EE 121 - Introduction to Digital Communications
Homework 7 Solutions

April 27, 2008

1. (a) The total number of bits reliably communicated in T time slots is RT. The total energy expended is ET. The energy per bit is then $ET/RT = E/R$ joules per bit. As the linear code from class was able to communicate reliably at a strictly positive rate $R > 0$, the energy per bit $E/R < \infty$, i.e. we are able to communicate each bit reliably using a bounded amount of energy.

(b) We can communicate reliably with the code from class, at rates $R < R^*$, where

$$R^* = \log_2 \left(\frac{2}{1 + e^{-E/2\sigma^2}} \right).$$

Thus the energy per bit of the code is no more than

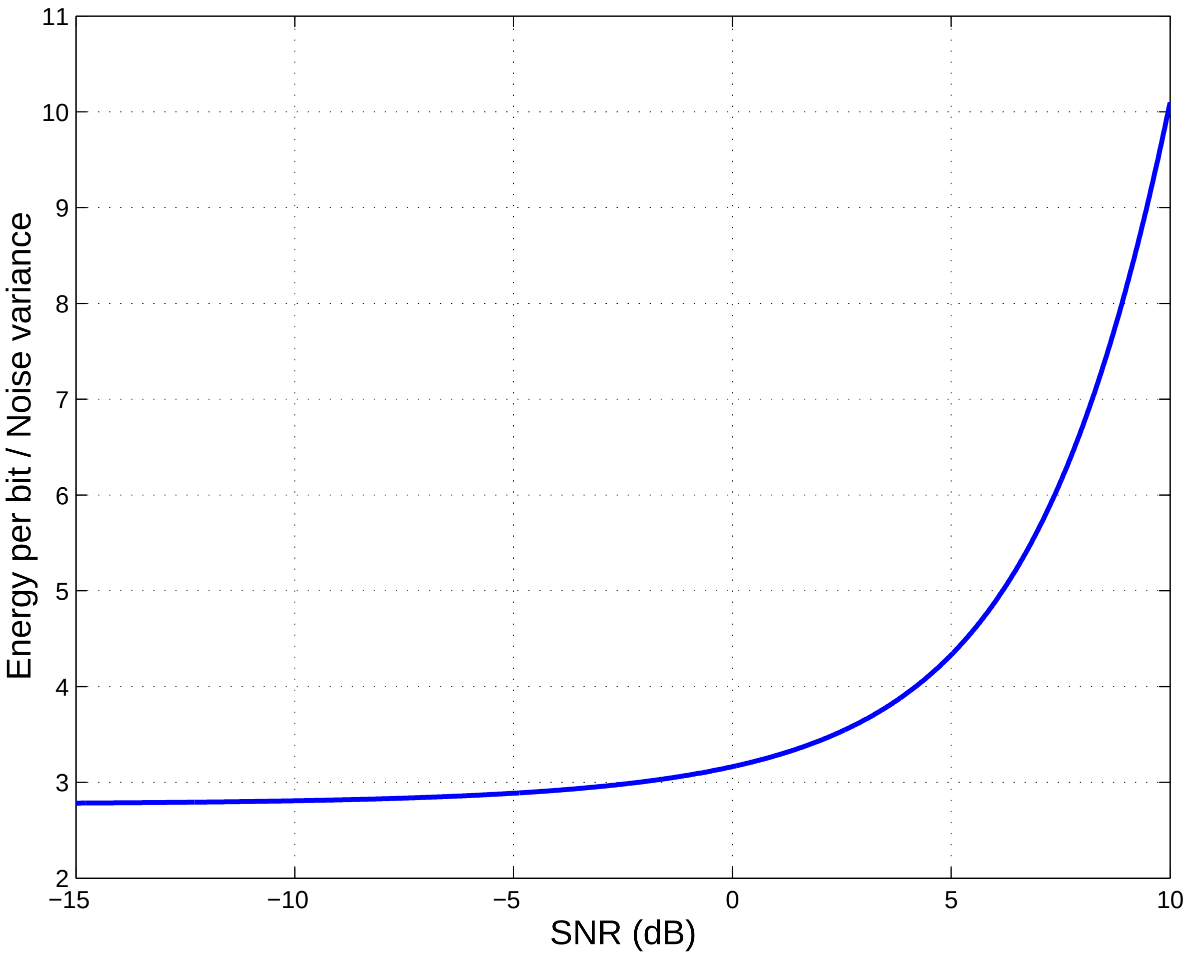
$$\frac{E}{\log_2 \left(\frac{2}{1+e^{-E/2\sigma^2}} \right)}.$$

(c) The ratio of the energy per bit to the noise variance is

$$\frac{E}{\sigma^2 \log_2 \left(\frac{2}{1+e^{-E/2\sigma^2}} \right)} = \frac{\mathsf{SNR}}{\log_2 \left(\frac{2}{1+e^{-\mathsf{SNR}/2}} \right)}.$$

This expression is minimizes for $\mathsf{SNR} = 0$.

(d)



2. (a) Denote the ith column of $\mathbf{G}'$ by $\mathbf{g}_i$. As $\mathbf{G}'$ has 2 columns, it is full rank if its rank is 2. Let $r(\mathbf{G}')$ denote the rank of $\mathbf{G}'$.

$$\begin{aligned}
\Pr(r(\mathbf{G}') = 2) &= \Pr(r([\mathbf{g}_1, \mathbf{g}_2]) = 2) \\
&= \Pr(r([\mathbf{g}_1, \mathbf{g}_2] = 2 | r(\mathbf{g}_1) = 1) \Pr(r(\mathbf{g}_1) = 1) \\
&\quad + \Pr(r([\mathbf{g}_1, \mathbf{g}_2] = 2 | r(\mathbf{g}_1) = 0) \Pr(r(\mathbf{g}_1) = 0) \\
&= \Pr(\mathbf{g}_2 \neq \mathbf{g}_1, \mathbf{g}_2 \neq \mathbf{0} | r(\mathbf{g}_1) = 1) \Pr(\mathbf{g}_1 \neq \mathbf{0}) + 0
\end{aligned}$$

As the entries of $\mathbf{G}'$ are i.i.d. Bernoulli(1/2) random variables, the probability that $\mathbf{g}_1$ is equal to the all zeros vector is 2^{-3}. As there are two ways of $\mathbf{g}_2$ being a linear combination of $\mathbf{g}_1$, given that $\mathbf{g}_1$ is full rank (i.e. $\mathbf{g}_1 \neq \mathbf{0}$), namely, $\mathbf{g}_2 = \mathbf{g}_1$ and $\mathbf{g}_2 = \mathbf{0}$, the probability of $[\mathbf{g}_1, \mathbf{g}_2]$ being full rank given $\mathbf{g}_1$ is full rank, is $2/2^{-3}$. Thus

$$\begin{aligned}
\Pr(r(\mathbf{G}') = 2) &= (1 - 2 \times 2^{-3})(1 - 2^{-3}) \\
&= 21/32
\end{aligned}$$

(b) $\mathbf{G}'$ has $(1 - \epsilon)T$ rows and RT columns. In order for it to be full rank we must have

$R \leq 1-\epsilon$, in which case $\mathbf{G}'$ is full rank if its rank is RT.

$$\begin{aligned}
&\Pr(r(\mathbf{G}') = RT) \\
&\quad = \Pr(r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT}]) = RT) \\
&\quad = \Pr(r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT}]) = RT \mid r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT-1}]) = RT-1) \Pr(r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT-1}]) = RT-1) \\
&\qquad + \Pr(r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT}]) = RT \mid r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT-1}]) < RT-1) \Pr(r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT-1}]) < RT-1) \\
&\quad = \Pr(\mathbf{g}_{RT} \neq [\mathbf{g}_1, \ldots, \mathbf{g}_{RT-1}]\mathbf{a} \text{ for some } \mathbf{a} \in \{0,1\}^{RT-1} | r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT-1}]) = RT-1) \\
&\qquad \times \Pr(r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT-1}]) = RT-1) + 0 \\
&\quad = (1 - 2^{RT-1}/2^{(1-\epsilon)T}) \Pr(r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT-1}]) = RT-1) \\
&\quad = (1 - 2^{(R-1+\epsilon)T-1}) \Pr(r([\mathbf{g}_1, \ldots, \mathbf{g}_{RT-1}]) = RT-1)
\end{aligned}$$

Solving this recursion we get

$$\Pr(r(\mathbf{G}') = RT) = \prod_{j=0}^{RT-1} (1 - 2^{j-(1-\epsilon)T})$$

(c)

$$\begin{aligned}
\Pr(r(\mathbf{G}') < RT) &= 1 - \exp\left(\log_e \prod_{j=0}^{RT-1} (1 - 2^{j-(1-\epsilon)T})\right) \\
&= 1 - \exp\left(\sum_{j=0}^{RT-1} \log_e (1 - 2^{j-(1-\epsilon)T})\right) \\
&\leq - \sum_{j=0}^{RT-1} \log_e (1 - 2^{j-(1-\epsilon)T}) \\
&\leq \sum_{j=0}^{RT-1} 2 \times 2^{j-(1-\epsilon)T} \\
&= 2^{1-(1-\epsilon)T} \sum_{j=0}^{RT-1} 2^j \\
&= 2^{1-(1-\epsilon)T} (2^{RT} - 1) \\
&\leq 2^{1+T(R-(1-\epsilon))}
\end{aligned}$$

where we have used the inequalities from the hint in steps 3 and 4. This expression goes to zero as $T \to \infty$ if $R < 1-\epsilon$.

(d) The probability of more than ϵT erasures occurring is

$$\sum_{k=\epsilon T+1}^{T} \binom{T}{k} p^k (1-p)^{T-k}$$

If $p < \epsilon$, this probability goes to zero as $T \to \infty$ by the law of large numbers.

(e) In part (c) we showed that the probability of decoding error goes to zero as $T \to \infty$ so long as the rate is less than one minus the fraction of erasures that occur. In part (d) we argued that the probability of more than a fraction p of erasures occurring goes to zero as $T \to \infty$. Thus if the rate is less than $1 - p$, reliable communication over the erasure channel is possible.

(f) From the law of large numbers we know that as $T \to \infty$, the fraction of erasures will be very close to p, with high probability. Consequently $\mathbf{G}'$ will have fewer rows than columns with high probability, and will therefore not be full rank if $R > 1 - p$. Thus reliable communication is not possible at rates greater than $1 - p$. Putting this together with the answer to part (e) we can conclude that the *capacity* of the erasure channel is $C = 1 - p$.

3. (a) We first condition the probability of error on the codeword transmitted

$$\Pr(\mathcal{E}) = \sum_{i=1}^{2^{RT}} \Pr(\mathcal{E}|\, \mathbf{u} = \mathbf{u}_i) \Pr(\mathbf{u} = \mathbf{u}_i).$$

Conditioned on transmitting codeword $\mathbf{u}_i$, we make an error if the received vector lies closer to a different codeword $\mathbf{u}_j$

$$\Pr(\mathcal{E}) = \sum_{i=1}^{2^{RT}} \Pr\left(\bigcup_{j=1, j\neq i}^{2^{RT}} \{\mathbf{u}_i \to \mathbf{u}_j\} \middle| \, \mathbf{u} = \mathbf{u}_i \right) \Pr(\mathbf{u} = \mathbf{u}_i).$$

Using the union bound we have

$$\Pr(\mathcal{E}) \leq \sum_{i=1}^{2^{RT}} \sum_{j=1, j\neq i}^{2^{RT}} \Pr\left(\mathbf{u}_i \to \mathbf{u}_j |\, \mathbf{u} = \mathbf{u}_i\right) \Pr(\mathbf{u} = \mathbf{u}_i).$$

(b) We first condition on number of codeword entries in which $\mathbf{u}_i$ and $\mathbf{u}_j$ differ, which we denote $d(\mathbf{u}_i, \mathbf{u}_j)$.

$$\Pr\left(\mathbf{u}_i \to \mathbf{u}_j |\, \mathbf{u} = \mathbf{u}_i\right) = \sum_{l=0}^{T} \Pr\left(\mathbf{u}_i \to \mathbf{u}_j |\, \mathbf{u} = \mathbf{u}_i, d(\mathbf{u}_i, \mathbf{u}_j) = l\right) \Pr(d(\mathbf{u}_i, \mathbf{u}_j) = l).$$

If codewords $\mathbf{u}_i$ and $\mathbf{u}_j$ differ in l entries then they are separated by l dimensions. In each dimension the minimum separation distance in that dimension is $2\sqrt{E}/(M-1)$. Thus the worst case distance between codewords differing in l entries is $d = 2l\sqrt{E}/(M-1)$. Then

$$\begin{aligned} \Pr\left(\mathbf{u}_i \to \mathbf{u}_j |\, \mathbf{u} = \mathbf{u}_i, d(\mathbf{u}_i, \mathbf{u}_j) = l\right) &\leq Q\left(\frac{d}{2\sigma}\right) \\ &= Q\left(\sqrt{\frac{lE}{(M-1)^2\sigma^2}}\right). \end{aligned}$$

As the codewords are generated by a random generator matrix with equiprobable entries, the distribution of the number of entries in which two length-T codewords differ is a binomial random variable $B(T, 1-1/M)$. Thus

$$\Pr(d(\mathbf{u}_i, \mathbf{u}_j) = l) = \binom{T}{l}\left(\frac{1}{M}\right)^{T-l}\left(1-\frac{1}{M}\right)^{l}.$$

Putting this together we have

$$\Pr\left(\mathbf{u}_i \to \mathbf{u}_j \middle| \, \mathbf{u} = \mathbf{u}_i\right) \leq \sum_{l=0}^{T}\binom{T}{l}\left(\frac{1}{M}\right)^{T-l}\left(1-\frac{1}{M}\right)^{l} Q\left(\sqrt{\frac{lE}{(M-1)^2\sigma^2}}\right).$$

(c) We use the bound $Q(x) \leq \frac{1}{2}e^{-x^2/2} < e^{-x^2/2}$ to get

$$\begin{aligned}
\Pr\left(\mathbf{u}_i \to \mathbf{u}_j \middle| \, \mathbf{u} = \mathbf{u}_i\right) &< \sum_{l=0}^{T}\binom{T}{l}\left(\frac{1}{M}\right)^{T-l}\left(1-\frac{1}{M}\right)^{l} e^{-\frac{lE}{2(M-1)^2\sigma^2}} \\
&= \left(\frac{1}{M}\right)^{T}\sum_{l=0}^{T}\binom{T}{l}\left(\frac{1-\frac{1}{M}}{\frac{1}{M}}\right)^{l} e^{-\frac{lE}{2(M-1)^2\sigma^2}} \\
&= \left(\frac{1}{M}\right)^{T}\sum_{l=0}^{T}\binom{T}{l}\left((M-1)e^{-\frac{E}{2(M-1)^2\sigma^2}}\right)^{l} \\
&= \left(\frac{1}{M}\right)^{T}\left(1+(M-1)e^{-\frac{\mathsf{SNR}}{2(M-1)^2}}\right)^{T}
\end{aligned}$$

Substituting back we get

$$\begin{aligned}
\Pr(\mathcal{E}) &< \sum_{i=1}^{2^{RT}}\sum_{j=1, j\neq i}^{2^{RT}}\left(\frac{1}{M}\right)^{T}\left(1+(M-1)e^{-\frac{E}{2(M-1)^2\sigma^2}}\right)^{T}\Pr(\mathbf{u} = \mathbf{u}_i) \\
&= 2^{RT}\left(\frac{1}{M}\right)^{T}\left(1+(M-1)e^{-\frac{E}{2(M-1)^2\sigma^2}}\right)^{T} \\
&= 2^{RT+T\log_2\left(\frac{1}{M}\right)+T\log_2\left(1+(M-1)e^{-\frac{E}{2(M-1)^2\sigma^2}}\right)} \\
&= 2^{T(R-R^*)}
\end{aligned}$$

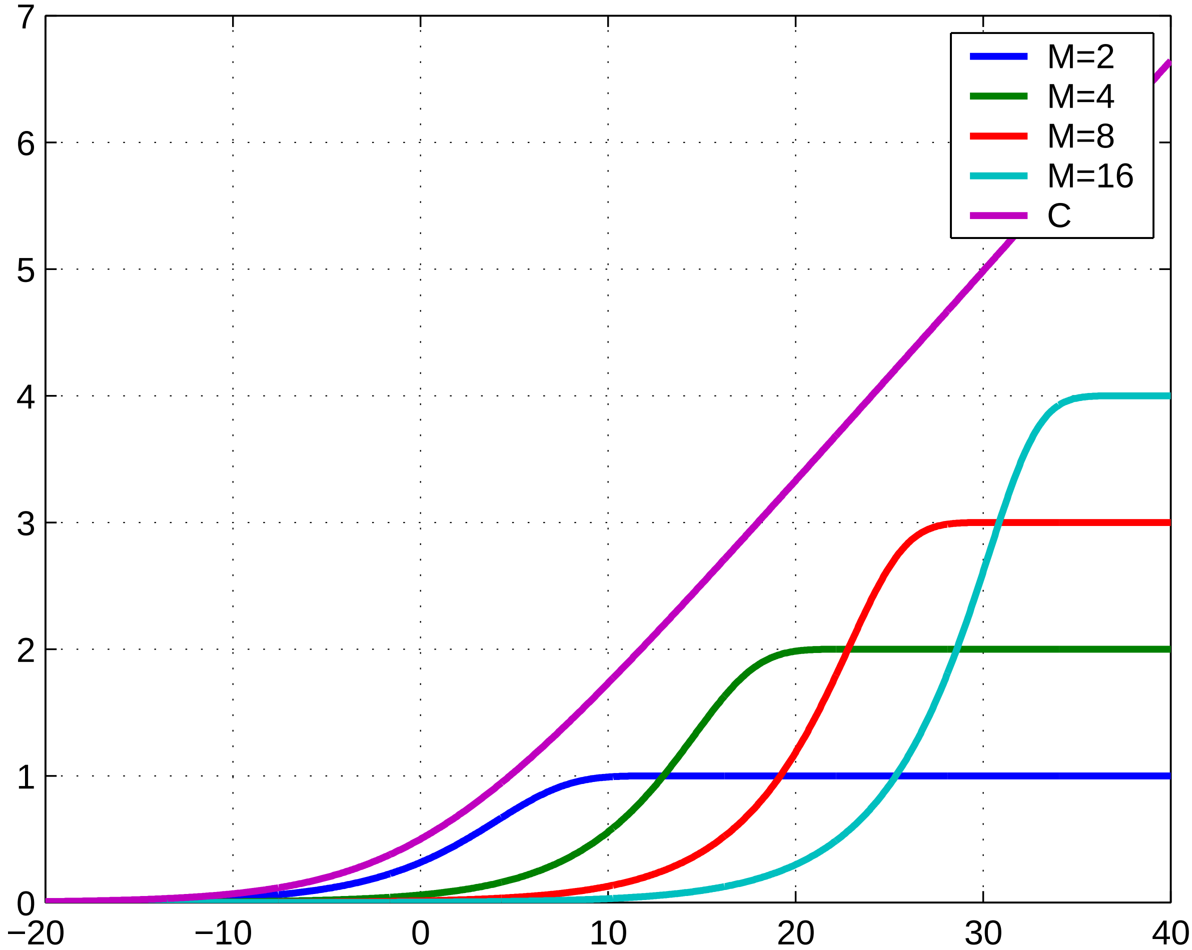
where $\mathsf{SNR} = E/\sigma^2$ and

$$R^* = \log_2\left(\frac{M}{1+(M-1)e^{-\frac{\mathsf{SNR}}{2(M-1)^2}}}\right).$$

Thus if $R < R^*$ the probability of making a decoding error goes to zero as $T \to \infty$.

(d) The answer to part (c) tells us that if we use a randomly generated code, then the error probability, averaged over the randomness of the generator matrix, goes to zero as $T \to \infty$, if $R < R^*$. This means that there exists at least one "good" generator matrix that has this property, i.e. that allows us to communicate reliably.

(e) The plot is shown below. R^*_M saturates at $\log_2 M$. Because we use an M-point constellation every time slot, we cannot hope to achieve a greater rate than this.



If the distance between constellation points is much less than the standard deviation of the noise, many entries of the transmitted codeword will be flipped and a large amount of redundancy will be required in order to ameliorate this effect. This will result in a low rate of reliable communication. On the other hand, if the distance between constellation points is much greater than the noise standard deviation, few if any of the codewords entries will be flipped and we could easily improve the rate without significantly effecting the number of flipped codeword entries, by increasing the constellation size. Thus in general we should space the constellation points at a distance roughly equal to the standard deviation of the noise. The conclusion is that in order to achieve higher rates, at low values of SNR we should use smaller constellation sizes, and at high values of SNR we should use larger constellation sizes.