Lessons Learned Validating 90B Entropy Sources

RBG Workshop
01 June 2023

Tim Hall (NIST)
Chris Celi (NIST)

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

Background

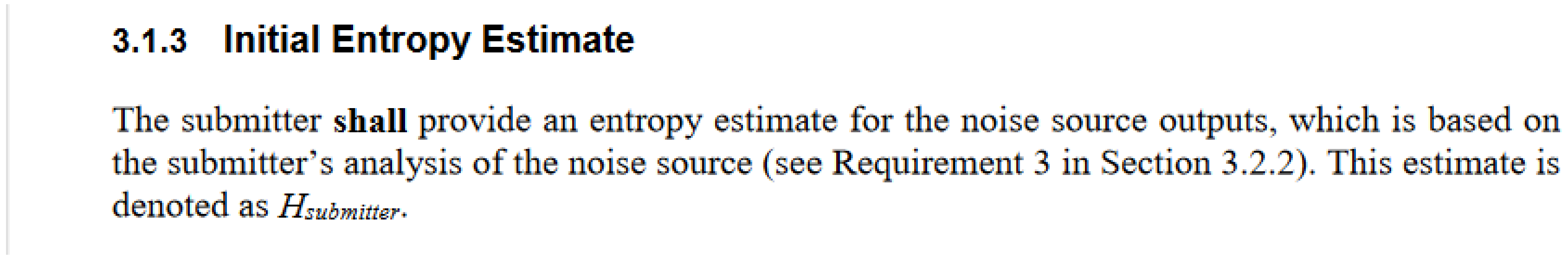
- Issues we have seen in NIST SP 800-90B Entropy Assessment Reports (EARs)
- Help entropy source vendors and CST labs in report preparation
- Inform guidance and future versions of standard
- Not a tutorial on how to address topics in an EAR

Topics

- Stochastic model or heuristic
- Digitization vs non-vetted conditioning
- Conditioning components
- More on health tests
- What about special cases?

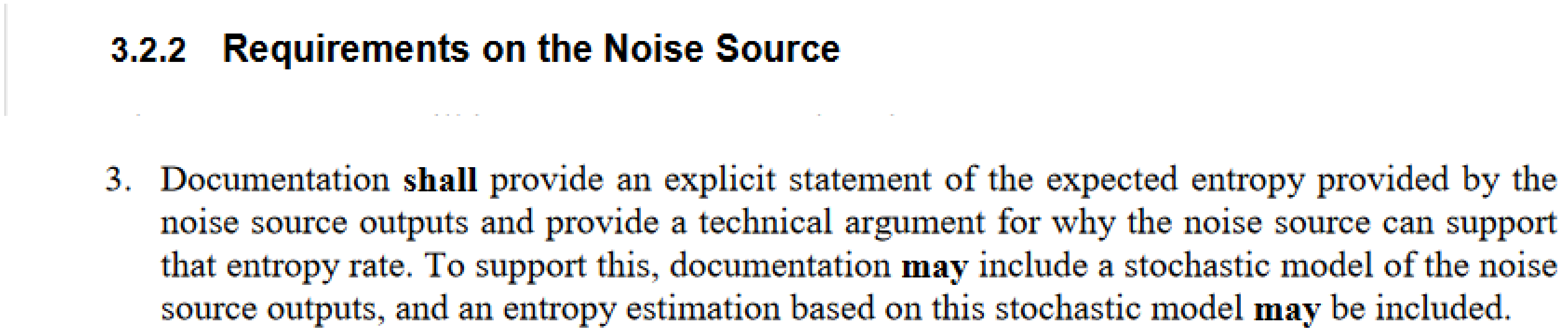
Stochastic model or heuristic

From SP 800-90B:



3.1.3 Initial Entropy Estimate

The submitter **shall** provide an entropy estimate for the noise source outputs, which is based on the submitter's analysis of the noise source (see Requirement 3 in Section 3.2.2). This estimate is denoted as $H_{submitter}$.



3.2.2 Requirements on the Noise Source

3. Documentation **shall** provide an explicit statement of the expected entropy provided by the noise source outputs and provide a technical argument for why the noise source can support that entropy rate. To support this, documentation **may** include a stochastic model of the noise source outputs, and an entropy estimation based on this stochastic model **may** be included.

Stochastic model or heuristic

- Must provide *technical support* for min-entropy estimate H_submitter
  - H_submitter: entropy estimate provided by the submitter
    - Based on the submitter’s analysis of the noise source
    - Not result of 90B Sec. 6 estimators (aka statistical tests)
  - Can be a lower bound

- If the report claims that the noise source output samples follow a well-known distribution, we like to see values of relevant parameters
  - Poisson – λ
  - Gaussian– μ and σ
  - Expect to see that parameter value used in derivation of H_submitter

Stochastic model or heuristic

- Cannot simply use a sample PMF from a source and perform a direct min-entropy calculation
  - Based on some sample of data, not analysis of the source and its design
    - You are already getting a version of this in Sec. 6.3.1 “The Most Common Value Estimate”
  - Requires some justification on why a particular sample PMF is representative of the true distribution

Digitization vs non-vetted conditioning

- SP 800-90B requires health tests to be run on raw noise samples

- Digitization vs non-vetted conditioning
    - ADC, D Flip-Flop
    - Input is digital, Processing (XORs, feedback, etc), compression

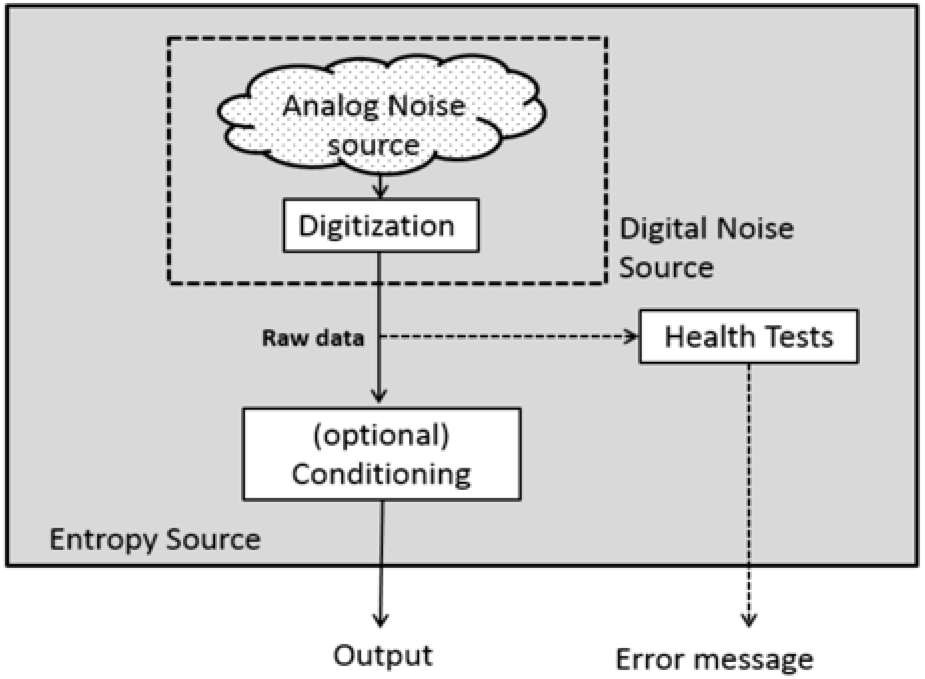
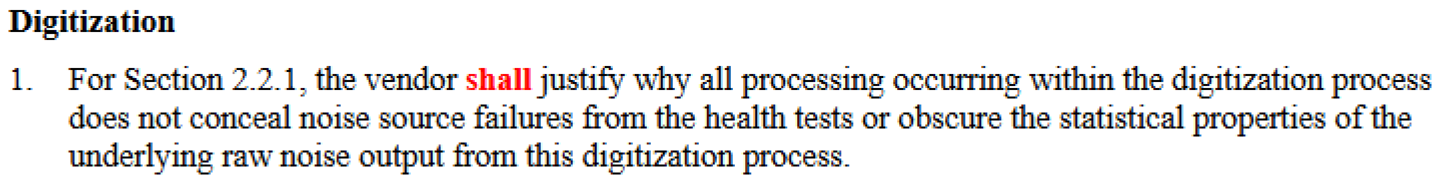


Figure 1 Entropy Source Model

Digitization vs non-vetted conditioning (2)

- FIPS 140-3 IG D.K, Resolution 1



Digitization

1. For Section 2.2.1, the vendor **shall** justify why all processing occurring within the digitization process does not conceal noise source failures from the health tests or obscure the statistical properties of the underlying raw noise output from this digitization process.

- CHTs can be run on output of non-vetted conditioner if report demonstrates that these tests catch the same errors as the APT and RCT on raw noise samples

- Can be shown by simulation or proof

DRBGs as Conditioning Components

- FIPS 140-3 IG D.K Resolutions 5 and 7
- Requires prediction resistance or reseed before each output
- Output number of bits <= to security strength of DRBG
- SP 800-90Ar1 DRBG considered vetted if meets IG D.K Res. 7 Note 1 (and IG D.K Res. 5)
  - Derivation function or at least *seedlen* bits of entropy
  - Seeded with h_in >= claimed security strength
  - Otherwise may be submitted as non-vetted

Full Entropy Output

- NIST IR 8427 (April 2023) “Discussion on the Full Entropy Assumption of the SP 800-90 Series”

- IG D.K Resolution 19 (March 2023)
    - Vetted conditioning component
    - h_in >= n_out + 64
    - n_out <= security strength of cryptographic function in conditioning component
    - Note: if n_in bits of *full entropy* provided, n_out maintains full entropy

- Bijective conditioning component, if full entropy input

Truncating output of a vetted conditioning component

- Outputs from a non-vetted conditioning component may not be truncated

- SP 800-90B Sec. 3.1.5.1.2 says

  "…it is acceptable to truncate the outputs from a vetted conditioning component. If this is done, *the entropy estimate is reduced to a proportion of the output*"

but that's not quite correct…

Truncating output of a vetted conditioning component (2)

- Appendix E defines the narrowest internal width
- Take case of SHA2-512 vetted conditioning component
  - n_in = 256 and h_in = 200
  - Output truncated to 256 bits
  - What is h_out?
    - Hint: it’s not 100
  - Using Output_Entropy() in Sec. 3.1.5.1.2, h_out = 199 to 200, depending on nw
- Will update this in IG

Developer-defined Health Tests

- Developer-defined health tests always permitted *in addition to* two approved tests (RCT and APT)
    - Explanation and rationale is always appreciated

- When used in place of RCT and/or APT, must meet the two criteria in SP 800-90B Section 4.5

- Also, provide “convincing evidence that the failure being considered will be reliably detected” by proof or simulation.

Entropy Estimate H for Health Tests

H used for APT and RCT, e.g.,

$$C = 1 + \left\lceil \frac{-\log_2 \alpha}{H} \right\rceil .$$

- Have seen H higher than and lower than H_i* used for CHTs
- If different than H_i, explain where it comes from and why it is used
- What is the effective alpha (false positive rate) using H_i?
- See Josh Hill’s presentation 30 May 2023 CMUF Entropy WG

* The initial entropy estimate of the noise source is calculated as $H_I = \min(H_{original},\ n \times H_{bitstring},\ H_{submitter})$ for non-binary sources and as $H_I = \min(H_{original},\ H_{submitter})$ for binary sources.

Special cases

- Sometimes existing designs do not line up neatly with structure from Figure 1 of SP 800-90B
  - Example from earlier: conditioning before CHTs

- We consider them on a case-by-case basis
  - Input from multiple reviewers and often SP 800-90B authors
  - Is there ambiguity or a gap in the standard?
  - Is there a legitimate security concern?

- Informs guidance and revision of standards

Questions?