Solutions 4

Exercise 4.1.

1. We first show that any pair of columns of H is linearly independent. Note that a linear dependency $ac_1 + bc_2 = 0$ for any two columns c_1 and c_2 would imply $a = b$, since the columns are all of the form $(1, *, *)^\top$. It is thus enough to check that the sum of any two columns cannot be zero. This is clearly the case if one of the two columns is either the first column of the second column. If the two columns are of the form $(1, \alpha, *)^\top$ and $(1, \alpha^2, *)^\top$, then their independence is clear too, since projection onto the first two coordinates gives two independent vectors. Otherwise, projection onto the last two coordinates gives two independent vectors.

   On the other hand,
   $$\alpha \begin{pmatrix} 1 \\ \alpha \\ 1 \end{pmatrix} + \alpha^2 \begin{pmatrix} 1 \\ \alpha \\ \alpha \end{pmatrix} + 1 \begin{pmatrix} 1 \\ \alpha \\ \alpha^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
   is a linear dependency among three columns. In other words, the weight-3 word $(00\alpha\alpha^2 1000)$ belongs to the code.

2. The weight distribution of this code involves only the nonnegative parameters A_0, A_3, A_4, A_5, A_6, A_7 and A_8. The objective function we want to maximize is
   $$A_0 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8.$$
   Recall the definition of the Krawtchouk polynomials
   $$K_k(x) := \sum_{j=0}^{k} (-1)^j \binom{x}{j} \binom{n-x}{k-j} (q-1)^{k-j}.$$
   Thus
   $$\begin{aligned} K_0(x) &= 1 \\ K_1(x) &= 3(8-x) - x = -4x + 24 \\ K_2(x) &= 9\binom{8-x}{2} - 3x(8-x) + \binom{x}{2} = 8x^2 - 92x + 252, \end{aligned}$$
   and the corresponding linear constraints are
   $$\begin{aligned} A_0 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8 &\geq 0 \\ 24\,A_0 + 12\,A_3 + 8\,A_4 + 4\,A_5 - 4\,A_7 - 8\,A_8 &\geq 0 \\ 252\,A_0 + 48\,A_3 + 12\,A_4 - 8\,A_5 - 12\,A_6 + 28\,A_8 &\geq 0. \end{aligned}$$
   The full linear program is
   $$\max A_0 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8 \text{ subject to}$$

$$\begin{aligned}
A_0 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8 &\geq 0 \\
24\,A_0 + 12\,A_3 + 8\,A_4 + 4\,A_5 - 4\,A_7 - 8\,A_8 &\geq 0 \\
252\,A_0 + 48\,A_3 + 12\,A_4 - 8\,A_5 - 12\,A_6 + 28\,A_8 &\geq 0 \\
1512\,A_0 + 44\,A_3 - 40\,A_4 - 28\,A_5 + 16\,A_6 + 28\,A_7 - 56\,A_8 &\geq 0 \\
5670\,A_0 - 150\,A_3 - 74\,A_4 + 50\,A_5 + 30\,A_6 - 70\,A_7 + 70\,A_8 &\geq 0 \\
13608\,A_0 - 252\,A_3 + 120\,A_4 + 44\,A_5 - 96\,A_6 + 84\,A_7 - 56\,A_8 &\geq 0 \\
20412\,A_0 + 216\,A_3 + 108\,A_4 - 144\,A_5 + 100\,A_6 - 56\,A_7 + 28\,A_8 &\geq 0 \\
17496\,A_0 + 324\,A_3 - 216\,A_4 + 108\,A_5 - 48\,A_6 + 20\,A_7 - 8\,A_8 &\geq 0 \\
6561A_0 - 243A_3 + 81A_4 - 27A_5 + 9A_6 - 3A_7 + A_8 &\geq 0 \\
A_0, A_3, A_4, A_5, A_6, A_7, A_8 &\geq 0
\end{aligned}$$

The linear program gives a solution

$$A_0 = 1, A_3 = 72, A_4 = 210, A_5 = 432, A_6 = 792, A_7 = \frac{4152}{7}, A_8 = \frac{1683}{7}.$$

The sum of these values is $16384/7$, so that $A_4(8,3) \leq \lfloor \log_4(16384/7) \rfloor = 5$, which shows the optimality of the code.

Exercise 4.2. The solution is similar to the proof of Gilbert-Varshamov bound for linear codes.

1. If the first k entries in $y = (y_1, \ldots, y_n)$ are zero, then the linear constraint $\langle H_i \mid y \rangle = 0$, where H_i is the ith row of H, simplifies to $y_{k+i} = 0$. As y is nonzero, this cannot happen for all i's, and thus, the probability of y being in the right kernel of H is zero.

   Now suppose $y_1, \ldots, y_k$ are not all zero and note that the linear map

   $$\begin{aligned}
   \phi_y : \mathbb{F}_q^k &\to \mathbb{F}_q \\
   (a_1, \ldots, a_k) &\mapsto \sum a_i y_i
   \end{aligned}$$

   being nontrivial, the preimage set of each $\alpha \in \mathbb{F}_q$ has the same size. Thus as the ith row A_i of A is picked uniformly in $\mathbb{F}_q^k$, $\langle A_i \mid (y_1, \ldots, y_k) \rangle$ will be uniformly distributed over $\mathbb{F}_q$, so that $\langle A_i \mid (y_1, \ldots, y_k) \rangle = y_{k+i}$ with probability $1/q$. But $Hy^\top = 0$ means that $\langle A_i \mid (y_1, \ldots, y_k) \rangle = y_{k+i}$ for each i. This happens with probability $(1/q)^{n-k} = q^{k-n}$.

2. We want to upper bound the probability that the code defined by the parity check matrix H contains a word of weight $d-1$ or less. For each fixed word y of weight $d-1$ or less, the probability that y is in the code (i.e., *bad event*) is given by the previous part. Now we use a union bound on all choices of y that have 1 as their first nonzero entry (this prevents overcounting since y belongs to the code if and only if αy belongs to the code for any nonzero scalar α); the number of such y is the volume of Hamming ball of radius $d-1$, that we denote by $V_q(n, d-1) = \sum_{i=1}^{d-1} \binom{n}{i}$, divided by $q-1$. Among these, $V_q(n-k, d-1)/(q-1)$ have zeros as their first k entries, and for these the bad event

probability is zero. Thus, the probability that we wish to compute is upper bounded by

$$q^{k-n} \cdot \frac{V_q(n, d-1) - V_q(n-k, d-1)}{q-1},$$

which is exactly ρ.

3. This is immediate from the previous part by observing that, for each i, H contains i dependent columns iff the linear code $\mathcal{C}_H$ for which it is a parity check matrix contains a codeword of weight i. Therefore

$$\Pr[\mathcal{C}_H \text{ has minimum distance } \leq d-1] \leq \Pr[H \text{ has } d-1 \text{ dependent columns}] \leq \rho.$$

Exercise 4.3.

1. Note that the support of y must be a subset of the support of c, otherwise $\operatorname{dist}(c, y) > t$. The number of such words y is thus $\binom{2t+1}{t+1} = \binom{2t+1}{t}$.

2. For every $c \in \mathcal{C}$, denote by Y_c the set

$$Y_c := \{y \in \mathbb{F}_q^n \colon \operatorname{wgt}(y) = t+1, \operatorname{dist}(y, c) = t\}.$$

Note that for each $c \neq c' \in \mathcal{C}$, we must have $Y_c \cap Y_{c'} = \emptyset$, as otherwise $\operatorname{dist}(c, c') < d$. Thus the number of $y \in \mathbb{F}_q^n$ of weight $t+1$ that are at distance t from some codeword of $\mathcal{C}$ is exactly $M\binom{2t+1}{t}$. But on the other hand the number of such words cannot exceed the total number of words of weight $2t+1$ in $\mathbb{F}_q^n$, which is $\binom{n}{t+1}(q-1)^{t+1}$. The bound follows.

Exercise 4.4. A *burst of length* ℓ is the event of having errors in a codeword such that the locations i and j of the first (leftmost) and last (rightmost) errors, respectively, satisfy $j - i = \ell - 1$. Let $\mathcal{C}$ be a linear $[n, k]$-code over $\mathbb{F}_q$ that is able to correct every burst of length t or less.

1. Consider a codeword $c = (c_1, \ldots, c_n)$ that contradicts this assumption. Then $w = (c_1, \ldots, c_{i+t-1}, 0, 0, \ldots, 0) = (0, \ldots, c_i, \ldots, c_{i+t-1}, 0, 0, \ldots, 0)$ can be either the zero codeword with a burst of length t starting at position i, or c with a burst of length t starting at position $i + t$, and is thus not uniquely decodable, a contradiction.

2. The proof is similar to that of the Singleton bound. Since the number of codewords is $q^k > q^{k-1}$, there must be at least two codewords that agree on their first $k-1$ coordinates, and thus, there is a nonzero codeword that has all zeros on its first $k-1$ coordinates, so that the position i of its first nonzero entry is such that $i \geq k$. On the other hand, the position j of its last nonzero entry satisfies $j \leq n$. Thus $j - i \leq n - k$. By the previous part, we have $j - i \geq 2t$, so that $2t \leq n - k$.

3. The proof is similar to the classical sphere-packing bound except that the shape of the "balls" are now different. For the sphere-packing bound we had to count the number of points that are at distance t from a given point, or the "volume" of the Hamming ball of radius t around each codeword. Here instead we only need to count the number of points within such a ball that are different from the word at the center (denoted by w) by a burst of size at most t. Denote this quantity by V. We have to distinguish the following cases and add up the numbers:

- The word w at the center,
- Words that are different from w in only one position. The number of such words is $n(q-1)$,
- Words that are different from w by a burst of size i, $2 \leq i \leq t$. The number of such words is $(n-i+1)(q-1)^2 q^{i-2}$.

Altogether, we will have

$$V = 1 + n(q-1) + (q-1)^2 \sum_{i=0}^{t-2} (n-i-1)q^i,$$

and similar to the sphere-packing bound, the "spheres" must be disjoint so that $q^k V \leq q^n$. The bound follows.

Exercise 4.5.

1. Let $\mathcal{C}$ be an $(n, M, 2r-1)$-code. By adding an overall parity check, we get an $(n+1, M, 2r)$-code $\mathcal{C}'$ (the fact that the minimum distance of $\mathcal{C}'$ is $2r$ follows from the fact that in $\mathcal{C}$, there exist two codewords at distance $2r-1$; the corresponding codewords in the extended code are at distance $2r$. If there were two codewords in $\mathcal{C}'$ at distance $2r-1$, they must differ in the parity check coordinate; but then there would exist two codewords in $\mathcal{C}$ at distance $2r-2$). Thus

$$A(n, 2r-1) \leq A(n+1, 2r).$$

Conversely, given an $(n+1, M, 2r)$-code, deleting one coordinate gives an $(n, M, d \geq 2r-1)$-code, thus

$$A(n, 2r-1) \geq A(n+1, 2r).$$

2. Given an (n, M, d)-code, divide the codewords into two classes, those beginning with 0 and those beginning with 1. One class must contain at least half of the codewords; deleting the first coordinate gives a code of length $n-1$ and minimum distance d. Thus

$$A(n-1, d) \geq \frac{1}{2} A(n, d).$$