A Unified NIST SP 800-90B Validation Framework for CMOS True Random Number Generators and Quantum Random Number Generators

Che-Ping Lin
Independent Researcher
Hsinchu City, Taiwan

ABSTRACT
Random number generators (RNGs) are foundational to modern cryptographic systems. While quantum random number generators (QRNGs) leverage inherently stochastic quantum measurements, practical deployments still exhibit implementation artifacts (e.g., detector dead time, afterpulsing, drift) that can introduce bias, correlation, and non-stationarity. This paper presents a unified validation framework for both CMOS true random number generators (TRNGs) and QRNGs under NIST SP 800-90B entropy source validation. Rather than simulating quantum physics, we model observable raw-output behaviors of practical implementations and evaluate them using a consistent pipeline: raw capture constraints, IID screening, conservative min-entropy bounding, and online health testing (repetition count test and adaptive proportion test). Using synthetic sources that emulate typical CMOS and QRNG failure modes, we demonstrate how similar artifacts drive the IID versus non-IID decision and tighten entropy bounds, motivating a vendor-agnostic, reproducible validation methodology.

General Terms
Security, Randomness, Statistical Validation, Cryptography, Hardware

Keywords
NIST SP 800-90B, min-entropy, QRNG, CMOS TRNG, IID assessment, non-IID estimation, health tests, entropy source validation

1. INTRODUCTION
Entropy quality directly affects cryptographic strength, key generation, nonces, and secure protocols. CMOS true random number generators (TRNGs) are widely integrated into SoCs and commonly rely on thermal noise, jitter, or metastability. Quantum random number generators (QRNGs) instead derive entropy from quantum measurement events (e.g., photon path selection, vacuum fluctuations). In practice, however, both classes of implementations include sensors, analog front-ends, and digitization stages that can introduce bias, correlation, and non-stationarity in the observable raw output bitstream.

Abbreviations: SP 800-90B (NIST entropy source validation), SP 800-90C (NIST DRBG constructions), QRNG (Quantum Random Number Generator), TRNG (True Random Number Generator), IID (independent and identically distributed), non-IID (non-independent/non-identically distributed), MCV (most common value), APT (adaptive proportion test), RCT (repetition count test), DRBG (deterministic random bit generator).

NIST SP 800-90B provides a uniform, data-driven procedure to validate entropy sources regardless of physical origin. Nevertheless, the literature often treats CMOS TRNGs and QRNGs in separate communities: circuit-centric CMOS work focuses on noise sources and design robustness, while many QRNG publications emphasize quantum principles or post-processing results, sometimes without a validation-boundary-first discussion aligned to SP 800-90B. This gap makes it difficult to compare entropy claims across technologies in a consistent and integration-ready manner [1].

This paper presents a unified SP 800-90B-aligned validation framework that enables fair comparison of CMOS TRNG and QRNG implementations at a clearly defined validation boundary. The key idea is not to simulate quantum physics in detail, but to model and analyze the observable raw-output behaviors that drive SP 800-90B decision points. The proposed framework connects practical implementation artifacts, such as detector dead time, afterpulsing, drift, efficiency mismatch (QRNG), and injection/coupling, recovery effects, and offset (CMOS), to measurable statistics (bias, lag-1 correlation, run-length anomalies, and non-stationarity), and then evaluates their impact on IID screening, conservative min-entropy bounds, and online health-test sensitivity.

The main contributions of this work are: (1) an artifact-to-validation mapping that links common failure modes to SP 800-90B outcomes (IID vs non-IID path selection, entropy bound tightening, and health-test triggers); (2) a reproducible synthetic evaluation suite with parameterized artifact models to benchmark entropy degradation trends under controlled conditions; and (3) practical design guidance on health tests by contrasting repetition count test (RCT) and adaptive proportion test (APT) sensitivity across representative artifact families. Together, these contributions provide a vendor-agnostic methodology for defensible entropy claims and integration-ready validation reporting.

2. BACKGROUND AND STANDARDS CONTEXT
NIST SP 800-90B specifies design principles, testing methodology, and entropy estimation procedures for entropy sources used by random bit generators. Public entropy validation reports from the NIST Cryptographic Module Validation Program (CMVP) demonstrate that commercial QRNG products can be evaluated under SP 800-90B, including ID Quantique Quantis devices (Entropy Certificate E63) and the Qrypt Atlas QRNG PCIe card (Entropy Certificate E246). These reports provide strong evidence that quantum origin does not exempt an implementation from boundary definition, raw data collection requirements, and conservative entropy bounding [3][4][6].

Beyond product reports, recent research analyzes statistical properties and practical limitations of SP 800-90B estimators, emphasizing that correlation, bias, and non-stationarity can substantially reduce conservative min-entropy bounds. Separately, QRNG research often focuses on physical mechanisms (beam splitters, phase noise, vacuum fluctuations) and may emphasize post-processed randomness tests, while CMOS TRNG literature frequently focuses on circuit techniques and robustness against environmental or adversarial influence [7].

This paper complements these directions by providing a unified, implementation-agnostic validation pipeline and an explicit artifact-to-validation mapping that connects QRNG and CMOS artifact families to SP 800-90B decision points. In addition, a reproducible synthetic benchmark is provided to support controlled comparisons and to communicate practical validation insights without requiring proprietary hardware details.

3. UNIFIED SP 800-90B VALIDATION FRAMEWORK

As shown in Figure 1, the proposed pipeline is consistent with the intent of SP 800-90B and organizes entropy validation into four main stages. First, raw data are captured without conditioning or post-processing so that the entropy assessment reflects the intrinsic statistical behavior of the source itself. Second, IID screening is performed to determine whether the observed sequence can reasonably satisfy the assumption of independence and identical distribution. When this assumption is supported, the IID path is applied and the min-entropy may be bounded using the corresponding H_{min} formulation. Otherwise, the analysis proceeds through the non-IID path, where multiple estimators, such as MCV, Markov, and compression-based methods, are used to obtain a conservative lower bound. Health tests, including RCT and APT, are then included as part of the operational validation flow to detect catastrophic degradation or abnormal source behavior before subsequent conditioning or DRBG use. In this way, the overall framework provides a standards-aligned basis for entropy evaluation and validated entropy claims across both CMOS TRNG outputs and QRNG-derived bitstreams [2].

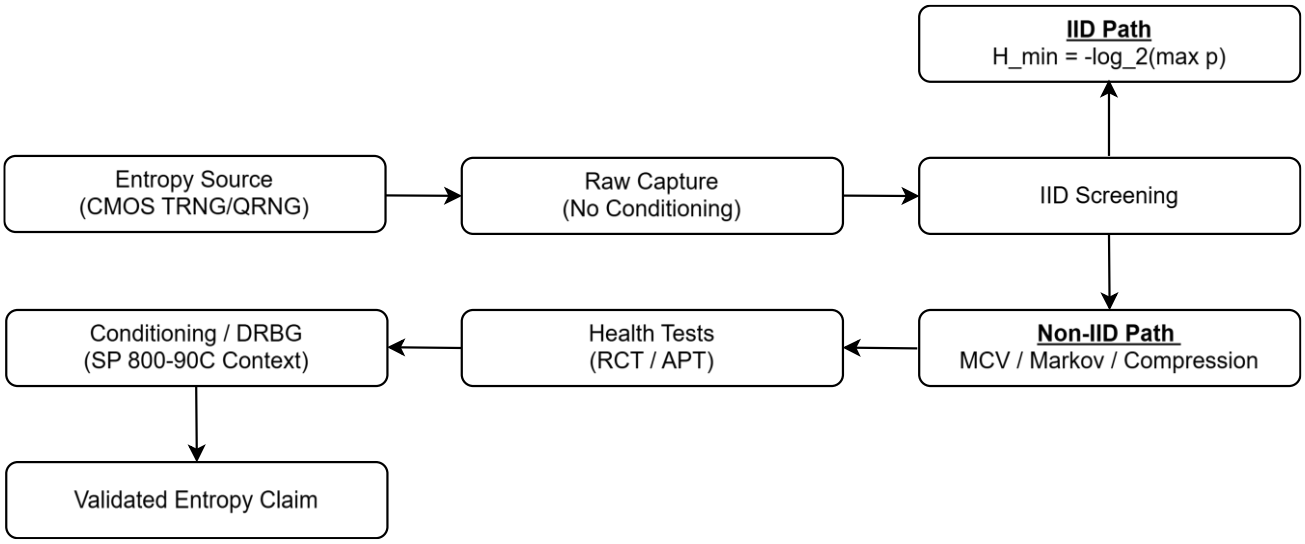

Figure 1. Unified entropy source validation flow aligned with NIST SP 800-90B

3.1 IID and non-IID concepts

IID denotes independent and identically distributed samples. In practice, entropy sources frequently violate IID due to short-term correlation (memory effects), bias, or non-stationarity (drift). SP 800-90B therefore requires an IID assessment stage; failing that stage implies using non-IID estimators that are generally more conservative.

3.2 Artifact-to-behavior mapping for CMOS TRNG and QRNG

At the validation boundary defined in Section 3.1, both CMOS TRNGs and QRNGs are reduced to observable stochastic processes. The purpose of artifact-to-behavior mapping is to translate implementation-level non-idealities into measurable statistical signatures that directly influence SP 800-90B entropy estimation and health-test outcomes.

For CMOS TRNGs, common artifacts include supply coupling, electromagnetic injection, metastability resolution bias, comparator offset, and temperature-induced drift. These effects typically manifest as: (i) static bias (increase of p_{max}), (ii) short-range temporal correlation (increase of conditional probability q_{max}), or (iii) slow non-stationary proportion shifts detectable within sliding windows. For QRNGs, implementation artifacts arise from detector efficiency mismatch, optical path imbalance, dead time, afterpulsing, analog front-end filtering, ADC quantization effects, and threshold offset. Although the entropy source is quantum mechanical, these non-idealities similarly translate into observable bias, Markov-style dependence, or drift at the digital output boundary. The mapping principle is therefore technology-agnostic: any artifact that increases the maximum observable probability term (either marginal p_{max} for IID paths or conditional p_{max} for non-IID paths) directly reduces the conservative min-entropy bound through $H_{min} = -\log_2(\max(\cdot))$. Correlation-driven artifacts often force the estimator to transition from IID to non-IID evaluation, resulting in substantially tighter entropy bounds.

This unified abstraction allows CMOS TRNG and QRNG implementations to be compared within the same entropy-degradation space, independent of physical origin. The following sections quantify these mappings using parameterized synthetic models and benchmark curves.

3.3 QRNG quantum-layer statistical model (at the validation boundary)

Although SP 800-90B is agnostic to the underlying physics, a QRNG system description is still valuable for explaining how practical variations translate into bias and dependence at the validation boundary. This section presents compact quantum-layer models and explicitly connects common implementation variations to conservative min-entropy bounds used later in Section 2 and Section 4.

3.3.1 Beam-splitter QRNG efficiency mismatch and the resulting min-entropy loss

An ideal 50/50 beam splitter produces a binary outcome $X \in \{0,1\}$ with $P(X=1)=\frac{1}{2}$. In practice, unequal detector efficiencies (η_0, η_1), unequal optical coupling, or comparator offset can bias the observed bit:

$$P(X=1)=\frac{\eta_1}{\eta_0+\eta_1}, \quad P(X=0)=\frac{\eta_0}{\eta_0+\eta_1}$$

This immediately maps to the IID min-entropy bound:

$$H_{min} = -log_2\left(max\big(P(X=0), P(X=1)\big)\right) = \\ = -log_2\big(max(\eta_0,\eta_1)/(\eta_0+\eta_1)\big)$$

Example: if $\frac{\eta_1}{\eta_0+\eta_1}=0.55$, then $H_{min} \approx -\log_2(0.55) = 0.862$ bits/sample (before any correlation effects). This explains why "quantum origin" can still yield < 1 bit/sample at the boundary when imbalance exists.

3.3.2 Correlation and non-IID bounds induced by dead time and afterpulsing in photon-arrival QRNGs

For a coherent optical source, photon arrivals over an interval t are often modeled as a Poisson process:

$N(t) \sim \text{Pois}(\lambda t)$

Digitization frequently maps parity of counts or quantized inter-arrival time to bits. However, detector dead time and afterpulsing introduce memory: (i) dead time suppresses events immediately after a detection; (ii) afterpulsing increases the likelihood of a near-future detection. Both effects create lag-1 dependence that is well captured by a first-order Markov abstraction:

Let $X_t \in \{0,1\}$ denote the digitized bit at time t, and model $\{X_t\}$ as a two-state Markov chain with transition matrix

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix},$$

Where

$$a = P(X_{t+1}=1 \mid X_t=0), \quad b = P(X_{t+1}=0 \mid X_t=1)$$

Under this model, a conservative per-sample conditional min-entropy bound is

$$H_{min}^{Markov} \approx -\log_2(\max\{P(0|0), P(1|0), P(0|1), P(1|1)\})$$

Interpretation:
Detector artifacts increase short-range dependence by pushing at least one transition probability closer to 1, which enlarges $\max\{\cdot\}$ and lowers H_{min}^{Markov}. In particular, dead time suppresses detections immediately after a '1', so it typically reduces $P(1 \mid 1) = 1-b$ (equivalently, it increases b). Conversely, afterpulsing increases the likelihood of a near-future detection after a '1', so it typically increases $P(1 \mid 1) = 1-b$ (equivalently, it reduces b). In both cases, the maximum transition probability can become large, tightening the conservative bound. This is why QRNG streams with detector artifacts often fail IID screening and require non-IID estimators.

3.3.3 Continuous-variable vacuum and phase-noise QRNG noise ratio and threshold effects on bias and correlation

Many high-speed QRNGs sample a continuous variable Y formed by quantum noise plus electronics noise:

$$Y = Q + E, \quad Q \sim \mathcal{N}(0, \sigma_Q^2), \quad E \sim \mathcal{N}(0, \sigma_E^2)$$

A common digitizer is a sign quantizer $X = 1\{Y \geq \theta\}$. If $\theta \neq 0$ (offset) or if σ_E dominates σ_Q, the probability $P(X=1)$ deviates from $\frac{1}{2}$, reducing H_{min} via the IID bound. Moreover, analog front-end filtering, ADC saturation, or clock coupling can introduce temporal correlation, which again moves validation to the non-IID path.

Summary: these compact models clarify where "quantum" enters (event/noise origin), and, more importantly for validation, how practical variations map to either (i) bias-driven H_min loss (IID) or (ii) correlation-driven H_min loss (non-IID). Table 1 summarizes common QRNG implementation variations and the primary SP 800-90B validation impact observed at the digital boundary.

Table 1. Typical QRNG implementation variations and their primary impact on SP 800-90B entropy bounds.

Variation (QRNG)	Primary bound affected	Effect on H_{min} (intuition)
Efficiency mismatch / offset ($\eta_0 \neq \eta_1$, $\theta \neq 0$)	IID min-entropy (MCV/IID)	$H_{min} = -log_2(p_{max})$
Dead time (refractory)	Markov/non-IID bound	$H_{min}^{Markov} = -log_2\left(max(1-\hat{a}, \hat{a}, \hat{b}, 1-\hat{b})\right)$
Afterpulsing	Markov/non-IID bound	$H_{min}^{Markov} = -log_2\left(max(1-\hat{a}, \hat{a}, \hat{b}, 1-\hat{b})\right)$
Optical power drift	APT + conservative bound	$P_{detect} \approx 1-(1-\gamma)^W$

4. SYNTHETIC SIMULATION METHODOLOGY

4.1 Simulation Setup

Table 2 lists the synthetic case definitions used in the evaluation suite. A synthetic evaluation suite is constructed to generate raw bitstreams under controlled artifact parameters. Each case produces $N = 200000$ raw bits across multiple

random seeds. The artifact families include (i) IID bias ($p \neq 0.5$) to emulate imbalance/offset, (ii) first-order Markov dependence to emulate dead-time/afterpulsing and coupling-induced memory, and (iii) windowed drift to emulate slow non-stationarity. A lightweight IID-screening proxy (bias threshold and lag-1 correlation threshold) is applied to select an IID versus non-IID analysis path, consistent with the SP 800-90B intent of separating IID and non-IID behavior before entropy bounding.

Reproducibility:
All synthetic experiments use $N = 200000$ raw bits per case and report averages over 5 random seeds. IID screening proxy uses $|p_1 - 0.5| \leq 0.02$ and $|lag - 1| \leq 0.05$. APT proxy uses window length W = 1024 with z = 4.0, and the RCT proxy uses a false-alarm target $\alpha = 10^{-6}\alpha$. All scripts to regenerate tables and benchmark outputs are included in the accompanying reproducibility package.

4.2 Entropy Estimator Formulations

To conservatively bound entropy at the validation boundary, three estimators are used for binary streams: (1) a most-common-value (MCV) bound for IID-dominant behavior, (2) a first-order Markov conditional-probability bound for short-range dependence, and (3) a compression-based proxy that captures structural regularity.

Most Common Value (MCV):
Let N0 and N1 be the counts of 0 and 1 in a sample of length N. Define the maximum marginal probability:

$$p_{max} = max\left(\frac{N_0}{N}, \frac{N_1}{N}\right), \quad H_{MCV} = -log_2(p_{max})$$

The IID min-entropy bound is then:

$$H_{MCV} = -log_2(p_{max})$$

First-order Markov bound:

Let N01 denote the number of 0→1 transitions and N10 the number of 1→0 transitions. We estimate transition probabilities as:

$$\hat{a} = \frac{N_{01}}{N_0}$$

$$\hat{b} = \frac{N_{10}}{N_1}$$

The conservative conditional min-entropy bound is:

$$H_{Markov} = -log_2\left(max\left(1 - \hat{a}, \hat{a}, \hat{b}, 1 - \hat{b}\right)\right)$$

This form directly reflects that any artifact increasing the maximum conditional probability term (e.g., dead time suppressing transitions or afterpulsing enhancing immediate repeats) lowers H_{min}.

Compression proxy bound:

As an auxiliary proxy for structured regularity, we approximate an entropy bound using the normalized compressed length L compressed:

$$H_{LZ} \approx \frac{L_{compressed}}{N}$$

Although this proxy is not a replacement for SP 800-90B estimators [5], it provides an interpretable indicator when repetitive structure exists beyond first-order dependence.

4.3 Health-Test Sensitivity Analysis

Finally, online health-test sensitivity is evaluated using the SP 800-90B repetition count test (RCT) and adaptive proportion test (APT). For each artifact family, the alarm rate is measured under parameter sweeps (bias magnitude, transition imbalance, and drift rate) to illustrate which failure signatures are more likely to trigger RCT versus APT at runtime.

Table 2 lists the synthetic cases used throughout the sensitivity study, covering IID bias (C1–C2), first-order Markov dependence (Q1–Q2), and windowed drift (D1). These cases were selected to represent distinct statistical degradation patterns, allowing the relative sensitivity of entropy bounds and runtime health tests to be compared under controlled and reproducible conditions.

Table 2. Synthetic case definitions and artifact parameters used in the evaluation suite.

ID	Family	Artifact (parameter)	Interpretation
C1	IID_UNBIAS	p1=0.500	Bernoulli IID with no bias (unbiased baseline)
C2	IID_BIAS	p1=0.550	Bernoulli IID with static bias as a proxy for hardware imbalance or offset
Q1	MARKOV	a=0.050, b=0.800	First-order Markov dependence as a proxy for dead time and afterpulsing
Q2	MARKOV	a=0.200, b=0.600	First-order Markov dependence as a proxy for dead time and afterpulsing
D1	DRIFT	p_start=0.500, p_end=0.580	Windowed drift as a proxy for non-stationarity

Note: IDs denote synthetic test cases: C for IID cases, Q for Markov dependence cases, and D for drift or non-stationarity cases.

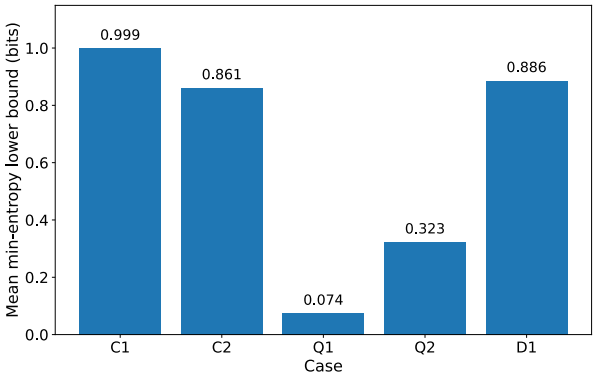

Figure 2. Mean min-entropy lower bound by case.

As shown in Figure 2, the unbiased baseline case C1 remains closest to 1 bit, while the biased IID case C2 shows a moderate reduction in the lower bound. The Markov-dependent cases Q1 and Q2 exhibit substantially lower values, indicating that

temporal dependence is penalized more strongly than simple static bias in the present benchmark setting. The drift case D1 remains higher than the Markov cases, suggesting that gradual non-stationarity degrades the conservative bound less severely than persistent short-range dependence.

5. RESULTS

5.1 Experimental Cases and Reporting Metrics

Aggregated proxy validation results are reported over five random seeds for representative cases spanning IID-like behavior, biased IID behavior, and non-IID behavior.

5.2 Results Analysis and Discussion

Table 3 summarizes the aggregated proxy results over five seeds for representative cases. Case C1 is close to IID with p1 near 0.5 and negligible lag-1 dependence, so it passes IID screening and yields an H_{min} bound close to 1 bit per sample. Case C2 is biased while remaining nearly uncorrelated; the Hmin bound drops according to p_{max}, and the APT alarm rate becomes non-zero under the chosen window and z parameters. Markov-like cases (Q1, Q2) show strong dependence, which forces non-IID analysis and tightens the bound through the maximum conditional probability term. The drift case (D1) illustrates non-stationarity: lag-1 can remain small while windowed proportion shifts increase APT and RCT sensitivity. These trends match the artifact-to-behavior mapping in Section 3.3, where increases in marginal p_{max} degrade IID bounds and increases in conditional terms degrade non-IID bounds.

5.3 Health-Test Analytical Bounds and Design Guidance

To connect the simulation outcomes to deployable runtime monitors, compact analytical expressions are provided to explain why APT tends to respond to windowed proportion shifts (bias/drift), while RCT responds to unusually long identical runs (including stuck-like failures). These expressions are used as qualitative guidance for selecting thresholds and interpreting alarms, rather than as replacements for the exact SP 800-90B test procedures.

Adaptive Proportion Test (APT) bounds (normal approximation):

For a window length W and expected proportion p under the nominal operating condition, an approximate two-sided acceptance band can be written as:

$$L = Wp - z\sqrt{Wp(1-p)}$$

$$U = W \cdot p + z\sqrt{W \cdot p \cdot (1-p)}$$

An APT alarm is likely when the observed count of ones in the window falls outside $[L, U]$. This makes APT sensitive to bias and slow drift that accumulates over the window.

Repetition Count Test (RCT) threshold (geometric tail proxy):

For an upper-bound maximum probability term p_max, an approximate run-length threshold C satisfying false-alarm target α can be expressed as:

$$C = \left\lceil \frac{\log(\alpha)}{\log(p_{max})} \right\rceil$$

This relation highlights why RCT is effective for detecting stuck-like behavior: when repeated identical values become more likely (p_{max} increases), the acceptable maximum run length decreases rapidly.

Injection / anomaly detection probability (simple independence proxy):

If an abnormal event occurs independently with probability γ per window, then the probability of observing at least one event over W windows is:

$$P_{detect} \approx 1 - (1-\gamma)^W$$

This back-of-the-envelope expression is useful to reason about monitoring horizon versus anomaly rate, and motivates continuous runtime monitoring as a complement to certification-style evaluation.

Table 3 highlights a clear separation between IID-like behavior and non-IID artifacts. The unbiased baseline case C1 remains close to the ideal IID condition, with an IID pass rate of 1.000, a mean proportion of ones near 0.500, and a mean min-entropy bound close to 1 bit. In contrast, C2 shows that even static bias can reduce the min-entropy bound and cause failure of IID screening under the present benchmark setting. The Markov-style cases Q1 and Q2 also fail IID screening and therefore rely on conservative non-IID bounds, which drop sharply relative to the IID baseline. This behavior is consistent with their positive lag-1 means and indicates that temporal dependence is penalized more strongly than simple static bias in the current proxy framework. The drift case D1 exhibits an intermediate min-entropy bound together with elevated health-test alarm rates, suggesting that gradual non-stationarity causes less severe entropy degradation than persistent Markov-type dependence, while still producing sustained deviations detectable at runtime. The alarm-rate patterns further show that APT and RCT respond differently to different artifact types. Q1 and Q2 strongly trigger APT but not RCT, whereas D1 produces elevated activity in both tests, with a comparatively stronger RCT response. APT and RCT alarm rates are computed over sliding windows (W = 1024) and aggregated across all windows and seeds.

Table 3. Aggregated proxy validation results (N=200k bits per case, 5 seeds).

ID	IID pass rate	Lag1 mean	p1 mean	H_{min} bound mean	APT alarm rate	RCT alarm rate
C1	1.000	-0.000	0.500	0.999	0.000	0.000
C2	0.000	-0.000	0.550	0.861	0.202	0.200
Q1	0.000	0.149	0.059	0.074	1.000	0.000
Q2	0.000	0.201	0.251	0.323	1.000	0.000
D1	0.000	0.003	0.540	0.886	0.228	0.600

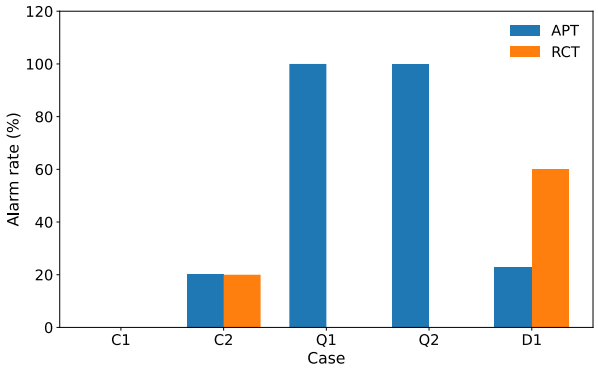

Figure 3. Comparison of APT and RCT alarm rates across cases

Figure 3 compares the APT and RCT alarm rates across the synthetic cases and shows that the two health tests respond differently to different artifact types. The baseline case C1 produces zero APT and RCT alarms, which is consistent with its role as an unbiased IID reference case under the present benchmark setting. Case C2 exhibits moderate alarm rates in both APT and RCT, consistent with a static bias condition. The Markov-dependent cases Q1 and Q2 produce consistently high APT alarm rates but negligible RCT response in the present benchmark setting, indicating that these sequences are more readily detected through window-level statistical deviation than through unusually long repeated runs. In contrast, the drift case D1 produces elevated alarm activity in both tests, with a notably stronger RCT response. Overall, the figure shows that APT and RCT provide complementary runtime sensitivity to different forms of entropy-source degradation.

6. DISCUSSION AND PRACTICAL GUIDANCE

For a QRNG, the fundamental entropy may be quantum, but the validated entropy claim depends on the measured raw-output behavior at the validation boundary. Designers should document the entropy source boundary and demonstrate stability across operating conditions. For CMOS TRNG, the same principle applies: the validation boundary should be specified clearly, raw samples should be collected across PVT corners, and health-test thresholds should be justified in relation to expected source behavior. A unified framework enables consistent comparison and reduces ambiguity in security claims. The results further suggest that APT and RCT should be treated as complementary monitors rather than interchangeable tests, since different artifact classes may produce different alarm patterns under the same validation setting. In addition, continuous runtime entropy monitoring can complement certification by providing early-warning diagnostics during field operation, and recent implementations demonstrate that APT- and RCT-style monitors can be embedded in firmware for online quality control [8][9].

Limitations: This work is a reproducible, simulation-driven framework study and does not claim device-specific certification values. The synthetic generators are used to isolate bias, dependence, and drift signatures in a controlled manner at the validation boundary. Future work will apply the same mapping to raw traces collected from representative CMOS TRNG and QRNG implementations to quantify real-world parameters and confirm the predicted estimator and health-test responses. Further extension may also include threshold calibration under deployment constraints, broader non-IID source models, and comparison against more complete SP 800-90B estimator flows.

7. CONCLUSION

This paper presented a unified SP 800-90B-aligned validation framework for CMOS TRNG and QRNG implementations. Using a reproducible synthetic evaluation suite, the study showed how practical artifacts map to IID screening outcomes, conservative min-entropy bounds, and health-test sensitivity. The results support a vendor-agnostic methodology: quantum origin does not bypass entropy validation, and robust security claims still require careful boundary definition, conservative bounding, and continuous health monitoring. The synthetic models used in this work are not intended to replace full SP 800-90B estimators or detailed hardware characterization; rather, they provide a controlled and interpretable framework for explaining how observable artifacts influence the IID decision, entropy bounds, and health-test behavior.

The results also indicate that different artifact classes affect validation outcomes in different ways. IID-like cases remain close to the ideal entropy bound, while non-IID artifacts, especially Markov-style dependence, lead to substantially lower conservative bounds. In addition, health-test responses differ across artifact types, showing that offline entropy estimation and runtime monitoring provide complementary information rather than interchangeable indicators. This distinction is important for practical entropy-source evaluation, since a source may not exhibit the same degree of degradation in conservative entropy bounds and online alarm behavior.

The future scope of this work includes extending the framework to measured hardware datasets from real CMOS TRNG and QRNG implementations, incorporating broader non-IID source models, and refining the proxy cases to better reflect implementation-level effects such as dead time, afterpulsing, drift, and environmental variation. Further study may also include calibration of health-test thresholds for deployment-oriented monitoring, as well as comparison with more complete SP 800-90B estimator flows. Overall, the proposed framework provides a practical and interpretable basis for studying how different artifacts influence entropy validation outcomes and runtime health monitoring under a unified SP 800-90B-aligned methodology.

Data Availability

Data and code supporting the findings of this study are publicly available. The reproducibility package includes the synthetic generators, evaluation scripts, and the CSV outputs used to regenerate the reported tables. The repository is hosted on GitHub at:

https://github.com/acelin1981/qrng-cmos-90b-repro

8. REFERENCES

[1] NIST, “Recommendation for the Entropy Sources Used for Random Bit Generation,” NIST SP 800-90B, 2018.

[2] NIST, “Recommendation for Random Bit Generator (RBG) Constructions,” NIST SP 800-90C, 2022.

[3] NIST CMVP, “Entropy Certificate #E63 Public Use: IDQ Quantis IID QRNG,” 2023.

[4] NIST CMVP, “Entropy Certificate #E246 Public Use: Qrypt Atlas Quantum Random Number Generator (QRNG) PCIe Card,” 2025.

[5] NIST, “SP 800-90B Entropy Assessment Tool (reference implementation),” GitHub repository: usnistgov/SP800-90B_EntropyAssessment, accessed 2026-02-19.

[6] ID Quantique, "Quantis QRNG Chip receives NIST Entropy Source Validation," 2023.

[7] M. Layat-Kevin et al., "Observations on NIST SP 800-90B entropy estimators," International Journal of Information Security, 2025.

[8] C.-P. Lin, "Continuous Entropy Monitoring of TRNGs via a CNN Autoencoder with Residual-Based Diagnostics," Zenodo Preprint, 2026. DOI: 10.5281/zenodo.18338590.

[9] C. Caratozzolo, V. Rossi, K. Witek, A. Trombetta, M. Baszczyk, P. Dorosz, W. Kucewicz, and M. Caccia, "Entropy measurement and online quality control of bit streams by a true random bit generator," Frontiers in Computer Science, vol. 7, art. 1642566, Oct. 2025. doi: 10.3389/fcomp.2025.1642566.