Notes on Optimal Multiplicity Matrices for RS Decoding

Jonathan Harel and Robert J. McEliece

December 18, 2003

1 Preliminaries

Throughout, we assume an RS code with paramaters $(n, k = \nu + 1)$. We assume this code consists of codewords whose symbols C are members a finite field $F = GF(q)$. A codeword $\mathcal{C} = (C_1, C_2, ..., C_n)$ is transmitted. We are given a Guruswam-Sudan (GS) decoder, whose input takes the form of a nonnegative integer[1] "multiplicity matrix" M. The aim of our discussion is to optimally map the channel output to a matrix M, such that the probability of decoding error for this particular codeword is minimized. We say there is a decoder error if the causal codeword is not on the decoder's list. We would also like to know the probability of error at this optimum.

We start by representing the channel output as a probability matrix χ with n columns $\{\chi_i\}$ (one for each C_i) and q rows (for each possible value of C_i). The entry $\chi_i(b)$, $b \in F$, is the *a posteriori* probability that the ith symbol in $\mathcal{C}$ is b. We can alternatively think of this matrix as a received random vector $\mathbf{X} = (X_1, X_2, ..., X_n)$, where the distribution on each X_i is given by the column χ_i, i.e.

$$\Pr\{X_i = b\} = \chi_i(b).$$

Our desired multiplicity matrix M, as χ, has a column M_i for each C_i and a row b for every element in F. We define the "cost" of M as:

$$C(M) \triangleq \frac{1}{2} \sum_{i=1}^{n} \sum_{b \in F} M_i(b) \cdot (M_i(b) + 1),$$

and the "score" of M with respect to $\mathcal{C}$ as:

$$S(\mathcal{C}, M) \triangleq \sum_{i=1}^{n} M_i(C_i)$$

Lastly, we assume that the GS decoder succeeds iff[2]

$$S(\mathcal{C}, M) > \sqrt{2\nu C(M)}.$$

[1] I use the convention that an "S matrix", where S is a set (e.g., integers, real numbers), is simply a matrix whose elements are all members of S.

[2] In fact, this is only an approximation to the sufficient condition for large cost.

2 Searching for optimal M

Alas, $S(\mathcal{C}, M)$ is not known since $\mathcal{C}$ itself is unknown, and we cannot know with certainty whether the decoder will succeed for a given matrix M. Thus, we define a new "score" of the received vector $\mathbf{X}$ with respect to M as:

$$S(\mathbf{X}, M) \triangleq \sum_{i=1}^{n} M_i(X_i).$$

(Often, we assume the probability matrix χ and multiplicity matrix M are known, so we suppress the dependence on $\mathbf{X}$ and M, referring to the score as simply S.) The quantity $M_i(X_i)$ is a discrete random variable with distribution

$$\Pr\{M_i(X_i) = M_i(b)\} = \Pr\{X_i = b\} = \chi_i(b), \text{ for each } b \in F,$$

and so our new "score" $S = S(\mathbf{X}, M)$ is also a random quantity, with distribution

$$\begin{aligned} \Pr\{S = v\} &= \sum_{\mathcal{X}_v} \prod_{i=1}^{n} \chi_i(x_i) \\ \textit{where } \mathcal{X}_v &\triangleq \{X = (x_1, x_2, ..., x_n) \ : \ S(X, M) = v\} \end{aligned}$$

Therefore, we can express the probability of decoder failure as[3]:

$$P_E(\chi, M) \triangleq \Pr\left\{S(\mathbf{X}, M) \leq \sqrt{2\nu C(M)}\right\}.$$

Our first problem is to find an M of cost no greater than γ that minimizes P_E for a given χ. That is, perform the following minimization:

$$\widetilde{M_\gamma} \triangleq \underset{M:C(M)\leq\gamma}{\operatorname{argmin}} \ P_E(\chi, M),$$

and we call the probability of error at this optimum point $P_E(\chi, \gamma)$, i.e.:

$$P_E(\chi, \gamma) \triangleq P_E(\chi, \widetilde{M_\gamma}).$$

The probability of decoder failure for this particular codeword can be no lower, given the posteriori probability matrix χ and cost constraint γ.

If we disregard the cost constraint, we have the lowest possible probability of decoder failure for a particular codeword given by:

$$P_E(\chi) = \lim_{\gamma\to\infty} P_E(\chi, \gamma),$$

[3] To compute $P_E(\chi, M)$, we can resort to the following exhaustive approach:

$$\begin{aligned} P_E(\chi, M) &= \sum_{A} \Pr\{\mathbf{X}\} \\ \text{where } A &= \left\{\mathbf{X} : S(\mathbf{X}, M) \leq \sqrt{2\nu C(M)}\right\} \\ \text{and } \Pr\{\mathbf{X} = (b_1, b_2, ..., b_n)\} &= \prod_{i=1}^{n} \Pr\{X_i = b_i\} \end{aligned}$$

3 Settling for optimal Q

Optimizing over the space of nonnegative integer matrices $M \in \mathcal{M}$ turns out to be a difficult problem. We will instead optimize over a much larger space, that of nonnegative real matrices $Q \in \mathcal{Q}$. Note that $\mathcal{M} \subset \mathcal{Q}$. As with M, we label the columns of Q as $(Q_1, Q_2, ..., Q_n)$. We define a cost function as before:

$$C(Q) = \frac{1}{2}\sum_{i=1}^{n}\sum_{b\in F} Q_i(b)\cdot(Q_i(b)+1),$$

and a score with respect to Q of X:

$$S = S(X,Q) = \sum_{i=1}^{n} Q_i(X_i),$$

where X_i, $Q_i(X_i)$, and S are all random quantities $\forall i$. And we define:

$$\begin{aligned}
P_E(\chi,Q) &\triangleq \Pr\left\{S(\mathbf{X},Q) \le \sqrt{2\nu C(Q)}\right\}\\
\widetilde{Q}_\gamma &\triangleq \underset{Q:C(Q)\le\gamma}{\operatorname{argmin}}\ P_E(\chi,Q)\\
P_E^*(\chi,\gamma) &\triangleq P_E(\chi,\widetilde{Q}_\gamma)\\
P_E^*(\chi) &= \lim_{\gamma\to\infty} P_E^*(\chi,\gamma).
\end{aligned}$$

Because $\mathcal{M} \subset \mathcal{Q}$, we are guaranteed that:

$$\begin{aligned}
P_E^*(\chi,\gamma) &\le P_E(\chi,\gamma)\\
P_E^*(\chi) &\le P_E(\chi).
\end{aligned}$$

One nice property of $\mathcal{Q}$ is that if $Q \in \mathcal{Q}$, then $\lambda Q \in \mathcal{Q}$, where $\lambda \in \mathbb{R}^+ \cup \{0\}$. Thus, for example, we can compare the cost of Q and λQ.

$$\begin{aligned}
C(\lambda Q) &= \frac{1}{2}\sum_{i,b}(\lambda Q_{ib}+1)\lambda Q_{ib}\\
&= \frac{1}{2}\sum_{i,b}(\lambda^2 Q_{ib}+\lambda^2+(\lambda-\lambda^2))Q_{ib}\\
&= \lambda^2 C(Q) + \frac{1}{2}\sum_{i,b}(\lambda-\lambda^2)Q_{ib}.
\end{aligned}$$

For $\lambda \ge 1$, $(\lambda-\lambda^2) < 0$, which implies[4]:

$$C(Q) \le C(\lambda Q) \le \lambda^2 C(Q), \tag{1}$$

with equality iff $\lambda = 1$. We can compare the probability of decoder failure for these two matrices:

$$\begin{aligned}
P_E(\chi,\lambda Q) &= \Pr\{\lambda S \le \sqrt{2\nu C(\lambda Q)}\} \text{ since } S(X,\lambda Q) = \lambda S(X,Q)\\
&= \Pr\{S \le \frac{1}{\lambda}\sqrt{2\nu C(\lambda Q)}\}\\
&= \Pr\{S \le \frac{1}{\lambda}\sqrt{2\nu C(\lambda Q)} \le \sqrt{2\nu C(Q)}\} \text{ by (1)},\\
&\le \Pr\{S \le \sqrt{2\nu C(Q)}\}\\
&= P_E(\chi,Q),
\end{aligned}$$

[4] The first half of the inequality, $C(Q) \le C(\lambda Q)$, is trivial.

which, written another way, is:

$$P_E(\chi, Q/\lambda) \geq P_E(\chi, Q) \text{ for all } \lambda \geq 1.$$

Thus, we are guaranteed that for every Q such that $C(Q) = \gamma$ the set $Q_{<\gamma} \equiv \{Q/\lambda : \lambda > 1\}$, has $P_E(\chi, Q') > P_E(\chi, Q)$ $\forall Q' \in Q_{<\gamma}$. It can be shown that[5]

$$\bigcup_{\{Q:C(Q)=\gamma\}} Q_{<\gamma} = \{Q : C(Q) < \gamma\}.$$

Thus,

$$\min_{C(Q)<\gamma} P_E(\chi, Q) > \min_{C(Q)=\gamma} P_E(\chi, Q),$$

and:

$$\begin{aligned} P_E^*(\chi, \gamma) &= \min_{C(Q)\leq\gamma} P_E(\chi, Q) \\ &= \min_{C(Q)=\gamma} P_E(\chi, Q). \end{aligned}$$

Theorem 1 $P_E(\chi) = P_E^*(\chi)$.

Proof. Theorem 2 implies that for any Q there exists a nonnegative integer matrix $M = f(Q)$ such that $P_E(\chi, M) \leq P_E(\chi, Q)$ (see below). Thus, for any minimizing matrix Q':

$$\begin{aligned} Q' &= \operatorname*{argmin}_{Q:C(Q)=y} \Pr\{S(Q) \leq \sqrt{2\nu y}\}, \\ \text{with } P_E^*(\chi, y) &= P_E(\chi, Q') = \Pr\{S(Q') \leq \sqrt{2\nu y}\}, \end{aligned}$$

there is some $M = f(Q')$ for which

$$P_E(\chi, M) \leq P_E(\chi, Q') = P_E^*(\chi, y).$$

But y is completely arbitrary, so we can never find a y large enough so that $P_E^*(\chi, y) < P_E(\chi, M)$ for all M. This implies

$$P_E^*(\chi) = \lim_{y\to\infty} P_E^*(\chi, y) \geq P_E(\chi) = \lim_{y\to\infty} P_E(\chi, y)$$

However, $\mathcal{M} \subset \mathcal{Q}$, so $P_E^*(\chi) \leq P_E(\chi)$ and $P_E(\chi) = P_E^*(\chi)$. ■

Theorem 2 *For any nonnegative real matrix Q there exists a nonnegative integer matrix $M = f(Q)$ such that $P_E(\chi, M) \leq P_E(\chi, Q)$.*

Proof. Choose

$$\begin{aligned} M &= f(Q) = \lceil AQ \rceil, \text{ where } A \text{ is any scalar such that} \\ A &> \frac{3z + \sqrt{9z^2 + 8znq}}{2z}, \text{ where} \qquad (2) \\ z &= \sum_{i,b} Q_i(b). \qquad (3) \end{aligned}$$

[5] how?

Then,

$$\begin{aligned} S(X,M) &= \sum_i \lceil AQ_i \rceil, \text{ where } Q_i = Q_i(x_i) \\ &\geq \sum_i AQ_i \\ &= A \cdot S(X,Q), \text{ for any } X = (x_1, x_2, ..., x_n) \end{aligned}$$

This implies

$$\Pr\{S(X,M) \leq \sqrt{2\nu C(M)}\} \leq \Pr\{A \cdot S(X,Q) \leq \sqrt{2\nu C(M)}\}, \tag{4}$$

since for any distribution on X, the event $\{S(X,M) \leq \sqrt{2\nu C(M)}\} \Rightarrow \{A \cdot S(X,Q) \leq \sqrt{2\nu C(M)}\}$.

Now, Theorem 3 implies

$$C(M) \leq A^2 C(Q).$$

Therefore,

$$\begin{aligned} & \Pr\{A \cdot S(X,Q) \leq \sqrt{2\nu C(M)}\} \\ = \ & \Pr\{A \cdot S(X,Q) \leq \sqrt{2\nu C(M)} \leq A\sqrt{2\nu C(Q)}\} \\ \leq \ & \Pr\{S(X,Q) \leq \sqrt{2\nu C(Q)}\} \\ = \ & P_E^*(\chi, Q). \end{aligned}$$

Combining this result with equation (4), we have:

$$\begin{aligned} & \Pr\{S(X,M) \leq \sqrt{2\nu C(M)}\} \leq P_E^*(\chi, Q) \\ \Rightarrow \ & P_E(\chi, M) \leq P_E(\chi, Q). \end{aligned}$$

■

Theorem 3 $C(M) \leq A^2 C(Q)$

Proof.

$$\begin{aligned} C(M) &= \frac{1}{2} \sum_{i,b} \lceil AQ_{ib} \rceil \left(\lceil AQ_{ib} \rceil + 1 \right) \\ &\leq \frac{1}{2} \sum_{i,b} (AQ_{ib} + 1)(AQ_{ib} + 2) \\ &= \frac{1}{2} \left(\sum_{i,b} (AQ_{ib} + A)(AQ_{ib} + 2) - (A-1) \sum_{i,b} (AQ_{ib} + 2) \right) \\ &= \frac{1}{2} \left(\sum_{i,b} AQ_{ib}(AQ_{ib} + A) + 2A \sum_{i,b} (Q_{ib} + 1) - (A-1) \sum_{ib} (AQ_{ib} + 2) \right) \\ &= A^2 C(Q) + \frac{1}{2} \left(2A(z + nq) - (A-1)(Az + 2nq) \right), \text{ recall } z = \sum_{i,b} Q_{ib}, \text{ and } \sum_{i,b} 1 = nq \\ &\leq A^2 C(Q) \end{aligned}$$

iff

$$2A(z + nq) - (A-1)(Az + 2nq) \leq 0,$$

which is satisfied when

$$A > \frac{3z + \sqrt{9z^2 + 8znq}}{2z},$$

which is true by assignment (2). ■

4 Defining an equivalent problem: $G(\chi, B)$

Let

$$\begin{aligned} F(\chi, x, y) &= \min_{C:C(Q)=y} \Pr\left\{\sum_i Q_i(X_i) \leq x\right\} \\ G(\chi, B) &= \min_{R:||R||^2=1} \Pr\left\{\sum_i R_i(X_i) \leq B\right\}, \end{aligned}$$

where $R_i(b)$ is a nonnegative real number for all i, b.

$F(\cdot)$ relates to our previous minimization problems as follows (substitution):

$$\begin{aligned} P_E^*(\chi, y) &= F(\chi, \sqrt{2\nu y}, y) \\ P_E(\gamma) &= \lim_{y \to \infty} F(\chi, \sqrt{2\nu y}, y) \end{aligned}$$

Theorem 4 *$lim_{y\to\infty} F(\chi, B\sqrt{2y}, y) = G(\chi, B)$ for all $B \geq \dot{0}$.*

Proof. Given a matrix Q with $C(Q) = y$, we define a new matrix R of the same dimensions with

$$\begin{aligned} R_i(b) &= K(y)\left(Q_i(b) + \frac{1}{2}\right), \\ \text{where } K(y) &= \frac{1}{\sqrt{2y + nq/4}}. \end{aligned}$$

Then,

$$\sum_i R_i(X_i) = K(y) \sum_i Q_i(X_i) + \frac{1}{2} K(y) n, \text{ for any choice of } X_i\text{'s},$$

and

$$\begin{aligned} ||R||^2 &= \sum_{i,b} R_i(b)^2 \\ &= \sum_{i,b} \frac{\left(Q_i(b) + \frac{1}{2}\right)^2}{2y + nq/4} \\ &= \frac{1}{2y + nq/4} \left(\sum_{i,b} Q_i(b)(Q_i(b) + 1) + nq/4\right) \\ &= \frac{2y + nq/4}{2y + nq/4} \\ &= 1, \text{ for all } y. \end{aligned}$$

Thus

$$\begin{aligned} \Pr\left\{\sum_i R_i(X_i) \leq K(y)(x + n/2)\right\} &= \Pr\left\{K(y)\sum_i Q_i(X_i) + \frac{1}{2}K(y)n \leq K(y)(x + n/2)\right\} \\ &= \Pr\left\{\sum_i Q_i(X_i) \leq x\right\}. \end{aligned}$$

Now, because $C(Q) = y \Rightarrow ||R||^2 = 1$, then

$$G\left(\chi, K(y)(x + \frac{n}{2})\right) \leq F(\chi, x, y),$$

since we are minimizing the same quantity on both sides of the inequality, but the search space on the left is at least as big as that on the right. This immediately implies

$$\begin{aligned} & \lim_{y\to\infty} G\left(\chi, K(y)(B\sqrt{2y} + \frac{n}{2})\right) \leq \lim_{y\to\infty} F(\chi, B\sqrt{2y}, y) \\ \Leftrightarrow \quad & G(\chi, B) \leq \lim_{y\to\infty} F(\chi, B\sqrt{2y}, y). \end{aligned} \tag{5}$$

Conversely, if we are given a matrix R of positive real numbers satisfying $||R||^2 = 1$, we can choose a cost y, and define a new matrix Q such that

$$Q_i(b) = \frac{R_i(b)}{K(y)} - \frac{1}{2},$$

which will give us $C(Q) = y$. Then,

$$\begin{aligned} \Pr\left\{\sum_i R_i \leq B\right\} &= \Pr\left\{K(y)\sum_i Q_i + \frac{n}{2}K(y) \leq B\right\} \\ &= \Pr\left\{\sum_i Q_i \leq B\sqrt{2y + nq/4} - \frac{n}{2}\right\}, \text{.for arbitrary } y. \end{aligned}$$

Thus,

$$\min_{R:||R||^2=1} \Pr\left\{\sum_i R_i \leq B\right\} \geq \lim_{y\to\infty} \min_{Q:C(Q)=y} \Pr\left\{\sum_i Q_i \leq B\sqrt{2y + nq/4} - \frac{n}{2}\right\}$$

since now $||R||^2 = 1 \Rightarrow C(Q) = y$. Then,

$$\begin{aligned} & \min_{R:||R||^2=1} \Pr\left\{\sum_i R_i \leq B\right\} \geq \lim_{y\to\infty} \min_{Q:C(Q)=y} \Pr\left\{\sum_i Q_i \leq B\sqrt{2y}\right\} \\ \Leftrightarrow \quad & G(\chi, B) \geq \lim_{y\to\infty} F(\chi, B\sqrt{2y}, y). \end{aligned}$$

■

The matrix:

$$Q_y = \operatorname*{argmin}_{Q:C(Q)=y} \Pr\left\{\sum_i Q_i(X_i) \leq \sqrt{2\nu y}\right\}$$

represents the optimal real multiplicity matrix, for a cost constraint of y. We could use

$$M_y = round(Q_y)$$

as an integer approximation to Q_y, with cost approximately equal to y, and score approximately equal to $S(Q_y)$. Then

$$\begin{aligned} \Pr\left\{S(M_y) \leq \sqrt{2\nu C(M_y)}\right\} &\approx \Pr\left\{S(Q_y) \leq \sqrt{2\nu C(Q_y)}\right\} \\ &= P_E^*(\chi, y) \\ &= F(\chi, \sqrt{2\nu y}, y). \end{aligned}$$

Let R be the minimizing matrix on the RHS of:

$$G(\chi, \sqrt{\nu}) = \min_{R:||R||^2=1} \Pr\left\{\sum_i R_i(X_i) \leq \sqrt{\nu}\right\},$$

i.e., the optimal real multiplicity matrix for cost tending to infinity. Then the matrix

$$Q_R(y) = R(2y + nq/4)^{1/2} - \frac{1}{2}$$

represents a good approximation for Q_y when $y >> nq$. Finally, the minimum probability of decoder error for our transmitted codeword, given probability matrix χ, and regardless of the cost of the multiplicity matrix, is:

$$P_E(\chi) = G(\chi, \sqrt{\nu}).$$

5 Chernoff Bounds

The quantities:

$$\begin{aligned} F(\chi, \sqrt{2\nu y}, y) &= \min_{Q:C(Q)=y} \Pr\left\{\sum_i Q_i(X_i) \leq \sqrt{2\nu y}\right\} \\ G(\chi, \sqrt{\nu}) &= \min_{R:||R||^2=1} \Pr\left\{\sum_i R_i(X_i) \leq \sqrt{\nu}\right\} \end{aligned}$$

are in general difficult to compute. We circumvent calculation of full probability distributions by using the Chernoff bound. We let

$$\begin{aligned} g_i(s) &= E\left\{\exp(-sQ_i)\right\} \\ &= \sum_{b\in F} \chi_i(b) \exp(-sQ_i(b)) \end{aligned}$$

be the moment generating function for Q_i. Then the moment generating function for $S = \sum_i Q_i(X_i)$ is:

$$\begin{aligned} g_Q(s) &= E\left\{\exp\left(-s\sum_{i=1}^n Q_i\right)\right\} \\ &= E\left\{\prod_{i=1}^n \exp(-sQ_i)\right\} \\ &= \prod_{i=1}^n g_i(s), \text{ by independence assumed in the Koetter-Vardy paradigm.} \end{aligned}$$

The Chernoff bound in this case is

$$\Pr\{S \leq x\} \leq \tilde{K}_Q(\chi, x),$$

where

$$\tilde{K}_Q(\chi, x) \equiv \min_{s\geq 0} g_Q(s)e^{sx},$$

which we will assume, for our purposes, is quite tight[6]. Now, if we define

$$\begin{aligned}\tilde{F}(\chi, x, y) &\equiv \min_{Q:C(Q)=y} \tilde{K}_Q(\chi, x)\\ \tilde{G}(\chi, x) &\equiv \min_{Q:||Q||^2=1} \tilde{K}_Q(\chi, x)\end{aligned}$$

then

$$\begin{aligned}P_E^*(\chi, y) &\leq \tilde{F}(\chi, \sqrt{2\nu y}, y)\\ P_E(\chi) &\leq \tilde{G}(\chi, \sqrt{\nu}).\end{aligned}$$

Thus, we have bounded the probability of error for finite and infinite cost multiplicity matrices.

For finite cost,

$$\tilde{Q}_y = \operatorname*{argmin}_{Q:C(Q)=y} \tilde{K}_Q(\chi, \sqrt{2\nu y})$$

achieves a probability of error close to $P_E^*(\chi, y)$. We set our multiplicity matrix $\tilde{M}_y = round(\tilde{Q}_y)$, as before. The cost of $\tilde{M}_y$ is approximately y, and the score of $\tilde{M}_y$ is approximately $S(\tilde{Q}_y)$, so we hope that:

$$\begin{aligned}\Pr\{S(\tilde{M}_y) \leq \sqrt{2\nu C(\tilde{M}_y)}\} &\approx \Pr\{S(\tilde{Q}_y) \leq \sqrt{2\nu C(\tilde{Q}_y)}\}\\ &= \tilde{F}(\chi, \sqrt{2\nu y}, y)\\ &= P_E^*(\chi, y) + \varepsilon\end{aligned}$$

with $\varepsilon > 0$ small, since the bound is tight.

For infinite cost,

$$\tilde{G}(\chi, \sqrt{\nu}) = \min_{R:||R||^2=1} \tilde{K}_R(\chi, \sqrt{\nu}) = \lim_{y\to\infty} \tilde{F}(\chi, \sqrt{2\nu y}, y)$$

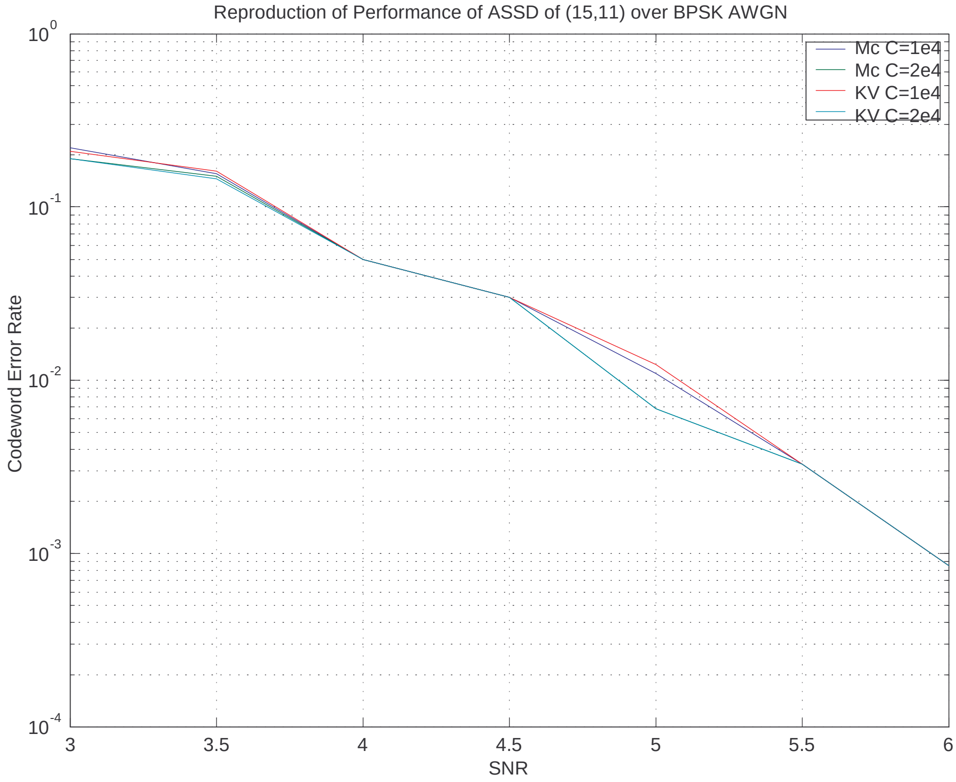
represents an approximation to the probability $P_E(\chi)$, with

$$\tilde{G}(\chi, \sqrt{\nu}) \geq G(\chi, \sqrt{\nu}) = P_E(\chi).$$

6 Experiments

[6] Because, in practice, we will be dealing with $\Pr\{S \leq x\}$ very close to zero.

6.1



References

[1] McEliece, Robert J. "Musings on Multiplicity Matrices", 9 Nov 2003. (unpublished)

[2] M. El-Khamy, J. Harel, and R. J. McEliece. "Performance Limits for Algebraic Soft-Decision Decoding of Reed-Solomon Codes." 1 Dec 2003. Abstract submitted to ISIT 2004.