Challenges in Generating, Transporting, and Verifying High-Entropy Sequences

Darren Hurley-Smith[1] & Julio Hernandez-Castro[2]

Abstract

Manufacturers make claims regarding RNG output randomness entropy based on statistical tests: NIST SP800-22, Dieharder, FIPS 140-2, and Ent being common tests. However, NIST and Dieharder tests require >1.5GB of data (often > 12Gb depending of configuration) and days of time to test reliably. Users therefore prefer the lightweight (deprecated) FIPS 140-2 tests. TestU01 provides an alternative, but is little known outside academic circles.

We identify where manufacturer claims fail to correctly quantify the randomness of their products, due to test limitations and misuse. We also identify how lightweight tests can be fooled by trivially biased sequences.

Methods

- FIPS 140-2 used to test 100 x 24kB biased sequences: sigma counter and epsilon hole
- IDQ Quantis biased identified using Chi-sq over 100 x 1GB samples
- DESFIRE EV1 bias identified using Chi-sq over 64 x 1MB samples
- EV1 bias modelled and characterised.

The Unbearable Lightness of FIPS 140-2

In collaboration with Constantinos Patsakis, we identified that overwriting an Independent and Identically Distributed (IID) sequence produced by urandom with a counter, with a byte being overwritten with probability sigma, resulted in sequences that FIPS 140-2 could not detect as non-random. For sigma <=0.62, FIPS 140-2 will identify biased, structured sequences as random. FIPS 140-2 is still used by manufacturers and end users even though it has been deprecated by NIST and is not used by BSI for any EAL evaluations involving RNG components.

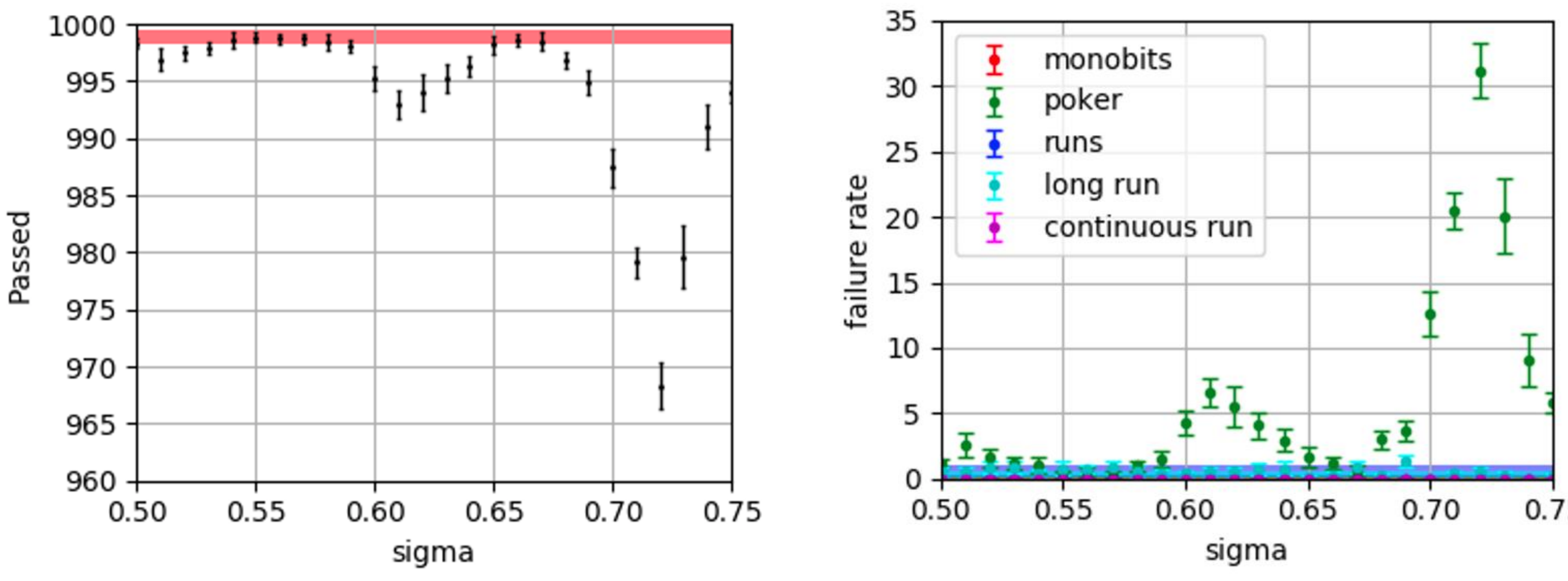

Fig.1 – FIPS 140-2 total pass rate (left) and individual test pass-rate (right) for sigma

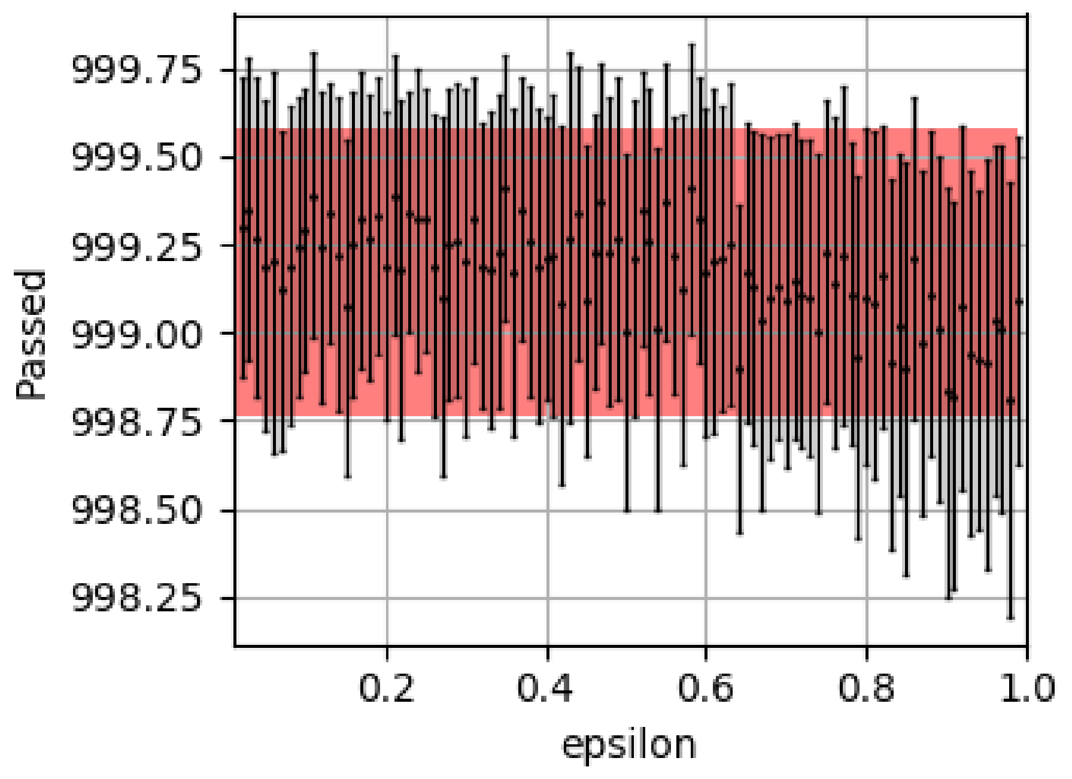

Fig.2 – FIPS 140-2 total pass rate for epsilon

For epsilon, where epsilon is the probability of a given byte value (int 255 here) is suppressed, FIPS 140-2 never falls below its confidence threshold (c=0.05). Though not as grossly effective at introducing bias as sigma, epsilon can be used to encode-by-omission, adding subtle structure that can be observed through informed frequency analysis.

Great Expectations: Bias Despite Manufacturer Testing & Claims

IDQ Quantis (now legacy since 2020) is an optical quantum entropy source with integrated RNG and postprocessor.

Despite passing Dieharder and SP800-22, Quantis devices express significant inherent thermal bias in their entropy source. User documentation claims post-processing is optional, when it is in fact required.

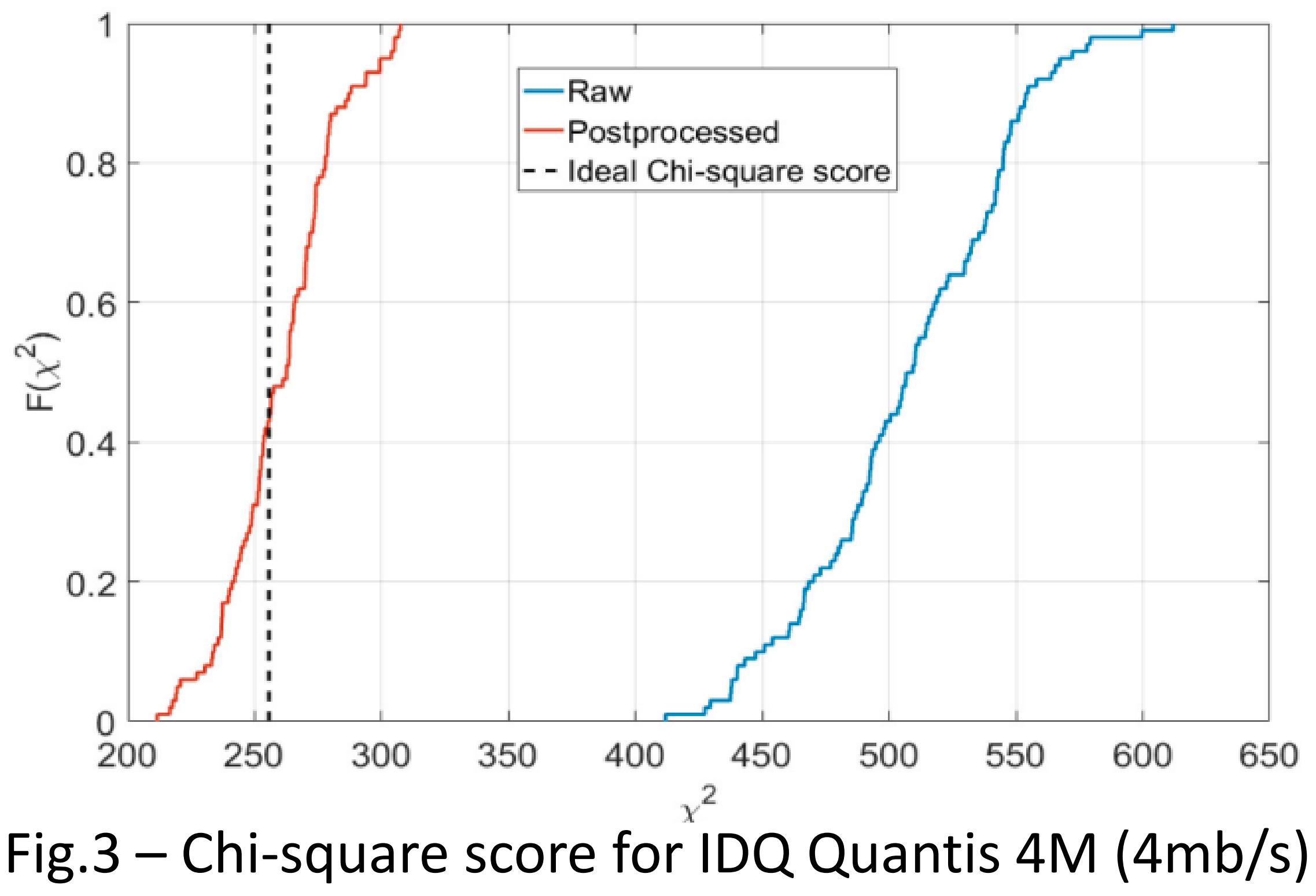

Fig.3 – Chi-square score for IDQ Quantis 4M (4mb/s) (mean 100 x 1GB sequences)

The Mifare DESFire Ev1 is a 2nd generation NXP smartcard, used as a programmable auth token and cash alternative. Oyster uses this chip.

DESFire EV1's integrated RNG passes NIST SP800-22 but trivial Chi-square tests identify underlying biases.

Tab.1 – DESFIRE EV1 NIST SP800-22 Results

Test	Card 1	Card 2	Card 3
Frequency	198/200	200/200	197/200
Block frequency	196/200	199/200	194/200
Cumulative sums	2/2	2/2	2/2
Longest run	196/200	198/200	198/200
Rank	198/200	199/200	197/200
FFT	197/200	199/200	198/200
Non-overlapping template	**147/148**	148/148	148/148
Overlapping template	198/200	198/200	198/200
Universal	198/200	198/200	198/200
Approximate entropy	197/200	198/200	196/200
Random excursions	8/8	8/8	8/8
Random excursions variant	18/18	18/18	18/18
Serial	2/2	2/2	2/2
Linear complexity	199/200	197/200	199/200

DESFire EV1 is rated EAL4+ by BSI, but displays predictable periodicity over multiple batches of cards. Key prediction isn't trivial, but entropy is decreased.

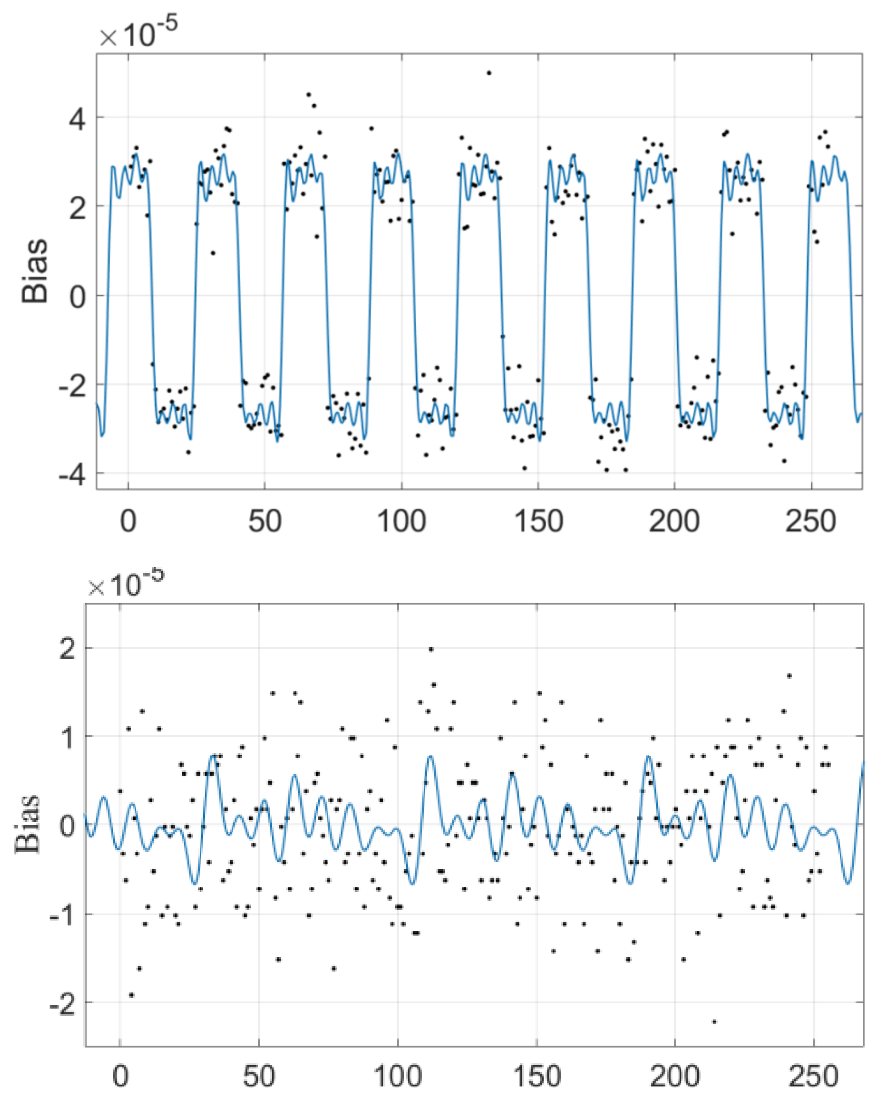

Fig.4 – DESFIRE EV1 (top) Byte value bias Compared with URANDOM (bottom).

Conclusions

- Over-reliance on statistical testing is a recognised issue in academic and advisory circles, but remains and issue for manufacturers and end-users
- Statistical tests that are both memory and time efficient cannot identify even trivial biases and encoded/steganographic content in manipulated sequences often appears random.

1. Senior Lecturer, InfoSec Group, Royal Holloway University of London
2. Professor, Cyber Security Research Group, University of Kent

This work was supported by the Horizon2020 funded RAMSES Project https://ramses2020.eu // https://cordis.europa.eu/project/id/700326

Adapted from previously published works:

Hurley-Smith, D., & Hernandez-Castro, J. (2021). Challenges in Certifying Small-Scale (IoT) Hardware Random Number Generators. In *Security of Ubiquitous Computing Systems* (pp. 165-181). Springer, Cham.

Hurley-Smith, D., & Hernandez-Castro, J. (2018, November). Great expectations: A critique of current approaches to random number generation testing & certification. In *International Conference on Research in Security Standardisation* (pp. 143-163). Springer, Cham.