You are encouraged to work on this homework together as long as you write up your own solutions. If you do so, write the names of your collaborators. Please refrain from looking up solutions to the homework on the internet or in other sources. If you must, state the source.

Problem 1

Recall the Gilbert-Varshamov bound says that linear $[n, k, d]$ codes exist whenever $k/n \leq 1 - H(d/n)$. Argue that such a code can be found in time $\mathrm{poly}(n)2^n$. Assume $k/n \leq 1 - H(d/n)$.

Hint: First show that for every collection of vectors $h_1, \ldots, h_t \in \{0,1\}^{n-k}$, $d - 1 \leq t \leq n$, there exists a vector $h \in \{0,1\}^{n-k}$ that is linearly independent from any subset of $d - 1$ of the vectors $h_1, \ldots, h_t$.

Problem 2

For $\{0,1\}$-valued random variables $Y_1, \ldots, Y_n$, we showed that the conditions (1) $Y_1, \ldots, Y_n$ are t-wise independent and (2) $\Pr[a_1 Y_1 \oplus \cdots \oplus a_n Y_n = 0] = 1/2$ for every nonzero $a = (a_1, \ldots, a_n)$ of hamming weight at most t are equivalent.

(a) State and prove an analogous condition for random variables taking values in $\{0, 1, \ldots, p-1\}$. (You may find it easier to deal with the case when p is a prime number first.)

(b) Can you give an approximate version of part (a) in the case when $Y_1, \ldots, Y_n$ are almost t-wise independent?

Problem 3

Recall the Plotkin bound: In an $[n, k, (1 + \varepsilon)n/2]$ code, $k \leq \log(1/\varepsilon + 1)$. In this prolem you will show the tightness of this bound and derive some of its consequences.

(a) Show that there exists a $[n, \log(n + 1), (n + 1)/2]$ code for infinitely many n. (**Hint:** Modify the Hadamard code a bit.)

(b) Show that for every $n_0, \varepsilon_0 > 0$ there exists a $[n, \log(1/\varepsilon + 1), (1 + \varepsilon)n/2]$ code for some $n \geq n_0$ and $\varepsilon \leq \varepsilon_0$. (**Hint:** Take the code in part (a) and replace 0 by 0^t, 1 by 1^t.)

(c) Show that if there exists an $[n, k, d]$ code, then there exists an $[n - 1, k - 1, d]$ code.

(d) Use part (c) and the Plotkin bound to show that there are no $[n, k, n/2]$ codes with $k > \log(n + 1)$.

(e) Show that in an $[n, k, (1 + \varepsilon)/2]$ code, we must have $k = O(\varepsilon n)$. (This bound is not optimal.)

Problem 4

An (n, k)-*affine hitter* is a function $f\colon \{0, 1\}^n \to \{0, 1\}^k$ that is not constant on any affine subspace of $\{0, 1\}^n$ of dimension k. (An affine subspace of dimension k consists of all the points of the form $Az + b$, where A is an $n \times k$ full rank matrix, b is a vector in $\{0, 1\}^n$, and z ranges over $\{0, 1\}^k$.)

(a) Show that$(n, \log n + \log\log n + O(1))$-affine hitters exist.

(b) How much time does it take to find an affine hitter with the parameters in part (a)?