

Lecture 4

Almost k -wise Independent Variables

Mårten Trolin

1 The Method

We will now consider the problem of creating a set of bitstrings such that the bits are almost k -wise independent. Our method consists of two steps. In the first step we create a set of bitstrings such that parity on every fixed index set is almost even. In the second step we present a method to convert such a set into a set where the bits are almost k -wise independent.

For presentational reasons we give the steps in reverse order, i.e., we start by showing how to create almost k -wise independent bitstrings from a set with almost even parity.

2 The Vandermonde Matrix

Consider the following $n \times k$ -matrix over a finite field $\mathbb{F}$:

$$\mathbf{M} = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{k-1} \\ 1 & \alpha_2 & \alpha_2^2 & \dots & \alpha_2^{k-1} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \dots & \alpha_n^{k-1} \end{pmatrix}.$$

This matrix can be used for evaluation of the polynomial $p(x) = \sum_{i=0}^{k-1} a_i x^i$ at $\alpha_1, \alpha_2, \dots, \alpha_n$ by computing the product

$$\mathbf{M} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{k-1} \end{pmatrix} = \begin{pmatrix} p(\alpha_1) \\ p(\alpha_2) \\ \vdots \\ p(\alpha_n) \end{pmatrix}.$$

When $n = k$ this matrix is called the Vandermonde matrix.

Proposition 1. *Every $k \times k$ submatrix of $\mathbf{M}$ is non-singular if $\alpha_i \neq \alpha_j$ for $i \neq j$.*

Proof. Since a $k - 1$ degree polynomial is uniquely determined by values at k distinct points, the Vandermonde matrix is invertible.

Now consider the determinant of the Vandermonde matrix. By definition the determinant is a polynomial of degree at most $k - 1$ in α_i . Obviously the determinant is zero when $\alpha_j = \alpha_i$ for some $i \neq j$, which implies that $(\alpha_j - \alpha_i)$ is a factor for every $i \neq j$, and hence $\prod_{1 \leq i < j \leq k} (\alpha_j - \alpha_i)$ divides the determinant. Considering that the definition gives degree of the monomials of the determinant of $1 + 2 + \dots + (k - 1)$, this expression gives the determinant up to constant factors. It can be shown that this constant is indeed 1.

3 Bits that Are k -wise Independent

Now consider the problem of producing a set of bit strings of length l where the bits are k -wise independent bits. As shown in the previous lecture this can be achieved by setting $q = 2^{\lceil \log n \rceil}$ where n is such that $l = n \log n$ and form an $n \times k$ matrix $\mathbf{M}$ over $GF[q]$. This matrix can be converted into a matrix $\mathbf{M}'$ over $\{0, 1\}$ by using the observation that multiplication of an element in $GF[q]$ by a constant $\alpha \in GF[q]$ can be performed by multiplying the corresponding bitstring by a $\log n \times \log n$ matrix $\mathbf{M}_\alpha$. Now we can produce our sample space by computing $\mathbf{M}'\mathbf{b}$ for each $\mathbf{b} \in \{0, 1\}^l$.

Our next goal is to make the sample space smaller by iterating not over all possible bitstrings, but only over a subset.

4 Bits that Are Almost k -wise Independent

In order to decrease the size of the sample space, we investigate a variant of k -wise independence, namely *almost k -independence*.

Definition 1. The $\{0, 1\}$ -vector $\mathbf{x}$ is said to be almost k -wise independent in the first sense with bias ϵ if for any choice of k indices $i_1, i_2, \dots, i_k$ it holds that

$$\sum_{b \in \{0, 1\}^k} |\Pr[\wedge (x_{i_j} = b_j)] - 2^{-k}| \leq \epsilon .$$

Intuitively this definition says that a distribution is almost k -wise independent if each constellation of k bits takes every possible value approximately equally often. An alternative definition is to say that k -wise independence means that the expected parity of each subset of k bits or less is approximately $1/2$:

Definition 2. The $\{0, 1\}$ -vector $\mathbf{x}$ is said to be almost k -wise independent in the second sense with bias δ if for any index set S , $|S| \leq k$, it holds that

$$\left| \Pr[(\oplus_{i \in S} x_i) = 1] - \frac{1}{2} \right| \leq \delta .$$

It can be shown that a distribution that is k -wise independent in the first sense with bias ϵ is k -wise independent in the second case the bias δ . The converse also holds with $\epsilon = 2^k \delta$. The proof will be presented at a later lecture.

Now the idea is to sample $\mathbf{b}$ from a set such each parity over a fixed index set is almost even, and then compute $\mathbf{M}\mathbf{b}$. The following lemma shows that this method gives almost k -wise independent bits.

Lemma 1. *If $\mathbf{M}\mathbf{b} = \mathbf{x}$ and each set of k rows of $\mathbf{M}$ is linearly independent, then for S such that $|S| \leq k$ it holds that $\oplus_{i \in S} x_i = \oplus_{j \in T} b_j$ some non-empty set T .*

Proof. Since $x_i = \oplus_j M_{ij} b_j$ we can write

$$\oplus_{i \in S} x_i = \oplus_{i \in S} \oplus_j M_{ij} b_j = \oplus_j ((\oplus_{i \in S} M_{ij}) b_j) \quad .$$

Now it remains to show that $\oplus_{i \in S} M_{ij} \neq 0$ for at least one j . Suppose this is not the case. Then we have found a $k \times k$ submatrix that is singular, which contradicts the assumption that every set of k rows is linearly independent.

Obviously this gives a set of bit strings of length n that are k -wise independent if the set from which $\mathbf{b}$ is chosen is $\{0, 1\}^k$ and $\mathbf{M}$ is $n \times k$. Since we only require $\mathbf{b}$ to be chosen from a set where every parity is almost even one might hope to find a smaller set to choose $\mathbf{b}$ from and still get almost k -wise independence. In the following section we show how to create such a set.

5 Linear Feedback Shift Registers

We now use Linear Feedback Shift Registers (LFSRs) to construct bit strings where each parity is almost even. For this description let m be the length of the strings we create. Given coefficients $a_0, a_1, \dots, a_{r-1}$ and start values $b_0, b_1, \dots, b_{r-1}$ a LFSR defines a bit sequence

$$b_i = \oplus_{j=0}^{r-1} a_j b_{i-r+j} \quad .$$

Since the cycles of a shift register are disjoint and the all zero vector produces a singleton cycle, the longest cycle we can hope for has length $2^r - 1$. Let us investigate the cycle length a little more closely.

For a shift register we can define the characteristic polynomial as $f(t) = \sum_{j=0}^r a_j t^j$ where we define $a_r = 1$. Now the period of the shift register is the smallest s such that $f(t)$ divides $t^s + 1$. The following proposition generalizes this observation:

Proposition 2. *Let G be a linear feedback shift register with characteristic polynomial $f(t)$. Then $f(t)$ divides $\sum_{j=1}^l t^{i_j}$ precisely when $\oplus_{j=1}^l b_{i_j} = 0$.*

Proof. $f(t)$ divides $\sum_{j=1}^l t^{i_j}$ when there exists a polynomial $g(t)$ such that $f(t)g(t) = \sum_{j=1}^l t^{i_j}$. The feedback rule of G directly implies that $\oplus_{j=0}^r a_j b_{i-r+j} = 0$ for any i . Now consider the corresponding rule defined by the polynomial $t^d f(t)$, $\oplus_{j=0}^r a_j b_{i-r+j+d} = 0$. It can also be verified that if the rule holds for two polynomials $g(t)$ and $h(t)$, it also holds for their sum $g(t) + h(t)$. Hence it holds for the product $f(t)g(t)$.

We now sketch a proof of the converse, that the polynomial $g(t)$ exists if $\oplus_{j=1}^l b_{i_j} = 0$. The parity of the bits b_{i_j} can be recursively rewritten as parity of the bits $\{b_0, b_1, \dots, b_r\}$ using the feedback relation. Since this corresponds exactly to polynomial division by $f(t)$, it gives $\oplus_{j=1}^l b_{i_j} = 0 = f(t)g(t) + r(t)$. The only choice of bits $b_0, b_1, \dots, b_r$ with even parity is bits defined by $f(t)$, which implies $r(t) = 0$. Hence $\sum_{j=1}^l t^{i_j}$ is divisible by $f(t)$.

We are now ready to give our construction. We let the sample space consist of the output on all start values of all shift registers of length r whose characteristic polynomial is irreducible. There are approximately $2^r/r$ such polynomials and 2^r start values, which gives a total of about $2^{2r}/r$ samples of length m . To show that each parity is almost even we examine $\oplus b_{i_j}$. Since this can be treated as a polynomial of degree at most m , it is divisible by at the most m/r irreducible polynomials. This gives

$$\left| \Pr [\oplus b_{i_j} = 1] - \frac{1}{2} \right| \leq \frac{m}{r} \bigg/ \frac{2^r}{r} = m2^{-r}$$

Now we can conclude that by setting $r = \log(m\delta^{-1})$ we achieve a bias of at most δ . The size of the sample space is approximately $2^{2r} \approx \left(\frac{m}{\delta}\right)^2$.