Sphere packing bound for constant composition codes

Nir Elkayam and Meir Feder
Department of Electrical Engineering - Systems
Tel-Aviv University, Israel
Email: nirelkayam@post.tau.ac.il, meir@eng.tau.ac.il

I. INTRODUCTION

General achievable and converse bounds were derived in [1] for general channels and mismatched decoding. These bounds were derived through the use of random coding, and connections between the error probability and error exponent, and the CDF of the pairwise error probability was given. After optimizing over the prior distribution, a converse bound was given, which coincides with the minimax-meta converse of Polyanskiy et al. [2, Theorem 27]. Moreover, the binary hypothesis argument for proving a converse theorem, which always starts with guessing an output distribution, is avoided as the optimal output distribution arises naturally in the calculations. This reveals a deep connection between the pairwise error probability of a general random coding and the best codes possible.

Choosing the optimal prior was addressed by Polyanskiy [3] and applies here also. For channels with high symmetry, the optimal prior is given by uniform distribution over the possible inputs, e.g., constant composition codes over Discrete Memoryless Channels (DMC) and uniform distribution on the power sphere for Additive white Gaussian noise (AWGN).

In this note we focus on constant composition code over DMC and we prove a refined version of the sphere packing bound [4]. It seems that the recent refining by Altug and Wagner [5] is tighter, but the analysis provided here can be the starting point for finer results by using a stronger method for evaluation of the pairwise error probability than the type-based method we used here. Moreover, our analysis is valid for any block length and the constants are well defined.

II. NOTATION

This paper uses bold lower case letters (e.g., \(x\)) to denote a particular value of the corresponding random variable denoted in capital letters (e.g., \(X\)). Calligraphic fonts (e.g., \(X\)) represent a set. Throughout this paper \(\log\) will be defined to base \(e\) and rates are expressed in nats. \(\Pr\{A\}\) will denote the probability of the event \(A\). \(\mathbb{E}_{Q(x)}(\cdot)\) denotes expectation with respect to the distribution \(Q(x)\). We assume the reader is familiar with the method of types and follow the standard notations, e.g., [6].

III. PRELIMINARIES FROM [7]

In the random coding setting the pairwise error probability is crucial for the evaluation of the random coding performance. By carefully handling ties, the exact pairwise error probability of any metric based decoder, given that \(x\) was transmitted and \(y\) was received, is:

\[
p_{e,m,x,y} = Q(m(X,y) > m(x,y)) + U \cdot Q(m(X,y) = m(x,y))
\]

where \(Q(x)\) is the prior distribution, \(m : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}\) is the decoding metric, and \(U\) is a random variable uniformly distributed over \([0,1]\). Notice that the pairwise error probability is a random variable that corresponds to the tie-breaking strategy.

The Cumulative Distribution Function (CDF) of the R.V. \(-\log(p_{e,m,x,y})\) was denoted by \(F(R)\), i.e.:

\[
F(R) = \Pr\{-\log(p_{e,m,x,y}) \leq R\}
\]

where the probability is with respect to \(Q(x)W(y|x)\), i.e. the channel input and output distribution. In the matched case, i.e. when the decoding metric \(m(x,y) = W(y|x)\) it is shown in [7, Theorem 6] that:

\[
F(R) = \sup_{Q(y)} \beta_1 - e^{-R} (Q(x)Q(y), Q(x)W(y|x))
\]

With [3] the minimax meta-converse by Polyanskiy et al. [2, Theorem 27] takes the form:

\[
\epsilon \geq \inf_{Q(x)} F(R).
\]
IV. SUMMARY OF THIS NOTE

In this note we focus on constant composition code over DMC and evaluate the bound \[ (5) \) using the method of types \[ (6) \). This evaluation results in a finite length (i.e., valid for all \( n \)) sphere packing bound which is tighter asymptotically than \[ (5) \).

Let \( Q_n(x) \) be the uniform distribution over the type class \( T_{p_n} = \mathcal{X}^n \) and \( W_n(y|x) = W^n(y|x) \) denote the \( n \) replicas of the memoryless channel \( W(b|a) \). Notice that \( x, y \) are \( n \)-length vectors in \( \mathcal{X}^n \) and \( \mathcal{Y}^n \), respectively.

From [2] Theorem 21 we know that \( Q_n(x) \) achieves the minimum in \( (3) \) when the minimization is over distributions on the type class \( T_{p_n} \). Thus we deduce that for any rate \( R \)-code of length \( n \) with composition \( P_n \):

\[
\epsilon \geq F_n(nR)
\]

where \( F_n(nR) \) is \( \Pr \{ -\log(p_{e,m,x,y}) \leq nR \} \) and the prior \( Q_n(x) \) is used in the evaluation. We next lower bound \( F_n(nR) \).

We first want to evaluate \( Q_n(m(x,y) \geq m(x,y)) \) instead of \( p_{e,m,x,y} \). By conditioning on \( U \geq \frac{1}{2} \):

\[
p_{e,m,x,y} = Q_n(m(x,y) > m(x,y)) + U \cdot Q_n(m(x,y) = m(x,y))
\]

and it follows that:

\[
F_n(nR) \geq \frac{1}{2} \Pr \{ -\log(Q_n(m(x,y) \geq m(x,y))) \leq nR - \log(2) \}
\]

By using the method of types we can further upper bound \( -\log(Q_n(m(x,y) \geq m(x,y))) \) by \( n \cdot I(x,y) + 2n \cdot \lambda_n \), where \( \lambda_n = \frac{\log(n+1)}{n} \). Plugging into \( (6) \) and dividing by \( n \), we get:

\[
F_n(nR) \geq \frac{1}{2} \Pr \left\{ I(x,y) \leq R - 2 \cdot \lambda_n - \frac{\log(2)}{n} \right\}.
\]

Again, by using the method of types we have:

\[
\Pr \{ I(x,y) \leq z \} \geq \max_{V \in \mathcal{P}(\mathcal{X})} \min_{f(x,v)} \frac{2^{-n(D(V||W|P_x)+\lambda_n)}}{I(x,v) \leq z}
\]

By a continuity argument, see \[ (6) \] Lemma 1.2.7, we can take the maximum in \( (8) \) on the entire set of conditional distributions \( \mathcal{P}(\mathcal{Y}|\mathcal{X}) \) at the expense of an extra \( \lambda_n \) at the exponent, i.e.

\[
\Pr \{ I(x,y) \leq z \} \geq \max_{V \in \mathcal{P}(\mathcal{X})} \frac{2^{-n(D(V||W|P_x)+2\lambda_n)}}{I(x,v) \leq z}
\]

where:

\[
E(W, P_x, z) = \min_{V \in \mathcal{P}(\mathcal{Y}|\mathcal{X})} D(V||W|P_x).
\]

Plugging \( (9) \) into \( (7) \) we get:

\[
\epsilon \geq F_n(nR) \geq \frac{1}{2} e^{-n(E(W, P_x, z) - 2 \lambda_n - \frac{\log(2)}{n}) + 2\lambda_n}
\]

which is the desired lower bound on the error probability. The bound is valid for any \( n \). Note that the dependence of both exponential corrections is of order \( \frac{\log(n+1)}{n} \), whereas in the original sphere packing bound outer correction is of order \( \frac{1}{\sqrt{n}} \).

REFERENCES


