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

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

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

September 2022

37   Certain commercial entities, equipment, or materials may be identified in this document in order to describe an
38   experimental procedure or concept adequately. Such identification is not intended to imply recommendation or
39   endorsement by the National Institute of Standards and Technology (NIST), nor is it intended to imply that the
40   entities, materials, or equipment are necessarily the best available for the purpose.

41   There may be references in this publication to other publications currently under development by NIST in
42   accordance with its assigned statutory responsibilities. The information in this publication, including concepts and
43   methodologies, may be used by federal agencies even before the completion of such companion publications. Thus,
44   until each publication is completed, current requirements, guidelines, and procedures, where they exist, remain
45   operative. For planning and transition purposes, federal agencies may wish to closely follow the development of
46   these new publications by NIST.

47   Organizations are encouraged to review all draft publications during public comment periods and provide feedback
48   to NIST. Many NIST cybersecurity publications, other than the ones noted above, are available at
49   https://csrc.nist.gov/publications.

50   **NIST Technical Series Policies**
51   Copyright, Fair Use, and Licensing Statements
52   NIST Technical Series Publication Identifier Syntax

59   **Author ORCID iDs**
60   Author 1: 0000-0000-0000-0000
61   Author 2: 0000-0000-0000-0000
62   Author 3: 0000-0000-0000-0000
63   Author 4: 0000-0000-0000-0000

64   **Public Comment Period**
65   September 7, 2022 – October 31, 2022

66   **Submit Comments**
67   rbg_comments@nist.gov
68
69   National Institute of Standards and Technology
70   Attn: Computer Security Division, Information Technology Laboratory
71   100 Bureau Drive (Mail Stop 8930) Gaithersburg, MD 20899-8930

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

73 **Reports on Computer Systems Technology**

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

82 **Abstract**

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

97 **Keywords**

98 entropy; min-entropy; random number generation.

99

100

101

102 **Table of Contents**

111 **List of Tables**

113

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 $\frac{Pr[B|REAL]}{Pr[B|IDEAL]}$ 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 jth 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 ith observed output in the REAL case, the following is true:

154
$$X = \log_2\left(\frac{Pr[B|\text{REAL}]}{Pr[B|\text{IDEAL}]}\right)$$

155
$$= \log_2(Pr[B|\text{REAL}]) - \log_2(Pr[B|\text{IDEAL}])$$

156
$$= \log_2\left(\prod_{i=1}^{W} p_{b_i}\right) - \log_2(2^{-nW})$$

157
$$= \sum_{i=1}^{W}\left(n + \log_2 p_{b_i}\right)$$

158 The statistic X is a random variable that depends on the set B of observed n-bit outputs b_i and the
159 probabilities p_{b_i} of those outputs in the REAL case. To assess the adversary's distinguishing
160 success probability, the probability distribution of X in both the REAL and IDEAL cases is
161 required. Note that X is the sum of W individual random variables $x_i = n + \log_2 p_{b_i}$. We will
162 assume that these variables, being determined by the generation of independent outputs b_i, are
163 independent and identically distributed. (In the IDEAL case, this assumption is clearly valid. In
164 the REAL case, it is a reasonable assumption given the generation of the outputs b_i from separate
165 entropy source sequences.) As determined below, an appropriate value of W for our purposes is
166 2^{48}. It is reasonable to assume that this value of W is sufficiently large to satisfy the Central Limit
167 Theorem, so X is approximately normally distributed.

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

181 Consider p_j, treated as a random variable. Suppose that there are M possible inputs to the
182 conditioning function, with probabilities $\{q_1, q_2, \ldots, q_M\}$. (Note that no assumptions are made on
183 the input probability distribution.) This analysis treats the conditioning function as a mapping that
184 uniformly assigns an n-bit output to each input in the input space so that, *a priori*, any specific
185 output value is assigned to a given input value with probability 2^{-n} (note that multiple input values
186 can be assigned a given output value). The output probability p_j can then be written as, $p_j =$
187 $\sum_{i=1}^{M} q_i I_{i,j}$, where $I_{i,j} = 1$ if the conditioning function maps the ith input to the jth output, and $I_{i,j} =$
188 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,

189
$$VAR[p_j] = \sum_{i=1}^{M} VAR[q_i I_{i,j}]$$

190
$$= \sum_{i=1}^{M} \left(E\left[(q_i I_{i,j})^2\right] - (E[q_i I_{i,j}])^2 \right)$$

191
$$= \sum_{i=1}^{M} (2^{-n} q_i^2 - 2^{-2n} q_i^2)$$

192
$$= (2^{-n} - 2^{-2n}) \sum_{i=1}^{M} q_i^2$$

193 The value of M, the number of possible inputs to the conditioning function and the number of
194 terms in this sum, is dependent on the characteristics of the entropy-source outputs and the
195 conditioning function input bit length used. However, it will be determined below that in order to
196 satisfy the definition of full entropy specified above, the input min-entropy H must be such that is
197 that $H \geq n + 64$. Therefore, M must be at least 2^{n+64}. It is reasonable to assume that this is large
198 enough to satisfy the Central Limit Theorem, so that p_j, being the sum of this large number of
199 individual random variables $q_i I_{i,j}$, is approximately normally distributed. Now write p_j as $p_j =$
200 $2^{-n}(1 + \theta_j)$. Then $\theta_j = 2^n p_j - 1$, so $E[\theta_j] = 2^n E[p_j] - 1 = 0$ and $VAR[\theta_j] = 2^{2n} VAR[p_j] =$
201 $(2^n - 1) \sum_{i=1}^{M} q_i^2$. Since the input collision entropy $H_2 = -\log_2 \sum_{i=1}^{M} q_i^2$, $VAR[\theta_j] = (2^n -$
202 $1)2^{-H_2}$ holds. Note that $\theta_j = 2^n p_j - 1$ is also approximately normally distributed.

203 The mean and variance of X depend on whether the source is REAL or IDEAL. Let $\mu_R =$
204 $E[x_i|REAL]$, $\mu_I = E[x_i|IDEAL]$, $\sigma_R^2 = VAR[x_i|REAL]$, and $\sigma_I^2 = VAR[x_i|IDEAL]$.

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

208
$$E[x_i|IDEAL] = E\left[n + \log_2 p_{b_i}|IDEAL\right]$$

209
$$= \sum_{j=1}^{2^n} (n + \log_2 p_j) 2^{-n}$$

210
$$= \sum_{j=1}^{2^n} \left(n + \frac{\ln\left(2^{-n}(1 + \theta_j)\right)}{\ln 2} \right) 2^{-n}$$

211
$$= \sum_{j=1}^{2^n} \frac{\ln(1 + \theta_j)}{\ln 2} 2^{-n}$$

212 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
213 that for cases of interest, $|\theta_j|$ is on the order of 10^{-8} or smaller. For such values of θ_j, $\ln(1 + \theta_j) \cong$

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

216
$$\sum_{j=1}^{2^n} \frac{\theta_j - \frac{\theta_j^2}{2}}{\ln 2} 2^{-n} = \frac{1}{\ln 2} \frac{\sum_{j=1}^{2^n} \theta_j}{2^n} - \frac{1}{2\ln 2} \frac{\sum_{j=1}^{2^n} \theta_j^2}{2^n}.$$

217 The first sum in this expression is zero by definition of θ_j. To evaluate the second sum, note that
218 the sum is computed over the 2^n values of θ_j. Each of these 2^n values can be considered as a
219 specific value of the corresponding random variable. Since these random variables have the same
220 distribution, the 2^n values can also be treated as a sample of any one of these random variables.

221 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,
222 therefore, approximately $VAR[\theta_j] + E[\theta_j]^2$. Substituting the values of $E[\theta_j]$ and $VAR[\theta_j]$ found
223 above, the following is obtained:

224
$$E[x_i | IDEAL] \cong -\frac{1}{2\ln 2}\left(VAR[\theta_j] + E[\theta_j]^2\right)$$

225
$$= -\frac{1}{2\ln 2}(2^n - 1)2^{-H_2}$$

226 The derivation of $E[x_i | REAL]$ is similar and is as follows.

227
$$E[x_i | REAL] = E\left[n + \log_2 p_{b_i} | REAL\right]$$

228
$$= \sum_{j=1}^{2^n}(n + \log_2 p_j)p_j$$

229
$$= \sum_{j=1}^{2^n}\left(n + \frac{\ln\left(2^{-n}(1 + \theta_j)\right)}{\ln 2}\right)p_j$$

230
$$= \sum_{j=1}^{2^n} \frac{\ln(1 + \theta_j)}{\ln 2} p_j$$

231
$$= \sum_{j=1}^{2^n} \frac{\ln(1 + \theta_j)}{\ln 2} 2^{-n}(1 + \theta_j)$$

232
$$\cong \sum_{j=1}^{2^n} \frac{\theta_j - \frac{\theta_j^2}{2}}{\ln 2} 2^{-n}(1 + \theta_j)$$

233
$$\cong \frac{1}{\ln 2} \frac{\sum_{j=1}^{2^n} \theta_j}{2^n} + \frac{1}{2\ln 2} \frac{\sum_{j=1}^{2^n} \theta_j^2}{2^n}$$

234
$$\cong \frac{1}{2\ln 2}\left(VAR[\theta_j] + E[\theta_j]^2\right)$$

235
$$= \frac{1}{2\ln 2}(2^n - 1)2^{-H_2}$$

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

237
$$VAR[x_i|\text{IDEAL}] = E\left[\left(n + \log_2 p_{b_i}\right)^2 |\text{IDEAL}\right] - E\left[n + \log_2 p_{b_i} |\text{IDEAL}\right]^2$$

238
$$\cong \sum_{j=1}^{2^n} \left(\frac{\theta_j - \frac{\theta_j^2}{2}}{\ln 2}\right)^2 2^{-n} - \left(-\frac{1}{2\ln 2}(2^n - 1)2^{-H_2}\right)^2$$

239
$$\cong \frac{1}{(\ln 2)^2} \frac{\sum_{j=1}^{2^n} \theta_j^2}{2^n} - \left(\frac{1}{2\ln 2}(2^n - 1)2^{-H_2}\right)^2$$

240
$$= \frac{1}{(\ln 2)^2}(2^n - 1)2^{-H_2} - \left(\frac{1}{2\ln 2}(2^n - 1)2^{-H_2}\right)^2$$

241
$$= \frac{1}{(\ln 2)^2}(2^n - 1)2^{-H_2}\left(1 - \frac{1}{4}(2^n - 1)2^{-H_2}\right)$$

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

243
$$VAR[x_i|\text{REAL}] = E\left[\left(n + \log_2 p_{b_i}\right)^2 |\text{REAL}\right] - E\left[n + \log_2 p_{b_i} |\text{REAL}\right]^2$$

244
$$\cong \sum_{j=1}^{2^n} \left(\frac{\theta_j - \frac{\theta_j^2}{2}}{\ln 2}\right)^2 2^{-n}(1 + \theta_j) - \left(\frac{1}{2\ln 2}(2^n - 1)2^{-H_2}\right)^2$$

245
$$\cong \frac{1}{(\ln 2)^2} \frac{\sum_{j=1}^{2^n} \theta_j^2}{2^n} - \left(\frac{1}{2\ln 2}(2^n - 1)2^{-H_2}\right)^2$$

246
$$= \frac{1}{(\ln 2)^2}(2^n - 1)2^{-H_2} - \left(\frac{1}{2\ln 2}(2^n - 1)2^{-H_2}\right)^2$$

247
$$= \frac{1}{(\ln 2)^2}(2^n - 1)2^{-H_2}\left(1 - \frac{1}{4}(2^n - 1)2^{-H_2}\right)$$

248   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
249    $\frac{1}{2\ln 2}2^{n-H_2}$, respectively. Also, assuming that H_2 will need to exceed n by at least a moderate
250   amount in order to satisfy the definition of full entropy, $\sigma_I^2 = \sigma_R^2$ can be closely approximated as
251    $\sigma^2 = \frac{1}{(\ln 2)^2}2^{n-H_2}$. The log likelihood ratio statistic X is therefore approximately normally
252   distributed with means and variance as follows:

253
$$E[X|\text{REAL}] = -E[X|\text{IDEAL}] = -W\mu_I \cong \frac{W}{2\ln 2}2^{n-H_2}$$

254
$$VAR[X|\text{REAL}] = VAR[X|\text{IDEAL}] = W\sigma^2 \cong \frac{W}{(\ln 2)^2}2^{n-H_2}$$

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

257    $$Pr[\text{Correct}] = Pr[\text{IDEAL}]Pr[\text{Correct}|\text{IDEAL}] + Pr[\text{REAL}]Pr[\text{Correct}|\text{REAL}]$$

258    $$= \frac{1}{2}Pr[X < 0|\text{IDEAL}] + \frac{1}{2}Pr[X > 0|\text{REAL}]$$

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

262    $$Pr[\text{Correct}] = Pr[X < 0|\text{IDEAL}]$$

263    $$= Pr\left[\frac{X - W\mu_I}{\sqrt{W\sigma^2}} < \frac{0 - W\mu_I}{\sqrt{W\sigma^2}}\Big|\text{IDEAL}\right]$$

264   Since in the IDEAL case, X is normally distributed with mean $W\mu_I$ and variance $W\sigma^2$, the value
265   $z = \frac{X - W\mu_I}{\sqrt{W\sigma^2}}$ 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
266   CDF of the standard normal distribution. $F(x) \leq \frac{1}{2} + \frac{1}{2}\sqrt{1 - e^{-2x^2/\pi}}$ when $x > 0$ (see Section
267   26.2.24 of [4]). Thus, $Pr[\text{Correct}] = F\left(\frac{-W\mu_I}{\sqrt{W\sigma^2}}\right) \leq \frac{1}{2} + \delta$ if the following inequality is satisfied:

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

269   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:

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

271    $$1 - e^{-\frac{1}{2}W\cdot 2^{n-H_2}/\pi} \leq 4\delta^2$$

272    $$1 - 4\delta^2 \leq e^{-\frac{1}{2}W\cdot 2^{n-H_2}/\pi}$$

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

274    $$-2\pi \ln(1 - 4\delta^2) \geq W \cdot 2^{n-H_2}$$

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

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

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

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

280 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
281 can be written as:

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

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

284

285 Table 1. Minimum value of $H - n$ for various values of W and δ

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

286

287 It is assumed in SP 800-90C that there is an upper bound of 2^{64} bits on the amount of output that
288 an adversary attempting a distinguishing attack can request. Consider the combination $W = 2^{48}$
289 and $\delta = 2^{-10}$. Given $W = 2^{48}$ n-bit RBG outputs, each output can be up to $2^{16} = 65536$ bits
290 long without exceeding the 2^{64} data-quantity bound. Note that
291 10 000 random bit generators, each producing 1000 outputs per second, would require nearly a
292 year to produce $W = 2^{48}$ outputs. According to the table above, an adversary who obtains $W =$
293 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
294 H, the conditioning function input min-entropy for each n-bit output, is at least $n + 63.3$. This
295 minimum value, rounded up to $n + 64$, is used in this document as the condition for satisfying the
296 full entropy definition.

2.2. Justification of Claim on θ_j

298 In order to derive the conditions for full entropy, sums of powers of θ_j higher than θ_j^2 were omitted.
299 This did not affect the validity of the conclusion if θ_j is sufficiently near zero. This is established
300 as follows. Recall that there are 2^n values of θ_j, each of which is approximately normally
301 distributed with mean zero and variance approximately 2^{n-H_2}. Consider the largest θ_j, $\theta_{max} =$
302 $max_j\{\theta_j\}$. 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
303 value of z is expected to be such that in a collection of 2^n standard normal random variables,
304 approximately one is greater than or equal to this value of z. If $f(z)$ and $F(z)$ are the density
305 function and the CDF of the standard normal distribution, respectively, then for large z, $1 -$
306 $F(z) \cong \frac{f(z)}{z}$ (see Section 26.2.12 of [4]). The desired value of z, therefore, gives $(1 - F(z))2^n \cong$

307 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

308 side of this equation, the desired value of z is approximately $\sqrt{2n\ln 2 - \ln(2\pi)}$. The value of

309 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

310 of n and a value of H_2 given by the lower bound computation above, $H_2 \geq n + 64$, so $2^{\frac{n-H_2}{2}} \leq$

311 2^{-32}, and it can be calculated that θ_{max} is a positive value that with high likelihood is less than

312 10^{-8}. A similar argument leads to θ_{min} being approximately $-\theta_{max}$, so it is expected that $|\theta_j| \leq$

313 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

314 2.2 that doing so has a negligible effect.

2.3. Derivation of Full Entropy Threshold

316 The second result following from the above definition of full entropy is the derivation of an

317 estimate of the min-entropy of an n-bit output, given that the input to the conditioning function

318 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

319 implies that the corresponding value $p_{max} = max_j\{p_j\}$ is approximately $2^{-n}\Big(1 +$

320 $2^{\frac{n-H_2}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\Big)$. If the min-entropy of the input to the conditioning function is H, then

321 $H_2 \geq H$, so

322
$$p_{max} \leq 2^{-n}\left(1 + 2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\right).$$

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

324
$$-\log_2 p_{max} \geq n - \log_2\left(1 + 2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\right)$$

325
$$= n - \frac{ln\left(1 + 2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\right)}{ln2}$$

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

327 $2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}\Big) \cong 2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}$, giving

328
$$-\log_2 p_{max} \geq n - \frac{2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}}{ln2}$$

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

330
$$1 - \frac{2^{\frac{n-H}{2}}\sqrt{2n\ln 2 - \ln(2\pi)}}{nln2}$$

331 When $H \geq n + 64$, a per-bit entropy of at least $1 - 2^{-32}c$ is obtained, where $0 < c < 1$ for all

332 the values of n of interest. Therefore, when $H \geq n + 64$, the average per-bit min-entropy in the n-

333 bit conditioning function output is at least $1 - 2^{-32}$.

334

References

335

336 [1] Barker EB, Kelsey JM (2015) *Recommendation for Random Number Generation Using*
337 *Deterministic Random Bit Generato*rs. (National Institute of Standards and Technology,
338 Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-90Ar1.
339 https://doi.org/10.6028/NIST.SP.800-90Ar1
340 [2] Sönmez Turan M, Barker EB, Kelsey JM, McKay KA, Baish ML, Boyle M (2018*)*
341 *Recommendation for the Entropy Sources Used for Random Bit Generati*on. (National
342 Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP)
343 NIST SP 800-90B. https://doi.org/10.6028/NIST.SP.800-90B
344 [3] Barker EB, Kelsey JM, McKay K, Roginsky A, Sönmez Turan M (2022) *Recommendation*
345 *for Random Bit Generator (RBG) Constructions*. (National Institute of Standards and
346 Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-90C 3pd, Third
347 public draft. https://doi.org/10.6028/NIST.SP.800-90C.3pd
348 [4] Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables,
349 Milton Abramowitz and Irene A. Stegun, editors, Dover Publications, New York, 1972.
350

351 **Appendix A. List of Symbols, Abbreviations, and Acronyms**

352 **CDF**
353 Cumulative Distribution Function

354 **NIST**
355 National Institute of Standards and Technology

356 **RBG**
357 Random Bit Generator

358 **SP**
359 (NIST) Special Publication

360 **0^x**
361 A string of x zeroes

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

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

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

368 **Log$_2$(x)**
369 Base-2 logarithm of X

370 **Ln(x)**
371 Natural logarithm of X

372 **Var(x)**
373 Variance of random variable X

374

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.