4-bit Linearity Test

We have

$$
\begin{aligned}
2\varepsilon \le 2\mathbf{Pr}_{x,y,z}[\text{accepts}] - 1 &= 2\mathbf{E}_{x,y,z}\left[\tfrac{1}{2}(1 + f(x)f(y)f(z)f(xyz))\right] - 1 \\
&= \mathbf{E}_{x,y,z}[f(x)f(y)f(z)f(xyz)] \\
&= \mathbf{E}_{x,y,z}\Big[\sum_{\alpha,\beta,\gamma,\delta} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \hat{f}_\delta \chi_\alpha(x)\chi_\beta(y)\chi_\gamma(z)\chi_\delta(xyz)\Big] \\
&= \sum_{\alpha,\beta,\gamma,\delta} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \hat{f}_\delta \mathbf{E}_{x,y,z}\Big[\chi_{\alpha\oplus\delta}(x)\chi_{\beta\oplus\delta}(y)\chi_{\gamma\oplus\delta}(z)\Big] \\
&= \sum_{\alpha,\beta,\gamma,\delta} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \hat{f}_\delta \mathbf{E}_{x}\Big[\chi_{\alpha\oplus\delta}(x)\Big]\mathbf{E}_{y}\Big[\chi_{\beta\oplus\delta}(y)\Big]\mathbf{E}_{z}\Big[\chi_{\gamma\oplus\delta}(z)\Big] \\
&= \sum_{\delta} \hat{f}_\delta^4 \le \max_\delta \hat{f}_\delta^2,
\end{aligned}
$$

and thus there exists δ such that $|\hat{f}_\delta| \ge \sqrt{2\varepsilon}$.

3LIN to 3SAT

Let's use additive notation, instead. Note that

$$x \oplus y \oplus z = 1$$

if and only if

$$(x, y, z) \notin \{(0,0,0), (0,1,1), (1,0,1), (1,1,0)\},$$

i.e., if and only if

$$(x \vee y \vee z) \wedge (x \vee \neg y \vee \neg z) \wedge (\neg x \vee y \vee \neg z) \wedge (\neg x \vee \neg y \vee z).$$

This gives 3LIN with n equations $\to$ 3SAT with $4n$ clauses,

- $(1-\varepsilon)n$ equations satisfied $\to$ $(1-\varepsilon)4n$ clauses satisfied
- at least $1/2 - \varepsilon$ equations not sat. $\to$ at least $(1/2-\varepsilon)n$ clauses not sat.

Hence, $\text{Gap3LIN}_{1-\varepsilon,1/2+\varepsilon}$ reduces to $\text{Gap3SAT}_{1-\varepsilon,7/8+\varepsilon/4}$.

Graph Linearity Test, $k = 3$

For $T = (1+f(x)f(y)f(xy))(1+f(x)f(z)f(xz))(1+f(y)f(z)f(yz))-1$, we have

$$
\begin{aligned}
8\varepsilon &\leq 8\mathbf{Pr}_{x,y,z}[\text{accepts}] - 1 = 8\mathbf{E}_{x,y,z}\big[(T+1)/8\big] - 1 = \mathbf{E}_{x,y,z}[T] \\
&= \mathbf{E}[f(x)f(y)f(xy) + \cdots f(y)f(z)f(xy)f(xz) + \cdots f(xy)f(xz)f(yz)] \\
&= \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \mathbf{E}[\chi_\alpha(x)\chi_\beta(y)\chi_\gamma(xy)] + \cdots \\
&\quad + \sum_{\alpha,\beta,\gamma,\delta} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \hat{f}_\delta \mathbf{E}[\chi_\alpha(y)\chi_\beta(z)\chi_\gamma(xy)\chi_\delta(xz)] + \cdots \\
&\quad + \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \mathbf{E}[\chi_\alpha(xy)\chi_\beta(xz)\chi_\gamma(yz)] \\
&= 4\sum_\alpha \hat{f}_\alpha^3 + 3\sum_\alpha \hat{f}_\alpha^4 \leq \max_\alpha(4\hat{f}_\alpha + 3\hat{f}_\alpha^2)
\end{aligned}
$$

and thus exists α such that $8\varepsilon \leq 4\hat{f}_\alpha + 3\hat{f}_\alpha^2$, implying $\hat{f}_\alpha \geq \frac{2}{3}(\sqrt{1+6\varepsilon} - 1) > \varepsilon$.

Optimality of Graph Linearity Test, Fourier Coefficients

Let us switch to the multiplicative notation, so

$$g(y) = \prod_{i=1}^{n} (-1)^{(1-y_{2i-1})(1-y_{2i})/4}$$

for $y \in \{-1, 1\}^{2n}$. We have

$$\begin{aligned}
|\hat{g}_\alpha| &= \left| \mathbf{E}_y [g(y)\chi_\alpha(y)] \right| \\
&= \left| \mathbf{E}_y \left[\prod_{i=1}^{n} (-1)^{(1-y_{2i-1})(1-y_{2i})/4} \prod_{i=1}^{n} \chi_{\alpha \cap \{y_{2i-1}, y_{2i}\}}(y) \right] \right| \\
&= \prod_{i=1}^{n} \left| \mathbf{E}_{y_{2i-1}, y_{2i}} \left[(-1)^{(1-y_{2i-1})(1-y_{2i})/4} \chi_{\alpha \cap \{y_{2i-1}, y_{2i}\}}(y_{2i-1}, y_{2i}) \right] \right| \\
&= 2^{-n}.
\end{aligned}$$

Optimality of Graph Linearity Test, Linearity in a Subspace

Random k-dimensional subspace L: choose vectors $v_1, \ldots, v_k$ independently and take L as their linear span.

- Prob. $v_1, \ldots, v_k$ not linearly independent $\to 0$ as $n \to \infty$.
- Coordinates $(x_1, \ldots, x_k) \in \mathbb{F}_2^k$: $y = \sum_{i=1}^k x_i v_i$.

$$g(y) = \sum_{a=1}^{n} y_{2a-1} y_{2a} = \sum_{a=1}^{n} \Big(\sum_{i=1}^{k} x_i v_{i,2a-1} \Big) \Big(\sum_{j=1}^{k} x_j v_{j,2a} \Big)$$
$$= \sum_{i=1}^{k} \Big(\sum_{a=1}^{n} v_{i,2a-1} v_{i,2a} \Big) x_i + \sum_{i<j} \Big(\sum_{a=1}^{n} v_{i,2a-1} v_{j,2a} + v_{i,2a} v_{j,2a-1} \Big) x_i x_j$$

Linear if the coefficient at $x_i x_j$ is 0 for every $i < j$.

$v_1, \ldots, v_k$ independent, want probability that for all $i < j$,

$$c_{i,j} = \sum_{a=1}^{n} v_{i,2a-1} v_{j,2a} + v_{i,2a} v_{j,2a-1} = 0.$$

- Imagine choosing the odd coordinates, $b_i = (v_{i,1}, v_{i,3}, \ldots, v_{i,2n-1})$ for all i, first.
  - A.a.s. as $n \to \infty$, for $i = 1, \ldots, k$, there exists a_i such that $b_{i,a_i} = 1$ and $b_{j,a_i} = 0$ for $j \neq i$.
  - Consider fixed $b_1, \ldots, b_k$ with this property ($\star$).
- Let $u_i = (v_{i,2}, v_{i,4}, \ldots, v_{2n})$ be the yet unchosen coordinates: $c_{i,j} = \langle b_i | u_j \rangle + \langle b_j | u_i \rangle$.
- Due to ($\star$), for any $t \leq k-1$ and any $d_1, \ldots, d_t$,

  $$\mathbf{Pr}[c_{1,t+1} = c_{2,t+1} = \ldots = c_{t,t+1} = 0 | u_1 = d_1, \ldots, u_t = d_t] = 2^{-t};$$

  hence,

  $$\mathbf{Pr}[c_{1,t+1} = c_{2,t+1} = \ldots = c_{t,t+1} = 0 | c_{ij} = 0 \text{ for } i < j \leq t] = 2^{-t}.$$

$$\mathbf{Pr}[c_{ij} = 0 \text{ for all } i < j] = 2^{-1} 2^{-2} \cdots 2^{-(k-1)} = 2^{-\binom{k}{2}}$$