Consider the following problem. We are given three independent, uniformly distributed Bernoulli random variables $Y_1$, $Y_2$, and $Y_3$. From these, we would like to generate seven pair-wise independent, uniformly distributed Bernoulli random variables. That is, we would like a set $X_1, \ldots, X_7$ of uniformly distributed Bernoulli random variables such that any two are independent.

We can do this as follows. With each $i$ from 1 to 7, associate different a non-empty subset of $\{1, 2, 3\}$. Then form $X_i$ by taking the exclusive-or of the random variables $Y$ picked out by the corresponding set. For example, if we map $3 \mapsto \{1, 2\}$, then $X_3 = Y_1 \oplus Y_2$.

It is not difficult to see that each $X_i$ is distributed uniformly on $\{0, 1\}$. Moreover, since each $X_i$ is generated from a different set of $Y$, any two distinct $X_i$ and $X_j$ must be independent. For example, if we also map $2 \mapsto \{2\}$, then $X_3 = Y_1 \oplus Y_2 = Y_1 \oplus X_2$, so whatever value $X_2$ assumes, $X_3$ assumes the opposite value with probability $\frac{1}{2}$.

We can create such a mapping by taking the binary representation of $i$. That is, we map $i$ to a subset $S \subseteq \{1, 2, 3\}$ such that $i = \sum_{s \in S} 2^{s-1}$.

Another way of looking at this is to consider a binary matrix $H$ whose columns are all the non-zero vectors of size 3. That is,

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$ 

Then each $i$ maps to the set characterized by the $i$th column of $H$. The matrix $H$ is called the Hamming matrix of order three. We will consider this approach again shortly, but first, we prove a generalization of the above.

**Theorem 1.** Given $m$ independent random variables uniformly distributed on $\GF(q)$, a finite field of order $q$, and a mapping $f : [n] \to \GF(q)^m$ such that any $k$ vectors $f(i_1), f(i_2), \ldots, f(i_k)$ are linearly independent, we can generate $n$ random variables uniformly distributed on $\GF(q)$ such that any $k$ are independent.

**Proof.** Let $\vec{y} \in \GF(q)^m$ be the vector of $m$ independent random variables uniformly distributed on $\GF(q)$. Then form $\vec{x} \in \GF(q)^n$ using $x_i = f(i)^T \vec{y}$. It is trivial to verify that the entries of $\vec{x}$ are uniformly distributed on $\GF(q)$. We will now show that any $k$ entries $x_{i_1}, \ldots, x_{i_k}$ of $\vec{x}$ are independent.

Let $z_j = x_{i_j} = f(i_j)^T \vec{y}$. We will compute the probability

$$\Pr[z_1 = b_1 \land z_2 = b_2 \land \cdots \land z_k = b_k],$$

Toward this end, form matrix $A \in \GF(q)^{k \times m}$ by taking the $f(i_j)^T$ as it’s rows. Then $\vec{z} = A \vec{y}$, and we are interested in

$$\Pr[A \vec{y} = \vec{b}],$$
over all the choices of $\vec{y}$. Because $A$ has full rank, we know, from linear algebra, that there are $q^{m-k}$ choices of $\vec{y}$ for every $\vec{b}$, so

$$\Pr[z_1 = b_1 \land z_2 = b_2 \land \cdots \land z_k = b_k] = \frac{q^{m-k}}{q^m} = \frac{1}{q^k} = \left(\frac{1}{q}\right)^k$$

$$= \Pr[z_1 = b_1] \Pr[z_2 = b_2] \cdots \Pr[z_k = b_k].$$

One way to find such mappings $f$ is to let $f(i)$ be the $i^{th}$ column of the parity check matrix for an error correcting code. Then $k$ is one less than the minimum distance of the code. For example, the Hamming matrix used in the example above is the parity check matrix of a Hamming code with block length 7, which has distance 3. Using a BCH code, we can design a mapping for any $k$ and $n$ such that $m = \Theta(k \log n)$. 

\[ \square \]