DENSITY EVOLUTION AND LDPC CODE OPTIMIZATION FOR INTERLEAVER DIVISION MULTIPLE ACCESS

Nedeljko Varnica†, Aleksandar Kavčić†, Xiao Ma* and Li Ping*
†Division of Engineering and Applied Sciences, Harvard University
*Department of Electronic Engineering, City University of Hong Kong

Abstract— **We consider multi-stage decoding in the interleaver division multiple access (IDMA) scheme with identical low-density parity-check (LDPC) component codes. We use density evolution to show that, with a proper choice of codes, the LDPC-IDMA scheme nearly achieves the capacities of Gaussian non-fading and fading channels. We discuss several power allocation methods suitable for this scheme and propose a new method based on the density evolution algorithm. When this power allocation is employed in systems in which the overall code rate is given and fixed, we show that it is beneficial to use low-rate component codes.**

I. Introduction

Recently, interleaver division multiple access (IDMA) was introduced as a method to implement multiple access in wireless channels. The separation of users is achieved by assigning different interleavers to different users. The receiver performs an iterative (turbo-like) decoding of different users, thereby utilizing the knowledge of each user's interleaver to separate it form other users. The performance of decoding (multiuser detection) using IDMA is comparable (and in some instances better) than multiuser detection in CDMA, at a complexity cost that is only linearly proportional to the number of users. In CDMA this cost is quadratically proportional to the number of users.

Interestingly, the same principle of IDMA can be used ubiquitously in communication systems. For example, we can use the IDMA principle to shape the input constellation of a Gaussian channel with or without memory. We can use the same principle to shape the input of a fading channel. Even further, the same principle can be used to create optimal channel inputs for a multiple-input multiple-output (MIMO) wireless channel, i.e., we can utilize IDMA to construct space-time codes. The range of applications of IDMA is very broad: single user communications, multiuser communications, low rate communications, high rate communications, single antenna, multiple antennas. Yet, despite this diversity of applications, the underlying principle remains the same: different component codes (each using a different interleaver) are superimposed and transmitted simultaneously.

This research was supported in part by NSF Grant No. CCR-0118701

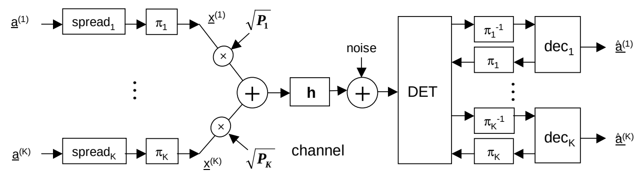

Fig. 1. IDMA principle.

The simplicity of the design and the ubiquity of applications, makes IDMA an ideal technology for 4G systems that will need to transmit data and voice over the same network, for various channel conditions. In this paper, we conduct a theoretical study of capacity achievability limits of the IDMA technology. We utilize the belief propagation method to investigate the IDMA performance bounds. We make a slight generalization of interleaver-separated codes in that we utilize random low-density parity-check (LDPC) codes [1]. These codes have a built-in random interleaver in the form of a random LDPC matrix and are a generalization of codes with random interleavers.

We show in this paper that, with a proper choice of codes, IDMA can nearly achieve the capacities of Gaussian non-fading and fading channels. This is done by developing a density evolution method for LDPC-IDMA systems with the multi-stage decoding (MSD) schemes [2]. We then propose a convenient power allocation method obtained within the density evolution algorithm. Finally, under Gaussian approximations, we derive a closed form expression for the gap between the channel capacity and the performance of the LDPC-IDMA scheme with the proposed power allocation. Asymptotic results based on this expression show that it is more beneficial to use optimized low-rate codes in order to achieve the capacities of systems in which the system code rate is given and fixed.

II. IDMA background

In Fig. 1 we illustrate the IDMA principle. The number of users (or component codes) in this system is denoted by K. At the transmitter, each bit of the input vector $\underline{a}^{(i)}$ for the i-th user ($i = 1, \cdots, K$) is spread. The spreader (or a short code [3]) is typically the same for all users. Then the user-specific chip-interleaver π_i is applied to obtain the transmitted vector $\underline{x}^{(i)}$. The vectors $\underline{x}^{(1)}, \cdots, \underline{x}^{(K)}$ are superimposed and sent through the channel.

We consider the additive white Gaussian noise (AWGN) channel and the Rayleigh fading channel. These channels are described by the probabilistic law

$$Y_t = h \cdot \sum_{i=1}^{K} \sqrt{P_i} X_t^{(i)} + W_t \ , \qquad (1)$$

where $X_t^{(i)} \in \{+1, -1\}$ and $Y_t \in \mathbb{R}$ respectively denote the channel input (for component code i) and channel output random variables at discrete time instant t. The additive term W_t represents the AWGN with zero mean and the standard deviation σ. The power allocated to the i-th component code is denoted by P_i. For simplicity we assume that the channel coefficient h is real. (The extension to channels with complex channel coefficients is straightforward.) For the normalized fading channel, the coefficient h is a random variable with the probability density function (pdf) $f(h) = 2h \cdot e^{-h^2}$; $h > 0$. For the AWGN channel, the coefficient h is a constant (i.e., $h = 1$). To simplify the notation in the sequel, we note that the channel coefficient h for the AWGN channel can also be represented as a random variable with the pdf $f(h) = \delta(h-1)$.

The capacity of the channel defined by (1) with the noise standard deviation σ and the power constraint $P = \sum_i P_i$ is given by $C = \frac{1}{2}\mathrm{E}_h \left[\log_2\left(1 + h^2 \frac{P}{\sigma^2}\right)\right]$. The SNR (defined as $10\log_{10}\frac{P}{\sigma^2}$) at which the channel capacity $C = r$ is achieved is denoted by SNR_r. The capacity of the channel with BPSK signaling input is denoted by C_{BPSK}.

The receiver operates through the detector DET in an iterative turbo-type fashion. The description of the detector DET appears in [4] and is omitted here due to space limitations.

III. LDPC-IDMA Coding/Decoding Scheme

A slightly more general IDMA system is obtained when the spreaders and interleavers from Fig. 1 are replaced by LDPC codes. This is shown in Fig. 2. In this paper we assume the following MSD scheme in the LDPC-IDMA system. We first perform iterative decoding of LDPC code 1 by employing LDPC decoder 1 and the channel detector DET. The information from other $K-1$ codes is in this step considered as noise. (The channel detector operates in the same fashion as the detector in the receiver of IDMA system of Fig. 1; see [4] for details.) Then the decoded vec-

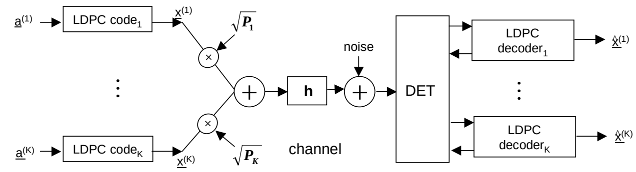

Fig. 2. IDMA scheme with LDPC component codes.

tor $\hat{\underline{x}}^{(1)} = \left[x_1^{(1)}, x_2^{(1)}, \cdots, x_N^{(1)}\right]$ is passed to the channel detector DET. The decoding of LDPC code 2 is performed in the same iterative fashion by employing LDPC Decoder 2 and the channel detector. At this decoding stage the information from codes $3, \cdots, K$ is considered as noise. The decoded vectors $\hat{\underline{x}}^{(1)}$ and $\hat{\underline{x}}^{(2)}$ from Decoders 1 and 2 are then passed to the channel detector, and so on.

Note that this coding scheme inherently performs shaping due to the underlying IDMA principle. (By the central limit theorem the channel input approaches the Gaussian distribution as the number of component codes increases.) Hence, this scheme is much simpler than the schemes that incorporate the multi-level coding combined with additional shaping methods [5], [6].

A. Density Evolution for LDPC-IDMA

For iterative belief-propagation (BP) decoders we use density evolution [7] to evaluate the code noise tolerance threshold σ^*. Here σ^* denotes the maximal noise standard deviation for which the given code can be decoded with arbitrarily low probability of error.

We now describe the density evolution for LDPC decoder i; $1 \leq i \leq K$. Let the edge degree vectors [8] $\underline{\lambda} = [\lambda_2, \cdots, \lambda_L]$ and $\underline{\rho} = [\rho_2, \cdots, \rho_R]$ of this code be given. Here by L and R we respectively denote the maximum variable node degree and the maximum check node degree [8] in the LDPC code graph. The vectors $\underline{\lambda}$ and $\underline{\rho}$ define an ensemble of LDPC codes. Assuming infinite block length and averaging over all LDPC codes in the given ensemble, we denote the average pdf of the log-likelihood-ratio (LLR) message emitted from a check node to a variable node after ℓ BP decoding iterations by $f_c^{(\ell)}$. The average pdf of the LLR message emitted from a variable node to a check node is denoted by $f_v^{(\ell)}$. The average pdf of the extrinsic information that the detector DET sends to the LDPC decoder is denoted by $f_D^{(\ell)}$. Denoting the convolution operator by $\otimes$ we have [9]

$$f_v^{(\ell)}(\xi) = f_{c_e}^{(\ell-1)} \otimes f_D^{(\ell-1)}(\xi) \ ,$$

where $f_{c_e}^{(\ell-1)} = \sum_{m=2}^{L} \lambda_m \left(\otimes_{k=1}^{m-1}\left(f_c^{(\ell-1)}(\xi)\right)\right)$ and

$\otimes_{k=1}^{m-1}$ denotes the convolution of m–1 pdf's. The update of the density $f_c^{(\ell)}$ is the same as in [7].

The density $f_D^{(\ell)}$ is obtained via a Monte Carlo simulation. In general (when the channel has memory), to obtain the $f_D^{(\ell)}$ we need the average pdf $f_e^{(\ell)}$ of the extrinsic information that the LDPC decoder sends to the channel detector. This pdf is given by $f_e^{(\ell)} = \sum_{m=2}^{L} \mu_m \left(\otimes_{k=1}^{m} \left(f_c^{(\ell)}(\xi) \right) \right)$, where $\mu_m = \frac{\lambda_m}{\sum_j \lambda_j}$. In a less general case considered in this paper, the channel is memoryless and the pdf $f_e^{(\ell)}$ need not be computed. The density $f_D^{(\ell)}$ is obtained as follows.

For LDPC code i we assume that all the bits from the codes $1, \cdots, i-1$ are known. We run the Monte Carlo simulation by generating a different channel coefficient h for each of N bit slots (where N is a very large integer, say $N = 10^6$) according to the pdf f_h. The unknown bits (of codes $i+1, \cdots, K$) are generated randomly with probabilities $P(x_t^{(j)} = +1) = P(x_t^{(j)} = -1) = 1/2$, for $j = i+1, \cdots, K$ and $t = 1, 2, \cdots, N$. By employing the decoder DET for this block, we obtain a histogram that represents the pdf $f_D^{(\ell)}$.

The average probability that a variable-to-check edge carries an erroneous message after ℓ iterations equals

$$p_\ell = \int_{-\infty}^{0} f_v^{(\ell)}(\xi) d\xi \ .$$

The thresholds are found for each individual LDPC decoder (i.e., for each i) as the maximum noise levels at which $\lim_{\ell \to \infty} p_\ell = 0$. We denote the threshold of the code i by σ_i^*.

B. Power Allocation

One of the central issues in the proposed LDPC-IDMA coding/decoding scheme is the power distribution between component codes. For related work on CDMA systems see [10] and references therein.

B.1 Power Allocation for Capacity-Achieving Codes

One approach is to split the system power $P = \sum_i P_i$ into the component code powers so that the maximal information rates (under the Gaussian approximation) are identical for all K subchannels, where each subchannel corresponds to a different component code [11]. In other words, we can select the powers P_i based on the subchannel capacities

$$\frac{1}{2} \mathrm{E} \left[\log_2 \left(1 + h^2 \frac{P_K}{\sigma^2} \right) \right] = r, \qquad (2)$$

and (for $i = 1, \cdots, K-1$)

$$\frac{1}{2} \mathrm{E} \left[\log_2 \left(1 + h^2 \frac{P_i}{\sum_{j=i+1}^{K} P_j + \sigma^2} \right) \right] = r \ , \qquad (3)$$

where σ satisfies $10 \log_{10} \frac{P}{\sigma^2} = SNR_{K \cdot r}$.

Assume that the threshold σ_i^* $(1 \le i \le K)$ for each individual LDPC code is computed as described in Section III-A. Then the threshold of the entire LDPC-IDMA coding scheme is $\sigma^* = \min_{1 \le i \le K} \sigma_i^*$. Denote the distance from the threshold $SNR_r^* = 10 \log_{10} \frac{P_K}{\sigma^{*2}}$ to SNR_r by $\Delta SNR_r = SNR_r^* - SNR_r$. Let $SNR_{K \cdot r}^* = 10 \log_{10} \frac{P}{\sigma^{*2}}$ and $\Delta SNR_{K \cdot r} = SNR_{K \cdot r}^* - SNR_{K \cdot r}$.

Fact 1. Assuming that the power is distributed according to Equations (2)-(3), the distance ΔSNR_r is preserved, i.e., $\Delta SNR_r = \Delta SNR_{K \cdot r}$.

This observation is useful from the code construction perspective. Namely, if we could find an optimal LDPC code that achieves the capacity C of the subchannel K, we would automatically obtain optimal codes for multi-user detection (or optimal shaping component codes) on the Rayleigh fading and AWGN channels. Furthermore, we can use *the same* code for all component codes. This makes the code optimization procedure much simpler, especially when the number of component codes is large.

Unfortunately, given that the BPSK capacity C_{BPSK} is strictly smaller than the channel capacity C, binary LDPC codes can not achieve the channel capacities. However, these codes can very closely approach the BPSK channel capacities, which for sufficiently low rates r are very near the channel capacities. Therefore, when powerful low-rate component codes are employed the above power allocation can still be used to obtain a code construction method that closely approaches the capacities of the single-user or multi-user communication systems.

B.2 Power Allocation for Nearly Optimal Codes

Since the capacities can only be closely approached (but not achieved), we propose a more practical way to choose power allocation below. (We opt to use the power allocation proposed below, rather than the optimal simulation-based power allocation from [11], as our power allocation leads to a tractable theoretical analysis.) For a given $\overline{\sigma}$, we perform the density evolution algorithm on each individual code as described in Section III-A. We compute the threshold $\overline{\sigma}^*$ as the maximal $\overline{\sigma}$ for which the probability of error p_ℓ for each LDPC decoder tends to 0. Unlike the power allocation from Section III-B.1, the power allocation here is dependent on $\overline{\sigma}$ in the density evolution algorithm.

For a power allocation $\overline{P}_1, \cdots, \overline{P}_K$, we consider the information rates

$$R = \frac{1}{2}\mathrm{E}\left[\log_2\left(1 + h^2\frac{P}{\overline{\sigma}^2}\right)\right] \tag{4}$$

$$R_i = \frac{1}{2}\mathrm{E}\left[\log_2\left(1 + h^2\frac{\overline{P}_i}{\sum_{j=i+1}^{K}\overline{P}_j + \overline{\sigma}^2}\right)\right] \;\; ; \; i < K \tag{5}$$

and

$$R_K = \frac{1}{2}\mathrm{E}\left[\log_2\left(1 + h^2\frac{\overline{P}_K}{\overline{\sigma}^2}\right)\right] \; . \tag{6}$$

We select $\overline{P}_i$ so that $R_i = \frac{R}{K}$ for each i. Clearly we have that the constraint $\sum_i \overline{P}_i = P$ is satisfied. Therefore, this power allocation is matched to a given $\overline{SNR}_{K\cdot r} = 10\log_{10}\frac{P}{\overline{\sigma}^2}$ and is different for different $\overline{\sigma}$ in the density evolution algorithm.

Denote the distance from the threshold $\overline{SNR}_r^* = 10\log_{10}\frac{\overline{P}_K}{\overline{\sigma}^{*2}}$ to SNR_r by $\Delta\overline{SNR}_r = \overline{SNR}_r^* - SNR_r$. Let $\Delta\overline{SNR}_{K\cdot r} = 10\log_{10}\frac{P}{\overline{\sigma}^{*2}} - SNR_{K\cdot r}$.

Theorem 1. Assume that the power is distributed according to Equations (4)-(6) and that the channel is AWGN. Then, under the Gaussian approximation, the distance $\Delta\overline{SNR}_{K\cdot r}$ is given by

$$\Delta\overline{SNR}_{K\cdot r} = 10\log_{10}\frac{\left(1 + 10^{\frac{\Delta\overline{SNR}_r}{10}}\cdot\left(2^{2r}-1\right)\right)^K - 1}{2^{2Kr}-1} \; .$$

Asymptotically as $K \to \infty$, if $\Delta\overline{SNR}_r > 0$ and $r > 0$, the distance $\Delta\overline{SNR}_{K\cdot r}$ increases linearly with K, i.e., $\Delta\overline{SNR}_{K\cdot r} = \alpha K$, where

$$\alpha = 10\log_{10}\frac{1 + 10^{\frac{\Delta\overline{SNR}_r}{10}}\cdot\left(2^{2r}-1\right)}{2^{2r}} \; .$$

Proof. Proof is given in the appendix.

Corollary 1. Under the following assumptions

1) $2^{2Kr} \gg 1$ and

2) $\Delta\overline{SNR}_r \ll 4$ dB,

we have that

$$\Delta\overline{SNR}_{Kr} \approx K\cdot\left(1 - 2^{-2r}\right)\cdot\Delta\overline{SNR}_r \; . \tag{7}$$

Furthermore, if $r \ll 1$,

$$\Delta\overline{SNR}_{Kr} \approx 2(\ln 2)\cdot(K\cdot r)\cdot\Delta\overline{SNR}_r \; . \tag{8}$$

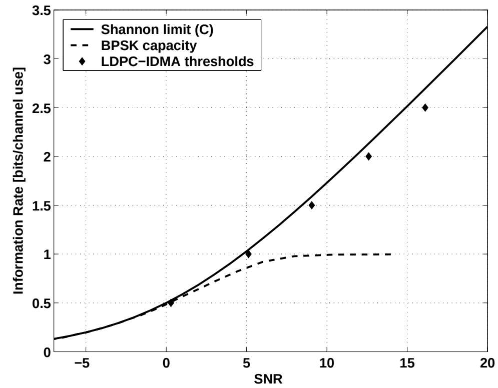

Fig. 3. Capacities and thresholds $\overline{\sigma}^*$ of IDMA schemes with $K = 1, 2, 3, 4, 5$ for AWGN channel. The component codes are identical irregular codes with $r = \frac{1}{2}$.

Proof. Proof is given in the appendix.

One interesting implication of this result is that it is worth optimizing codes at low code rates r for systems for which the total rate $K\cdot r$ is given and fixed. To see this, note that by optimizing *binary* codes we can not make $\Delta\overline{SNR}_r$ very small unless r is relatively low. That is, for example, for AWGN channel, if $r = 0.5$ the gap between the SNR at which $C_{BPSK} = r$ and the SNR at which $C = r$ is 0.2dB. This gap can not be bridged by a linear binary code. Hence, we have from (7) that the gap at, say, $K\cdot r = 5$ (note that our approximations are valid for $K\cdot r = 5$) would be at least 1dB if we used $K = 10$ binary codes with rate $r = 0.5$.

On the other hand, for low r ($r \ll 1$) the gap between the SNRs corresponding to $C_{BPSK} = r$ and $C = r$ can be made arbitrarily small. Denote this gap by δ_{SNR}. Then, from (8) we have that the gap at $K\cdot r = 5$ is still only $10\cdot(\ln 2)\cdot\delta_{SNR} \approx 7\cdot\delta_{SNR}$. (Naturally, optimization of LDPC codes at very low rates r is very difficult, but other near-optimal low-rate codes may be used.)

IV. LDPC-IDMA Performance Bounds

In Fig. 3 and 4 we plot the results of the LDPC-IDMA scheme for AWGN and Rayleigh fading channels when the density evolution is used for $K = 1, 2, 3, 4$ and 5 component codes. The power allocation matched to the density evolution (Section III-B.2) was used. In the presented results, we used the optimized irregular LDPC edge degree vectors with the maximum variable node degree $L = 50$ from [12] (for AWGN channel) and from [13] (for fading channel). We can see that the capacities of both channels are closely approached for all shown values of K. An ap-

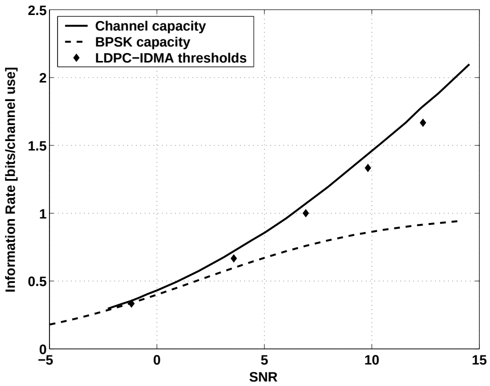

Fig. 4. Capacities and thresholds $\overline{\sigma}^*$ of IDMA schemes with $K = 1, 2, 3, 4, 5$ for Rayleigh fading channel. The component codes are identical irregular codes with $r = \frac{1}{3}$.

parent advantage of this scheme is that, regardless of the value of K, all component codes can be the same. Therefore, we need to optimize only *one* LDPC code rather than K codes, which is necessary when the code rates are different [6].

As predicted by Theorem 1 we observe a trend of increased distance from the thresholds $\overline{\sigma}^*$ to the capacities C as the number of component codes increases while the code rates of the component codes remain identical and unchanged. Nevertheless, even for $K = 5$ component codes the gaps to the channel capacities are approximately only 0.85dB and 1.2dB for the fading and the AWGN channel, respectively.

V. Conclusion

We conclude with several remarks regarding the relation between the proposed power allocation and the optimal power allocation. First note that for multilevel schemes decoded in multi-stage fashion [2] we can find the optimal power allocation using the simulation based approach from [11]. In this approach the optimal power for the K-th component code is found first. Then, with the knowledge of the power of the K-th component code, the optimal power for the $(K-1)$-st code is found. The knowledge of the optimal powers from the component codes K and $K-1$ is then utilized to find the optimal power for the code $K-2$, and so on. Interestingly, our preliminary results show that, for good LDPC codes, the gap between the performance of the proposed power allocation and the optimal power allocation does not increase with K, for larger values of K. This is demonstrated in Table I where we compare the performances for the LDPC-IDMA system with the regular-(3,6) and irregular component codes with 1) the optimal power allocation and 2) the power allocation from Section III-B.2. We see that the per-

	Irregular (Fig. 3)			Regular (3,6)		
	Threshold [dB]		Gap	Threshold [dB]		Gap
	opt	III-B.2		opt	III-B.2	
K = 1	0.29	0.29	0	1.11	1.11	0
K = 2	5.23	5.30	0.07	6.24	6.28	0.04
K = 3	9.11	9.24	0.13	10.36	10.41	0.05
K = 4	12.62	12.75	0.13	14.16	14.23	0.07
K = 5	15.99	16.12	0.13	17.84	17.91	0.07

TABLE I

Thresholds (in dB) of IDMA schemes with $K = 1, 2, 3, 4, 5$ regular and irregular rate-1/2 component codes for AWGN channel. The results for the optimal power allocation and the power allocation from Section III-B.2 are shown.

formance gaps between the two power allocation methods are very small for both regular and irregular codes. Furthermore, the gap "saturates" for larger values of K. This (together with Corollary 1) indicates that, if the total code rate $K \cdot r$ of an IDMA-LDPC scheme with MSD is given and fixed, it is necessary to employ low-rate component codes in order to closely approach the capacity – even when the optimal power allocation is utilized.

Although the presented results give the performance of different power allocation methods for the multilevel LDPC coding schemes with MSD, the question of optimal power allocation for decoders that do not employ the MSD approach remains. Nevertheless, when the component codes are powerful turbo-like codes, the proposed approaches yield near optimal power allocations in the instances when the overall code rate is not very high and in the instances when the component code rates are low.

APPENDIX

Sketch of the proof of Theorem 1: Recall the Shannon's formula for the minimum power required to transmit r bits of information

$$P_{K_{\min}} = \left(2^{2r} - 1\right) \sigma^2 \ .$$

Assume that for a given (practical) code the required power at the K-th level is $P_K = a \cdot P_{K_{\min}}$, where $a = 10^{\frac{\Delta \overline{SNR}_r}{10}} > 1$. Now consider superimposing the $(K–1)$-st layer. In decoding the $(K–1)$-st layer signal we treat the K-th layer signal as AWGN. Thus, the required power level for the $(K-1)$-st layer is

$$P_{K-1} = a \cdot \left(2^{2r} - 1\right) \cdot \left(\sigma^2 + P_K\right) =$$
$$a \cdot \left(2^{2r} - 1\right) \cdot \left(1 + a(2^{2r} - 1)\right) \sigma^2 \ .$$

Similarly, for $k = 2, \cdots, K-1$ we obtain

$$P_{K-k} = a \cdot \left(2^{2r} - 1\right) \cdot \left(1 + a(2^{2r} - 1)\right)^k \sigma^2 \ .$$

Therefore, the total power used is

$$\sum_{i=1}^{K} P_i = \left(\left(1 + a(2^{2r} - 1)\right)^K - 1\right) \sigma^2$$

and the total supported rate is $K \cdot r$.

On the other hand, the minimum power required to transmit $K \cdot r$ bits of information is

$$P_{\min} = (2^{2Kr} - 1)\sigma^2 \ .$$

Hence, the difference (in dB) in efficiency between the $\sum_{k=1}^{K} P_k$ and $P_{\min}$ is

$$\Delta\overline{SNR}_{K \cdot r} = 10 \log_{10} \frac{\left(1 + 10^{\frac{\Delta\overline{SNR}_r}{10}} \cdot \left(2^{2r} - 1\right)\right)^K - 1}{2^{2Kr} - 1} ,$$

which completes the proof.

Sketch of the proof of Corollary 1: If $2^{2Kr} \gg 1$, we can ignore the constant term -1 in both the numerator and the denominator of the above expression for $\Delta\overline{SNR}_{K \cdot r}$. To simplify notation, let

$$x = \frac{\Delta\overline{SNR}_r}{10}$$

and

$$\mathcal{R} = 2^{2r} - 1 \ .$$

When $x \ll 1$ we have $10^x \approx 1 + x \cdot \ln(10)$, and thus

$$\Delta\overline{SNR}_{Kr} \approx$$

$$K \cdot [10 \log_{10} \left(1 + \mathcal{R} \left(1 + x \ln(10)\right)\right) - 10 \log_{10}(1 + \mathcal{R})] =$$

$$K \cdot 10 \log_{10} \left(1 + y\right) \ ,$$

where $y = \frac{\mathcal{R}}{\mathcal{R}+1} \cdot \ln(10) \cdot x$.

Hence, if $y \ll 1$ (which holds if $\mathcal{R} \ll 1$ or if $x \ll \frac{1}{\ln(10)}$) we have $\log_{10}(1 + y) \approx \frac{y}{\ln(10)}$, and

$$\Delta\overline{SNR}_{Kr} \approx K \cdot \frac{10}{\ln(10)} \cdot y = 10 \cdot K \cdot \frac{\mathcal{R}}{\mathcal{R} + 1} \cdot x,$$

i.e.,

$$\Delta\overline{SNR}_{Kr} \approx K \cdot \left(1 - 2^{-2r}\right) \cdot \Delta\overline{SNR}_r \ .$$

To complete the proof, we note that if $r \ll 1$ we have $1 - 2^{-2r} \approx 2 \cdot r \cdot (\ln 2)$ and hence

$$\Delta\overline{SNR}_{K \cdot r} \approx 2(\ln 2) \cdot (K \cdot r) \cdot \Delta\overline{SNR}_r \ .$$

Remark: The Gaussian approximation is assumed in Theorem 1 and, thereby, throughout these derivations. This approximation is very accurate when $r \ll 1$.

To see this, first note that for $r \ll 1$ we have $\overline{P}_K \ll \sigma^2$. Also note that $K \gg 1$, because $r \ll 1$ and $K \cdot r$ has to be at least 3 or 4 for our assumption $2^{2Kr} \gg 1$ to hold. Recall that Decoder i considers the symbols from codes $i+1, i+2, \cdots, K$ as noise. For large values of i (e.g., $i \in \{K-3, K-2, K-1\}$) we have $\overline{P}_j \ll \sigma^2$ (for $j \in \{i+1, \cdots, K\}$) and the Gaussian approximation on the subchannel i is accurate. For smaller values of i the Gaussian approximation on the subchannel i is also accurate, because the central limit theorem "kicks in".

References

[1] R. G. Gallager, *Low-Density Parity-Check Codes.* Cambridge, MA: MIT Press, 1962.

[2] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes," *IEEE Trans. Inform. Theory*, vol. 23, pp. 371–377, May 1977.

[3] L. Ping, L. Liu, and W. K. Leung, "A simple approach to near-optimal multiuser detection: interleave-division multiple access," in *IEEE Wireless Comm. and Networking Conference*, pp. 391–396, 2003.

[4] W. K. Leung, L. Liu, and L. Ping, "Interleaving-based multiple access and iterative chip-by-chip multi-user detection," *IEICE Trans. on Commun.*, vol. E86-B, pp. 3634–3637, Dec. 2003.

[5] U. Wachsmann, R. F. Fischer, and J. B. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Transactions on Information Theory*, vol. 45, pp. 1361–1391, July 1999.

[6] N. Varnica, X. Ma, and A. Kavčić, "Capacity-approaching codes for partial response channels." To appear in Handbook of Coding and Signal Processing for Recording Systems, CRC Press 2004.

[7] T. Richardson and R. Urbanke, "The capacity of low-density parity check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 599–618, February 2001.

[8] M. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. Spielman, "Improved low-density parity-check codes using irregular graphs," *IEEE Trans. Inform. Theory*, vol. 47, pp. 585–598, February 2001.

[9] A. Kavčić, X. Ma, and M. Mitzenmacher, "Binary intersymbol interference channels: Gallager codes, density evolution and code performance bounds," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1636–1652, July 2003.

[10] G. Caire, S. Guemghar, A. Roumy, and S. Verdu, "Maximizing the spectral efficiency of coded CDMA under successive decoding," *IEEE Trans. Inform. Theory*, vol. 50, pp. 152–164, Jan. 2004.

[11] X. Ma and L. Ping, "Power allocations for multilevel coding with sigma mapping," *IEE Electronics Letters*, vol. 40, pp. 609–610, May 2004.

[12] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of capacity-approaching low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 619–637, February 2001.

[13] J. Hou, P. H. Siegel, and L. B. Milstein, "Performance analysis and code optimization of low density parity-check codes on Rayleigh fading channels," *IEEE Journal on Sel. Areas in Communications*, vol. 19, pp. 924–934, May 2001.