Bounds on the ϵ-covering radius of linear codes with applications to self-noise in nested Wyner-Ziv coding

Uri Erez and Ram Zamir

April 2002

Abstract

This report shows that the "self-noise" of a good linear source code under Hamming distortion is not much worse (with respect to a good linear channel code over BSC) than a Bernoulli noise. This proves that the random ensemble of linear nested codes is good. Such codes can be used, e.g., for structured binning for the binary Wyner-Ziv problem (rate-distortion with side information at the decoder).

Let $\mathcal{C}$ be an n-dimensional parity check code with decision cell Ω_0 (see [1]). We now show that if $\mathcal{C}$ is a "good source code" in the sense of [1, sec. II.B], a uniform probability distribution on Ω_0 is "not far" from an i.i.d. Bernoulli process.

Let $\mathcal{B}(r)$ denote a Hamming ball of radius r, i.e., $\mathcal{B}(r) = \{\mathbf{x} : w(\mathbf{x}) \leq r\}$. Further, refer to any set containing all point of weight less than r, no points of weight greater than r, and some points of weight r as a wide sense Hamming ball of radius r. In other words, a wide sense Hamming ball of radius r is the union of a Hamming ball of radius $r-1$ with a not necessarily full shell of radius r. Let d_ϵ be the smallest integer such that

$$|\mathcal{B}(d_\epsilon + 1) \cap \Omega_0| > \lceil (1-\epsilon)|\Omega_0| \rceil. \tag{A.1}$$

Lemma 1 *Let $\mathcal{C}^{(n)}$ be any sequence of δ-good codes for source coding. Then for any $0 < \epsilon < 1$, $\lim_{n\to\infty} d_\epsilon / n = \delta$.*

The lemma says, basically, that most of the volume of the decision cell of a δ-good source code is contained inside a Hamming ball of radius $n\delta$. Hence the probability that a noise vector uniformly distributed over the cell will exceed this radius goes to zero. This implies, along the lines of the analysis in the lattice-code / quadratic-Gaussian case [1], that the effect of this noise (alone, or as the self-noise component in a mixture with a Bernulli noise), on a random BSC-code ensemble, would not be much worse than that of a (pure) Bernulli noise with parameter δ.

Proof: Let d_l denote the "effective radius" of Ω_0, that is, the "radius" of a wide sense Hamming ball having the same cardinality as Ω_0. Otherwise stated, d_l is the largest integer such that

$$|\mathcal{B}(d_l - 1)| \leq |\Omega_0|. \tag{A.2}$$

Likewise, let d_l^ϵ be the largest integer such that

$$|\mathcal{B}(d_l^\epsilon - 1)| \leq \lceil (1-\epsilon)|\Omega_0| \rceil. \tag{A.3}$$

Partition Ω_0 into two disjoint sets, Ω_{in} and $\Omega_{\text{out}} = \Omega_0 \setminus \Omega_{in}$, as follows. Ω_{in} consists of all $\mathbf{x} \in \Omega_0$ such that $w_H(\mathbf{x}) < nd_\epsilon$ and no $\mathbf{x}$ with $w_H(\mathbf{x}) > nd_\epsilon$. Further, the partition is such that $|\Omega_{\text{in}}| = \lceil (1-\epsilon)|\Omega_0| \rceil$ and consequently $|\Omega_{\text{out}}| = \lfloor \epsilon |\Omega_0| \rfloor$. Write the average distortion associated with $\mathcal{C}$ as

$$\frac{1}{n} E\{w_H(\mathbf{X})\} = \frac{1}{n|\Omega_0|} \Big\{ \sum_{\mathbf{x}\in\Omega_{in}} w_H(\mathbf{x}) + \sum_{\mathbf{x}\in\Omega_{\text{out}}} w_H(\mathbf{x}) \Big\} \tag{A.4}$$

Now since $|\Omega_{\text{in}}| \geq (1-\epsilon)|\Omega_0| \geq |\mathcal{B}(d_l^\epsilon)|$ and since a Hamming ball has the smallest average Hamming weight for a given cardinality (the "isoperimetric" inequality)

$$\sum_{\Omega_{\text{in}}} w_H(\mathbf{x}) \geq \sum_{\mathcal{B}(d_l^\epsilon)} w_H(\mathbf{x}). \tag{A.5}$$

Furthermore,

$$\sum_{\Omega_{\text{out}}} w_H(\mathbf{x}) \geq d_\epsilon \cdot |\Omega_{\text{out}}| \geq (\epsilon \cdot |\Omega_0| - 1) \cdot d_\epsilon. \tag{A.6}$$

Thus we have

$$\begin{aligned}
\frac{1}{n} E\{w_H(\mathbf{X})\} &= \Pr(w_H(\mathbf{X}) \leq d_\epsilon) \cdot \frac{1}{n} E\{w_H(d(\mathbf{X})) | w_H(\mathbf{X}) \leq d_\epsilon\} + && \text{(A.7)}\\
& \quad \Pr(w_H(\mathbf{X}) > d_\epsilon) \cdot \frac{1}{n} E\{w_H(d(\mathbf{X})) | w_H(\mathbf{X}) > d_\epsilon\} && \text{(A.8)}\\
&\geq (1-\epsilon)(d_\epsilon^l/n + o(1)) + (\epsilon + o(1)) \cdot \frac{d_\epsilon}{n} && \text{(A.9)}\\
&= (1-\epsilon)d_\epsilon^l/n + \epsilon \cdot d_\epsilon/n + o(1) && \text{(A.10)}
\end{aligned}$$

By definition of d_ϵ^l, and by the Stirling approximation for the volume of a Hamming ball, we have

$$2^{nh(d_\epsilon^l/n)} = (1-\epsilon) 2^{n[h(d_l/n)+o(1)]}. \tag{A.11}$$

Taking the logarithm and dividing by n this gives

$$h(d_\epsilon^l/n) = \frac{1}{n}\log(1-\epsilon) + h(d_l/n) + o(1). \tag{A.12}$$

Since the first term on the r.h.s. goes to zero as n goes to infinity we have $\lim_{n\to\infty} d_\epsilon^l/d_l = 1$. Since by assumption $\mathcal{C}$ is a δ-good source code, we have that $E\{w_H(\mathbf{X})\} = \delta$ and $\lim_{n\to\infty} d_l/n = \delta$, and consequently

$$\lim_{n\to\infty} d_\epsilon^l/n = \delta. \tag{A.13}$$

Taking into account again that $\lim_{n\to\infty} \frac{1}{n} E\{w_H(\mathbf{X}\} = \delta$, by (A.10) and (A.13) the lemma follows.

References

[1] R. Zamir, S. Shamai and U. Erez Nested linear/lattice codes for sturctured multiterminal binning *to appear in IEEE trans. on info. theory, June 2002.*