Uniform distribution of noise

Contents

Abstract

The statistical distribution of the photon numbers has different signatures, depending on the energy source providing their energy. Visible or infrared light from the stars commonly present a Bose-Einstein statistical distribution in the number of photons. The cosmic background at microwave wavelengths follows Planck's blackbody spectral curve, that also fits into the Bose-Einstein statistical distribution of thermal photons.

Electromagnetic field constrained within boundaries manifest themselves in specific field modes and may present specific photon number distributions. For example, the statistical distribution of photon numbers in a coherent laser field (for example, created by an optical reflecting cavity) is Poisson, *within* the coherence time of the field. This statistical distribution of photon numbers associated with the coherent field is a signature of photon numbers in the field.

In the process of detection of these photons, by atoms, molecules, or condensed matter, the interactions may produce a stream of electrons, named photoelectrons. These resulting photoelectron statistics is a Gaussian. These recorded Gaussian events can be used to create random bits. For example, by detecting random *instances* above or below the average number of photoelectrons: **number fluctuations above the average defines a bit 1 and below a bit 0**. This is the entropy source of the KeyBITS random number generator.

Detection within coherence times demand a fast detector, amplification electronics, analog to digital converters and recording processing. The light field statistical properties extend to frequencies well above the used electronics. Therefore, just the electronic recording speed delimitations define the maximum useful bit generation rate: this speed can increase with the technology advances.

Sequences or blocks of random bits represent random numbers. These random numbers, generated by a good entropy source, have a **uniform distribution**, within the limitations imposed by filtering effects due to the electronic circuitry. The random quality of the bits is tested by all randomness tests available. This note gives a practical example of the uniformity level of this uniform distribution.

The electromagnetic field quantum fluctuations inherent to the Universe are the cause of the obtained random numbers … and therefore, also guarantee the secure command/control communication signals by the KeyBITS technology, including protection for IoT and IIoT.

Distribution of noise from recorded quantum fluctuations of a laser beam

The KeyBITS physical random generator continuously extracts random bits from the light intensity fluctuations from a laser beam[1]. Bit sequences can be separated in blocks of length m. Distinct arrangements of these bits in a block gives a random number from 0 to $M-1$, where $M = 2^m$. For example, with $m = 8$ bits, $M = 256$ numbers can be generated.

To exemplify these distributions, using a 3 Gbit sequence of bits from the KeyBITS generator, a set of 21 independent sample sequences were taken, with their sizes being systematically doubled. For each sample, the number's N mean and variance were obtained and compared with the mean and variance of a theoretical uniform distribution.

The PDF of the Uniform Distribution is

$$p(N) = \begin{cases} \dfrac{1}{N_{Max} - N_{min}} \,, & \text{for } N_{min} \le N \le N_{Max} \\ 0\,, & \text{otherwise.} \end{cases} \tag{1}$$

A multiplicative constant α, not shown, may represent the detecting system's efficiency in the raw data set. The Mean of the Uniform Distribution is

$$\text{Mean} = \tfrac{1}{2}(N_{Max} + N_{\text{min}}) \,, \tag{2}$$

and the Variance is given by

$$\text{Variance} = \tfrac{1}{12}(N_{Max} - N_{min})^2 \,. \tag{3}$$

Figure 1 shows δ_{Mean} = (**Mean of data** minus the **Mean of a uniform distribution**) for the 21 samples used. As the number of numbers in each sample increases exponentially, this deviation between the mean of the data and the theoretical value also decreases also exponentially, indicating that the data approaches the theoretical distribution very fast.

[1] US Patent 7,831,050 B2 (2010): "Fast Multi-Photon Key Distribution Scheme Secured by Quantum Noise".

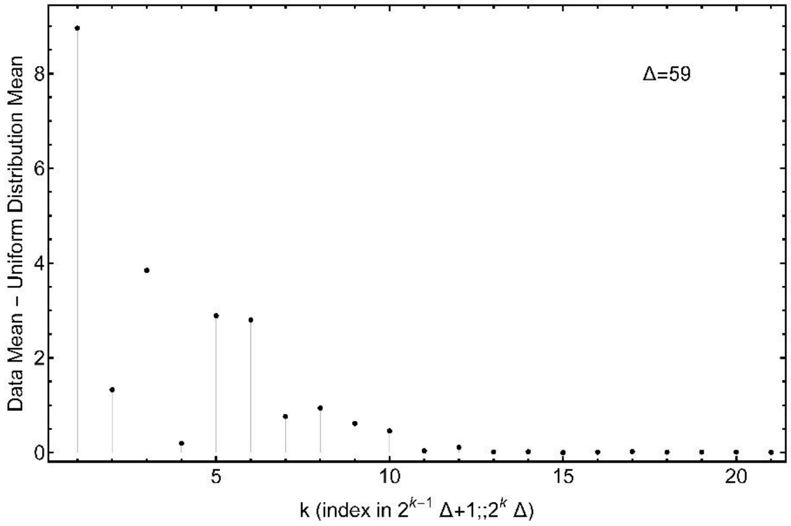

Figure 1 – δ_{Mean} for 21 samples with sizes doubling for each index increase.

Figure 2 shows the $\mathrm{Log}_2\ \delta_{Mean}$, indicating the exponential behavior of δ_{Mean}.

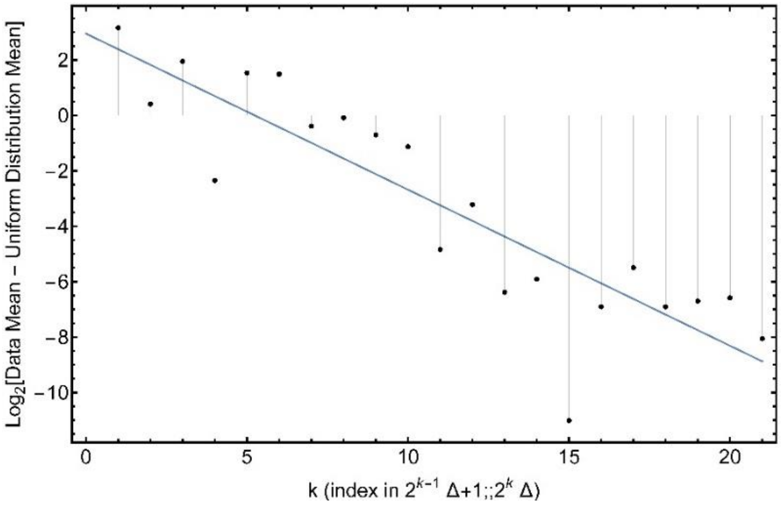

Figure 2 - $\mathrm{Log}_2\ \delta_{Mean}$ for 21 samples with sizes doubling for each index increase. The solid straight line is a fit with $2.950 - 0.563\ k$.

Figure 3 shows δ_{Variance} = (**Variance of data** minus the **Variance of a uniform distribution**).

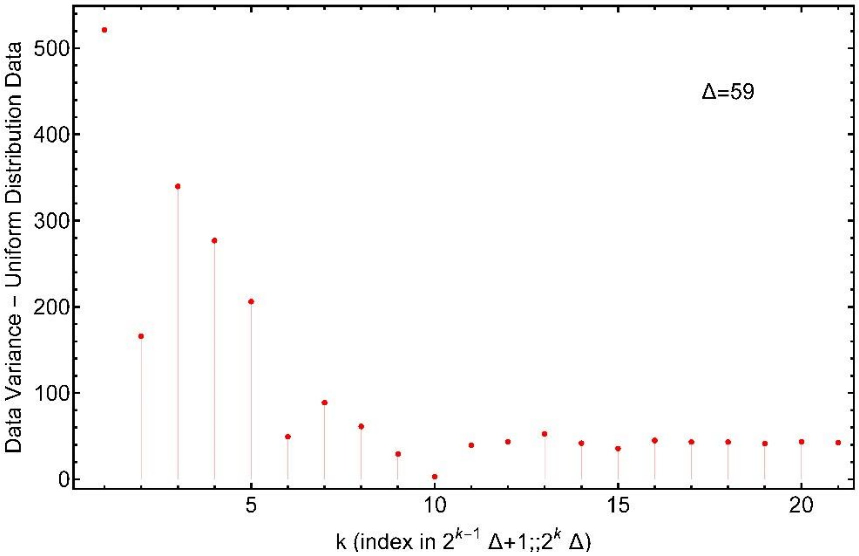

Figure 3 – δ_{Variance} for the 21 samples used. A non-zero limit process is apparent.

The ratio of the limiting value above ($\cong 42.30$) to the average of the **Variance of data** and the **Variance of a uniform distribution** gives ≈ 0.008 , a small deviation.

Practically, the number distribution can be approximated by a Uniform Distribution of numbers.

M-ary levels and the adversary chance to hit the right bit

The KeyBITS technology uses an M-ary coding used to transform signal bits 0 or 1 onto M possible levels. Besides this coding, recorded noise is added to the coded signal to produce random projections around each one of the possible coding levels. The noise amplitude makes the identification of a coding level impossible.

The prediction capability for an adversary to hit the right coding level will be $1/M$. However, as adjacent levels carry opposite bit value, the adversary may hit the bit send by chance, even without knowing the coded level used. With the **uniform distribution** probability of numbers and the added noise, his chance of getting the right bit is $1/2$. For a sequence of n encoded and noisy signal bits, his chance of getting the right signal bit is $1/2^n$, due to the independence of each signal. That means a negligible chance of success. This negligible chance is applicable both for the key distribution stage as well as for the distilled bits used for the one-time-pad encryption.

The Privacy Amplification (PA) protocol used (see "*Generalized Privacy Amplification",* by C. H. Bennett G. Brassard, C. Crepeau, U. M. Maurer. IEEE Transactions on Information Theory, v. 41, pp.1915, 1995. doi: 10.1109/18.476316) establishes that if a certain number t of bits could be leaked to the adversary, this same number over the distilled bits must be discarded. In this case, t is negligible. The Mutual Information I between the adversary and the legitimate users is given by

$$I = \frac{1}{2^{n-(t+r)} \times \ln 2}\,. \tag{4}$$

As discussed, t is negligible and r is the total number of distilled bits. As shown, the adversary has a negligible chance to get the bit sequence and, therefore, t can be made 0. The result is that the Mutual Information is well represented by

$$I = \frac{1}{2^{n-r} \times \ln 2} \ . \tag{5}$$

The difference $n - r$ is the safety parameter, adjusted to reach a specified Mutual Information level. The remaining distilled bits r, are the encryption bits ($\equiv z$).

Figure 4 and Figure 5 shows the Mutual Information I for the cases of starting bits in number $n = 256$ and 512. The referred information is the information that an attacker could have obtained during transmission of the coded noisy bits after the PA protocol. The conjunction of the noisy signals and the shuffling stage of the PA protocol guarantees that even a known plaintext attack cannot discover the encryption bit sequences used before or after the attack. This is true even if quantum resources are used because no deterministic patterns exist in the raw key generation process.

The process is post-quantum secure as I can be made infinitesimally small in practice and no formation rule for the encryption bits exists. Therefore, one-time-pad encryption with these bits fulfills all conditions for practical secrecy and indicates the path for $I \to 0$. For large starting sequences, one can reach negligible values of I. See Figure 4 and Figure 5.

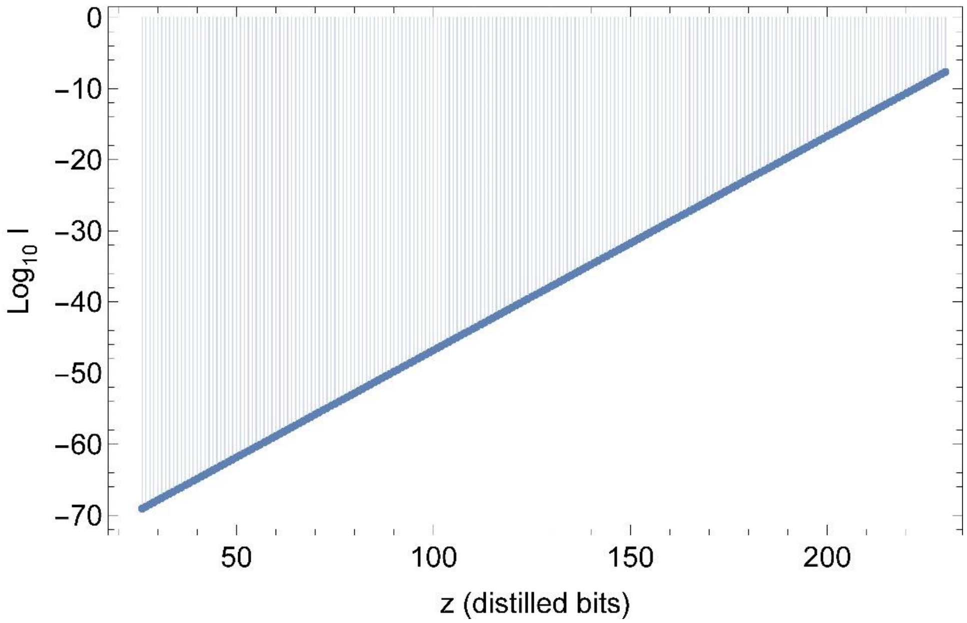

Figure 4 - Logarithm of the Mutual Information as a function of the number of distilled bits z, for a starting sequence of 256 bits.

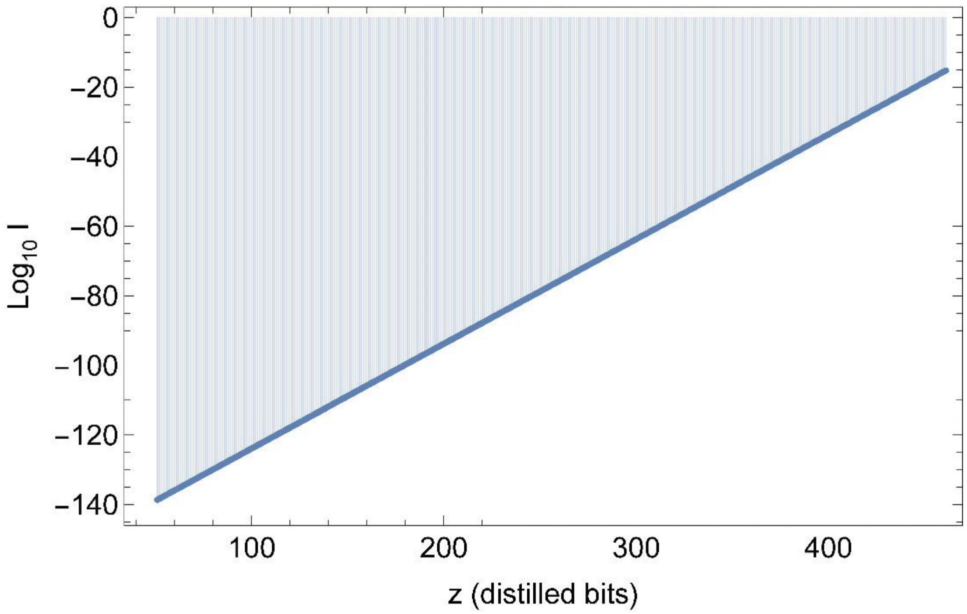

Figure 5 - Logarithm of the Mutual Information as a function of the number of distilled bits z, for a starting sequence of 512 bits.

Other models for light channels

Probability of errors by an attacker have also been done for the case when the signal carrier is a light channel. In that case, the utilized noise to provide protection for the transmitted signals (raw bits) is the intrinsic quantum noise in the carrier itself (laser beam). See "*Untappable key distribution system: a one-time-pad booster*, G. A. Barbosa and J. van de Graaf, arXiv:1406.1543v2 [cs.CR] 8 July 2015; Enigma (Brazilian Journal of Information Security and Cryptography), Vol. **1**, No. 2, 16 (2015)" and "*Fast and secure key distribution using mesoscopic coherent states of light;* G. A. Barbosa, Phys. Rev. A **68**, 052307 (2003)".

For the current treatment of security in **digital** channels, the noise was extracted from an optical source as well but, differently, instances of that noise were **recorded** in short time windows. A record outcome belongs to the classical realm. As such, records can be faithfully copied – as many times one wants, differently from quantum signals. The random property now utilized for the digital signals is that each record is a "projection" of the quantum signal onto the classical world. Being generated by distinct instances of the quantum world they are *also* random by Nature, with no generation algorithm.

The recorded bits originated from a "Gaussian" distribution of signals, and **sequences** of these bits represent **numbers**, and these numbers are uniformly distributed within the range they were digitally created. For example, a sequence with m bits in different orderings can generate $M = 2^m$ numbers, uniformly distributed.

In other words, the randomness of these numbers is a direct reflect from the quantum world, but they are classical. **The central point is that no generation rule is associated with them**.

====================================