Информационные процессы, Том 20, № 2, 2020, стр. 127–132.

Simple Asymptotic Bounds on Statistical Decoder Error Rate[1]

Alexey A. Kreshchuk*

Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute)

Received June 11, 2020

Abstract—Statistical decoders are one of the most robust decoders for positional modulations like FSK and PPM. As we show in this work they are applicable to any (unknown) channels that have non-zero distance between received signals. This makes it possible to use statistical decoders in NOMA random access communication systems with bad Channel State Information. In this work we consider the problem of data transmission over unknown memoryless channels. To the author's knowledge this problem was not studied in literature till now. We propose repetition Kautz-Singleton codes and statistical decoders as a solution to this problem.
To estimate the performance of the proposed solution we propose a lower and an upper asymptotic bounds on error rate for statistical decoder. These bounds are evaluated for Kolmogorov-Smirnov goodness-of-fit criteria and compared to a computer simulation. The lower bound seems to be close to the simulation result. The upper bound is not that close to the simulation result but it still holds. To the author's knowledge this is the first technique to derive bounds on error rate for any distance-based statistical decoder.
These bounds for the Kolmogorov-Smirnov goodness-of-fit criterion also show that the error rate should be inversely proportional to the square root of code distance. Other goodness-of-fit criteria might yield asymptotically better results.

KEYWORDS: coded FSK, goodness-of-fit criterion, unknown channel, Kautz-Singleton code.

1. INTRODUCTION

Special decoders for coded FSK (Frequency Shift Keying) constructions intended have being developed for a long time (one of the first might be [1]). In 2008 Zyablov and Osipov [2] have described a multiple access system with each user transmitting an FSK symbol with randomly selected frequencies under the name DHA FH OFDMA (Dynamic Hopset Allocation Frequency Hopping Orthogonal Frequency Division Multiple Access). No coding was defined in that paper though. In 2012 upper bounds on erroneous decoding [3] and denial [4] probabilities of coded DHA FH OFDMA in presence of multitone jamming.

The first statistical decoder was proposed in [1] although it didn't mention it was in fact using Mann-Whitney U test as a goodness-of-fit criteria. However, it had bounds on error rate for this specific statistical decoder.

The general framework of statistical decoders was first presented in 2012 in [5,6]. Two goodness-of-fit criteria were used in these works: Kolmogorov-Smirnov criterion and Mann-Whitney U test (also known as Wilcoxon rank-sum test). These results were later expanded and published in [7]. Statistical decoders based on other two-sample goodness-of-fit criteria were studied in [8]. Statistical decoders based on one-sample goodness-of-fit criteria were introduced in [9]. All these works have the same decoder scheme but different goodness-of-fit criteria.

[1] The reported study was funded by RFBR according to the research projects № 18–37–00322 and № 18–07–01409.

In [10, 11] a reliability metric was introduced for statistical decoder based on Mann-Whitney U test, but the likelihood values were not deduced. A simplified metric was proposed in [12].

In this work we will be studying a slightly different statistical decoders that were described earlier in [13]. These decoders are much easier to reason about as they can output correct likelihoods. They can also be used in binary channels unlike the ones described earlier.

We must also mention other kinds of universal decoders to show the difference between them and the one under consideration. There are a lot of works on universal decoding [14, 15, 16] where decoders with error correction performance close to maximum likelihood decoder are proposed. Unfortunately all of them deal with parametric channels, i.e. the set of all channels (or channel parameters) is countable. In this work the set of all channels is uncountable, i.e. we include all channels for which an inequality (similar to nonzero capacity) holds.

Another similar problem is the problem of mismatched decoding [17]. The main difference in the problem statement is that in mismatched decoding encoder knows the channel model and can select the codebook based on this knowledge. In this work we use a fixed codebook (parameterized by its length) independent on the channel model.

2. CHANNEL MODEL

Let us define a general channel model:

$$x \to y : \{0,1\} \to [0,1]$$

$$\Pr\{y_t < y | x_t = 1\} = F(y) \qquad \Pr\{y_t < y | x_t = 0\} = y$$

where y_t is channel output at time t, x_t is channel input at time t, $F(y)$ is a cumulative distribution function (c.d.f.). All received symbols are statistically independent. $F(y)$ is not known either at the receiver or the transmitter. In this system the conventional decoding doesn't apply as we cannot compute

$$\mathrm{LLR} = \log \frac{\Pr\{y|x=1\}}{\Pr\{y|x=0\}}. \tag{1}$$

So alternative decoders such as statistical decoders are necessary.

In this work we assume that the transmitted symbols are coded and that only two codewords are possible: $\mathbf{x}_0 = \{0,1,0,1,\ldots,0,1\}$ and $\mathbf{x}_1 = \{1,0,1,0,\ldots,1,0\}$. Let us denote the length of these codewords by $2T$. Therefore, each codeword has T zeros and T ones and Hamming distance between them is $2T$. These two codewords can be regarded as Kautz-Singleton codewords with an outer repetition code.

Remark 1. This channel is based on the generalized channel mentioned in remark in [13]. In this work we give that channel a mathematical description. The difference is that in [13] receiver has measured $\Pr\{y|x=0\}$ with the help of pilot symbols while in this work we suppose that the receiver has perfect knowledge of this probability.

3. STATISTICAL DECODER

Let $\mathbf{y}_j = \{y_{2t+j}\}_{t=1}^{T}$. Then instead of computing (1) we will compute

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \Pr\left\{ |g(\mathbf{y}_j) - g(\xi)| < \epsilon \middle| \xi \sim \mathbb{U}_{[0,1]}^{T} \right\} \tag{2}$$

for some $g(y)$. $\mathbb{U}_{[0,1]}$ denotes a uniform distribution on $[0,1]$. In case of statistical decoders we choose $g(y)$ to be a statistic of a goodness-of-fit criterion. For most criteria $g\left(\{y_{2t+j}\}_{t=1}^{T}\right)$ tends to infinity for $\mathbf{x} \neq \mathbf{x}_j$. Let us define the statistical decoder output LLRs as

$$\mathrm{LLR} = \lim_{\epsilon \to 0} \log \frac{\Pr\left\{|g(\xi) - g(\mathbf{y}_1)| < \epsilon \middle| \mathbf{x} = \mathbf{x}_1,\, \xi \sim \mathbb{U}_{[0,1]}^T\right\}}{\Pr\left\{|g(\xi) - g(\mathbf{y}_0)| < \epsilon \middle| \mathbf{x} = \mathbf{x}_0,\, \xi \sim \mathbb{U}_{[0,1]}^T\right\}}$$

This value is known to decoder as the distribution of a goodness-of-fit statistic converges to a known distribution when sample size (code length) goes to infinity.

To compute error rate of this decoder rate we have to compute

$$P_{err} = \Pr\left\{g(\mathbf{y}_0) > g(\mathbf{y}_1) \middle| \mathbf{x} = \mathbf{x}_1\right\}. \tag{3}$$

Unfortunately the distribution of goodness-of-fit statistic in the mismatched case is not widely researched in literature. Therefore, we propose bounds for this distribution for one class of criteria: ones based on a distance between c.d.f.

Remark 2. It is worth noting the symmetries of this decoder. Most of goodness-of-fit criteria don't change if any strictly monotonic transform is applied to their input. But using more popular Kullback-Leibler divergence (in coding theory mostly known as a product of likelihoods) breaks this symmetry for continuous distributions. The ratio of these divergences also breaks it.

4. DISTANCE-BASED CRITERIA

Let us define empirical c.d.f. as

$$\tilde{F}_i(y) = \frac{1}{2T}\left|\{t : y_{2t+i} < y\}\right|.$$

In this work we only consider goodness-of-fit criteria with statistic of form

$$g\left(\mathbf{y}_i\right) = \|\tilde{F}_i(y) - y\|,$$

where $\|F(y)\|$ is some norm of function $F(y)$. We will often skip the argument y in later equations. This c.d.f. is only valid for large samples, although there are some good estimates for samples of small size [18].

Now we can rewrite (3) as

$$\begin{aligned} P_{err} &= \Pr\left\{\|\tilde{F}_0(y) - y\| > \|\tilde{F}_1(y) - y\|\right\} \\ &= \Pr\left\{\|\tilde{F}_0 - y\| > \|\tilde{F}_1 - F + F - y\|\right\} \end{aligned} \tag{4}$$

Let us derive lower bound from (4) using a simple inequality:

$$\Big|\, \|x\| - \|y\| \,\Big| \leq \|x + y\| \leq \|x\| + \|y\|. \tag{5}$$

$$\begin{aligned} P_{err} &\geq \Pr\left\{\|\tilde{F}_0 - y\| > \|\tilde{F}_1 - F\| + \|F - y\|\right\} \\ &= \Pr\left\{\xi_1 - \xi_0 > \|F(y) - y\|\right\}, \end{aligned} \tag{6}$$

where ξ_i are i.i.d. random variables with distribution defined by the goodness-of-fit criterion used (for large samples). This bound can be easily computed for any channel.

$$P_{err} \geq 1 - \int_0^{+\infty} p_\xi(\xi) F_\xi\left(\|F(y) - y\| + \xi\right) d\xi,$$

The bound (6) also gives a metric of channel that defines error rate of statistical decoder: $\|F(y)-y\|$. It depends on the function norm and therefore the goodness-of-fit criterion, but it doesn't depend on the code length and distance. The part of this bound defines how error rate decreases with length: in the same fashion as ξ. For example, for Kolmogorov-Smirnov criterion $F_\xi(x) = \sqrt{n}D(x)$ for sufficiently large n, where $D(x)$ doesn't depend on sample size n [19]. Therefore, for Kolmogorov-Smirnov statistical decoder $P_{err} = \Omega(T^{-0.5})$.

Now let us use the other part of (5) to derive the upper bound:

$$\begin{aligned} P_{err} &\le \Pr\left\{\|\tilde{F}_0 - y\| > \left| \|\tilde{F}_1 - F\| + \|F - y\| \right| \right\} \\ &= \Pr\left\{\xi_0 > \xi_1 - \|F-y\|, \xi_0 > -\xi_1 + \|F-y\|\right\} \\ &= \Pr\left\{\xi_1 - \xi_0 < \|F-y\|, \xi_0 + \xi_1 > \|F-y\|\right\} \\ &= \Pr\left\{\|F-y\| - \xi_0 < \xi_1 < \|F-y\| + \xi_0\right\} \end{aligned}$$

This probability can be computed as

$$P_{err} \le \int_0^{+\infty} p_\xi(\xi_0)\Big(F_\xi(d+\xi_0) - F_\xi(d-\xi_0)\Big)d\xi_0,$$

where $d = \|F-y\|$ and $p_\xi(x) = \frac{d}{dx}F_\xi(x)$ is the p.d.f. defined by the chosen goodness-of-fit criterion. Most of the conclusions for the lower bound also apply for the upper bound. Therefore, for Kolmogorov-Smirnov statistical decoder $P_{err} = O(T^{-0.5})$. Combining it with the lower bound we get $P_{err} = \Theta(T^{-0.5})$.

5. SIMULATION

To test the proposed bound a computer simulation was made. The parameters were selected from [13]. As it is impossible to study the decoder performance for a non-parametric class of channels we have limited our simulation to a single class of channels. Let describe the channel model

$$y_t = \left| x_t\alpha_t + (\delta_t\sqrt{P_{\text{strong}}} + \sqrt{P_{\text{weak}}})\eta_t \right|^2,$$

where α_t, η_t are i.i.d. complex standard normal variables $\mathcal{CN}(0,1)$; P_{strong}, P_{weak} are the powers of interference and noise respectively and $\delta_t \sim B(1,p)$ is a Bernoulli random variable. This channel approximates DHA FH OFDMA channel [2] with Rayleigh fading with collision probability p. Channel input x_t were described earlier. The following parameter values were chosen:

- $p = 0.5$, $T = 16$,
- $P_{\text{strong}} = -20$ dB,
- Kolmogorov-Smirnov criterion was used in decoder.

Kolmogorov-Smirnov criterion is defined as follows:

$$D_T = \sup|\tilde{F}(y) - y|$$

$$\Pr\left\{\frac{(6TD_T+1)^2}{18T} < \xi\right\} = K\left(\sqrt{\frac{\xi}{2}}\right) + O\left(\frac{1}{T}\right)$$

$$K(\xi) = \sum_{k=-\infty}^{\infty}(-1)^k e^{-2k^2\xi^2}$$

From the second equation we can conclude that the error rate should be inversely proportional to the square root of length (or the square root of Hamming distance if we consider codes other than the repetition codes). Though this result might seem very weak it is the only bound for unknown memoryless channels without feedback.

Figure 1 shows the simulation result. The lower bound is very close to the simulation results, but the upper bound is not. Nevertheless, both bounds hold.

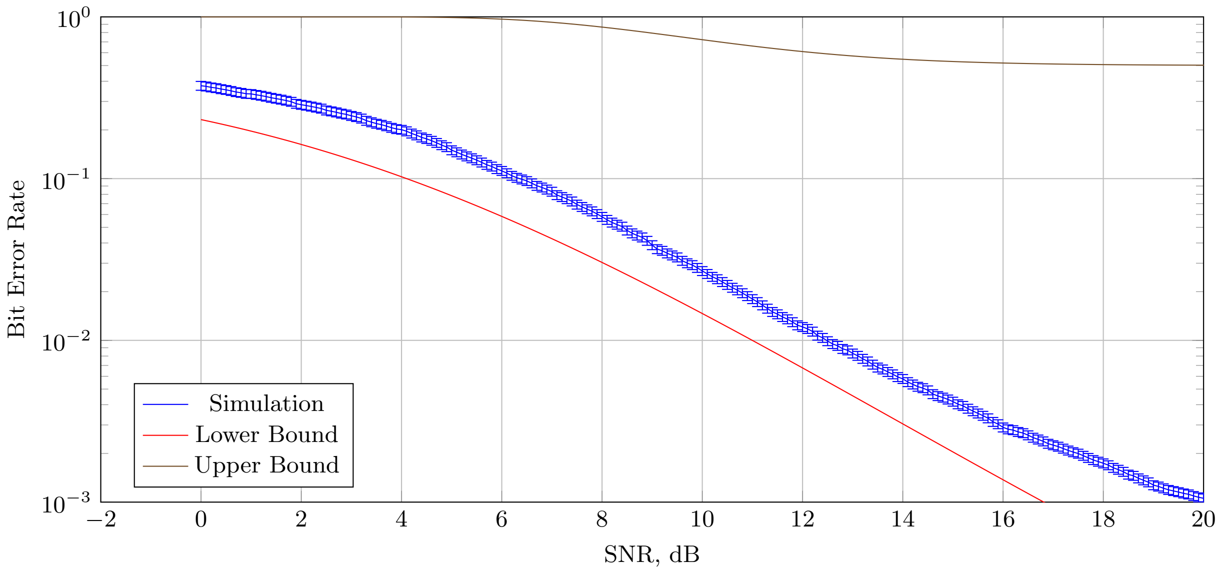

Figure 1. Error Rate and Bounds.

6. FUTURE WORK

The proposed bounds could be improved for the decoder based on the Kolmogorov-Smirnov criterion by using the distribution of one-sided difference between c.d.f.

Generalization to non-binary case and non-repetition codes requires careful handling of dependent random variables but should be possible with manipulations of function norms similar to the ones presented in this work.

7. CONCLUSION

In this paper we consider data transmission over unknown memoryless channels without feedback. We are not aware of any other papers on this problem so this work lacks any comparisons with other bounds. We propose repetition Kautz-Singleton codes and statistical decoders and study their error correction performance as a function of distance between transition probability functions and distance between codewords.

In this work we propose lower and upper asymptotic bounds for statistical decoder and binary repetition code. Computer simulation to test these bounds was performed for Kolmogorov-Smirnov criterion. It shows that the lower bound is close to actual code error rate. The upper bound just holds.

REFERENCES

1. R. Viswanathan and S. Gupta, "Nonparametric receiver for FH-MFSK mobile radio," *IEEE Transactions on Communications*, vol. 33, no. 2, pp. 178–184, 1985.

2. V. Zyablov and D. Osipov, “On the optimum choice of a threshold in a frequency hopping OFDMA system,” *Problems of Information Transmission*, vol. 44, pp. 91–98, 2008.

3. D. Osipov, “A coded DHA FH OFDMA system with a noncoherent ML detector under multitone jamming,” in *Multiple Access Communications*, pp. 37–48, Springer Berlin Heidelberg, 2012.

4. D. Osipov, “Upper bounds for a coded DHA FH OFDMA system with noncoherent threshold reception,” in *2012 XIII International Symposium on Problems of Redundancy in Information and Control Systems*, IEEE, 9 2012.

5. A. Kreshchuk and V. Potapov, “New coded modulation for the frequency hoping OFDMA system,” in *Proc. of the Thirteenth International Workshop on Algebraic and Combinatorial Coding Theory, Pomorie, Bulgaria, June*, pp. 209–212, 2012.

6. K. Kondrashov and V. Afanassiev, “Ordered statistics decoding for semi-orthogonal linear block codes over random non-gaussian channels,” in *Proc. of the Thirteenth International Workshop on Algebraic and Combinatorial Coding Theory, Pomorie, Bulgaria, June*, pp. 15–21, 2012.

7. A. A. Kreshchuk and V. G. Potapov, “Statistical demodulators for frequency shift keying with fast frequency hopping,” *Autom Remote Control*, vol. 74, pp. 1688–1695, 10 2013.

8. A. Kreshchuk and V. Potapov, “Better goodness-of-fit statistics for coded FSK decoding,” *Electronic Notes in Discrete Mathematics*, vol. 57, pp. 139–145, 3 2017.

9. A. Kreshchuk and V. Potapov, “On applying one-sample goodness-of-fit statistics to coded FSK decoding,” in *2016 XV International Symposium Problems of Redundancy in Information and Control Systems (REDUNDANCY)*, IEEE, 9 2016.

10. D. Osipov, “Inner convolutional codes and ordered statistics decoding in a multiple access system enabling wireless coexistence,” in *Multiple Access Communcations*, pp. 33–38, Springer International Publishing, 2013.

11. D. Osipov, “Order statistics-based decoder with reliability information in a multiple access system under interference,” in *Fourteenth International Workshop on Algebraic and Combinatorial Coding Theory ACCT2014*, pp. 283–288, 9 2014.

12. D. Osipov, “Reduced-complexity robust detector in a DHA FH OFDMA system under mixed interference,” in *Multiple Access Communications*, pp. 29–34, Springer International Publishing, 2014.

13. A. Kreshchuk, “Soft-decision statistical decoder for coded DHA FH OFDMA,” in *2018 Engineering and Telecommunication (EnT-MIPT)*, IEEE, 11 2018.

14. J. Ziv, “Universal decoding for finite-state channels,” *IEEE Transactions on Information Theory*, vol. 31, pp. 453–460, 7 1985.

15. M. Feder and A. Lapidoth, “Universal decoding for channels with memory,” *IEEE Transactions on Information Theory*, vol. 44, no. 5, pp. 1726–1745, 1998.

16. N. Merhav, “Universal decoding for memoryless gaussian channels with a deterministic interference,” *IEEE Transactions on Information Theory*, vol. 39, pp. 1261–1269, 7 1993.

17. I. Telatar, A. Lapidoth, and A. Ganti, “Mismatched decoding revisited: general alphabets, channels with memory, and the wide-band limit,” *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2315–2328, 2000.

18. A. Kobzar, *Applied mathematical statistics. For engineers and science workers.* PhysMatLit, 2012. [In Russian].

19. L. Bolshev and N. Smirnov, *Tables of mathematical statistics.* Nauka, Moscow, 1983. [In Russian].