8. Some remarks on coset enumeration and syndrome decoding

Let C be an $[n, k]$-linear code over some field F. In class we described a "naive" method of enumerating cosets of C: start with $C = 0 + C$ itself as our first coset, then choose some $v_2 \notin C$ and compute its coset $v_2 + C$, then choose v_3 which does not lie in the union of the first two cosets etc. When the number of cosets is large or when C is large, this procedure is quite ineffective and hard to apply regardless of whether you are trying to perform enumeration by hand or using a computer. Below by a *system of coset representatives for* C we will mean any subset T of F^n which contains exactly one element from each coset of C.

Question 8.1. *Is there an efficient way to find a system of coset representatives T? More specifically, can we write down elements of T right away, without computing any cosets of C beforehand.*

A positive answer to this question is provided by the following simple algorithm. Choose some basis $\{v_1, \dots, v_k\}$ for C and extend it to a basis of F^n, that is, find elements $v_{k+1}, \dots, v_n \in F^n$ such that $\{v_1, \dots, v_n\}$ is a basis of F^n.

Lemma 8.2. *Let $T = Span(v_{k+1}, \dots, v_n)$. Then T is a system of coset representatives for C.*

Proof. Exercise. Note that since $v_{k+1}, \dots, v_n$ are linearly independent, T is a subspace of dimension $n - k$ and hence $|T| = |F|^{n-k}$. We also know that the number of cosets of C is $\frac{|F|^n}{|C|} = \frac{|F|^n}{|F|^k} = |F|^{n-k}$. This means that T has the right size to be a system of coset representatives, and to prove the lemma it suffices either to show that every coset of C contains at LEAST one element of T or show that every coset of C contains at MOST one element of T (neither statement is hard to proof). □

Note that if we already happened to know a PCM H for C, then we can use the rows of PCM as vectors $v_{k+1}, \dots, v_n$ above. However, there are many possible choices for $v_{k+1}, \dots, v_n$ that do not come from PCMs; in particular, one can always choose $v_{k+1}, \dots, v_n$ as a subset of the standard basis $e_1, \dots, e_n$, so computing PCM just for the purpose of finding $v_{k+1}, \dots, v_n$ is usually not a good idea.

Question 8.1 address the problem of efficient coset enumeration, but for the purpose of efficient decoding we would like to answer a different question:

Question 8.3. *Is there an efficient way to find* **a leader** *for a given coset of C? More specifically, given $w \in F^n$, can we find a leader of the coset $w + C$ without having to write down every element of $w + C$?*

This question is harder to answer algorithmically, but there is a simple ad-hoc procedure involving the syndrome function that works well in suitable examples. Let H be a PCM for C. Recall that the syndrome function $S : F^n \to F^{n-k}$ is defined by $S(v) = Hv$ (we treat v as a column vector here) and has the property that $S(x) = S(y) \iff x$ and y lie in the same coset of C. Thus, to answer Question 8.3 for a given w, we need to find $u \in F^n$ of smallest possible weight such that $S(w) = S(u)$.

Let u be an arbitrary vector in F^n, and let $u_1, \ldots, u_n$ be the coordinates of u. Then a direct computation shows that if u is considered as a column vector, then $S(u) = \sum_{i=1}^{n} u_i col_i(H)$ where $col_i(H)$ is the i^{th} column of H (we already used this observation in the proof of Theorem 7.3 from class). Since the weight of u is the number of indices i such that $u_i \neq 0$, this leads to the following procedure for answering question 2.

Start with the given $w \in F^n$, compute its syndrome $S(w)$ and then write $S(w)$ as a linear combination of the columns of H with the smallest number of nonzero coefficients. Note that there is no simple procedure for this task, but if $S(w)$ can be represented as linear combination of, say, at most 3 columns, it usually does not take much work to find an optimal linear combination. Once this is done, the vector whose coordinates are the coefficients in the obtained linear combination is a coset leader for $w + C$.

Computing the Syndrome Look-Up Table. One can also use the same idea to compute the full syndrome look-up table for C. To do this consider all vectors in F^{n-k} (these are all possible values of the syndrome function). For each $s \in F^{n-k}$ write s as a linear combination of the columns of H with the smallest number of nonzero coefficients. This will determine a leader of the coset C_s consisting of all elements $w \in F^n$ with $S(w) = s$.

Note that using this procedure we will construct the syndrome look-up table "from right to left". In the usual procedure we put coset leaders in the left column and then compute their syndromes and enter the results in the right column. In the above procedure we start with all possible syndromes and for each syndrome s compute a coset leader u with $S(u) = s$. While determining the left column of the table from the right column is much harder than going from left to right, the new procedure has a major advantage in that it does not require any coset computation prior to filling the table.