4.2 The BBS Generator and Quadratic Residuosity

Given a seed $s \in \mathbb{M}_n^+$ the BBS generator outputs a bit sequence $(b_1(s), \ldots, b_r(s))$—by the way the same sequence as the seed $s' = \sqrt{s^2} \bmod n$ that is a quadratic residue. A probabilistic circuit (see Appendix B of Part III)

$$C\colon \mathbb{F}_2^r \times \Omega \longrightarrow \mathbb{F}_2$$

has an ε-advantage for **BBS extrapolation** with respect to n if

$$P(\{(s,\omega) \in \mathbb{M}_n \times \Omega \mid C(b_1(s), \ldots, b_r(s), \omega) = \mathrm{lsb}(\sqrt{s^2} \bmod n)\}) \geq \frac{1}{2} + \varepsilon.$$

In other words: The algorithm implemented by C “predicts” (or extrapolates) the bit preceding a given subsequence with ε-advantage.

> If we seed the generator with a quadratic residue s, then C outputs the parity of s (with ε-advantage). If fed with a later segment $(b_{i+1}, \ldots, b_{i+r})$ (with $i \geq 1$) of a BBS output C extrapolates the preceding bit b_i.

In the following lemmas and proposition let τ_t be the maximum expense of the operation $xy \bmod n$ where n is a t-bit integer and $0 \leq x, y < n$. We know that $\tau_t = \mathrm{O}(t^2)$ (and even know an exact upper bound for the circuit size).

Lemma 16 *Let n be a* Blum *integer $< 2^t$. Assume the probabilistic circuit $C\colon \mathbb{F}_2^r \times \Omega \longrightarrow \mathbb{F}_2$ has an ε-advantage for BBS extrapolation with respect to n. Then there is a probabilistic circuit $C' : \mathbb{F}_2^t \times \Omega \longrightarrow \mathbb{F}_2$ of size $\#C' \leq \#C + r\tau_t + 4$ that has an ε-advantage for deciding quadratic residuosity for $x \in \mathbb{M}_n^+$.*

Proof. First we compute the BBS sequence $(b_1, \ldots, b_r)$ for the seed $s \in \mathbb{M}_n^+$ at an expense of $r\tau_t$. Then C computes the bit $\mathrm{lsb}(\sqrt{s^2} \bmod n)$ with advantage ε. Therefore setting

$$C'(s,\omega) := \begin{cases} 1 & \text{if } C(b_1, \ldots, b_r, \omega) = \mathrm{lsb}(s), \\ 0 & \text{otherwise}, \end{cases}$$

we decide the quadratic residuosity of s with ε-advantage by the corollary of Proposition 24 in Appendix A.11 of Part III. The additional costs for comparing bits are at most 4 additional nodes in the circuit. ◇

Now let $C\colon \mathbb{F}_2^t \times \Omega \longrightarrow \mathbb{F}_2$ be an arbitrary probabilistic circuit. Then for $m \geq 1$ we define the **m-fold circuit** by

$$C^{(m)}\colon \mathbb{F}_2^t \times \Omega^m \longrightarrow \mathbb{F}_2,$$

$$C^{(m)}(s,\omega_1,\ldots,\omega_m) := \begin{cases} 1 & \text{if } \#\{i \mid C(s,\omega_i)=1\} \geq \frac{m}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

So this circuit represents the "majority decision". Its implementation consists of m parallel copies of C, one integer addition of m bits, and one comparision of $\lceil {}^2\log m \rceil$-bit integers, hence by Appendix B.3 of Part III its size is

$$\#C^{(m)} \leq r \cdot \#C + 2m^2.$$

Lemma 17 (Amplification of advantage) *Let $A \subseteq \mathbb{F}_2^t$, and let C be a circuit that computes the Boolean function $f : A \longrightarrow \mathbb{F}_2$ with an ε-advantage. Let $m = 2h+1$ be odd.*

Then $C^{(m)}$ computes the function f with an error probability of

$$\leq \frac{(1-4\varepsilon^2)^h}{2}.$$

For each $\delta > 0$ there is an

$$m \leq 3 + \frac{1}{2\delta\varepsilon^2}$$

such that $C^{(m)}$ computes the function f with an error probability δ.

Proof. The probability that C gives a correct answer is

$$p := P(\{(s,\omega) \in A \times \Omega \mid C(s,\omega) = f(s)\}) \geq \frac{1}{2} + \varepsilon.$$

Since enlarging ε tightens the assertion we may assume that $p = \frac{1}{2} + \varepsilon$. The complementary value $q := 1 - p = \frac{1}{2} - \varepsilon$ equals the probability that C gives a wrong answer. Hence the probability of getting exactly k correct answers from m independent invocations of C is $\binom{m}{k} p^k q^{m-k}$. Thus the error probability we search is

$$P(\{(s,\omega_1,\ldots,\omega_m) \in A \times \Omega^m \mid C^{(m)}(s,\omega_1,\ldots,\omega_m) = f(s)\})$$

$$\begin{aligned}
&= \sum_{k=0}^{h} \binom{m}{k} (\frac{1}{2}+\varepsilon)^k (\frac{1}{2}-\varepsilon)^{m-k} \\
&= (\frac{1}{2}+\varepsilon)^h (\frac{1}{2}-\varepsilon)^{h+1} \cdot \sum_{k=0}^{h} \binom{m}{k} (\frac{1}{2}+\varepsilon)^{k-h} (\frac{1}{2}-\varepsilon)^{h-k} \\
&= (\frac{1}{4}-\varepsilon^2)^h \cdot (\frac{1}{2}-\varepsilon) \cdot \underbrace{\sum_{k=0}^{h} \binom{m}{k} \underbrace{\left(\frac{\frac{1}{2}-\varepsilon}{\frac{1}{2}+\varepsilon}\right)^{h-k}}_{\leq 1}}_{\leq 2^{m-1}=4^h} \\
&\leq (1-4\varepsilon^2)^h
\end{aligned}$$

which proves the first statement.

For an error probability δ a sufficient condition is:

$$\begin{aligned} (1-4\varepsilon^2)^h &\le 2\delta, \\ h\cdot\ln(1-4\varepsilon^2) &\le \ln 2+\ln\delta, \\ h &\ge \frac{\ln 2+\ln\delta}{\ln(1-4\varepsilon^2)}. \end{aligned}$$

Therefore we choose

$$h := \left\lceil \frac{\ln 2+\ln\delta}{\ln(1-4\varepsilon^2)} \right\rceil. \tag{1}$$

Then the error probability of $C^{(m)}$ is at most δ, and

$$\begin{aligned} h &\le 1+\frac{\ln 2+\ln\delta}{\ln(1-4\varepsilon^2)} = 1+\frac{\ln\frac{1}{\delta}-\ln 2}{\ln\frac{1}{1-4\varepsilon^2}} \\ &\le 1+\frac{\frac{1}{\delta}-1-\ln 2}{4\varepsilon^2} \le 1+\frac{1}{4\delta\varepsilon^2}, \end{aligned}$$

proving the second statement. ◇

By the way the size of $C^{(m)}$ is

$$\#C^{(m)} \le \left[3+\frac{1}{2\delta\varepsilon^2}\right]\cdot\#C+2\cdot\left[3+\frac{1}{2\delta\varepsilon^2}\right]^2.$$

Merging the two lemmas we get:

Proposition 13 *Let n be a* BLUM *integer $< 2^t$. Assume the probabilistic circuit $C:\mathbb{F}_2^r\times\Omega\longrightarrow\mathbb{F}_2$ has an ε-advantage for BBS extrapolation with respect to n. Then for each $\delta>0$ there is a probabilistic circuit $C'\colon\mathbb{F}_2^t\times\Omega'\longrightarrow\mathbb{F}_2$ that decides quadratic residuosity in $\mathbb{M}_n^+$ with error probability δ and has size*

$$\#C' \le \left[3+\frac{1}{2\delta\varepsilon^2}\right]\cdot[\#C+r\tau_t+4]+2\cdot\left[3+\frac{1}{2\delta\varepsilon^2}\right]^2.$$

Note that the size of C' is polynomial in r, $\#C$, $\frac{1}{\delta}$, $\frac{1}{\varepsilon}$, and t, and we even could make this polynomial explicit. Thus:

> From an efficient probabilistic BBS extrapolation algorithm for the module n with ε-advantage we can construct an efficient probabilistic decision algorithm for quadratic residuosity for n with arbitrary small error probability.

This complexity bound becomes even more perspicuous, when we specify dependencies from the input complexity, measured by the bit size t. Thus we choose

- $r \leq f(t)$ with a polynomial $f \in \mathbb{Q}[T]$ (that is we generate only "polynomially many" pseudorandom bits),
- $\frac{1}{\delta} \leq g(t)$ (or $\delta \geq 1/g(t)$) with a polynomial $g \in \mathbb{Q}[T]$ (that is we don't choose δ "too small", not like an ambitious $\delta < 1/2^t$),
- $\frac{1}{\varepsilon} \leq h(t)$ (or $\varepsilon \geq 1/h(t)$) with a polynomial $h \in \mathbb{Q}[T]$ (that is ε is reasonably small, not only like a modest $\varepsilon \approx 1/\log(t)$).

Then

$$\begin{aligned}
\#C' &\leq \left[3 + \frac{1}{2}\, g(t)\, h(t)^2\right] \cdot [\#C + f(t)\, \tau_t + 4] + 2 \cdot \left[3 + \frac{1}{2}\, g(t)\, h(t)^2\right]^2 \\
&\leq \Phi(t) \cdot \#C + \Psi(t)
\end{aligned}$$

with polynomials Φ, $\Psi \in \mathbb{Q}[t]$. In the following section we'll see how this statement makes BBS a "perfect" pseudorandom generator.

The hypothetical decision algorithm for $s \in \mathbb{M}_n^+$ from Proposition 13 runs like this (assuming that n is a public parameter):

1. Construct the BBS-sequence $b_1(s), \ldots, b_r(s)$ (using the public parameter n).
2. Choose the desired error probability δ.
3. Choose $m = 2h + 1$ with h as in Equation 1.
4. Choose random elements $\omega_1, \ldots, \omega_m \in \Omega$ and determine $b_i = C(s, \omega_i) \in \mathbb{F}_2$ for $i = 1, \ldots, r$.
5. Count $z = \#\{i \mid b_i = \operatorname{lsb}(s)\}$.
6. If $z \geq m/2$ output 1 ("quadratic residue"), else output 0 ("quadratic nonresidue").