Journal of Tr. Math. Sci., Vol. 3, No. 3, 2026, pp. 42-54
E-ISSN: 3062-3138

Research Article

Entropy Extraction from Audio Signals Using Hilbert-Mod Transformation and Von Neumann Refinement

Cemil Karaçam[*1], Alper Vural[2], Mehmet Özükanar[3]

[1]*Corresponding Author, Department of Mathematics, Muğla Art and Science Center, Muğla, Türkiye
cemil-karacam@hotmail.com, ORCID: 0000-0001-7186-5114
[2]Department of Computer Engineering, Faculty of Engineering, Boğaziçi University, Türkiye
vuralalperalper2005@gmail.com, ORCID: 0000-0001-9985-3932
[3]Department of Mathematics, Yıldız Technical University, İstanbul, Türkiye
mehmet.ozknr13@gmail.com, ORCID: 0009-0004-8877-156X

Abstract
This paper introduces a multi-stage framework for extracting high-entropy binary sequences from audio signals. The procedure consists of generating the analytic signal via Hilbert transformation, applying thresholding and amplification, and performing modular reduction before binary encoding based on local amplitude comparisons. Experiments were carried out using several prime moduli and amplification parameters, with three-source XOR aggregation and optional Von Neumann debiasing as post-processing stages. Randomness quality was assessed using the NIST SP 800-22 test suite. The results indicate that configurations employing intermediate amplification factor yield statistically robust bitstreams, while smaller amplification tend to introduce detectable structural patterns. Overall, the method provides effective entropy enhancement, particularly when dispersion and debiasing parameters are appropriately tuned.

Keywords: Audio entropy, Hilbert transform, Von Neumann correction, TRNG, XOR post-processing

MSC: 94A60, 94A17, 94A12, 68P25, 42A38

Received: 25/03/2026 **Accepted:** 30/06/2026

1 INTRODUCTION

Random number generation is a foundational component in cryptography, simulation, and secure communication. While algorithmic pseudo-random generators are widely used, true randomness often requires physical sources of entropy. Among these, audio signals—rich in environmental noise, analog variability, and unpredictable fluctuations—have emerged as promising candidates for entropy extraction [1].

Early work by Morrison [2] demonstrated that standard audio inputs, such as microphones, can yield cryptographically strong random bits without specialized hardware.

Chen et al. [3] later showed that filtered audio from everyday devices can pass up to 99% of NIST SP 800-22 statistical tests, highlighting the feasibility of audio-based TRNGs in mobile contexts. K. Lee and M. Lee [4] extended this approach by utilizing FM radio signals as entropy sources for embedded systems, reinforcing the role of acoustic randomness in edge computing.

Recent studies have explored the integration of audio-derived entropy into cryptographic key generation [5], and the use of hyperchaotic post-processing to enhance statistical uniformity [6]. Safaryan et al. [7] proposed an acoustics-based encryption model driven by TRNG outputs, demonstrating real-world applicability in secure communication systems.

The Hilbert transform provides a powerful tool for extracting instantaneous amplitude and phase information, and it has been applied across domains including power-quality analysis [8] and holographic pattern reconstruction [9]. Related hybrid approaches, such as the Hilbert-Huang transform, have also proven effective in biomedical signal classification, particularly in identifying subtle ingestion sounds [10]. These works indicate that Hilbert based representations can emphasize fine temporal variations that are useful for entropy modeling and random bit extraction.

Post-processing techniques such as XOR aggregation and Von Neumann correction are widely employed to enhance the statistical quality of raw entropy sources. XOR operations help disperse local bit-level bias, while Von Neumann's classical debiasing procedure [11] removes correlated bit-pairs. Similar filtering has been shown to stabilize output distributions in hardware TRNGs [12], while recent theoretical studies have related Von Neumann entropy to quantum unitarity [13], offering deeper foundations for entropy interpretation.

Building on these foundations, we propose a modular transformation pipeline integrating Hilbert transformation, amplitude thresholding, scalar amplification, modular reduction, and comparative binary encoding. We further evaluate the effects of XOR aggregation and Von Neumann correction as optional post-processing stages. Experimental results confirm that randomness quality is highly sensitive to amplification factor selection, and that well-tuned configurations can yield statistically reliable bitstreams suitable for cryptographic applications.

2 METHODOLOGY

This section outlines the proposed transformation pipeline for extracting high-entropy binary sequences from audio signals. The method integrates signal processing, nonlinear transformation, and comparative encoding to enhance entropy. Each stage is described below. The algorithm consists of the following steps:

1. Apply Hilbert transform: $X_h = hilbert(X)$
2. Thresholding: $X' = \{x_i \in X_h : |x_i| > \varepsilon\}$
3. Amplitude amplification: $X'' = qX'$
4. Modular reduction: $Z = X''modp$
5. Bit generation: $b_n = \{1, \text{ if } Z_n < Z_{n+1};\ 0, \text{ otherwise}\}$
6. XOR combination (if multiple sources are used)

Parameter Selection Justification. The prime sampling parameters $p \in \{3, 7, 17\}$ and amplification factors $q \in \{93, 1009, 10007\}$ were selected to span a range of small, intermediate, and large values. This design enables the evaluation of entropy behavior under varying residue space sizes. The intermediate value $q = 1009$ was hypothesized to offer optimal dispersion without over-preserving structural continuity. Future work may include a broader sweep across primes and composite moduli to refine this selection.

2.1 Analytic Signal Construction via Hilbert Transform

Let $X = (x_1, x_2, ..., x_T) \in R^T$ denote a discrete-time real-valued audio signal. We compute its analytic representation $X_h \in C^T$ using the discrete Hilbert transform:

$$X_h = H(X) \tag{1}$$

The Hilbert transform maps a real-valued signal to a complex-valued representation, where the real part is the original signal and the imaginary part is its Hilbert pair. This analytic form enables phase-aware processing and exposes the instantaneous envelope of the waveform. The resulting magnitude allows for more effective thresholding and amplification, helping isolate high-energy components that contribute to entropy.

2.2 Amplitude-Based Thresholding

To suppress low-energy components and reduce noise, we apply a hard threshold $\varepsilon > 0$ to the magnitude of the analytic signal:

$$X' = \{x_i \in X_h : |x_i| > \varepsilon\} \tag{2}$$

In all experiments, $\varepsilon = 1/q$ was fixed to ensure consistency across datasets.

Threshold Sensitivity. The threshold $\varepsilon = 1/q$ is coupled to the amplification factor, ensuring that the retained samples are precisely those which, after scaling by q, produce values of magnitude at least unity. This adaptive design maintains consistent filtering behaviour across all parameter configurations, effect of varying ε is left for future study.

2.3 Scalar Amplification

The filtered signal X' is scaled by a positive real-valued amplification factor $q \in R^+$, yielding:

$$X'' = q.X' \tag{3}$$

This step increases the dynamic range of the signal prior to modular reduction, allowing finer granularity in the subsequent encoding.

2.4 Modular Reduction and Fractional Extraction

Each element of the amplified signal is reduced modulo a prime number $p \in Z^+$, and only the fractional part is retained:

$$Z_i = X''_i \bmod p \tag{4}$$

Here, the modulo operation is applied in the real domain, preserving sub-integer variations that contribute to entropy.

2.5 Comparative Binary Encoding

A binary sequence $B = (b_1, b_2, ...) \in \{0, 1\}^{T-1}$ is generated by comparing consecutive elements of Z:

$$b_n = \begin{cases} 1, & \text{if } Z_n < Z_{n+1} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

This encoding captures local fluctuations in the transformed signal and converts them into binary transitions.

2.6 Multi-Source Aggregation via XOR

When multiple audio sources are available, their respective binary sequences $B^{(1)}, B^{(2)}, ..., B^{(k)}$ are combined using bitwise XOR:

$$B_{final} = B^{(1)} \oplus B^{(2)} \oplus ... \oplus B^{(k)} \tag{6}$$

This aggregation step aims to further decorrelate the output and increase statistical independence.

2.7 Optional Bias Correction

To mitigate potential bias in the binary output, an optional von Neumann extractor is applied. This post-processing step discards repeated bits and retains only transitions (*i.e.*, $01 \to 1,\ 10 \to 0$), improving uniformity without assuming independence.

2.8 Empirical Evaluation

We conducted a series of experiments to evaluate the entropy of the generated sequences using the NIST SP 800-22 randomness test suite. A significance threshold of $\alpha = 0.05$ was adopted throughout; a test is considered failed when its reported p-value falls below this level. The following configurations were tested:

- **Single-Source Parameter Sweep:**

  Each of the four audio recordings was processed independently under all nine (p, q) combinations with $p \in \{3, 7, 17\}$ and $q \in \{93, 1009, 10007\}$, yielding 36 bitstreams. The same configurations were repeated with Von Neumann correction applied, producing an additional 36 corrected sequences. Test reports are provided for all 72 outputs.

- **Triple-Source XOR Aggregation (Primary Evaluation):**

  For all $\binom{4}{3} = 4$ triples of audio sources, the three corresponding bitstreams were combined via bitwise XOR under each of the nine (p, q) configurations, yielding 36 aggregated sequences. Von Neumann correction was subsequently applied to each, producing 36 additional corrected sequences. The resulting 72 bitstreams form the basis of the visual evaluation in Section 3.

- **Same Source, Varying Primes:**

  Using a single audio file, the algorithm was applied with $(p, q) \in \{(3,\ 1009),\ (7,\ 1009),\ (17,\ 1009)\}$ and the three resulting bitstreams were XORed. Both uncorrected and Von Neumann-corrected outputs were assessed. Test reports are provided.

- **Same Source, Varying Amplification:**

  Using a single audio file, the algorithm was applied with $(p, q) \in \{(7,\ 93),\ (7,\ 1009),\ (17,\ 1009)\}$ and the three resulting bitstreams were XORed. Both uncorrected and Von Neumann-corrected outputs were assessed. Test reports are provided.

3 Visual Evaluation of NIST Test RESULTS

This section presents graphical summaries of how different parameter settings affect randomness quality. The figures are grouped by theme: parameter sensitivity, histogram distributions, and comparative fail rates. Each group highlights specific aspects of entropy behavior under the proposed transformation.

3.1 Impact of Amplification Factor and Correction on Randomness

Figure 1 shows the amplification factor (q) and Von Neumann correction influence fail rates in NIST tests. Both small and large q values produce elevated fail rates, while the intermediate value $q = 1009$ consistently yields the lowest failure rate. Correction helps in most cases, though its effect depends on signal variability.

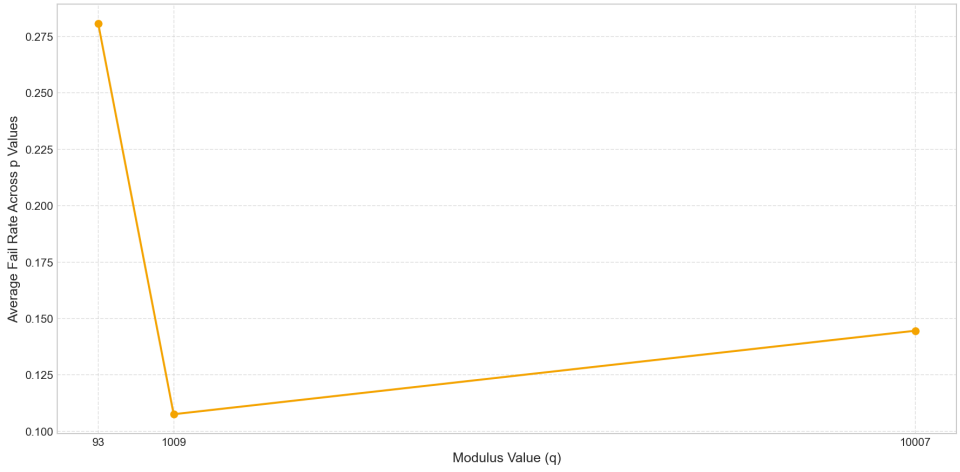

Figure 1: Effect of amplification factor (q) on Randomness Quality

Commentary: The fail rate is highest at low amplification factor ($q = 93$), indicating poor randomness. It drops significantly at $q = 1009$, suggesting optimal statistical uniformity. A slight increase at $q = 10007$ may reflect residual structural dependencies.

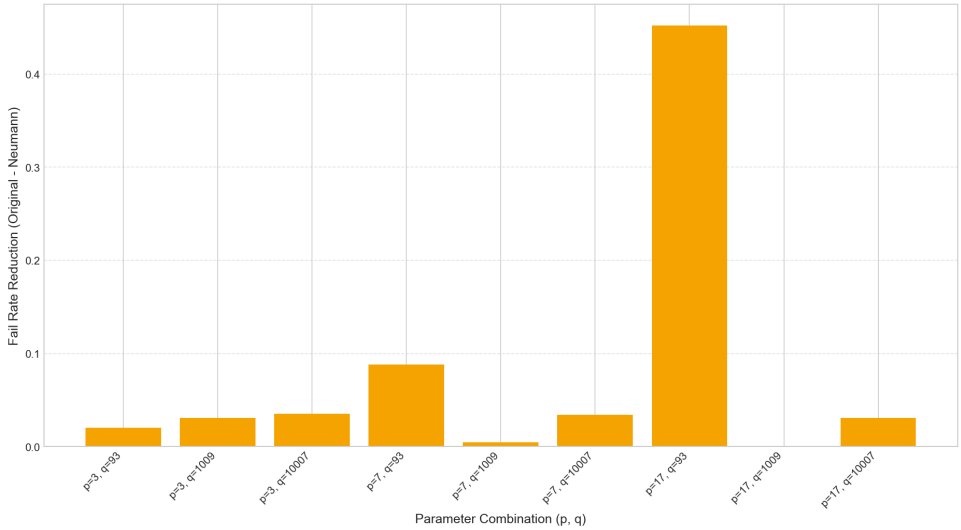

Figure 2: Impact of Neumann Correction on Randomness Quality

Commentary. Figure 2 shows the effect of the Neumann correction is parameter-dependent. Clear improvements are observed for $(p = 7,\ q = 93)$ and $(p = 17,\ q = 93)$, where short range correlations are effectively reduced. In contrast, a slight degradation appears at $(p = 17,\ q = 1009)$, indicating over-filtering. Thus, the correction operates best when the underlying sequence exhibits moderate structural dependence.

3.2 Histogram Distributions of p-values

Histograms in Figures 3–6 visualize the spread of p-values from the Frequency and Cumulative Sums tests, comparing original and corrected bitstreams. They help assess statistical balance and detect bias patterns.

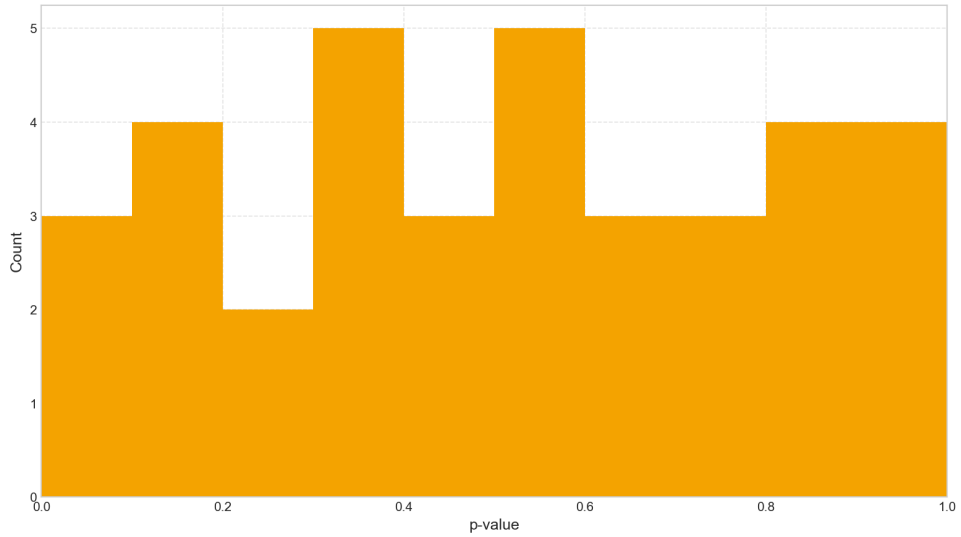

Figure 3: Histogram of p-values for Frequency test (Neumann)

Commentary. In Figure 3, we observe that p-values after Neumann correction are broadly distributed, with no clustering near zero. This indicates balanced bit frequencies and confirms that short-range dependencies have been effectively mitigated.

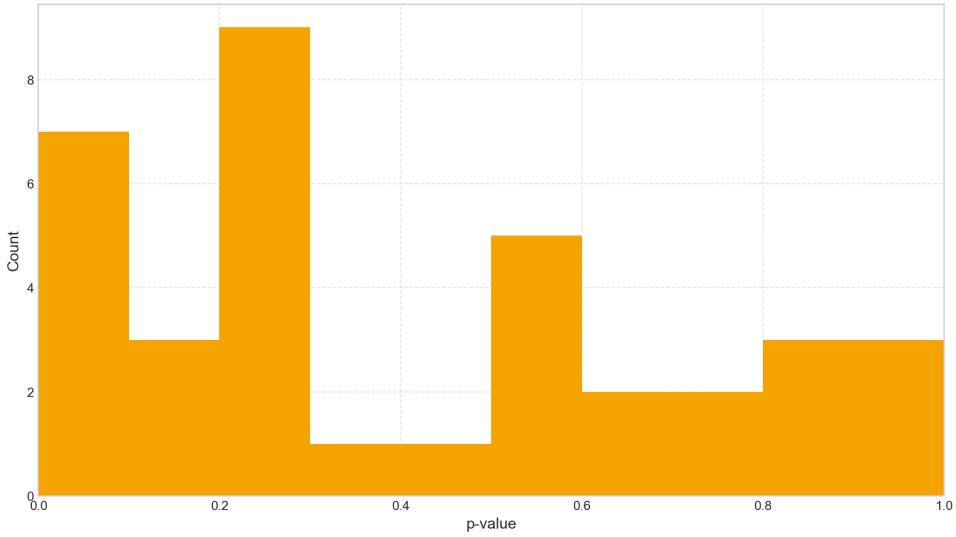

Figure 4: Histogram of p-values for Frequency test (Original)

Commentary. In Figure 4, we see that the original bitstream yields a broad p-value distribution, with both mid-range and extreme values. This suggests that the Frequency test is passed, but short-range dependencies remain prior to correction.

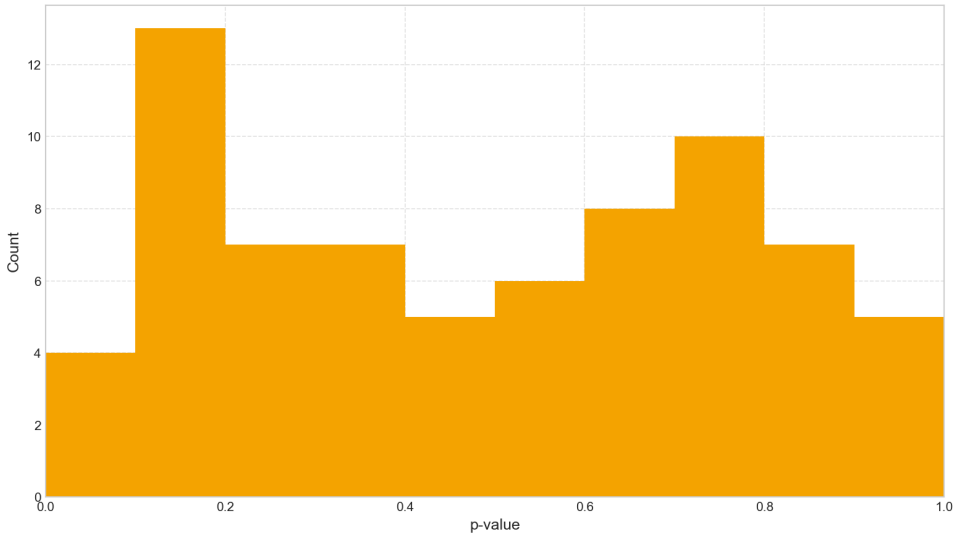

Figure 5: Histogram of p-values for CumulativeSums test (Neumann)

Commentary. Figure 5 shows that most p-values lie within the range $0.1 \leq p \leq 0.8$, with a noticeable concentration around $p \approx 0.2$. This distribution indicates stable cumulative fluctuations and confirms the absence of long-range imbalance or drift in the corrected bitstream.

Commentary. In Figure 6, we see a mild concentration of p-values in the range $0.1 \leq p \leq 0.6$, with more values near the extremes. Although the Cumulative Sums test is passed, this distribution suggests that short-range cumulative dependencies remain in the uncorrected bitstream.

3.3 Comparative Fail Rate Analysis

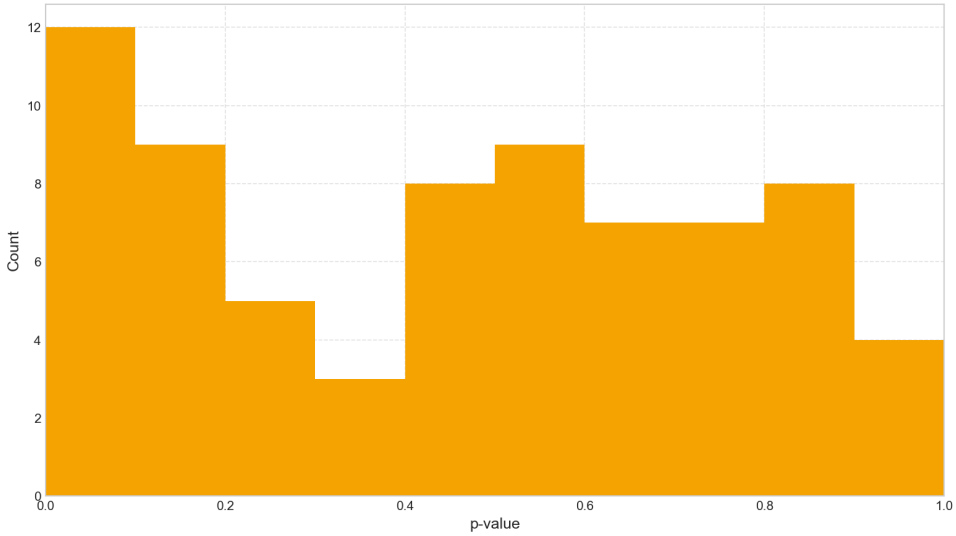

Figure 6: Histogram of p-values for CumulativeSums test (Original)

Figures 7–9 compare the fail rates across different (p, q) configurations and illustrate the impact of the Von Neumann correction on statistical randomness performance. Together, they allow identification of parameter ranges that yield stable entropy, as well as regions where correction either improves or degrades randomness quality.

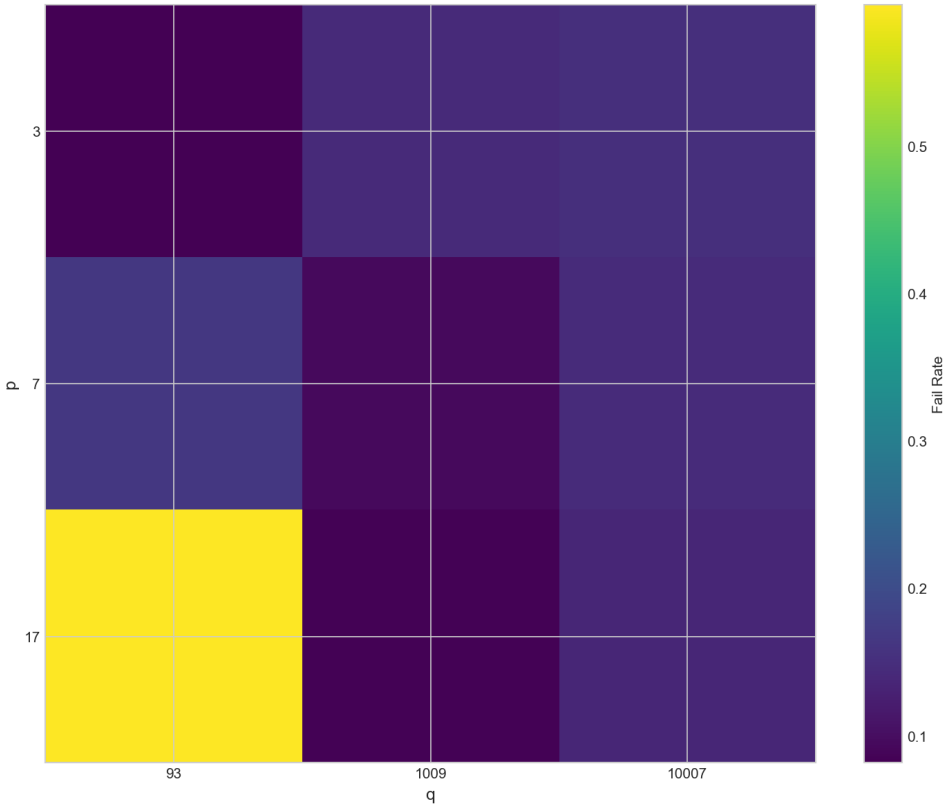

Figure 7: Fail Rate Heatmap (Original)

Commentary. Figure 7 shows fail rates vary jointly with both p and q. Across all three values of p, the lowest fail rates consistently occur when $q = 1009$, whereas both very small ($q = 93$) and very large ($q = 10007$) amplification factor result in increased failures. In addition, increasing p generally leads to higher fail rates, suggesting that larger sampling strides introduce additional correlation in the bitstream.

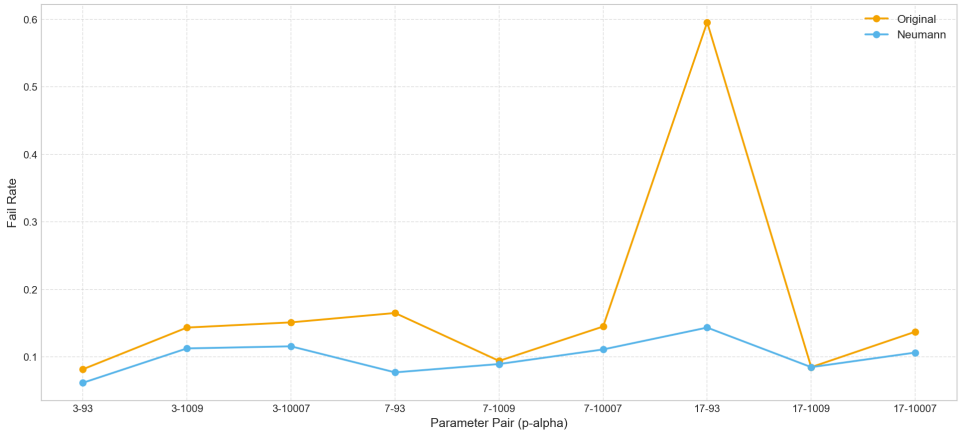

Figure 8: Fail Rate Comparison (Original vs Neumann)

Commentary. Figure 8 demonstrates that Von Neumann correction generally reduces the fail rate, with the strongest improvement observed when $q = 93$. For larger amplification factors, the benefit becomes weaker, indicating that the effectiveness of the correction depends on the structural balance of the original bitstream.

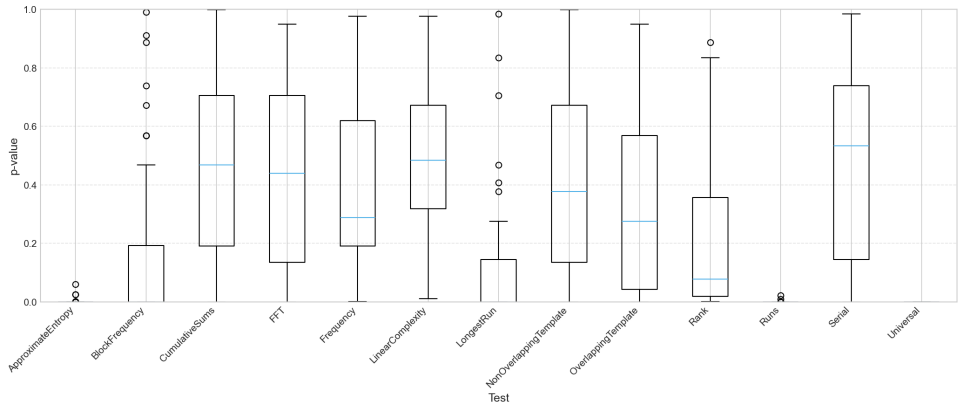

Figure 9: Boxplot of p-values by Test (Original)

Commentary. The expanded boxplot reveals pronounced stratification in NIST STS performance across the thirteen applicable statistical tests applied to the original, un-

corrected bitstreams. The Linear Complexity test emerges as the strongest performer (median ≈ 0.47, failure rate 5.6%), while the Frequency and Rank tests satisfy coarse-grained randomness criteria with moderate p-value distributions. Tests sensitive to local sequential structure — notably Runs (100%), Universal (100%), and Approximate Entropy (99.2%) — collapse to near-zero p-values, reflecting short-range serial dependencies intrinsic to the consecutive-comparison operation of the modular transform and, in the case of Universal, insufficient sequence length relative to the test's internal block parameter requirements. Block Frequency, Longest Run, and Overlapping Template exhibit analogous behaviour with failure rates exceeding 78%, while the Discrete Fourier Transform, Serial, and Non-overlapping Template tests occupy an intermediate regime characterised by right-skewed p-value distributions. Collectively, these results indicate that the raw modular transform output satisfies only macro-level randomness criteria; tests probing local entropy regularity and block-level uniformity expose systematic structural deficiencies, motivating the application of post-processing via Von Neumann correction.

4 Experimental Setup and Observations

To evaluate the entropy contribution of the proposed transformation, experiments were performed on multiple audio recordings. Each signal was processed as described in Section 2, using prime sampling parameters $p \in \{3, 7, 17\}$ and amplification factor $q \in \{93, 1009, 10007\}$. The threshold was fixed at $\varepsilon = 1/q$ throughout all trials to maintain consistent operating conditions.

Randomness of the resulting bitstreams was assessed using the NIST SP 800-22 statistical test suite. In addition to the direct outputs, two optional post-processing techniques were examined:

- XOR aggregation across variants of the transformation,
- Von Neumann correction to mitigate local frequency bias.

Audio Dataset Limitations. Experiments were conducted on four audio recordings, each downloaded from open source websites. While this setup ensures repeatability, it limits generalizability. Future evaluations should include diverse sources such as environmental noise, mobile devices, and varying acoustic environments to assess robustness across real-world scenarios.

Real-Time Feasibility. Although the proposed method is intended for cryptographic applications, its real-time performance has not been benchmarked. Bit generation rate, latency, and hardware requirements remain untested. Future work will address embedded implementation and throughput analysis to validate practical deployment.
Initial results showed that the Hilbert-mod transformation increases entropy, particularly when the amplification factor is chosen appropriately. XOR aggregation provided an additional improvement by reducing short-term correlation, while the Von Neumann correction contributed further balance when the underlying bitstream already exhibited adequate variability.

5 Results

The NIST evaluations reveal that randomness quality depends strongly on the amplification factor q. When q is small (*e.g.*, $q = 93$), the amplitude dispersion is insufficient, and low-energy components dominate the bitstream, leading to elevated failure rates in tests such as Frequency and Cumulative Sums.

The most favorable performance was observed when $q = 1009$, where the p-value distributions were the most stable overall, achieving the lowest aggregate fail rate across all tested values of p. Increasing the amplification further to $q = 10007$ did not continue to improve randomness; instead, the effectiveness plateaued and slight structural bias re-emerged.

XOR aggregation consistently reduced local dependencies, while the Von Neumann correction yielded the strongest improvements at $q = 93$.

Entropy Evaluation Scope. Randomness was assessed using the NIST SP 800-22 suite, which provides a baseline for statistical uniformity. However, cryptographic strength also depends on entropy concentration. Metrics such as min-entropy, Shannon entropy, and collision entropy were not included and should be considered in future evaluations for a more comprehensive assessment.

Post-Processing Behavior. Von Neumann correction improved randomness in most configurations, particularly when the underlying bitstream exhibited moderate variability. However, at $q = 93$, despite yielding the largest absolute reduction in fail rate, the corrected sequences still retain higher failure rates than those obtained at $q = 1009$, indicating that debiasing cannot fully compensate for insufficient initial amplitude dispersion. This highlights the importance of applying debiasing only after sufficient dispersion has been achieved.

6 Discussion

These findings suggest that the transformation works best when the amplification produces a residue space large enough to disperse amplitude variation without over-amplifying structural features of the original signal. At $q = 1009$, this balance appears to be optimal. When q is too small, deterministic structure dominates; when it becomes too large, waveform continuity is preserved in a way that reintroduces subtle dependencies.

The behavior of the Von Neumann correction is consistent with this interpretation. The correction step does not create randomness but rather filters existing bias, and therefore performs well only when a sufficient level of entropy is already present. This indicates that randomness extraction in this setting is inherently a two-stage process: first ensuring adequate dispersion via parameter selection, and then applying targeted bias reduction.

Overall, the results highlight that parameter choices are central to the method's cryptographic viability, rather than peripheral tuning considerations.

7 Conclusion and Future Work

This work introduced a hybrid approach for generating random bitstreams from audio signals, combining Hilbert-mod transformation with optional post-processing. The ex-

periments demonstrate that randomness quality depends sensitively on the amplification factor. Intermediate values, particularly $q = 1009$, produced statistically stable outputs that satisfied the NIST SP 800-22 criteria. Very small amplification failed to remove structural repetition, while very large moduli did not yield further improvement.

Future work will focus on:

1. Formal modeling of entropy propagation in the Hilbert-mod mapping,
2. Evaluation of alternative extractors such as universal hash-based post-processors,
3. Real-time and embedded implementation for cryptographic applications.
4. Adaptive or data-driven thresholding to better isolate high-energy components and suppress low-amplitude noise.

Future Work Expansion. Beyond the listed directions, future work should include comparative benchmarking against existing audio-based TRNGs, such as FM radio and hyperchaotic models. Additionally, integration of universal hash-based post-processors and entropy quantification under adversarial conditions (*e.g., controlled acoustic environments*) will be explored to validate security guarantees.

Author Contributions

All authors contributed to the conception, analysis, writing, revision, and approval of the manuscript.

Acknowledgments

The authors acknowledge their affiliated institutions for academic support.

Conflicts of Interest

The authors declare no conflicts of interest.

REFERENCES

[1] T. Menon and J. Harees, "Survey on using audio video for true random number generation," *2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)*, 2022.

[2] R. Morrison, "Design of a true random number generator using audio input," *Journal of Cryptology*, vol. 1, no. 1, pp. 1–4, 2001.

[3] T. Chen, J.-M. Tsai, and J. Tzeng, "Audio random number generator and its application," *2011 International Conference on Machine Learning and Cybernetics*, Guilin, China, pp. 1678–1683, 2011. doi: 10.1109/ICMLC.2011.6017002.

[4] K. Lee and M. Lee, "True random number generator (TRNG) utilizing FM radio signals for mobile and embedded devices in multi-access edge computing," *Sensors*, vol. 19, no. 19, Art. no. 4130, 2019. doi: 10.3390/s19194130.

[5] A. Krishnan Raghunath, D. Bharadwaj, and M. Prabhuram, "Designing a secured audio based key generator for cryptographic symmetric key algorithms," *Computer Science and Information Technologies*, vol. 2, no. 2, pp. 87–94, 2021. doi: 10.11591/csit.v2i2.p87-94.

[6] J. S. Teh, W. Teng, A. Samsudin, and J. Chen, "A post-processing method for true random number generators based on hyperchaos with applications in audio-based generators," *Frontiers of Computer Science*, vol. 14, no. 6, Art. no. 146405, 2020. doi: 10.1007/s11704-019-9120-2.

[7] O. Safaryan, L. Cherckesova, N. Gapon, I. Alferova, B. Akishin, and E. Semenishchev, "Self-generated model of acoustics encryption by means of true random number generator," *Proc. SPIE, Signal Processing, Sensor/Information Fusion, and Target Recognition XXXIII*, vol. 13057, Art. no. 1305713, Jun. 2024. doi: 10.1117/12.3017674.

[8] S. Karasu and Z. Saraç, "Classification of power quality events signals with pattern recognition methods by using Hilbert transform and genetic algorithms," *2018 26th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, 2018.

[9] M. Bayraktar and M. Özcan, "Hilbert transform based hologram filtering," *2010 IEEE 18th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, 2010.

[10] M. A. Tuğtekin Turan and E. Erzin, "Classification of ingestion sounds using Hilbert-Huang transform," *2017 25th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, 2017.

[11] Y. Peres, "Iterating von Neumann's procedure for extracting random bits," *The Annals of Statistics*, vol. 20, no. 1, pp. 590–597, 1992.

[12] V. B. Suresh and W. P. Burleson, "Entropy extraction in metastability-based TRNG," *2010 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*, pp. 100–105, 2010.

[13] P. Boes, D. Reeb, and M. M. Wolf, "Von Neumann entropy from unitarity," *Physical Review Letters*, vol. 122, no. 21, Art. no. 210402, 2019. doi: 10.1103/PhysRevLett.122.210402.