NIST Interagency Report
NIST IR 8427 ipd

Discussion on the Full Entropy Assumption of the SP 800-90 Series

Initial Public Draft (IPD)

Darryl Buller
Aaron Kaufer
Allen Roginsky
Meltem Sönmez Turan

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8427.ipd

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST Interagency Report
NIST IR 8427 ipd

Discussion on the Full Entropy Assumption of the SP 800-90 Series

Initial Public Draft (IPD)

Darryl Buller
Aaron Kaufer
National Security Agency

Allen Roginsky
Meltem Sönmez Turan
Computer Security Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8427.ipd

September 2022

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

Publication History
Approved by the NIST Editorial Review Board on YYYY-MM-DD [will be updated in the final publication]

How to Cite this NIST Technical Series Publication:
Buller D, Kaufer A, Roginsky A, Sönmez Turan M (2022) Discussion on the Full Entropy Assumption of the SP 800-90 Series. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) NIST IR 8427 ipd. https://doi.org/10.6028/NIST.IR.8427.ipd

Author ORCID iDs
Author 1: 0000-0000-0000-0000
Author 2: 0000-0000-0000-0000
Author 3: 0000-0000-0000-0000
Author 4: 0000-0000-0000-0000

Public Comment Period
September 7, 2022 – October 31, 2022

Submit Comments
rbg_comments@nist.gov

National Institute of Standards and Technology
Attn: Computer Security Division, Information Technology Laboratory
100 Bureau Drive (Mail Stop 8930) Gaithersburg, MD 20899-8930

All comments are subject to release under the Freedom of Information Act (FOIA).

Reports on Computer Systems Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) promotes the U.S. economy and public welfare by providing technical leadership for the Nation's measurement and standards infrastructure. ITL develops tests, test methods, reference data, proof of concept implementations, and technical analyses to advance the development and productive use of information technology. ITL's responsibilities include the development of management, administrative, technical, and physical standards and guidelines for the cost-effective security and privacy of other than national security-related information in federal information systems.

Abstract

NIST SP 800-90 series support the generation of high-quality random bits for cryptographic and non-cryptographic use. The security of a random number generator depends on the *unpredictability* of its outputs, which can be measured in terms of entropy. NIST SP 800-90 series uses *min-entropy* to measure entropy. A full-entropy bitstring has an amount of entropy equal to its length. Full-entropy bitstrings are important for cryptographic applications, as these bitstrings have ideal randomness properties and may be used for any cryptographic purpose. Due to the difficulty of generating and testing full-entropy bitstrings, SP 800-90 series assume that a bitstring has *full entropy* if the amount of entropy per bit is at least $1 - \varepsilon$, where ε is at most 2^{-32}. This report provides a justification for the selection of ε. This is accomplished as follows. The report begins by defining full entropy in terms of a hypothetical distinguishing game. The report then derives two results following from this definition. First, it is shown how output satisfying this definition can be generated using a conditioning function acting on data having a known entropy level. Second, the actual entropy level of output produced by such a process is computed, thereby providing support for the selected value of ε.

Keywords

entropy; min-entropy; random number generation.

Table of Contents

1. Introduction **1**
2. Full Entropy Definition **1**
2.1. Derivation of Conditions for Full Entropy 1
2.2. Justification of Claim on $\boldsymbol{\theta j}$ 7
2.3. Derivation of Full Entropy Threshold 8
References **9**
Appendix A. List of Symbols, Abbreviations, and Acronyms **10**
Appendix B. Glossary **11**

List of Tables

Table 1. Minimum value of $\boldsymbol{H-n}$ for various values of $\boldsymbol{W}$ and $\boldsymbol{\delta}$ 7

1. Introduction

The NIST SP 800-90 series [1][2][3] support the generation of high-quality random bits for cryptographic and non-cryptographic use. The security of a random number generator depends on the *unpredictability* of its outputs, which can be measured in terms of entropy. NIST SP 800-90 series uses *min-entropy* to measure entropy. A full-entropy bitstring has an amount of entropy equal to its length. Full-entropy bitstrings are important for cryptographic applications, as these bitstrings have ideal randomness properties and may be used for any cryptographic purpose. Due to the difficulty of generating and testing full-entropy bitstrings, SP 800-90 series assume that a bitstring has full entropy if the amount of entropy per bit is at least $1 - \varepsilon$, where ε is at most 2^{-32}. This report provides the foundation for the selection of this value of ε. This is accomplished as follows. The report begins by defining full entropy in terms of a hypothetical distinguishing game. The report then derives two results following from this definition. First, it is shown how output satisfying this definition can be generated using a conditioning function acting on data having a known entropy level. Second, the actual entropy level of output produced by such a process is computed, thereby providing support for the selected value of ε.

2. Full Entropy Definition

The definition of full entropy is based on a distinguishing game where an adversary attempts to distinguish between two cases – REAL and IDEAL. Assume that the adversary is provided with W n-bit outputs $b_{1,}, b_2, \ldots, b_W$. In the REAL case, the outputs are generated by a conditioning function applied to a specified quantity of raw entropy data. In the IDEAL case, the outputs are generated by an ideal randomness source. Each case has a probability of ½. n-bit outputs generated in the REAL case are defined as having *full entropy* with respect to W and δ (where $\delta > 0$) if the probability that a computationally unlimited adversary can correctly distinguish between the REAL and IDEAL cases is no more than $\frac{1}{2} + \delta$.

2.1. Derivation of Conditions for Full Entropy

Suppose that random output is generated by processing a quantity of entropy data using a conditioning function. The first result following from the above definition is that given values of W and δ, it is possible to find a threshold such that if the min-entropy of the input to the conditioning function meets or exceeds that threshold, the conditioning function output will satisfy the above definition of full entropy.

Let $B = \{b_{1,}, b_2, \ldots, b_W\}$ be the set of observed n-bit outputs and consider the likelihood ratio $\frac{Pr[REAL|B]}{Pr[IDEAL|B]}$. Clearly, the adversary will conclude that B was produced by the REAL case if this likelihood ratio is greater than one and by the IDEAL case otherwise. Since the REAL and IDEAL cases are equally likely, we can rewrite this likelihood ratio as $\frac{Pr[B|\text{REAL}]}{Pr[B|\text{IDEAL}]}$ using Bayes Theorem. For ease of computation, compute the base-2 log of the likelihood ratio and denote the resulting statistic as X. The adversary will conclude that B was produced by the REAL case if $X > 0$ and by the IDEAL case otherwise. If p_j denotes the probability of the j^{th} possible output from the conditioning function applied to the specified quantity of raw entropy data, so that p_{b_i} denotes the probability of the i^{th} observed output in the REAL case, the following is true:

$$X = \log_2\left(\frac{Pr[B|\text{REAL}]}{Pr[B|\text{IDEAL}]}\right)$$
$$= \log_2(Pr[B|\text{REAL}]) - \log_2(Pr[B|\text{IDEAL}])$$
$$= \log_2\left(\prod_{i=1}^{W} p_{b_i}\right) - \log_2(2^{-nW})$$
$$= \sum_{i=1}^{W}\left(n + \log_2 p_{b_i}\right)$$

The statistic X is a random variable that depends on the set B of observed n-bit outputs b_i and the probabilities p_{b_i} of those outputs in the REAL case. To assess the adversary's distinguishing success probability, the probability distribution of X in both the REAL and IDEAL cases is required. Note that X is the sum of W individual random variables $x_i = n + \log_2 p_{b_i}$. We will assume that these variables, being determined by the generation of independent outputs b_i, are independent and identically distributed. (In the IDEAL case, this assumption is clearly valid. In the REAL case, it is a reasonable assumption given the generation of the outputs $b_i$from separate entropy source sequences.) As determined below, an appropriate value of W for our purposes is 2^{48}. It is reasonable to assume that this value of W is sufficiently large to satisfy the Central Limit Theorem, so X is approximately normally distributed.

In the distinguishing scenario, the adversary has complete knowledge of the conditioning function and its input space, and therefore, being computationally unlimited, can determine the REAL case output probabilities p_j. These probabilities are determined by the interaction between the conditioning function used and the space of possible inputs to that function. For the purposes of this analysis, these probabilities cannot be precisely determined. However, it is possible and useful to consider the p_j as random variables rather than fixed values and use statistics associated with these random variables to find the probability distribution of X. The characteristics of the entropy source and the selected length of the entropy source sequences input to the conditioning function effectively result in a selection from a large number of possible input spaces for the conditioning function, each having a different set of probabilities for the input values. Since the conditioning function was designed to obscure any dependencies between inputs and outputs, there is no simple relationship between the output probabilities resulting from the many different input spaces. It is therefore reasonable to treat the conditioning function output probabilities p_j as random variables.

Consider p_j, treated as a random variable. Suppose that there are M possible inputs to the conditioning function, with probabilities $\{q_1, q_2, \ldots, q_M\}$. (Note that no assumptions are made on the input probability distribution.) This analysis treats the conditioning function as a mapping that uniformly assigns an n-bit output to each input in the input space so that, *a priori*, any specific output value is assigned to a given input value with probability 2^{-n} (note that multiple input values can be assigned a given output value). The output probability p_j can then be written as, $p_j = \sum_{i=1}^{M} q_i I_{i,j}$, where $I_{i,j} = 1$ if the conditioning function maps the i^{th} input to the j^{th} output, and $I_{i,j} = 0$ otherwise. Then $E[p_j] = \sum_{i=1}^{M} q_i E[I_{i,j}] = \sum_{i=1}^{M} 2^{-n} q_i = 2^{-n}$. Similarly,

$$VAR[p_j] = \sum_{i=1}^{M} VAR[q_i I_{i,j}]$$

$$= \sum_{i=1}^{M} \left(E\left[(q_i I_{i,j})^2 \right] - \left(E[q_i I_{i,j}] \right)^2 \right)$$

$$= \sum_{i=1}^{M} (2^{-n} q_i^2 - 2^{-2n} q_i^2)$$

$$= (2^{-n} - 2^{-2n}) \sum_{i=1}^{M} q_i^2$$

The value of M, the number of possible inputs to the conditioning function and the number of terms in this sum, is dependent on the characteristics of the entropy-source outputs and the conditioning function input bit length used. However, it will be determined below that in order to satisfy the definition of full entropy specified above, the input min-entropy H must be such that is that $H \geq n + 64$. Therefore, M must be at least 2^{n+64}. It is reasonable to assume that this is large enough to satisfy the Central Limit Theorem, so that p_j, being the sum of this large number of individual random variables $q_i I_{i,j}$, is approximately normally distributed. Now write p_j as $p_j = 2^{-n}(1 + \theta_j)$. Then $\theta_j = 2^n p_j - 1$, so $E[\theta_j] = 2^n E[p_j] - 1 = 0$ and $VAR[\theta_j] = 2^{2n} VAR[p_j] = (2^n - 1) \sum_{i=1}^{M} q_i^2$. Since the input collision entropy $H_2 = -\log_2 \sum_{i=1}^{M} q_i^2$, $VAR[\theta_j] = (2^n - 1) 2^{-H_2}$ holds. Note that $\theta_j = 2^n p_j - 1$ is also approximately normally distributed.

The mean and variance of X depend on whether the source is REAL or IDEAL. Let $\mu_R = E[x_i|REAL]$, $\mu_I = E[x_i|\text{IDEAL}]$, $\sigma_R^2 = VAR[x_i|\text{REAL}]$, and $\sigma_I^2 = VAR[x_i|\text{IDEAL}]$.

Now derive μ_R, μ_I, σ_R^2, and σ_I^2. Each of these values is computed by summing over the relevant expression using 2^{-n} or p_j as the probability weighting factors for the IDEAL and REAL cases, respectively. Thus,

$$E[x_i|IDEAL] = E\left[n + \log_2 p_{b_i} | \text{IDEAL}\right]$$

$$= \sum_{j=1}^{2^n} (n + \log_2 p_j) 2^{-n}$$

$$= \sum_{j=1}^{2^n} \left(n + \frac{\ln\left(2^{-n}(1+\theta_j)\right)}{\ln 2} \right) 2^{-n}$$

$$= \sum_{j=1}^{2^n} \frac{\ln(1+\theta_j)}{\ln 2} 2^{-n}$$

The Taylor series for $\ln(1+\theta_j)$ is $\theta_j - \frac{{\theta_j}^2}{2} + \frac{{\theta_j}^3}{3} - \frac{{\theta_j}^4}{4} + \cdots$. In Section 2.2. below, it is shown that for cases of interest, $|\theta_j|$ is on the order of 10^{-8} or smaller. For such values of θ_j, $\ln(1+\theta_j) \cong$

θ_j, and it can be shown that if the terms beyond ${\theta_j}^2$ are omitted, the relative error in $\ln\left(1+\theta_j\right)$ is on the order of 10^{-16}. The sum above is therefore approximately

$$\sum_{j=1}^{2^n} \frac{\theta_j - \frac{\theta_j^2}{2}}{\ln 2} 2^{-n} = \frac{1}{\ln 2} \frac{\sum_{j=1}^{2^n} \theta_j}{2^n} - \frac{1}{2\ln 2} \frac{\sum_{j=1}^{2^n} \theta_j^2}{2^n}.$$

The first sum in this expression is zero by definition of θ_j. To evaluate the second sum, note that the sum is computed over the 2^n values of θ_j. Each of these 2^n values can be considered as a specific value of the corresponding random variable. Since these random variables have the same distribution, the 2^n values can also be treated as a sample of any one of these random variables. By definition, $VAR[\theta_j] = E[\theta_j^2] - E[\theta_j]^2$. The term $\frac{\sum_{j=1}^{2^n} \theta_j^2}{2^n}$ is the sample mean of θ_j^2 and is, therefore, approximately $VAR\left[\theta_j\right] + E[\theta_j]^2$. Substituting the values of $E\left[\theta_j\right]$ and $VAR\left[\theta_j\right]$ found above, the following is obtained:

$$\begin{aligned} E[x_i|\text{IDEAL}] &\cong -\frac{1}{2\ln 2}\left(VAR\left[\theta_j\right] + E\left[\theta_j\right]^2\right) \\ &= -\frac{1}{2\ln 2}(2^n - 1)2^{-H_2} \end{aligned}$$

The derivation of $E[x_i|\text{REAL}]$ is similar and is as follows.

$$\begin{aligned} E[x_i|\text{REAL}] &= E\left[n + \log_2 p_{b_i}|\text{REAL}\right] \\ &= \sum_{j=1}^{2^n} \left(n + \log_2 p_j\right) p_j \\ &= \sum_{j=1}^{2^n} \left(n + \frac{\ln\left(2^{-n}\left(1+\theta_j\right)\right)}{\ln 2}\right) p_j \\ &= \sum_{j=1}^{2^n} \frac{\ln\left(1+\theta_j\right)}{\ln 2} p_j \\ &= \sum_{j=1}^{2^n} \frac{\ln\left(1+\theta_j\right)}{\ln 2} 2^{-n}\left(1+\theta_j\right) \\ &\cong \sum_{j=1}^{2^n} \frac{\theta_j - \frac{\theta_j^2}{2}}{\ln 2} 2^{-n}\left(1+\theta_j\right) \\ &\cong \frac{1}{\ln 2} \frac{\sum_{j=1}^{2^n} \theta_j}{2^n} + \frac{1}{2\ln 2} \frac{\sum_{j=1}^{2^n} \theta_j^2}{2^n} \\ &\cong \frac{1}{2\ln 2}\left(VAR\left[\theta_j\right] + E\left[\theta_j\right]^2\right) \\ &= \frac{1}{2\ln 2}(2^n - 1)2^{-H_2} \end{aligned}$$

Reusing portions of these calculations, the variance of x_i in the IDEAL case is obtained as follows:

$$
\begin{aligned}
VAR[x_i|\text{IDEAL}] &= E\left[\left(n+\log_2 p_{b_i}\right)^2|\text{IDEAL}\right] - E\left[n+\log_2 p_{b_i}\,|\text{IDEAL}\right]^2 \\
&\cong \sum_{j=1}^{2^n}\left(\frac{\theta_j-\frac{\theta_j^2}{2}}{\ln 2}\right)^2 2^{-n} - \left(-\frac{1}{2\ln 2}(2^n-1)2^{-H_2}\right)^2 \\
&\cong \frac{1}{(\ln 2)^2}\frac{\sum_{j=1}^{2^n}\theta_j^2}{2^n} - \left(\frac{1}{2\ln 2}(2^n-1)2^{-H_2}\right)^2 \\
&= \frac{1}{(\ln 2)^2}(2^n-1)2^{-H_2} - \left(\frac{1}{2\ln 2}(2^n-1)2^{-H_2}\right)^2 \\
&= \frac{1}{(\ln 2)^2}(2^n-1)2^{-H_2}\left(1-\frac{1}{4}(2^n-1)2^{-H_2}\right)
\end{aligned}
$$

Similarly, the variance of x_i in the REAL case is obtained as follows:

$$
\begin{aligned}
VAR[x_i|\text{REAL}] &= E\left[\left(n+\log_2 p_{b_i}\right)^2|\text{REAL}\right] - E\left[n+\log_2 p_{b_i}\,|\text{REAL}\right]^2 \\
&\cong \sum_{j=1}^{2^n}\left(\frac{\theta_j-\frac{\theta_j^2}{2}}{\ln 2}\right)^2 2^{-n}\left(1+\theta_j\right) - \left(\frac{1}{2\ln 2}(2^n-1)2^{-H_2}\right)^2 \\
&\cong \frac{1}{(\ln 2)^2}\frac{\sum_{j=1}^{2^n}\theta_j^2}{2^n} - \left(\frac{1}{2\ln 2}(2^n-1)2^{-H_2}\right)^2 \\
&= \frac{1}{(\ln 2)^2}(2^n-1)2^{-H_2} - \left(\frac{1}{2\ln 2}(2^n-1)2^{-H_2}\right)^2 \\
&= \frac{1}{(\ln 2)^2}(2^n-1)2^{-H_2}\left(1-\frac{1}{4}(2^n-1)2^{-H_2}\right)
\end{aligned}
$$

Note that for typical values of n, μ_I and μ_R are closely approximated as $-\frac{1}{2\ln 2}2^{n-H_2}$ and $\frac{1}{2\ln 2}2^{n-H_2}$, respectively. Also, assuming that H_2 will need to exceed n by at least a moderate amount in order to satisfy the definition of full entropy, $\sigma_I^2 = \sigma_R^2$ can be closely approximated as $\sigma^2 = \frac{1}{(\ln 2)^2}2^{n-H_2}$. The log likelihood ratio statistic X is therefore approximately normally distributed with means and variance as follows:

$$
E[X|\text{REAL}] = -E[X|\text{IDEAL}] = -W\mu_I \cong \frac{W}{2\ln 2}2^{n-H_2}
$$

$$
VAR[X|\text{REAL}] = VAR[X|\text{IDEAL}] = W\sigma^2 \cong \frac{W}{(\ln 2)^2}2^{n-H_2}
$$

Now consider the probability that the adversary correctly determines whether the REAL or IDEAL case produced the observed sample B. This probability is as follows:

$$\begin{aligned} Pr[\text{Correct}] &= Pr[\text{IDEAL}]Pr[\text{Correct}|\text{IDEAL}] + Pr[\text{REAL}]Pr[\text{Correct}|\text{REAL}] \\ &= \frac{1}{2}Pr[X < 0|\text{IDEAL}] + \frac{1}{2}Pr[X > 0|\text{REAL}] \end{aligned}$$

Note that because of the symmetry resulting from X having a normal distribution with variance $W\sigma^2$ in both the REAL and IDEAL cases and expected values that are negatives of each other in these two cases, $Pr[X < 0|\text{IDEAL}] = Pr[X > 0|\text{REAL}]$, which gives the following:

$$\begin{aligned} Pr[\text{Correct}] &= Pr[X < 0|\text{IDEAL}] \\ &= Pr\left[\frac{X - W\mu_I}{\sqrt{W\sigma^2}} < \frac{0 - W\mu_I}{\sqrt{W\sigma^2}}|\text{IDEAL}\right] \end{aligned}$$

Since in the IDEAL case, X is normally distributed with mean $W\mu_I$ and variance $W\sigma^2$, the value $z = \frac{X - W\mu_I}{\sqrt{W\sigma^2}}$ is a standard normal random variable, so this probability is $F\left(\frac{-W\mu_I}{\sqrt{W\sigma^2}}\right)$, where F is the CDF of the standard normal distribution. $F(x) \le \frac{1}{2} + \frac{1}{2}\sqrt{1 - e^{-2x^2/\pi}}$ when $x > 0$ (see Section 26.2.24 of [4]). Thus, $Pr[\text{Correct}] = F\left(\frac{-W\mu_I}{\sqrt{W\sigma^2}}\right) \le \frac{1}{2} + \delta$ if the following inequality is satisfied:

$$\frac{1}{2} + \frac{1}{2}\sqrt{1 - e^{-2\left(\frac{-W\mu_I}{\sqrt{W\sigma^2}}\right)^2/\pi}} \le \frac{1}{2} + \delta$$

From the derivations above, $\frac{-W\mu_I}{\sqrt{W\sigma^2}} = \frac{1}{2}\sqrt{W} \cdot 2^{\frac{n-H_2}{2}}$, giving the following sequence of inequalities:

$$\frac{1}{2}\sqrt{1 - e^{-2\left(\frac{1}{4}W \cdot 2^{n-H_2}\right)/\pi}} \le \delta$$

$$1 - e^{-\frac{1}{2}W \cdot 2^{n-H_2}/\pi} \le 4\delta^2$$

$$1 - 4\delta^2 \le e^{-\frac{1}{2}W \cdot 2^{n-H_2}/\pi}$$

$$\ln(1 - 4\delta^2) \le -\frac{1}{2}W \cdot 2^{n-H_2}/\pi$$

$$-2\pi \ln(1 - 4\delta^2) \ge W \cdot 2^{n-H_2}$$

$$\log_2(2\pi) + \log_2(-\ln(1 - 4\delta^2)) \ge \log_2 W + n - H_2$$

$$H_2 \ge n + \log_2 W - \log_2(2\pi) - \log_2(-\ln(1 - 4\delta^2))$$

Note that since collision-entropy H_2 is an upper bound on min-entropy H, the above inequality holds when H_2 is replaced by H. Thus, the inequality is as follows:

$$H \ge n + \log_2 W - \log_2(2\pi) - \log_2(-\ln(1 - 4\delta^2))$$

Since $4\delta^2 \cong 0$ when $\delta \cong 0$, $-\ln(1-4\delta^2)$ is closely approximated by $4\delta^2$, so the inequality can be written as:

$$H \geq n + \log_2\left(\frac{W}{\delta^2}\right) - (log_2\pi + 3)$$

The following table shows the minimum difference $H - n$ for various values of W and δ.

Table 1. Minimum value of $\boldsymbol{H - n}$ for various values of $\boldsymbol{W}$ and $\boldsymbol{\delta}$

W \ δ	2^{-20}	2^{-18}	2^{-16}	2^{-14}	2^{-12}	2^{-10}	2^{-8}
2^{32}	67.3	63.3	59.3	55.3	51.3	47.3	43.3
2^{40}	75.3	71.3	67.3	63.3	59.3	55.3	51.3
2^{48}	83.3	79.3	75.3	71.3	67.3	63.3	59.3
2^{56}	91.3	87.3	83.3	79.3	75.3	71.3	67.3

It is assumed in SP 800-90C that there is an upper bound of 2^{64} bits on the amount of output that an adversary attempting a distinguishing attack can request. Consider the combination $W = 2^{48}$ and $\delta = 2^{-10}$. Given $W = 2^{48}$ n-bit RBG outputs, each output can be up to $2^{16} = 65536$ bits long without exceeding the 2^{64} data-quantity bound. Note that 10 000 random bit generators, each producing 1000 outputs per second, would require nearly a year to produce $W = 2^{48}$ outputs. According to the table above, an adversary who obtains $W = 2^{48}$ n-bit outputs has a distinguishing probability no greater than $\frac{1}{2} + \delta = \frac{1}{2} + 2^{-10} \cong 0.501$ when H, the conditioning function input min-entropy for each n-bit output, is at least $n + 63.3$. This minimum value, rounded up to $n + 64$, is used in this document as the condition for satisfying the full entropy definition.

2.2. Justification of Claim on $\boldsymbol{\theta_j}$

In order to derive the conditions for full entropy, sums of powers of θ_j higher than θ_j^2 were omitted. This did not affect the validity of the conclusion if θ_j is sufficiently near zero. This is established as follows. Recall that there are 2^n values of θ_j, each of which is approximately normally distributed with mean zero and variance approximately 2^{n-H_2}. Consider the largest θ_j, $\theta_{max} = max_j\{\theta_j\}$. θ_{max} is $z = \frac{\theta_{max}}{2^{\frac{n-H_2}{2}}}$ standard deviations away from zero, which is the mean of θ_j. The value of z is expected to be such that in a collection of 2^n standard normal random variables, approximately one is greater than or equal to this value of z. If $f(z)$ and $F(z)$ are the density function and the CDF of the standard normal distribution, respectively, then for large z, $1 - F(z) \cong \frac{f(z)}{z}$ (see Section 26.2.12 of [4]). The desired value of z, therefore, gives $(1 - F(z))2^n \cong$

1, which leads to $\frac{2^n}{z\sqrt{2\pi}}e^{-\frac{z^2}{2}} = 1$, or $z^2 + 2\ln z = 2n\ln 2 - \ln(2\pi)$. Since z^2 dominates the left side of this equation, the desired value of z is approximately $\sqrt{2n\ln 2 - \ln(2\pi)}$. The value of θ_{max} is then expected to be approximately $2^{\frac{n-H_2}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}$. For any of the typical values of n and a value of H_2 given by the lower bound computation above, $H_2 \geq n + 64$, so $2^{\frac{n-H_2}{2}} \leq 2^{-32}$, and it can be calculated that θ_{max} is a positive value that with high likelihood is less than 10^{-8}. A similar argument leads to θ_{min} being approximately $-\theta_{max}$, so it is expected that $|\theta_j| \leq 10^{-8}$ for all j. Therefore, it is safe to omit powers of θ_j higher than θ_j^2, since it is shown in Section 2.2 that doing so has a negligible effect.

2.3. Derivation of Full Entropy Threshold

The second result following from the above definition of full entropy is the derivation of an estimate of the min-entropy of an n-bit output, given that the input to the conditioning function has a collision entropy of H_2. The above result gives $\theta_{max} \cong 2^{\frac{n-H_2}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}$, which implies that the corresponding value $p_{max} = max_j\{p_j\}$ is approximately $2^{-n}\left(1 + 2^{\frac{n-H_2}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\right)$. If the min-entropy of the input to the conditioning function is H, then $H_2 \geq H$, so

$$p_{max} \leq 2^{-n}\left(1 + 2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\right).$$

The output min-entropy corresponding to this value of p_{max} is:

$$-\log_2 p_{max} \geq n - \log_2\left(1 + 2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\right)$$

$$= n - \frac{ln\left(1 + 2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\right)}{ln2}$$

Since $H \geq n + 64$, $2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}$ is a very small positive number, so $ln\left(1 + 2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\right) \cong 2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}$, giving

$$-\log_2 p_{max} \geq n - \frac{2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}}{\ln 2}$$

Dividing this value by n gives an average per-bit min-entropy of at least

$$1 - \frac{2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}}{nln2}$$

When $H \geq n + 64$, a per-bit entropy of at least $1 - 2^{-32}c$ is obtained, where $0 < c < 1$ for all the values of n of interest. Therefore, when $H \geq n + 64$, the average per-bit min-entropy in the n-bit conditioning function output is at least $1 - 2^{-32}$.

References

[1] Barker EB, Kelsey JM (2015) *Recommendation for Random Number Generation Using Deterministic Random Bit Generato*rs. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-90Ar1. https://doi.org/10.6028/NIST.SP.800-90Ar1

[2] Sönmez Turan M, Barker EB, Kelsey JM, McKay KA, Baish ML, Boyle M (2018*) Recommendation for the Entropy Sources Used for Random Bit Generati*on. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-90B. https://doi.org/10.6028/NIST.SP.800-90B

[3] Barker EB, Kelsey JM, McKay K, Roginsky A, Sönmez Turan M (2022) *Recommendation for Random Bit Generator (RBG) Constructions.* (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-90C 3pd, Third public draft. https://doi.org/10.6028/NIST.SP.800-90C.3pd

[4] Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables, Milton Abramowitz and Irene A. Stegun, editors, Dover Publications, New York, 1972.

Appendix A. List of Symbols, Abbreviations, and Acronyms

CDF
Cumulative Distribution Function

NIST
National Institute of Standards and Technology

RBG
Random Bit Generator

SP
(NIST) Special Publication

0^x
A string of *x* zeroes

$\lceil x \rceil$
The ceiling of *x*; the least integer number that is not less than the real number *x*. For example, $\lceil 3 \rceil = 3$, and $\lceil 5.5 \rceil = 6$.

ε
A positive constant that is assumed to be smaller than 2^{-32}

E(*X*)
The expected value of the random variable *X*

$\mathrm{Log}_2(x)$
Base-2 logarithm of *X*

Ln(x)
Natural logarithm of *X*

Var(x)
Variance of random variable *X*

Appendix B. Glossary

adversary
A malicious entity whose goal is to determine, to guess, or to influence the output of an RBG.

bitstring
An ordered sequence (string) of 0s and 1s. The leftmost bit is the most significant bit.

conditioning function
A deterministic function used to reduce bias and/or improve the entropy per bit.

cryptographic boundary
An explicitly defined physical or conceptual perimeter that establishes the physical and/or logical bounds of a cryptographic module and contains all of the hardware, software, and/or firmware components of a cryptographic module.

entropy
A measure of the randomness or uncertainty of a random variable.

entropy source
The combination of a noise source, health tests, and optional conditioning component that produce bitstrings containing entropy. A distinction is made between entropy sources having physical noise sources and those having non-physical noise sources.

full-entropy bitstring
A bitstring with ideal randomness (i.e., the amount of entropy per bit is equal to 1). This Recommendation assumes that a bitstring has *full entropy* if the entropy rate is at least $1 - \varepsilon$, where ε is at most 2^{-32}.

ideal randomness source
The source of an ideal random sequence of bits. Each bit of an ideal random sequence is unpredictable and unbiased, with a value that is independent of the values of the other bits in the sequence. Prior to an observation of the sequence, the value of each bit is equally likely to be 0 or 1, and the probability that a particular bit will have a particular value is unaffected by knowledge of the values of any or all of the other bits. An ideal random sequence of n bits contains n bits of entropy.

likelihood ratio test
A statistical test aimed at distinguishing between two competing models that could have produced an observed event based on a comparison of the likelihoods of the observed event, given the two models.

min-entropy
A lower bound on the entropy of a random variable. The precise formulation for min-entropy is $(-\log_2 \max p_i)$ for a discrete distribution having probabilities $p_1, ..., p_k$. Min-entropy is often used as a measure of the unpredictability of a random variable.