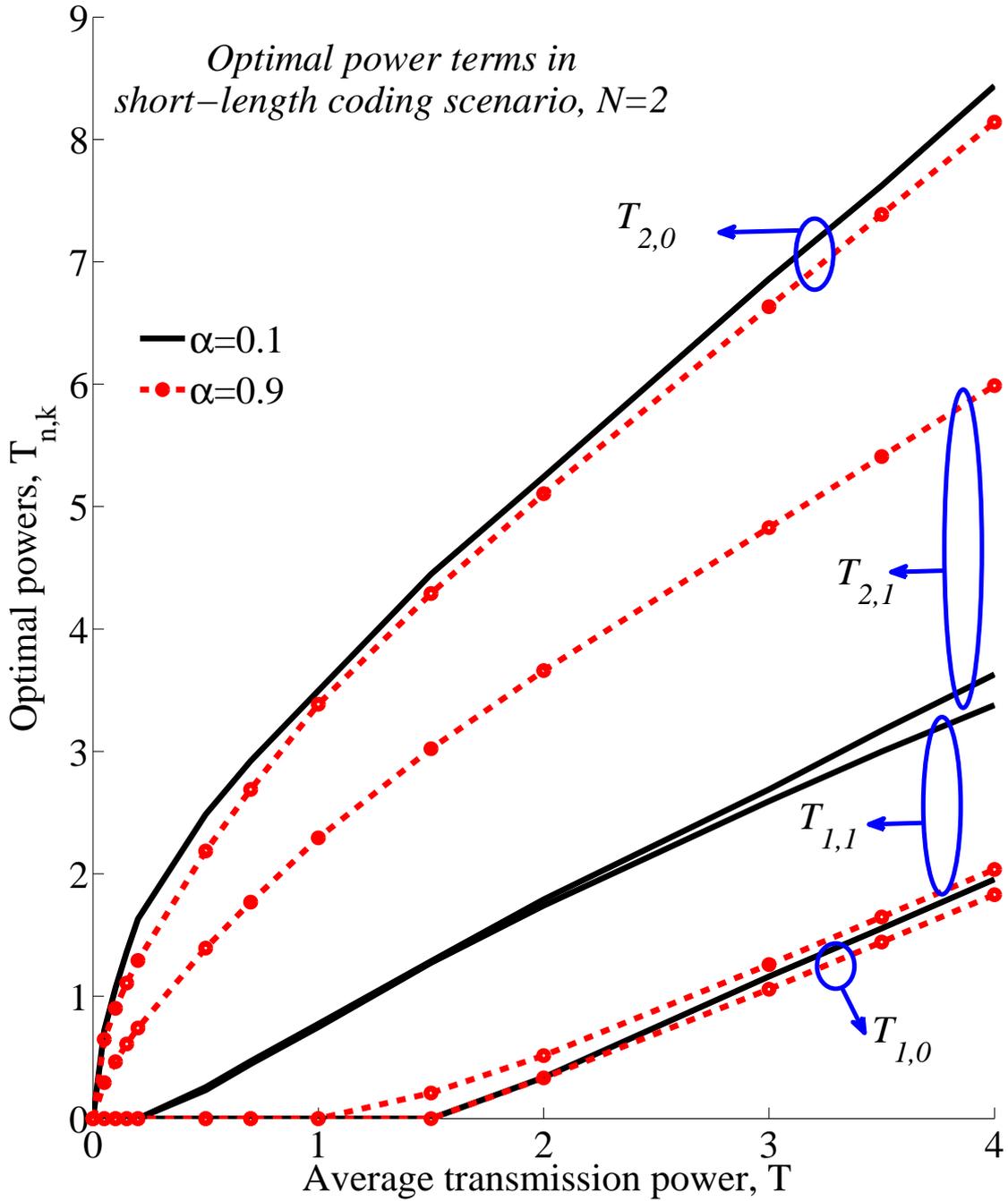


Fig. 3. Optimal powers $T_{n,k}$ versus the average transmission power T . The power $T_{n,k}$ is allocated to the k -th time slot if the quantized index n has been fed back at time slot $k = 0$. Long-term power constraint, short-length coding scenario. At low transmission powers the on-off approach (not transmitting at all in the time slots where the gain is estimated to be low) is the best data transmission scheme, maximizing the throughput (Remark 3). Also, low powers are allocated to the slots in which the channel is estimated to be weak (Remark 2).

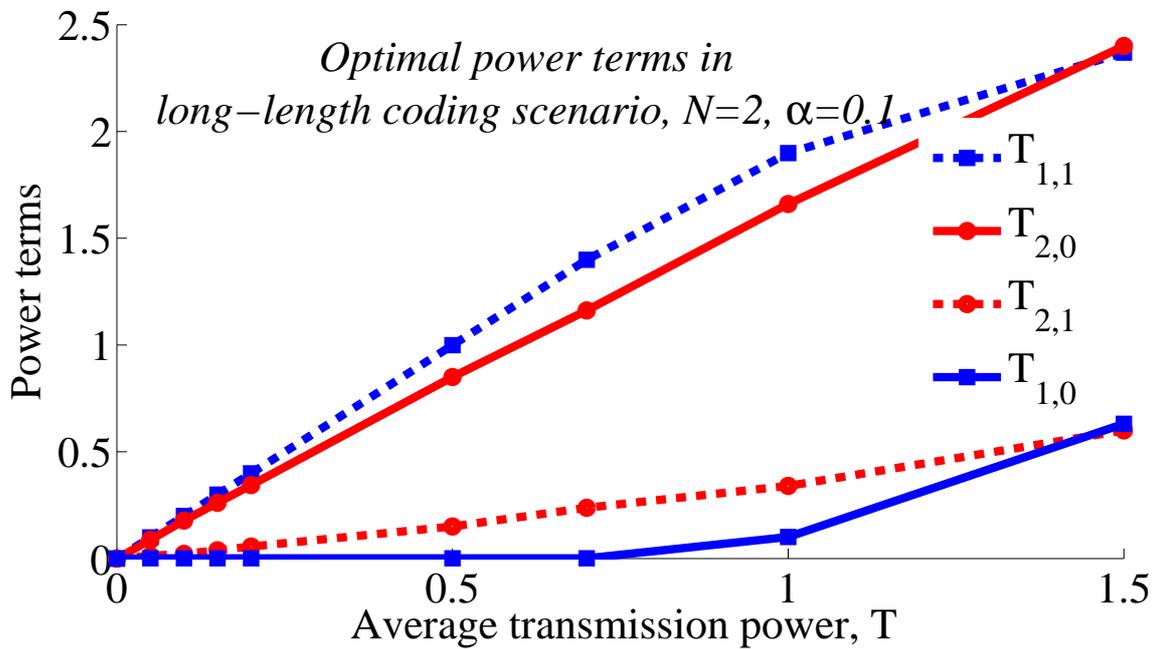
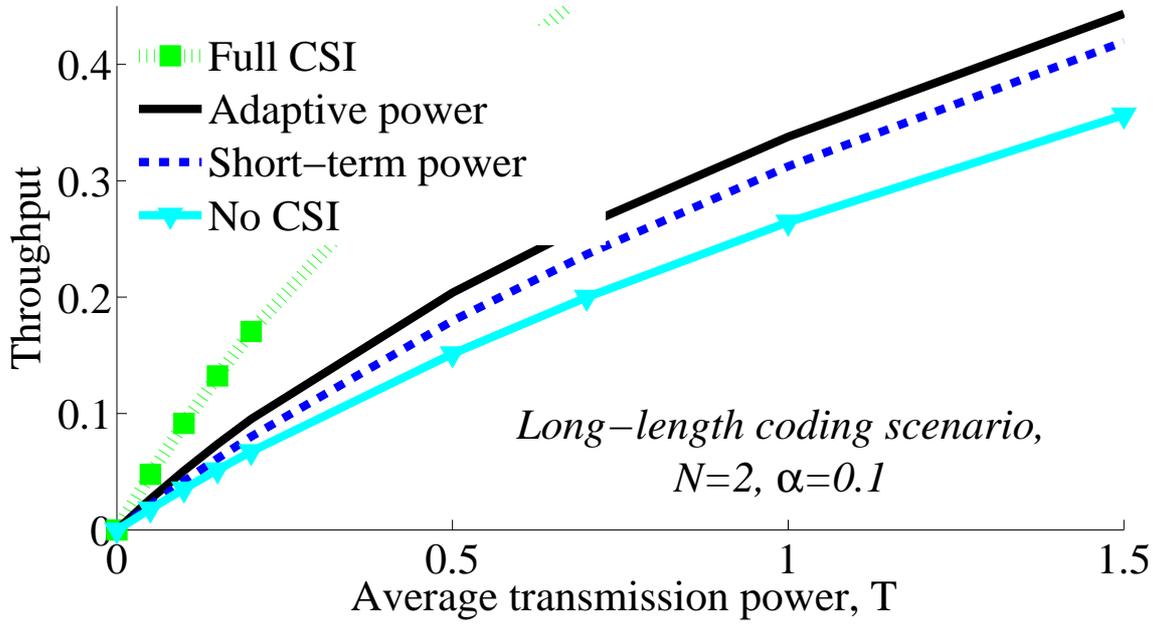


Fig. 4. (a): Throughput and (b): the optimal power terms versus the average transmission power T . Long-length coding scenario, $\alpha = 0.1$. The results are in harmony with Remark 5.

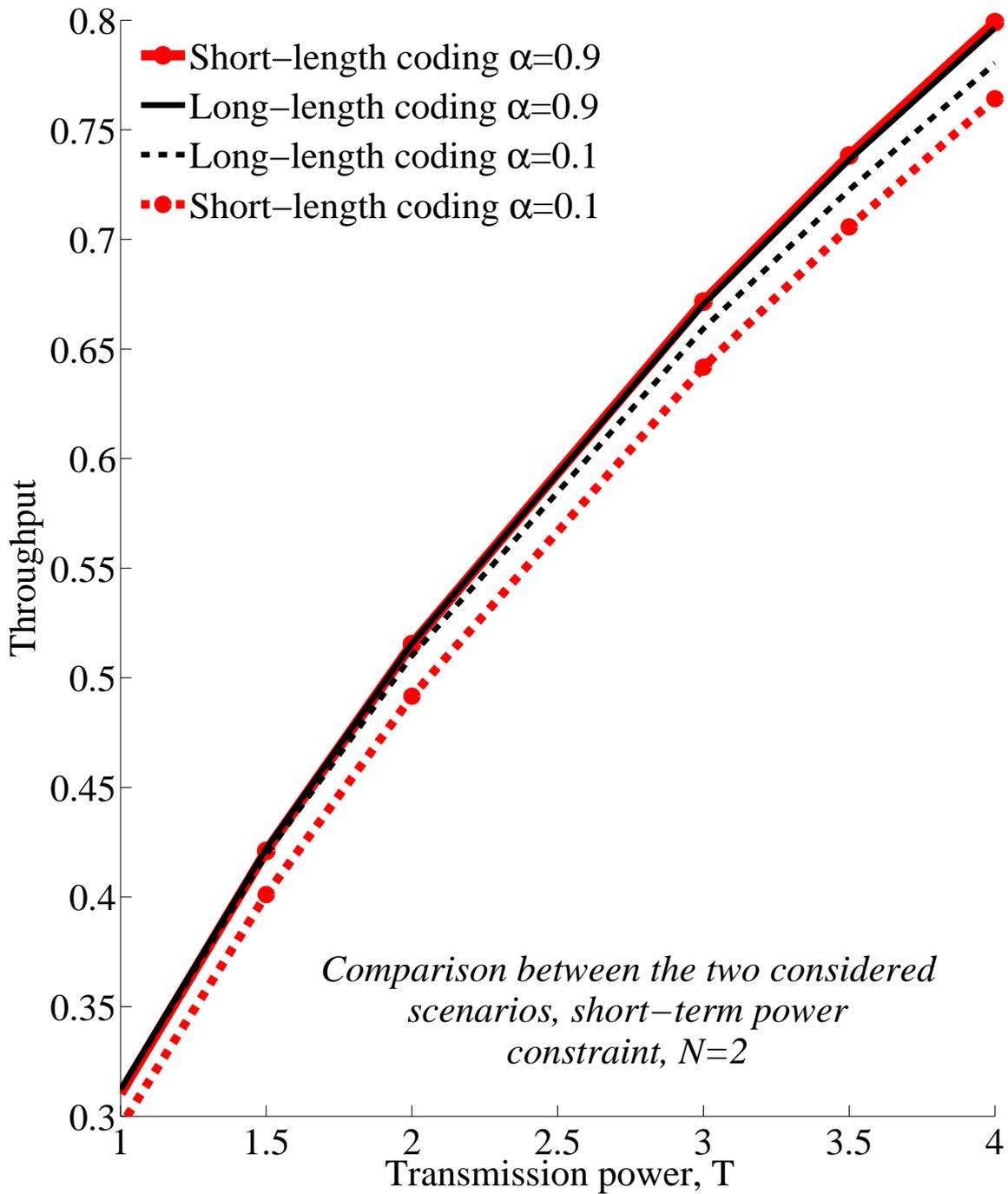


Fig. 5. Comparison between the short- and long-length coding scenarios. Short-term power constraint. For low correlation conditions, higher throughput is achieved by long-length coding scheme, compared to short-length coding model. However, the results converge when the temporal correlation increases.

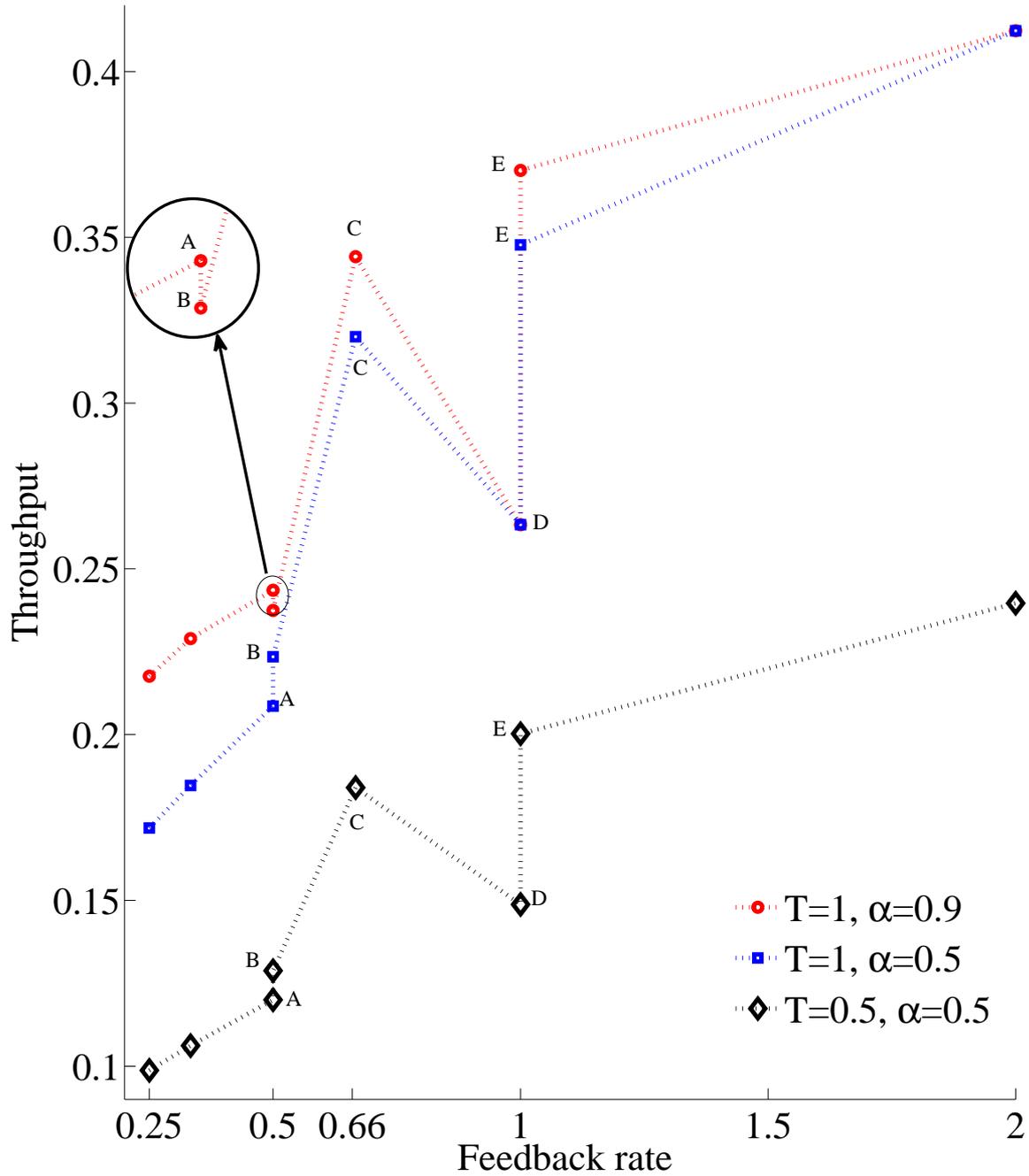


Fig. 6. Throughput vs the feedback rate $\frac{1}{K} \log_2 N$, equal-probability quantization, short-term power constraint. The points A, B, C, D and E correspond to the cases $(K = 2, N = 2)$, $(K = 4, N = 4)$, $(K = 3, N = 4)$, $(K = 1, N = 2)$ and $(K = 2, N = 4)$, respectively. Moreover, the feedback rates $\frac{1}{4}$, $\frac{1}{3}$ and 2 are obtained by setting $(K = 4, N = 2)$, $(K = 3, N = 2)$ and $(K = 1, N = 4)$, respectively.

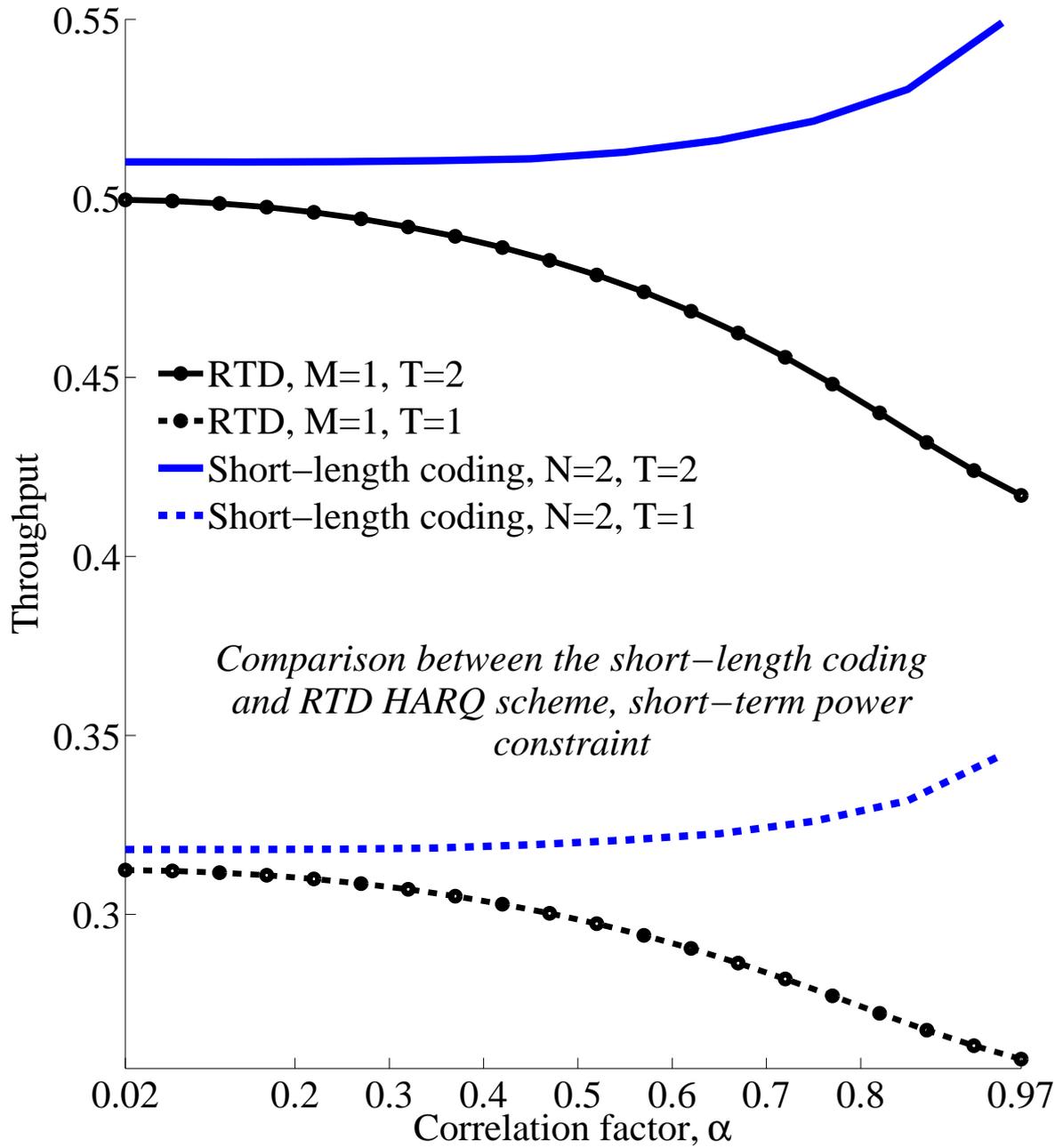


Fig. 7. Comparison between the short-length coding scheme and RTD HARQ feedback approach in the bursting communication model. Short-term power constraint. The quantized CSI approach outperforms the HARQ scheme, in terms of throughput.

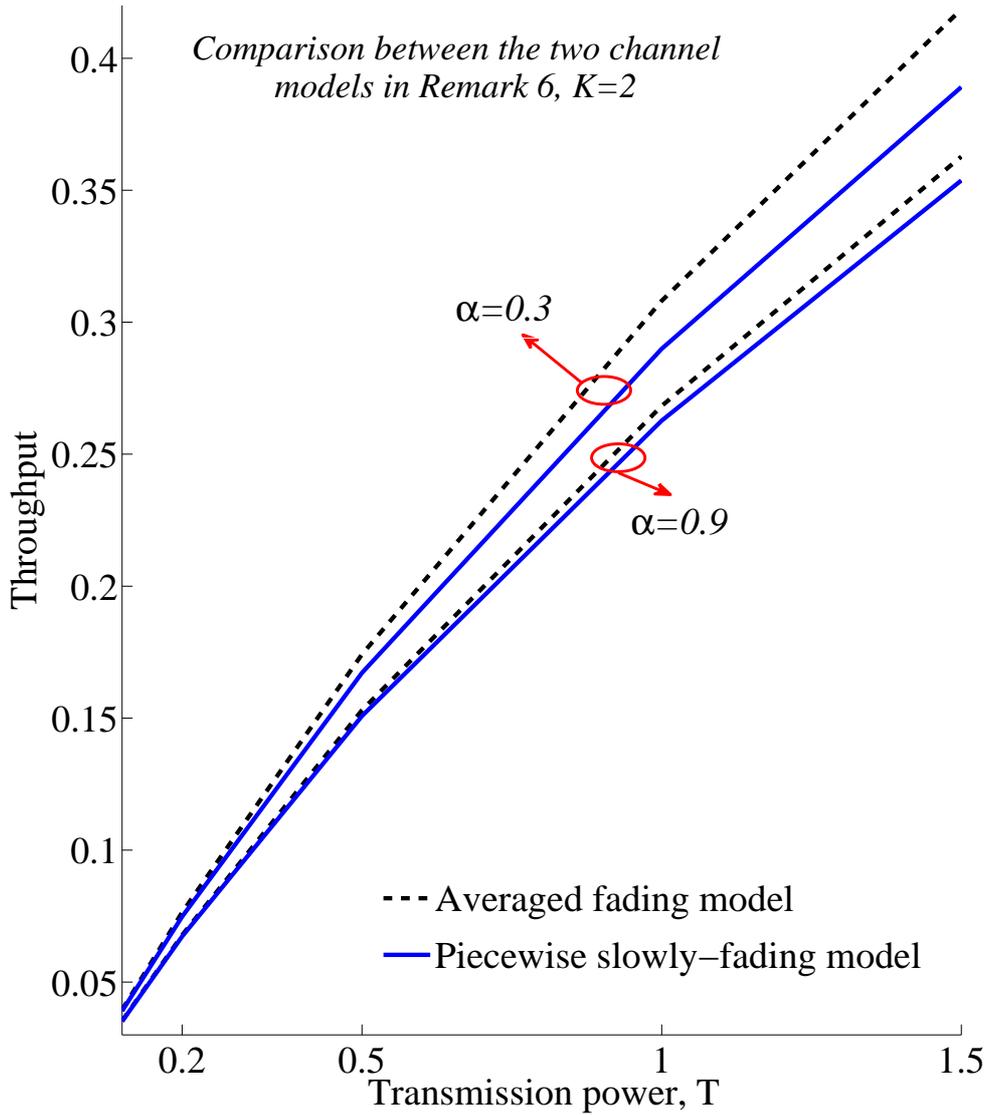


Fig. 8. Comparison between the piecewise slowly-fading and the equivalent averaged channel model in Remark 6, $K = 2$. Less time diversity is exploited when the temporal correlation increases. Therefore, with no CSI at the transmitter, the throughput decreases with the correlation parameter α .

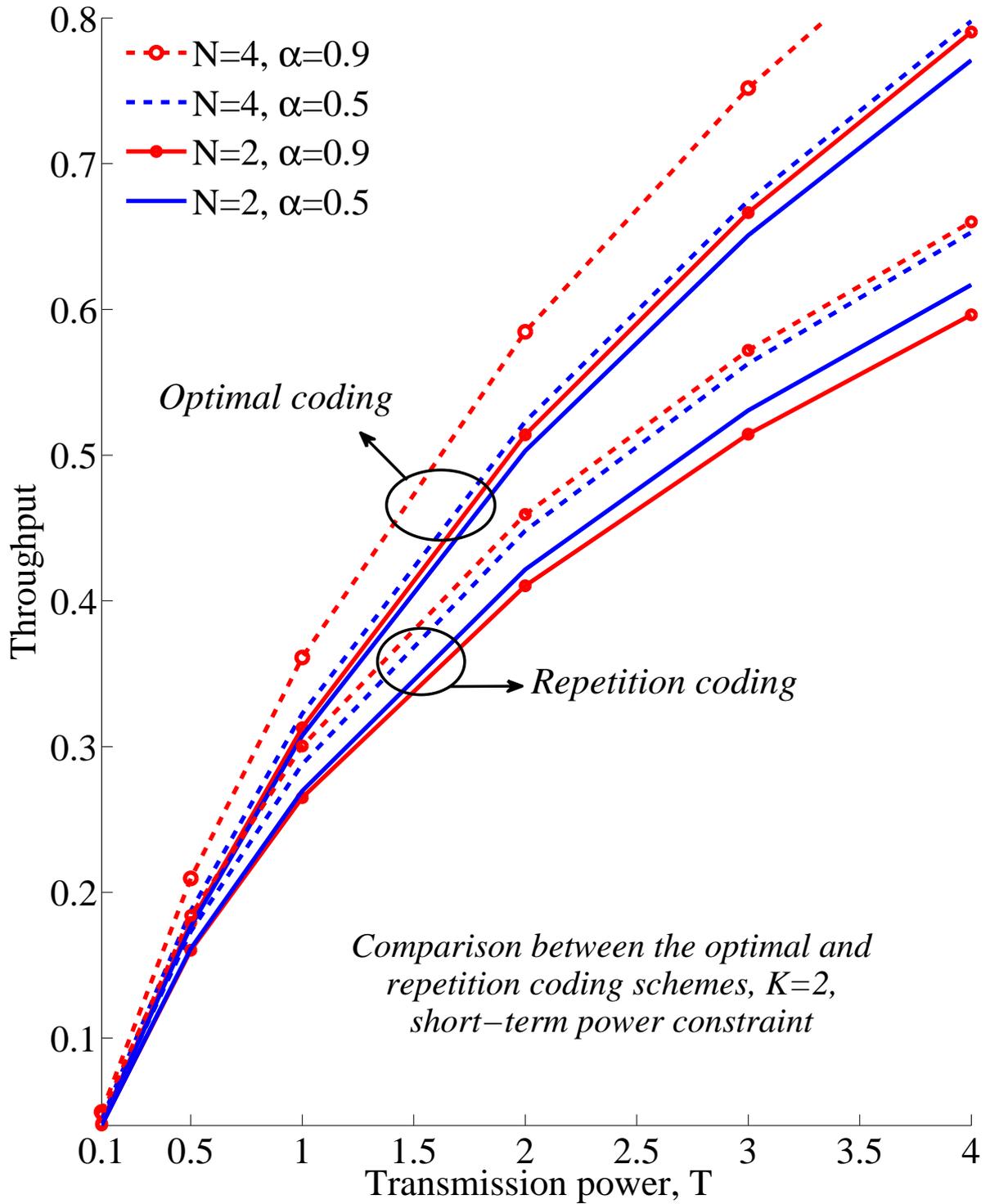


Fig. 9. Comparison between the optimal and repetition codes in the long-length coding scenario. The relative efficiency of the repetition codes, compared to the optimal coding approach, decreases with the transmission power.

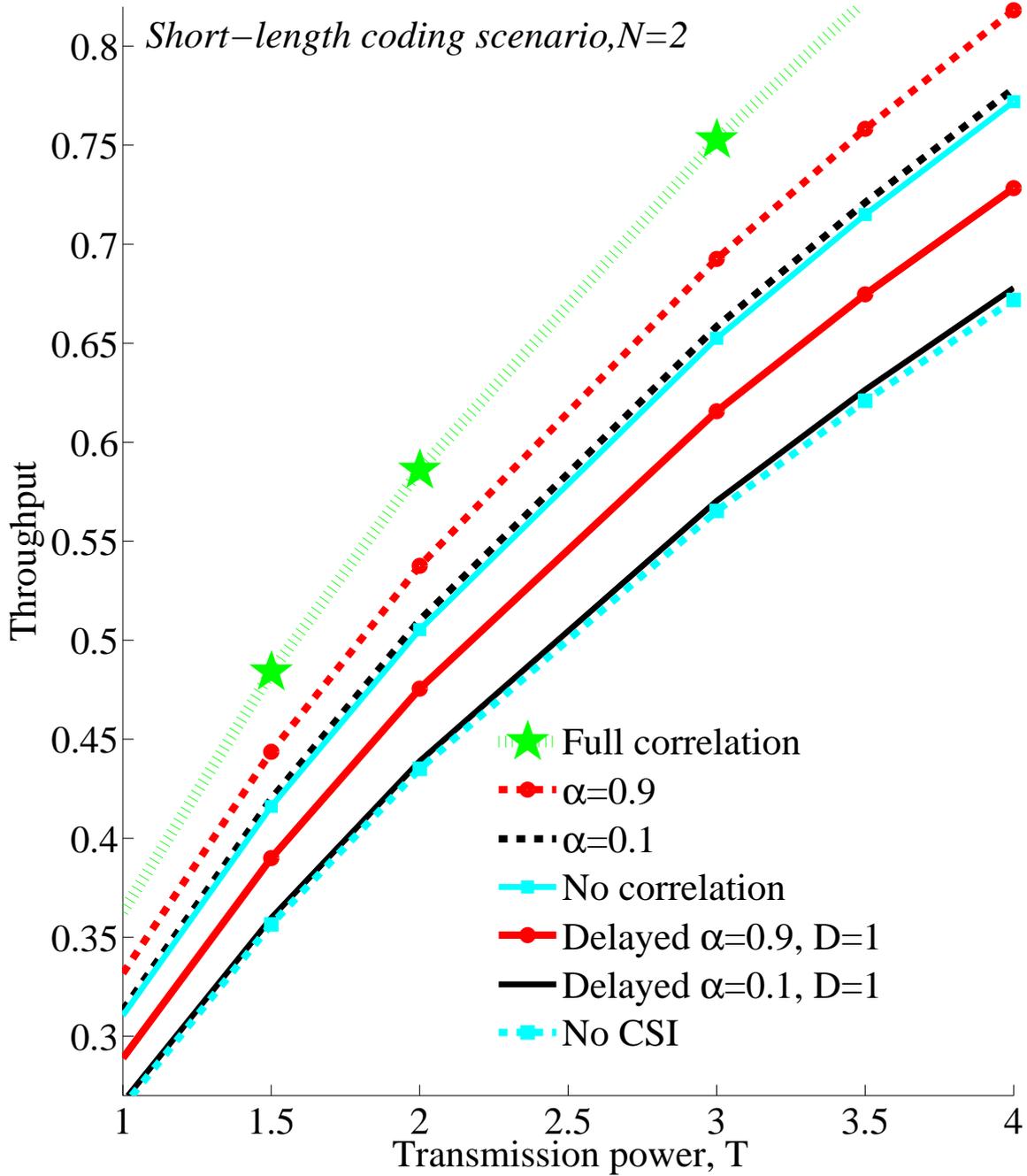


Fig. 10. Throughput versus the transmission power T and different feedback delay conditions. Short-term power constraint, short-length coding scenario. With small temporal correlation, feedback delay deteriorates the performance of the partial CSI scheme substantially.

LIST OF TABLES

I A big picture of the paper. 36

TABLE I
A BIG PICTURE OF THE PAPER.

	Short-length coding	Long-length coding
Assumption	A codeword spans a single block	A codeword spans multiple blocks
Problem	Power-limited throughput optimization	Power-limited throughput optimization
Studied in	Section III	Section IV
Comparison with	HARQ feedback, delayed feedback	Repetition codes