Communication Efficiency of Error Correction Mechanism Based on Retransmissions

Dragan Samardzija

Summary

- The SNR gap between the capacity and throughput stays constant as SNR increases.
- The relative efficiency is improving with SNR.
- A minor improvement when combining retransmitted frames.

I. System Outline

In this document we analyze efficiency of a communication system that applies the error correction mechanism using retransmissions. Namely, the transmitter sends a block consisting of N_{in} information and N_{oh} overhead bits, such that $N_b = N_{in} + N_{oh}$. No forward error correction coding is applied. The transmission is performed using digital modulation, where R bits are transmitted per each modulation symbol. For example, $R_{BPSK} = 1$, $R_{QPSK} = 2$, $R_{16-QAM} = 4$ and $R_{64-QAM} = 6$. Therefore the N_b bits are transmitted using a frame consisting of

$$N_s = \lceil N_b / R \rceil \tag{1}$$

symbols.

Upon the reception, the receiver estimates the transmitted symbols and then maps them back onto the corresponding bits, i.e., performs demodulation. The receiver determines whether or not the received block of bits is error-free[1]. If an error is present, the receiver sends a request for the retransmission of the block. While discarding the overhead, only successfully received frames contribute to the average data rate, i.e., throughput. Assuming that there are no missed errors, the throughput is

$$T = R\left(1 - \frac{N_{oh}}{N_b}\right)(1 - FER)\ \ [bits/symbol] \tag{2}$$

where FER is the frame error rate such that

$$FER = 1 - (1 - SER)^{N_s}. \tag{3}$$

SER is the symbol error rate, i.e., the probability that the receiver demodulator fails to correctly estimate the transmitted symbol.

In the above case, frames in error are simply discarded. However, the receiver may combine the corresponding retransmitted frames, thus improving probability of successful reception when retransmissions occur. Consequently, the throughput becomes

$$T^* = R\left(1 - \frac{N_{oh}}{N_b}\right)\left((1 - FER_1) + \sum_{j=2}^{\infty}\frac{1}{j}\left((1 - FER_j)\prod_{k=1}^{j-1} FER_k\right)\right) \tag{4}$$

where FER_j corresponds the jth transmission of a frame.

[1] In practice, to enable the receiver to determine a potential presence of errors, a check code word is added consisting of N_{cw} bits (e.g., for a cyclic redundancy code CRC, if an error is present, it is detected with probability $1 - 2^{-N_{cw}}$). The N_{cw} bits would contribute to the N_{oh} overhead bits.

II. Symbol Error Rate

In this analysis we focus on quadrature-amplitude modulation (QAM) schemes such as QPSK, 16-QAM and 64-QAM. The modulation alphabet consist of M constellation points, $M_{QPSK} = 4$, $M_{16-QAM} = 16$ and $M_{64-QAM} = 64$. In order to determine their symbol error rates we first consider their corresponding pulse-amplitude modulation (PAM) schemes, each with $\sqrt{M}$ uniformly-spaced levels. A QAM scheme may be viewed as two independent orthogonalized PAM schemes assigned to the inphase (i.e., cosine) and quadrature (i.e., sine) component, respectively. Consequently, the relationship between the symbol error rates of the two schemes is

$$SER_{QAM} = 1 - (1 - SER_{PAM})^2. \tag{5}$$

Therefore we will first determine the SER_{PAM}.

The PAM amplitude levels that are symmetric about the origin (i.e., antipodal) are

$$A_i = A\left((i-1) - \frac{1}{2}(\sqrt{M}-1)\right),\ \ i = 1, \cdots, \sqrt{M} \tag{6}$$

where A is the distance between the adjacent levels. The average PAM signal power is

$$P_{PAM} = \frac{1}{\sqrt{M}}\sum_{i=1}^{\sqrt{M}} A_i^2 = \underbrace{\frac{\sum_{i=1}^{\sqrt{M}}\left((i-1) - \frac{1}{2}(\sqrt{M}-1)\right)^2}{\sqrt{M}}}_{\Upsilon} A^2 = \Upsilon A^2. \tag{7}$$

In the case of additive white Gaussian noise (AWGN) channel, the probability of error when the ith amplitude level is transmitted is

$$P_e(i) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma}\int_{A/2}^{\infty} e^{-\frac{x^2}{2\sigma^2}}dx & i = 1, \sqrt{M} \\ \frac{2}{\sqrt{2\pi}\sigma}\int_{A/2}^{\infty} e^{-\frac{x^2}{2\sigma^2}}dx & i = 2, \cdots, \sqrt{M}-1 \end{cases} \tag{8}$$

where σ^2 is the AWGN variance. Since the symbols are equally probable,

$$\begin{aligned} SER_{PAM} &= \frac{1}{\sqrt{M}}\sum_{i=1}^{\sqrt{M}} P_e(i) = \frac{\sqrt{M}-1}{\sqrt{M}}\frac{2}{\sqrt{2\pi}\sigma}\int_{A/2}^{\infty} e^{-\frac{x^2}{2\sigma^2}}dx = \\ &= \frac{\sqrt{M}-1}{\sqrt{M}} erfc\left(\frac{A}{2\sqrt{2}\sigma}\right) = \frac{\sqrt{M}-1}{\sqrt{M}} erfc\left(\frac{1}{2}\sqrt{\frac{P_{PAM}}{2\sigma^2\Upsilon}}\right). \end{aligned} \tag{9}$$

Since two orthogonalized PAM schemes correspond to one QAM scheme, the signal and noise power relate as

$$P_{QAM} = 2\, P_{PAM} \text{ and } N_0 = 2\,\sigma^2. \tag{10}$$

Consequently, the signal to noise ratio (SNR) is

$$SNR = \frac{P_{QAM}}{N_0} = \frac{P_{PAM}}{\sigma^2}. \tag{11}$$

Based on the expressions (5), (7), (9) and (11)

$$SER_{QAM} = 1 - \left(1 - \frac{\sqrt{M}-1}{\sqrt{M}} erfc\left(\frac{1}{2}\sqrt{\frac{\sqrt{M}\, SNR}{2\sum_{i=1}^{\sqrt{M}}\left((i-1) - \frac{1}{2}(\sqrt{M}-1)\right)^2}}\right)\right)^2. \tag{12}$$

In the case of BPSK, $\sqrt{M} = 2$, the SER_{PAM} in (9) should be modified considering that the signal occupies only one dimension, i.e., component, thus doubling the power per component,

$$SER_{BPSK} = \frac{1}{2} erfc\left(\sqrt{\frac{P_{BPSK}}{2\sigma^2}}\right) = \frac{1}{2} erfc\left(\sqrt{SNR}\right). \tag{13}$$

Considering typical modulation schemes, the SER as a function of SNR is depicted in Figure 1.

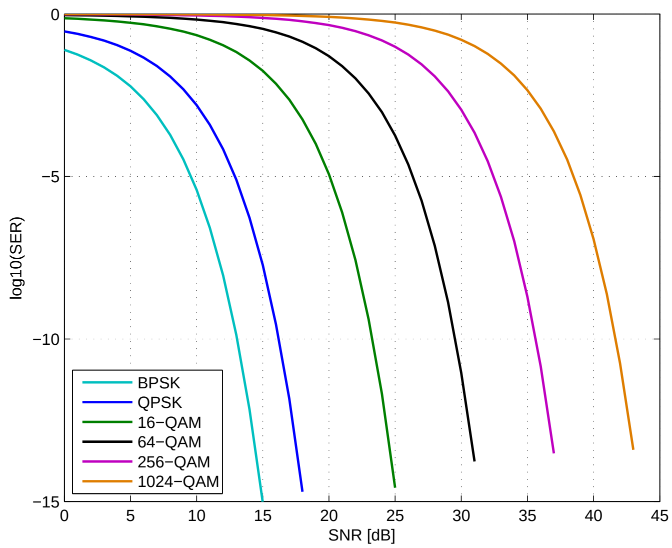

Fig. 1. The SER as a function of SNR.

III. THROUGHPUT

Using the SER that is derived in the previous section, in Figure 2 the throughput in (2) is depicted as a function of the block length N_b, for $SNR = 20$ dB, 64-QAM and $N_{oh} = 32$ bits[2]. The function presents a trade-off between a large block, that lowers the relative overhead, and a small block that lowers the FER. Depending on the SNR and modulation scheme, the block length N_b is selected such that the corresponding throughput is maximized.

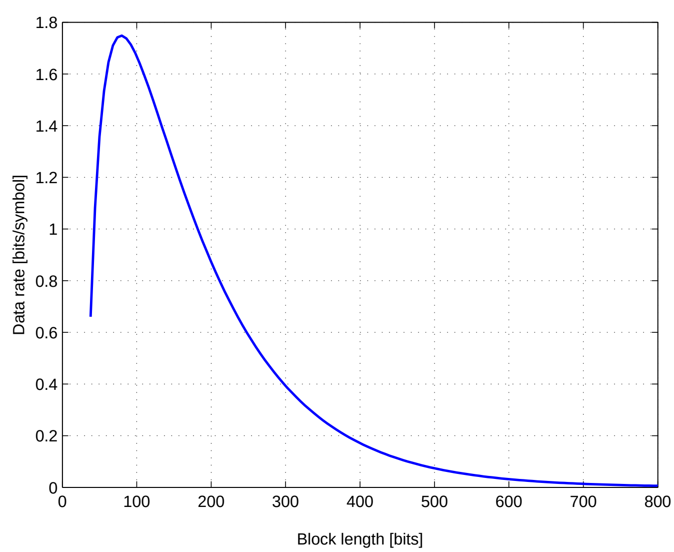

Fig. 2. The throughput as a function of the block length N_b, for $SNR = 20$ dB, 64-QAM and $N_{oh} = 32$ bits.

Once the optimal block length is selected, the throughout is compared against the AWGN channel capacity

$$C = \log_2(1 + SNR) \ \ [bits/symbol]. \tag{14}$$

[2]A typical CRC applies a 32-bit check code word, guaranteeing very low probability of missed errors.

For typical modulation schemes, the throughput and capacity, as functions of SNR, are depicted in Figure 3.

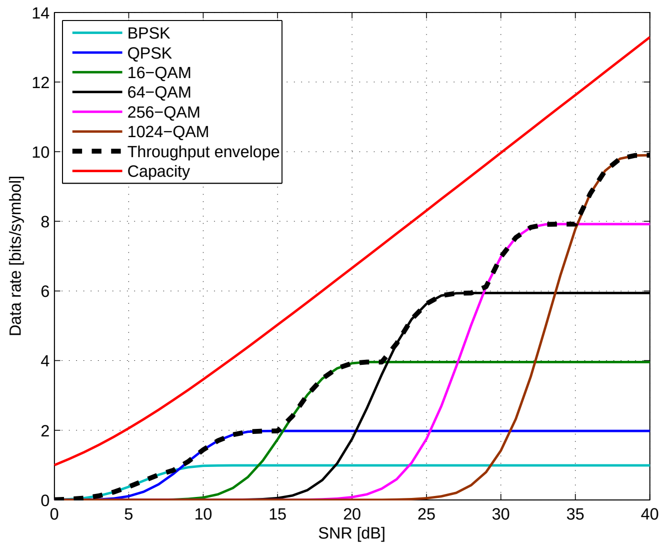

Fig. 3. The throughput and capacity as functions of SNR.

Assuming that the transmitter has an ability to select a modulation scheme that results in the highest throughput for a specific SNR (i.e., tracking the throughput envelope in Figure 3), the SNR gap between the capacity and throughput stays constant as SNR increases. In other words, the rate of capacity and throughput increase with SNR is the same. In the given example the SNR gap is approximately 8 dB.

However, in relative terms, the throughput of this communication scheme is approaching the channel capacity with SNR, i.e., its relative efficiency is improving. For example, based on Figure 3, at 10 dB and then 40 dB, the throughput is respectively reaching approximately 42% and 74% of the channel capacity.

Let us now consider the case when the receiver combines, i.e., adds the corresponding retransmitted frames. In that case, for a frame that is transmitted j times (because of errors detected in its previous $j-1$ transmissions), the SNR is

$$SNR_j = j\, SNR. \tag{15}$$

Using the above, the envelope of the throughput in (4) is depicted in Figure 4. It is compared to the case when the receiver does not combine retransmitted frames, i.e., the throughput envelope in (2). The block length N_b is optimized for each receiver scheme, individually. From the results we note a minor improvement when combining retransmitted frames.

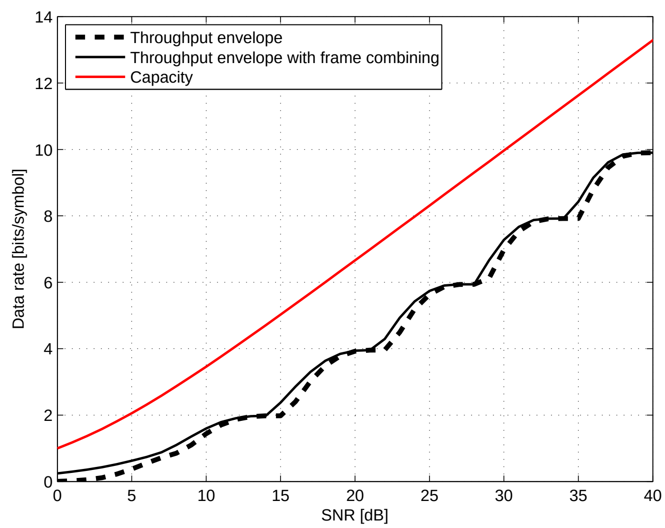

Fig. 4. The throughput envelopes and capacity as functions of SNR.