IACR Transactions on Cryptographic Hardware and Embedded Systems
ISSN 2569-2925, Vol. 2024, No. 4, pp. 285–306. DOI:10.46586/tches.v2024.i4.285-306

TRNG Entropy Model in the Presence of Flicker FM Noise

Adriaan Peetermans and Ingrid Verbauwhede

COSIC, KU Leuven, Leuven, Belgium, firstname.lastname@esat.kuleuven.be

Abstract. Flicker Frequency Modulated (FM) noise, which influences free-running Ring Oscillators (ROs), can make a substantial contribution to the entropy generated by RO-based True Random Number Generators (TRNGs). While current TRNG stochastic models predominantly concentrate on white FM noise, the addition of flicker FM noise could remarkably enrich the analysis of the TRNG entropy production rate. This paper introduces an entropy model for TRNGs, employing Gaussian processes, to estimate entropy generation from both white FM and flicker FM noise. We analytically derive the flicker FM noise Auto-Correlation Function (ACF), enabling assessment of entropy contributions conditioned on partial knowledge of the TRNG's internal state. Utilizing the developed model with commonly reported noise magnitudes found in literature, it is determined that flicker FM noise holds the potential to substantially enhance the TRNG's entropy rate. However, due to considerable variation in reported magnitudes across limited available research on flicker FM noise, it cannot yet be universally accepted as a dependable source of TRNG entropy.

Keywords: TRNG · Stochastic entropy model · White FM noise · Flicker FM noise

1 Introduction

As a fundamental building block capable of delivering a stream of fresh entropy, TRNGs play an indispensable role in modern cryptographic systems. ROs are a popular component in digital TRNG designs, as they are well studied, easily implemented and integrated into a digital architecture and allow for some frequency flexibility, by varying the number of stages. Consequently, numerous RO-based TRNG designs are found in literature, e.g. the Elementary Ring Oscillator (ERO) TRNG [BLMT11], the Transition Effect Ring Oscillator (TERO) TRNG [VD10], the Edge Sampling (ES) TRNG [YRG+18] or the Coherent Sampling (COSO) TRNG [KG04].

Stochastic models, capable of estimating the TRNG entropy production are mandatory by international standards such as those set by BSI [PS22], NIST [TBK+18] and ISO [ISO19]. Despite the abundance of TRNG stochastic models, many relying on the presence of only white FM noise as seen in [PV22, HFBN15, BLMT11], the significance of flicker FM noise was often disregarded or deemed less important. This is typically justified by the rationale that accounting solely for the entropy provided by the white FM noise component is adequate for establishing a lower bound on the entropy produced by the TRNG [LF24]. Any entropy generated by other independent noise components was regarded as a surplus that would not invalidate the entropy bound previously declared.

Increased attention for the flicker FM noise component [HTBF14, LB15, PV24] underscores the need to incorporate this type of noise into TRNG models. Particularly, recent research [BCF+24] has indicated that flicker FM noise can make a meaningful contribution to the TRNG's entropy production, thus emphasizing the importance of

Received: 2024-04-15 Accepted: 2024-06-15 Published: 2024-09-05

accurately describing it. Although [BCF+24] proposes a method for generating period length samples for an RO under the influence of white FM and flicker FM noise, it lacks a rigorous mathematical analysis to derive the TRNG output entropy density. This absence is particularly unfortunate given the intriguing claim that an increased flicker FM magnitude tends to reduce the observed correlation of the generated output bits. Instead, the study relies on simulation results fitted to empirical data. Moreover, it assumes that the absence of linear correlation in the output bits proves independence, which is crucial for the validity of the reported entropy results.

Precisely modeling flicker FM noise is challenging due to its inherent long-term dependencies, which arise from the physical nature of the charge carrier trapping and detrapping process in the transistors that comprise the oscillator [GRN+91]. A trapped charge carrier can affect the transistor's drive strength over multiple oscillation periods, leading to a sustained increase or decrease in the oscillating frequency. A notable effort was made by [PV24], where a time-based analysis revealed the dependency of an oscillator's excess phase variance on the accumulation time length. However, this work lacks a comprehensive description of the oscillator phase and a method for characterizing phase dependencies. Additionally, it does not provide a method for estimating the entropy induced by flicker FM noise. Instead [PV24] focuses on quantifying the magnitudes of the prevalent noise types in free-running oscillators.

Apart from [PV24], there are only a limited number of flicker FM magnitude estimates available: [HTBF14, LB15, FL14, BCF+24], which span multiple orders of magnitude. As demonstrated in this study, the specific magnitude employed, notably impacts the resulting entropy estimate, thereby dictating the extent to which the contribution of flicker FM noise outweighs that of white FM noise in the total TRNG entropy rate.

The primary contributions of this work are as follows:

- An analytical derivation of the oscillator excess phase ACF, influenced by flicker FM-shaped noise sources, is presented as a generalization of the phase variance derivation in [PV24].
- This work proposes a simulation method for the phase process, similar to the approach presented by [BCF+24]. However, our method offers a more robust mathematical foundation. We achieve this by developing a unified Gaussian process model that incorporates the ACF for both white FM and flicker FM noise.
- Simulation results using the constructed model compare the entropy produced in an ERO TRNG under the influence of white FM and flicker FM noise, conditioned on the oscillator's phase or previously produced bit values.

2 Excess phase process

The phase of a noisy oscillator, that starts running at time $t = 0$, is modeled as a real-valued stochastic process through continuous time $t \in \mathbb{R}_{\geq 0}$:

$$\Phi(t) = 2\pi f_n t + \phi_0 + \Phi_e(t) \qquad \text{for } t \in \mathbb{R}_{\geq 0},$$

with f_n, the nominal oscillator frequency, ϕ_0, the initial phase at time $t = 0$ and $\{\Phi_e(t)\}_{t\in\mathbb{R}_{\geq 0}}$, a real-valued stochastic process describing the oscillator's excess phase. The excess phase is assumed to be unbiased (equal probability to become positive as to become negative in value) through time. Therefore: $\forall t \in \mathbb{R}_{\geq 0} : \mathbf{E}[\Phi_e(t)] = 0$, with $\mathbf{E}$ the expectation operator.

The ACF is denoted as follows: $R_A : \mathbb{R}_T^2 \to \mathbb{R}$ by $R_A(t_i, t_j) = \mathbf{E}[A(t_i)A(t_j)]$, with $\{A(t)\}_{t\in T}$ any stochastic process on $t \in T$. Instead of the phase itself, properties of

the oscillator are often described using the instantaneous relative frequency deviation: $\forall t \in \mathbb{R}_{\geq 0} : Y(t) = \frac{\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t) - 2\pi f_n}{2\pi f_n}$. The phase ACF can be written in terms of the ACF for this relative frequency deviation, by generalizing Eq. (8) from [PV24], $\forall (t_i, t_j) \in \mathbb{R}^2_{\geq 0}$:

$$\begin{aligned} R_{\Phi_e}(t_i, t_j) &= \mathbf{E}\big[\Phi_e(t_i)\Phi_e(t_j)\big] = \mathbf{E}\left[2\pi f_n \int_0^{t_i} Y(\theta_i)\mathrm{d}\theta_i 2\pi f_n \int_0^{t_j} Y(\theta_j)\mathrm{d}\theta_j\right] \\ &= 4\pi^2 f_n^2 \int_0^{t_i}\int_0^{t_j} \mathbf{E}\big[Y(\theta_i)Y(\theta_j)\big]\mathrm{d}\theta_j\mathrm{d}\theta_i = 4\pi^2 f_n^2 \int_0^{t_i}\int_0^{t_j} R_Y(\theta_i, \theta_j)\mathrm{d}\theta_j\mathrm{d}\theta_i. \end{aligned} \tag{1}$$

2.1 Gaussian process

The excess phase: $\big\{\Phi_e(t)\big\}_{t\in\mathbb{R}_{\geq 0}}$ is assumed to behave as a Gaussian process, with zero mean function: $\forall t \in \mathbb{R}_{\geq 0} : \mu(t) = 0$. The ACF fully describes the behavior of the excess phase process. Sampling the excess phase at n time instances: $(t_0, t_1, \ldots, t_{n-1}) \in \mathbb{R}^n_{\geq 0}$, produces an n-dimensional multivariate normal distributed vector:

$$\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}} = \big(\Phi_e(t_0), \Phi_e(t_1), \ldots, \Phi_e(t_{n-1})\big)^{\intercal} \sim \mathcal{N}_n\big(\vec{\mathbf{0}}_{\boldsymbol{n}}, \boldsymbol{\Sigma_e}\big),$$

with mean vector: $\vec{\mathbf{0}}_{\boldsymbol{n}}$, an $n \times 1$ all-zero vector and covariance matrix given by

$$\boldsymbol{\Sigma_e} = \begin{bmatrix} R_{\Phi_e}(t_0,t_0) & R_{\Phi_e}(t_0,t_1) & \ldots & R_{\Phi_e}(t_0,t_{n-1}) \\ R_{\Phi_e}(t_1,t_0) & R_{\Phi_e}(t_1,t_1) & \ldots & R_{\Phi_e}(t_1,t_{n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ R_{\Phi_e}(t_{n-1},t_0) & R_{\Phi_e}(t_{n-1},t_1) & \ldots & R_{\Phi_e}(t_{n-1},t_{n-1}) \end{bmatrix}. \tag{2}$$

2.2 Source of noise

The Power Spectral Density (PSD) for the relative frequency deviation, denoted by $S_Y(f)$, with f the Fourier frequency, is assumed to be composed of a sum of noise contributions of different type α [HAB81]:

$$S_Y(f) = \sum_{\alpha=-2}^{2} h_\alpha |f|^\alpha = \sum_{\alpha=-2}^{2} S_{Y^\alpha}(f). \tag{3}$$

The noise magnitudes, h_α, represent the magnitudes of the five most prevalent noise types in oscillators: random walk FM ($\alpha = -2$), flicker FM ($\alpha = -1$), white FM ($\alpha = 0$), flicker PM ($\alpha = 1$), and white PM ($\alpha = 2$). The terms *white* and *flicker* refer to the shape of the oscillator's frequency or phase spectrum, hence the adjectives *Frequency Modulated (FM)* and *Phase Modulated (PM)* are used in this text. Additionally, the contributions are assumed mutually independent. Using the derivation outlined in Appendix A, the excess phase random process is shown to consist of a sum of individual excess phase noise contributions, $\big\{\Phi_e^\alpha(t)\big\}_{t\in\mathbb{R}_{\geq 0}}$: $\forall t \in \mathbb{R}_{\geq 0} : \Phi_e(t) = \sum_{\alpha=-2}^{2} \Phi_e^\alpha(t)$. Each of these contributions has a PSD following a power law: $S_{\Phi_e^\alpha}(f) = \big(\frac{f_n}{f}\big)^2 h_\alpha |f|^\alpha$.

As experiments in [PV24] have shown, the noise types of interest when studying an oscillator for a reasonable time frame are white FM ($\alpha = 0$) and flicker FM ($\alpha = -1$) noise. In this work, these two noise types are exclusively studied and indicated by the indices $\cdot^w$ and $\cdot^f$ for white FM and flicker FM components respectively. The constants $h_w = h_0$ and $h_f = h_{-1}$ are used to indicate the noise magnitudes.

2.3 Excess phase ACF

2.3.1 White FM noise

As assumed by [PV24], the relative frequency deviation, $\{Y(t)\}_{t\in\mathbb{R}_{\geq 0}}$, is a Wide-Sense Stationary (WSS) process. Its ACF is therefore only dependent on the time shift and a simplified notation is used: $\forall (t_i, t_j) \in \mathbb{R}^2_{\geq 0} : R_Y(t_i, t_j) = R_Y(t_j - t_i) = R_Y(\tau)$. Derived in [PV24], under the influence of only white FM noise, the relative frequency deviation ACF equals a scaled Dirac delta distribution function: $R_Y(\tau) = h_0\delta(\tau)$. Using Eq. (1), the excess phase ACF becomes

$$\begin{aligned}\forall\ (t_i, t_j) \in \mathbb{R}^2_{\geq 0} : R_{\phi_e}(t_i, t_j) &= 4\pi^2 f_n^2 h_w \int_0^{t_i}\int_0^{t_j} \delta(\theta_j - \theta_i)\mathrm{d}\theta_j\mathrm{d}\theta_i \\ &= 4\pi^2 f_n^2 h_w \min(t_i, t_j).\end{aligned} \tag{4}$$

2.3.2 Flicker FM noise

Generalizing the derivation made in Section II-D from [PV24], the excess phase ACF can be obtained. The same assumption is made here: a band-limited version of the relative frequency deviation, $\{Y(t)\}_{t\in\mathbb{R}_{\geq 0}}$, to the frequency interval $[f_l, f_h]$, under the action of flicker FM noise, is WSS. From Eq. (13) in [PV24], the relative frequency deviation ACF equals

$$R_Y(\tau) = 2h_f \int_{f_l}^{f_h} \frac{\cos(2\pi f\tau)}{f}\mathrm{d}f.$$

Using Eq. (1), the excess phase ACF can be obtained:

$$\forall (t_i, t_j) \in \mathbb{R}^2_{\geq 0} : R_{\Phi_e}(t_i, t_j) = 4\pi^2 f_n^2 \int_0^{t_i}\int_0^{t_j} 2h_f \int_{f_l}^{f_h} \frac{\cos\big(2\pi f(\theta_j - \theta_i)\big)}{f}\mathrm{d}f\mathrm{d}\theta_j\mathrm{d}\theta_i. \tag{5}$$

The details of working out the integral and assuming $f_h \to \infty$ are shown in Appendix B. The following could be obtained:

$$\begin{aligned}\forall (t_i, t_j) \in \mathbb{R}^2_{>0} : R_{\Phi_e}(t_i, t_j) = 4\pi^2 f_n^2 h_f t_i t_j \Bigg(&3 - 2\gamma - 2\ln\big(2\pi f_l |t_j - t_i|\big) \\ &+ \frac{t_i}{t_j}\ln\left(\frac{|t_j - t_i|}{t_i}\right) + \frac{t_j}{t_i}\ln\left(\frac{|t_j - t_i|}{t_j}\right)\Bigg).\end{aligned} \tag{6}$$

The ACF becomes zero whenever $t_i = 0$ or $t_j = 0$. When $t_i = t_j = t$, the $\ln|t_j - t_i|$ terms cancel out and Eq. (6) reduces to Eq. (19) in [PV24].

Example 1. Take the following realistic values: $f_n = 520\,\mathrm{MHz}$, $h_w = 18.9\,\mathrm{fs}$, $h_f = 1\times 10^{-10}$ and $f_l = 1\,\mathrm{mHz}$. Figure 1 illustrates an example white FM and flicker FM noise Gaussian process: $\{\Phi_e^w(t)\}_{t\in\mathbb{R}_{\geq 0}}$ (blue) and $\{\Phi_e^f(t)\}_{t\in\mathbb{R}_{\geq 0}}$ (red). The markers show four realizations: $\varphi_e^y(t_i)$ for $t_i \in \{0\,\mu\mathrm{s}, 1\,\mu\mathrm{s}, 3\,\mu\mathrm{s}, 4\,\mu\mathrm{s}\}$ and $y \in \{w, f\}$. The confidence region (one standard deviation above and below the mean containing 68 % of the samples) and mean through time: $\sqrt{\mathbf{Var}\big[\Phi_e^y(t)\big]}$ and $\mathbf{E}\big[\Phi_e^y(t)\big]$, given these four realizations, are represented by a shaded area and a solid line respectively. Note the linear increase versus the square root increase of the standard deviation through time for flicker FM versus white FM noise.

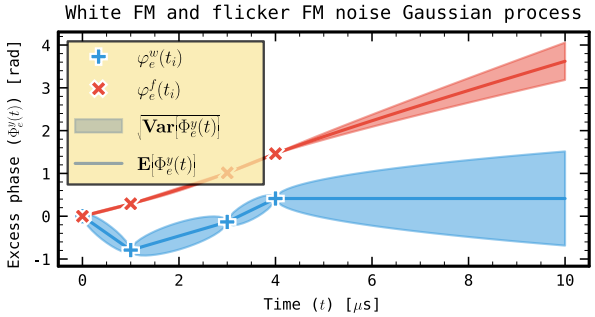

Figure 1: Gaussian process for white FM (blue) and flicker FM (red) noise in Example 1. The markers represent realized values for the excess phase process. The shaded area and the solid line represent the standard deviation and mean for the process at any time instant respectively, given the realizations: $\Phi_e^y(t_i) = \varphi_e^y(t_i)$ for $t_i \in \{0\,\mu s, 1\,\mu s, 3\,\mu s, 4\,\mu s\}$ and $y \in \{w, f\}$.

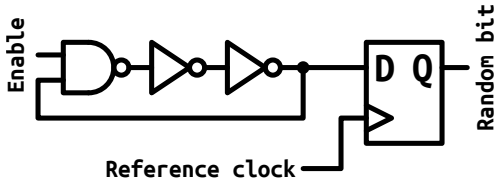

Figure 2: ERO TRNG reference architecture.

3 ERO TRNG entropy model

To study the effects of flicker FM noise on the entropy generated by a TRNG, the ERO TRNG is selected as the reference architecture in this work. As shown in Fig. 2, the ERO TRNG being studied consists of a single free-running RO, a reference clocking signal and a sampling flip-flop. Although this work assumes the reference clock is jitter-free, it is important to note that, in reality, the reference clock may also exhibit jitter. Techniques exist to transfer the jitter from the reference clock to the free-running oscillator under study [FL14], allowing the results presented here to remain valid. However, when the origin of the jitter in the reference clock is unknown, it should not be considered in the entropy estimate.

From Section 2, both the white and flicker excess phase components can be considered a Gaussian process. A random excess phase vector is defined as

$$\vec{\mathbf{\Phi}}_{\boldsymbol{e}} = \left(\Phi_e^w(t_1), \Phi_e^w(t_2), \ldots, \Phi_e^w(t_n), \Phi_e^f(t_1), \Phi_e^f(t_2), \ldots, \Phi_e^f(t_n)\right)^\intercal,$$

describing the sampled excess phase at time instances $(t_1, t_2, \ldots, t_n) \in \mathbb{R}^n_{\geq 0}$. The excess phase vector is then multivariate normal distributed, as the white and flicker noise components are independent: $\vec{\mathbf{\Phi}}_{\boldsymbol{e}} \sim \mathcal{N}_{2n}\left(\vec{\mathbf{0}}_{\boldsymbol{2n}}, \mathbf{\Sigma}_{\boldsymbol{e}}\right)$. The covariance matrix is constructed as

$$\mathbf{\Sigma}_e = \begin{bmatrix} \mathbf{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{w}} & \mathbf{0}_{\boldsymbol{n}\times\boldsymbol{n}} \\ \mathbf{0}_{\boldsymbol{n}\times\boldsymbol{n}} & \mathbf{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{f}} \end{bmatrix}, \tag{7}$$

with $\mathbf{0}_{\boldsymbol{n}\times\boldsymbol{n}}$, an $n \times n$ all zero matrix. The $n \times n$ matrices $\mathbf{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{w}}$ and $\mathbf{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{f}}$ are the covariance matrices for the individual white and flicker noise components respectively, constructed using the noise ACF, as illustrated in Eq. (2).

3.1 Bit distribution

In an ERO TRNG, the oscillator is sampled at regular time intervals: $t_i = it_{acc}$ for $i \in \{1, 2, \ldots, n\}$ and accumulation time, t_{acc}. The sampled value of the i-th bit, B_i, equals

$$B_i = \left\lfloor \frac{\Phi(t_i)}{\pi} \right\rfloor \bmod 2 = \left\lfloor \frac{2\pi f_n t_i + \phi_0 + \Phi_e(t_i)}{\pi} \right\rfloor \bmod 2 = b_i^d \oplus \left\lfloor \frac{\phi_i + \Phi_e(t_i)}{\pi} \right\rfloor \bmod 2, \quad (8)$$

with $b_i^d = \left\lfloor \frac{2\pi f_n t_i + \phi_0}{\pi} \right\rfloor \bmod 2$ and $\phi_i = (2\pi f_n t_i + \phi_0) \bmod \pi$, the number of completed half cycles modulo two and fractional part of the current oscillator half cycle, both deterministic quantities and $\oplus$, the binary XOR operator. XORing with b_i^d can be considered a form of post-processing, as it does not alter the entropy content of the bit B_i. Remove the XOR post-processing to obtain an adjusted bit with identical entropy content:

$$B_i' = \left\lfloor \frac{\phi_i + \Phi_e(t_i)}{\pi} \right\rfloor \bmod 2 = \left\lfloor \frac{\phi_i + \Phi_e^w(t_i) + \Phi_e^f(t_i)}{\pi} \right\rfloor \bmod 2.$$

B_i' is a discrete binary random variable, dependent on the oscillator excess phase, $\Phi_e(t_i)$, at sampling time t_i. The conditional Probability Mass Function (PMF), given the excess phase equals

$$f_{B_i'|\Phi_e(t_i)}(b \mid \varphi) = \begin{cases} 1 & \text{if } b = \left\lfloor \frac{\phi_i + \varphi}{\pi} \right\rfloor \bmod 2 \\ 0 & \text{otherwise.} \end{cases}$$

When considering m sample time instances: $\forall j \in \{0, 1, \ldots, m-1\}, \forall i_j \in \{1, 2, \ldots, n\}$: $(t_{i_0}, t_{i_1}, \ldots, t_{i_{m-1}})^\intercal \in \mathbb{R}_{\geq 0}^m$, define the random bit vector $\vec{\boldsymbol{B}} = (B_{i_0}', B_{i_1}', \ldots, B_{i_{m-1}}')^\intercal$, with conditional PMF equal to

$$f_{\vec{\boldsymbol{B}}|\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}}(\vec{\boldsymbol{b}} \mid \vec{\boldsymbol{\varphi}}) = \prod_{j=0}^{m-1} f_{B_{i_j}'|\Phi_e(t_{i_j})}(b_{i_j} \mid \varphi_{i_j - 1} + \varphi_{i_j - 1 + n}), \quad (9)$$

where b_{i_j}, φ_{i_j-1}, and φ_{i_j-1+n} are the elements at the j-th, $(i_j - 1)$-th, and $(i_j - 1 + n)$-th position in the vectors $\vec{\boldsymbol{b}}$, and $\vec{\boldsymbol{\varphi}}$, respectively. This PMF only equals one if for all m indexes, i_j, the bit b_{i_j} equals $\left\lfloor \frac{\phi_{i_j} + \varphi_{i_j-1} + \varphi_{i_j-1+n}}{\pi} \right\rfloor \bmod 2$, with φ_{i_j-1} and φ_{i_j-1+n}, the given excess white and flicker phase respectively for the index corresponding with sampling time t_{i_j}.

3.2 Conditional distributions

This work assumes that an entity has observed or will observe a certain amount of information about the ERO TRNG. This information could include the exact oscillator phase value, $\Phi_e^w(t_i)$ and $\Phi_e^f(t_i)$, or the produced output bit B_i' at specific sampling time instances: $t_i \in \mathbb{R}_{\geq 0}$. Take as an example: person A collecting future or previous TRNG output. The collection of already observed random variables is referred to as the observed part. The conditional distributions developed in this study, describe the distribution for an unobserved part (phase or bit values), from the perspective of that entity who already possesses knowledge of the observed part. For instance: the current bits' distribution for person A, given this person has the knowledge of different bits produced in the future or past.

The n sampling time instances of interest are now partitioned in an observed part $\{t_{i_0^o}, t_{i_1^o}, \ldots, t_{i_{n^o-1}^o}\}$ and an unobserved part $\{t_{i_0^u}, t_{i_1^u}, \ldots, t_{i_{n^u-1}^u}\}$, containing n^o and n^u time instances respectively and $n = n^o + n^u$, the total number of samples under study. A sample instance cannot be both observed and unobserved: $\forall (k, j) \in \{0, 1, \ldots, n^u - 1\} \times \{0, 1, \ldots, n^o - 1\} : i_k^u \neq i_j^o$.

The excess phase vector can be similarly partitioned (by reordering the indexes) into observed and unobserved parts $\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}} = \left(\boldsymbol{\Phi}_{\boldsymbol{e}}^{\vec{\boldsymbol{o}},\boldsymbol{w}^\intercal}, \boldsymbol{\Phi}_{\boldsymbol{e}}^{\vec{\boldsymbol{u}},\boldsymbol{w}^\intercal}, \boldsymbol{\Phi}_{\boldsymbol{e}}^{\vec{\boldsymbol{o}},\boldsymbol{f}^\intercal}, \boldsymbol{\Phi}_{\boldsymbol{e}}^{\vec{\boldsymbol{u}},\boldsymbol{f}^\intercal}\right)^\intercal$, with: $\forall (x,y) \in \{o,u\} \times \{w,f\} : \boldsymbol{\Phi}_{\boldsymbol{e}}^{\vec{\boldsymbol{x}},\boldsymbol{y}} = \left(\Phi_e^y(t_{i_0^x}), \Phi_e^y(t_{i_1^x}), \ldots, \Phi_e^y(t_{i_{n^x-1}^x})\right)^\intercal$. Similarly, the random bit vectors representing the observed and unobserved TRNG output equal $\forall x \in \{o,u\} : \vec{\boldsymbol{B}}^{\boldsymbol{x}} = (B'_{i_0^x}, B'_{i_1^x}, \ldots, B'_{i_{n^x-1}^x})^\intercal$.

3.2.1 Conditioned on the oscillator phase

The unobserved excess phase vector, $\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}} = \left(\boldsymbol{\Phi}_{\boldsymbol{e}}^{\vec{\boldsymbol{u}},\boldsymbol{w}^\intercal}, \boldsymbol{\Phi}_{\boldsymbol{e}}^{\vec{\boldsymbol{u}},\boldsymbol{f}^\intercal}\right)^\intercal$, given a realization of the observed excess phase, $\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o}} = \left(\boldsymbol{\Phi}_{\boldsymbol{e}}^{\vec{\boldsymbol{o}},\boldsymbol{w}^\intercal}, \boldsymbol{\Phi}_{\boldsymbol{e}}^{\vec{\boldsymbol{o}},\boldsymbol{f}^\intercal}\right)^\intercal = \left(\vec{\boldsymbol{\varphi}}^{\boldsymbol{w}^\intercal}, \vec{\boldsymbol{\varphi}}^{\boldsymbol{f}^\intercal}\right)^\intercal = \vec{\boldsymbol{\varphi}}^{\boldsymbol{o}}$, has a multivariate normal conditional Probability Density Function (PDF):

$$f_{\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}} | \vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o}}}\left(\vec{\boldsymbol{\varphi}}^{\boldsymbol{u}} \mid \vec{\boldsymbol{\varphi}}^{\boldsymbol{o}}\right) = \phi_{\mathcal{N}_{2n^u}}\left(\vec{\boldsymbol{\varphi}}^{\boldsymbol{u}}; \vec{\boldsymbol{\mu}}_{\boldsymbol{e}}^{\boldsymbol{u}|\boldsymbol{o}}, \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{u}|\boldsymbol{o}}\right), \tag{10}$$

with conditional covariance matrix equal to

$$\boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{u}|\boldsymbol{o}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uu},\boldsymbol{w}} - \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uo},\boldsymbol{w}} \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{oo},\boldsymbol{w}^{-1}} \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{ou},\boldsymbol{w}} & \boldsymbol{0}_{\boldsymbol{n^u} \times \boldsymbol{n^u}} \\ \boldsymbol{0}_{\boldsymbol{n^u} \times \boldsymbol{n^u}} & \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uu},\boldsymbol{f}} - \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uo},\boldsymbol{f}} \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{oo},\boldsymbol{f}^{-1}} \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{ou},\boldsymbol{f}} \end{bmatrix}, \tag{11}$$

and conditional mean vector equal to

$$\vec{\boldsymbol{\mu}}_{\boldsymbol{e}}^{\boldsymbol{u}|\boldsymbol{o}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uo},\boldsymbol{w}} \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{oo},\boldsymbol{w}^{-1}} \vec{\boldsymbol{\varphi}}^{\boldsymbol{w}} \\ \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uo},\boldsymbol{f}} \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{oo},\boldsymbol{f}^{-1}} \vec{\boldsymbol{\varphi}}^{\boldsymbol{f}} \end{bmatrix}.$$

The submatrices are derived from the excess phase covariance matrix in Eq. (7):

$$\boldsymbol{\Sigma}_{\boldsymbol{e}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{oo},\boldsymbol{w}} & \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{ou},\boldsymbol{w}} & & \\ \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uo},\boldsymbol{w}} & \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uu},\boldsymbol{w}} & \boldsymbol{0}_{\boldsymbol{n} \times \boldsymbol{n}} & \\ & & \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{oo},\boldsymbol{f}} & \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{ou},\boldsymbol{f}} \\ \boldsymbol{0}_{\boldsymbol{n} \times \boldsymbol{n}} & & \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uo},\boldsymbol{f}} & \boldsymbol{\Sigma}_{\boldsymbol{e}}^{\boldsymbol{uu},\boldsymbol{f}} \end{bmatrix},$$

Note that only the conditional mean is dependent on the phase realization.

3.2.2 Conditioned on the output bits

Using Bayes' rule, the conditional PDF for the random excess phase vector, $\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}$, given the observation of n^o bits $\vec{\boldsymbol{B}}^{\boldsymbol{o}} = \vec{\boldsymbol{b}}$, equals

$$f_{\vec{\boldsymbol{\Phi}}_e | \vec{\boldsymbol{B}}^o}(\vec{\boldsymbol{\varphi}} \mid \vec{\boldsymbol{b}}) = \frac{f_{\vec{\boldsymbol{B}}^o | \vec{\boldsymbol{\Phi}}_e}(\vec{\boldsymbol{b}} \mid \vec{\boldsymbol{\varphi}}) f_{\vec{\boldsymbol{\Phi}}_e}(\vec{\boldsymbol{\varphi}})}{f_{\vec{\boldsymbol{B}}^o}(\vec{\boldsymbol{b}})}, \tag{12}$$

with the unconditional excess phase PDF equal to the $2n$-dimensional multivariate normal distribution PDF, $f_{\vec{\boldsymbol{\Phi}}_e}(\vec{\boldsymbol{\varphi}}) = \phi_{\mathcal{N}_{2n}}(\vec{\boldsymbol{\varphi}}; \vec{\boldsymbol{0_{2n}}}, \boldsymbol{\Sigma}_{\boldsymbol{e}})$. The unconditional marginal PMF, for a bit vector, $f_{\vec{\boldsymbol{B}}}(\vec{\boldsymbol{b}})$ or equivalently $f_{\vec{\boldsymbol{B}}^o}(\vec{\boldsymbol{b}})$, can be found by integrating the unconditional excess phase PDF over the subspace where $f_{\vec{\boldsymbol{B}} | \vec{\boldsymbol{\Phi}}_e}(\vec{\boldsymbol{b}} \mid \vec{\boldsymbol{\varphi}})$, from Eq. (9), equals one:

$$f_{\vec{\boldsymbol{B}}}(\vec{\boldsymbol{b}}) = \int_{\mathbb{R}^{2n}} f_{\vec{\boldsymbol{B}} | \vec{\boldsymbol{\Phi}}_e}(\vec{\boldsymbol{b}} \mid \vec{\boldsymbol{\varphi}}) f_{\vec{\boldsymbol{\Phi}}_e}(\vec{\boldsymbol{\varphi}}) \mathrm{d}\vec{\boldsymbol{\varphi}}. \tag{13}$$

Example 2. Take the following realistic parameter values: $f_n = 520\,\mathrm{MHz}$, $t_{acc} = 4.11\,\mu\mathrm{s}$, $\phi_0 = 0\,\mathrm{rad}$ (initial oscillator phase), $h_w = 18.9\,\mathrm{fs}$ (white noise strength, as in [PV24]), and $h_f = 100 \times 10^{-12}$ (flicker noise strength, with a noise corner around $5\,\mu\mathrm{s}$). In this example,

two sampling time instances are considered: $t_1 = t_{acc}$ and $t_2 = 2t_{acc}$. The random excess phase vector becomes $\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}} = \left(\Phi_e^w(t_1), \Phi_e^w(t_2), \Phi_e^f(t_1), \Phi_e^f(t_2)\right)^\intercal$. The generated bit from the first sample is observed equal to one, $b_{i_0^o} = 1$ for $i_0^o = 1$ and $\vec{\boldsymbol{B}}^{\boldsymbol{o}} = [B'_{i_0^o}]$. The probability of sampling a one at the first sampling time instance, is obtained by integrating the unconditional total oscillator phase PDF over the area highlighted in Fig. 3 and equals

$$f_{\vec{\boldsymbol{B}}^{\boldsymbol{o}}}([1]) = \int_{\left\lfloor \frac{\phi_1+\varphi_0+\varphi_2}{\pi} \right\rfloor \bmod 2=1} \phi_{\mathcal{N}_4}(\vec{\boldsymbol{\varphi}}; \vec{\mathbf{0}}_{\mathbf{4}}, \boldsymbol{\Sigma}_{\boldsymbol{e}}) \mathrm{d}\vec{\boldsymbol{\varphi}} = 21.3\,\%,$$

with the unconditional excess phase covariance matrix equal to

$$\boldsymbol{\Sigma}_{\boldsymbol{e}} = \begin{bmatrix} R^w_{\Phi_e}(t_1,t_1) & R^w_{\Phi_e}(t_1,t_2) & 0 & 0 \\ R^w_{\Phi_e}(t_2,t_1) & R^w_{\Phi_e}(t_2,t_2) & 0 & 0 \\ 0 & 0 & R^f_{\Phi_e}(t_1,t_1) & R^f_{\Phi_e}(t_1,t_2) \\ 0 & 0 & R^f_{\Phi_e}(t_2,t_1) & R^f_{\Phi_e}(t_2,t_2) \end{bmatrix}.$$

Equation (12) determines that the knowledge of the first bit being equal to one, changes the distributions for the white- and flicker excess phase at the second sample, as illustrated by Figs. 4 and 5.

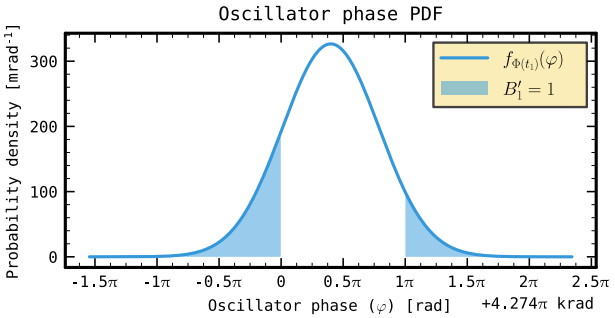

Figure 3: Oscillator total phase PDF and integration area for the first sample being equal to one.

3.3 Monte Carlo integration

Example 2 was generated by evaluating the four-dimensional multivariate normal PDF, $\phi_{\mathcal{N}_4}(\cdot)$, at a four-dimensional grid. While providing accurate results, this approach quickly

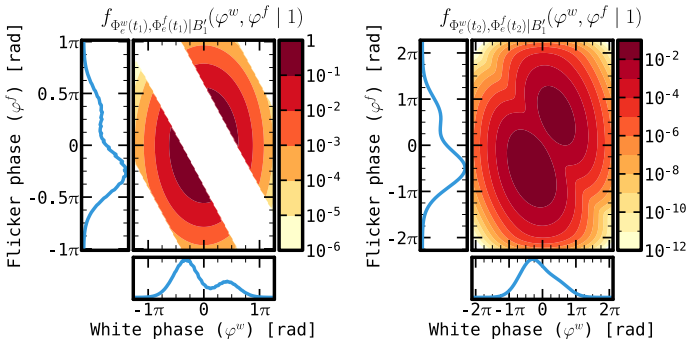

Figure 4: Joint PDFs for the white and flicker excess phase at the first sampling time instance: t_1 (left) and second sampling time instance: t_2 (right), conditioned on the first sampled bit being equal to one.

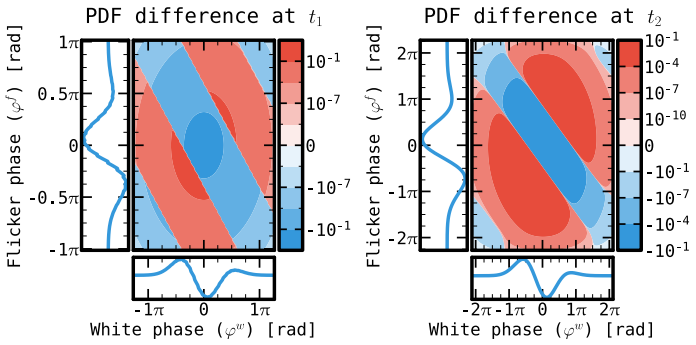

Figure 5: Difference between the unconditioned joint PDF and the conditioned joint PDF: $f_{\Phi_e^w(t_i),\Phi_e^f(t_i)|B_1'}(\varphi^w,\varphi^f \mid 1) - f_{\Phi_e^w(t_i),\Phi_e^f(t_i)}(\varphi^w,\varphi^f)$, for t_i equal to t_1 (left) or t_2 (right).

becomes intractable when dealing with more dimensions (i.e. the number of samples at interest increases). When evaluating Eq. (12), only the denominator poses a problem. The unconditional bit PMF, $f_{\vec{\boldsymbol{B}}}(\vec{\boldsymbol{b}})$, requires the integration of a multivariate normal PDF in a $2n$-dimensional subspace, as shown by Eq. (13). No closed form solution exists when there is correlation between the individual random variables or the integration boundaries depend on more than one random variable.

The knowledge of the generated bit at time t_{i_j}: $B'_{i_j} = b_{i_j}$, leads to an additional factor in Eq. (9). The excess phase becomes additionally constrained by the relation: $\left\lfloor \frac{\phi_{i_j} + \Phi_e(t_{i_j})}{\pi} \right\rfloor \bmod 2 = b_{i_j}$, which generates a periodic tilted bands integration region, with period 2π, in the $\Phi_e^w(t_{i_j})$, $\Phi_e^f(t_{i_j})$-plane, as could be seen in Fig. 4 (left).

For evaluating the integral in Eq. (13), a Monte Carlo integration method is used. Generate s random samples, $\vec{\boldsymbol{\varphi}}_{\boldsymbol{i}}$ for $i \in \{1, 2, \ldots, s\}$, from the excess phase distribution, $\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}$ (multivariate normal). The ratio between the number of samples that fall inside the integration region to the total number of generated samples is used as an approximation to the integral in Eq. (13):

$$f_{\vec{\boldsymbol{B}}}(\vec{\boldsymbol{b}}) \approx \frac{1}{s}\left|\left\{\vec{\boldsymbol{\varphi}}_{\boldsymbol{i}} \mid f_{\vec{\boldsymbol{B}}|\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}}(\vec{\boldsymbol{b}} \mid \vec{\boldsymbol{\varphi}}_{\boldsymbol{i}}) = 1, i \in \{1, 2, \ldots, s\}\right\}\right|. \tag{14}$$

For all the results presented in this work, the `multivariate_normal()` method from the `numpy.random` (version `1.14.3`) [SNR] Python library was used to generate one million (s) samples.

3.4 Entropy study

Using Eq. (13) or Eq. (14) as an approximation, to obtain the unconditional PMF for the bit vector $\vec{\boldsymbol{B}}^{\boldsymbol{u}}$, the unconditional Shannon entropy for the unobserved bits can be determined as usual:

$$\mathbf{H}[\vec{\boldsymbol{B}}^{\boldsymbol{u}}] = -\sum_{\vec{\boldsymbol{b}} \in \{0,1\}^{n^u}} f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}}}(\vec{\boldsymbol{b}}) \log_2\left(f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}}}(\vec{\boldsymbol{b}})\right). \tag{15}$$

3.4.1 Entropy conditioned on the oscillator phase

Given the realization of n^o phase values, at sampling time instances $(t_{i_0^o}, t_{i_1^o}, \ldots, t_{i_{n^o-1}^o})^\intercal \in \mathbb{R}_{\geq 0}^{n^o}$: $\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o}} = \vec{\boldsymbol{\varphi}}$ and combining Eqs. (9) and (10), the unobserved bits PMF, given the observed oscillator phase at n^o time instances equals

$$f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}}|\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o}}}(\vec{\boldsymbol{b}} \mid \vec{\boldsymbol{\varphi}}^{\boldsymbol{o}}) = \int_{\mathbb{R}^{2n^u}} f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}}|\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}}}(\vec{\boldsymbol{b}} \mid \vec{\boldsymbol{\varphi}}^{\boldsymbol{u}}) f_{\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}}|\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o}}}(\vec{\boldsymbol{\varphi}}^{\boldsymbol{u}} \mid \vec{\boldsymbol{\varphi}}^{\boldsymbol{o}}) \mathrm{d}\vec{\boldsymbol{\varphi}}^{\boldsymbol{u}}.$$

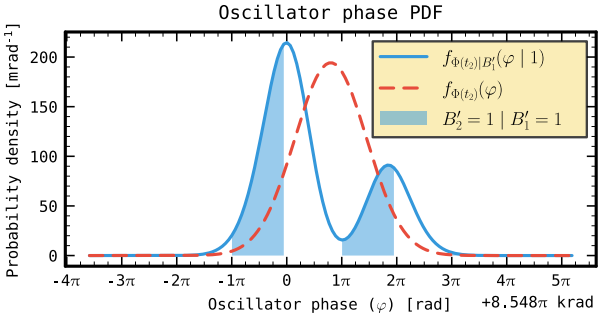

Figure 6: Oscillator total phase PDF and integration area for the second sample being equal to one, conditioned on the first sample being equal to one (solid blue) and unconditioned (dashed red).

From this conditional PMF, the conditional Shannon entropy, given the observed phase values, $\mathbf{H}[\vec{\boldsymbol{B}}^{\boldsymbol{u}} \mid \vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o}} = \vec{\boldsymbol{\varphi}}^{\boldsymbol{o}}]$, can be determined similarly to Eq. (15).

3.4.2 Entropy conditioned on the output bits

Given again n^o bit observations $\vec{\boldsymbol{b}}^{\boldsymbol{o}} = (b_0^o, b_1^o, \ldots, b_{n^o-1}^o)^\intercal \in \{0,1\}^{n^o}$, at sampling time instances $(t_{i_0^o}, t_{i_1^o}, \ldots, t_{i_{n^o-1}^o})^\intercal \in \mathbb{R}_{\geq 0}^{n^o}$, we therefore have a realization: $\vec{\boldsymbol{B}}^{\boldsymbol{o}} = \vec{\boldsymbol{b}}^{\boldsymbol{o}}$. Combining Eqs. (9) and (12), the conditional unobserved bits PMF, given the realization of $\vec{\boldsymbol{B}}^{\boldsymbol{o}}$ can be determined as follows:

$$f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}}|\vec{\boldsymbol{B}}^{\boldsymbol{o}}}(\vec{\boldsymbol{b}}^{\boldsymbol{u}} \mid \vec{\boldsymbol{b}}^{\boldsymbol{o}}) = \int_{\mathbb{R}^{2n}} f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}}|\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}}(\vec{\boldsymbol{b}}^{\boldsymbol{u}} \mid \vec{\boldsymbol{\varphi}}) f_{\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}|\vec{\boldsymbol{B}}^{\boldsymbol{o}}}(\vec{\boldsymbol{\varphi}} \mid \vec{\boldsymbol{b}}^{\boldsymbol{o}}) \mathrm{d}\vec{\boldsymbol{\varphi}}.$$

Using the Monte Carlo integration method from Section 3.3, this conditional PMF can be estimated by the following ratio:

$$f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}}|\vec{\boldsymbol{B}}^{\boldsymbol{o}}}(\vec{\boldsymbol{b}}^{\boldsymbol{u}} \mid \vec{\boldsymbol{b}}^{\boldsymbol{o}}) = \frac{f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}},\vec{\boldsymbol{B}}^{\boldsymbol{o}}}(\vec{\boldsymbol{b}}^{\boldsymbol{u}}, \vec{\boldsymbol{b}}^{\boldsymbol{o}})}{f_{\vec{\boldsymbol{B}}^{\boldsymbol{o}}}(\vec{\boldsymbol{b}}^{\boldsymbol{o}})},$$

with

$$f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}},\vec{\boldsymbol{B}}^{\boldsymbol{o}}}(\vec{\boldsymbol{b}}^{\boldsymbol{u}}, \vec{\boldsymbol{b}}^{\boldsymbol{o}}) \approx \frac{1}{s} \left| \left\{ \vec{\boldsymbol{\varphi}}_{\boldsymbol{i}} \mid f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}}|\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}}(\vec{\boldsymbol{b}}^{\boldsymbol{u}} \mid \vec{\boldsymbol{\varphi}}_{\boldsymbol{i}}) = 1, f_{\vec{\boldsymbol{B}}^{\boldsymbol{o}}|\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}}(\vec{\boldsymbol{b}}^{\boldsymbol{o}} \mid \vec{\boldsymbol{\varphi}}_{\boldsymbol{i}}) = 1, i \in \{1, 2, \ldots, s\} \right\} \right|,$$

and with $f_{\vec{\boldsymbol{B}}^{\boldsymbol{o}}}(\vec{\boldsymbol{b}}^{\boldsymbol{o}})$ from Eq. (14). Use s randomly generated samples from the unconditional multivariate normal $\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}$ distribution, $\vec{\boldsymbol{\varphi}}_{\boldsymbol{i}}$ for $i \in \{1, 2, \ldots, s\}$. From the conditional PMF, $f_{\vec{\boldsymbol{B}}^{\boldsymbol{u}}|\vec{\boldsymbol{B}}^{\boldsymbol{o}}}$, the conditional Shannon entropy for the unobserved bits, given the observed bits, $\mathbf{H}[\vec{\boldsymbol{B}}^{\boldsymbol{u}} \mid \vec{\boldsymbol{B}}^{\boldsymbol{o}} = \vec{\boldsymbol{b}}^{\boldsymbol{o}}]$, can be determined similar to Eq. (15).

Example 3. Given the scenario from Example 2, the total oscillator phase PDF is shown in Fig. 6. Both the unconditional PDF as the conditional PDF, given the first sample equaled one, are shown. The conditional probability of obtaining a one at the second sample equals 52.8 % and the conditional Shannon entropy is $\mathbf{H}[B_2' \mid B_1' = 1] = 0.998\,\mathrm{bit}$.

3.4.3 Worst-case entropy

Despite the bimodal shape of the conditioned total oscillator phase PDF from Fig. 6 at Example 3, the obtained Shannon entropy for the second sample, given the first sample equals one is significantly larger than the entropy for the first sample, $\mathbf{H}[B_1'] = 0.747\,\mathrm{bit}$

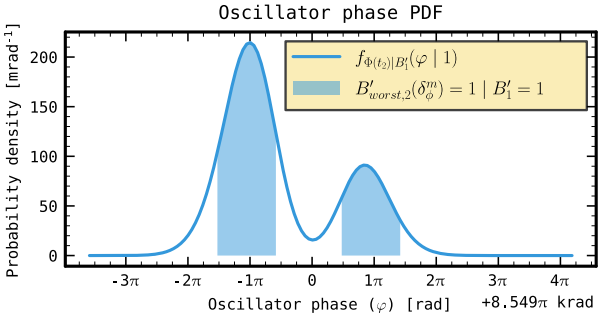

Figure 7: Oscillator total phase PDF and worst-case entropy integration area for the second sample being equal to one, given the first sample was one.

vs. $\mathbf{H}[B_2' \mid B_1' = 1] = 0.998\,\text{bit}$. The area under the PDF curve used for determining the bit probability is highly influenced by the horizontal position of the curve. This horizontal position is determined by the nominal phase at the sampling time: $2\pi f_n i t_{acc} + \phi_0 \bmod 2\pi$ for the i-th sample, related to ϕ_i in Eq. (8). To eliminate the influence of the nominal phase on the entropy, a worst-case entropy function is defined.

Definition 1. (Worst-case entropy) Given an oscillator total phase random variable at some time instance $\Phi : \Omega \to \mathbb{R}$ and its corresponding random bit, $B = \left\lfloor \frac{\Phi}{\pi} \right\rfloor \bmod 2$. A worst-case random bit function, $B_{worst}(\delta_\phi) : \Omega \times [0, 2\pi) \to \{0, 1\}$, is defined as

$$B_{worst}(\delta_\phi) = \left\lfloor \frac{\Phi + \delta_\phi}{\pi} \right\rfloor \bmod 2,$$

with δ_ϕ, a deterministic phase offset. The worst-case Shannon entropy for B, conditioned on an event in the event space, $E \in \mathcal{F}$, is equal to

$$\mathbf{H_{worst}}[B \mid E] = \mathbf{H}\big[B_{worst}(\delta_\phi^m) \mid E\big],$$

with: $\delta_\phi^m = \arg\max_{\delta_\phi \in [0,2\pi)} \mathbf{P}\big[B_{worst}(\delta_\phi) = 1\big]$.

As the worst-case entropy is independent from a phase shift, one can assign a worst-case entropy value to a phase random variable. Both the total oscillator phase as only the excess oscillator phase have an equal worst-case entropy content, as they only differ in a deterministic phase offset (nominal phase).

Definition 2. (Worst-case entropy for a phase random variable) Given an oscillator's total phase and excess phase random variables at some time instance $\Phi : \Omega \to \mathbb{R}$ and $\Phi_e : \Omega \to \mathbb{R}$. The worst-case Shannon entropy for Φ and Φ_e, conditioned on an event, $E \in \mathcal{F}$, is equal to

$$\mathbf{H_{worst}}[\Phi \mid E] = \mathbf{H_{worst}}[\Phi_e \mid E] = \mathbf{H_{worst}}[B \mid E],$$

with $B : \Omega \to \{0, 1\}$, a random bit extracted from that oscillator, $B = \left\lfloor \frac{\Phi + \delta_\phi}{\pi} \right\rfloor \bmod 2$, for any phase shift $\delta_\phi \in \mathbb{R}$.

Example 4. Given the scenario from Example 2, Fig. 7 provides the conditional PDF for the total oscillator phase at the second sampling moment, given the first sample obtained a one. The shaded area indicates the worst-case probability of obtaining a one, equaling 76.6 %. The worst-case Shannon entropy now becomes $\mathbf{H_{worst}}[B_2' \mid B_1' = 1] = 0.785\,\text{bit}$, which is significantly reduced compared to the regular Shannon entropy in Example 3, and independent from the nominal oscillator phase.

Table 1: Numerical values used for: h_w, h_f and t_{acc} and obtained white FM noise $\mathbf{H_{worst}}$.

Estimate	h_w	h_f	Corner	Sampling speed 25 %		100 %		400 %	
				t_{acc}	$\mathbf{H_{worst}}$	t_{acc}	$\mathbf{H_{worst}}$	t_{acc}	$\mathbf{H_{worst}}$
	[fs]	[1×10^{-12}]	[µs]	[µs]	[bit]	[µs]	[bit]	[µs]	[bit]
Low	-	6.22	100	25.0	0.992	100	> 0.999	400	> 0.999
Mid	18.9	104	5.00	1.25	0.0186	5.00	0.523	20.0	0.979
High	-	9480	0.0434	0.0109	< 0.001	0.0434	< 0.001	0.174	< 0.001

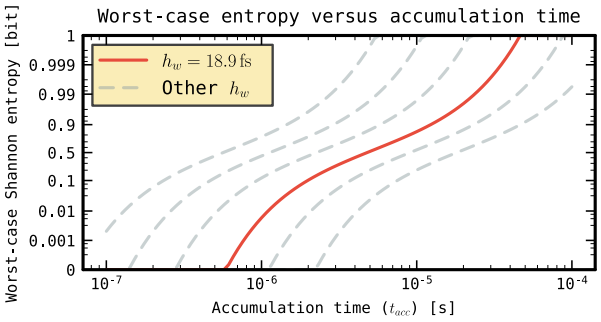

Figure 8: Worst-case Shannon entropy for a white FM noise source, versus accumulation time. Entropy curves for noise magnitudes $h_w = 18.9$ fs (solid red) and other magnitudes (dashed gray) for reference: $\{5\,\text{fs}, 10\,\text{fs}, 40\,\text{fs}, 80\,\text{fs}, 160\,\text{fs}\}$ are shown.

4 Model simulation

The Gaussian process model, developed in Section 3, will now be used to generate entropy estimates for an ERO TRNG affected by both white FM and flicker FM noise under realistic operating conditions.

4.1 Noise magnitude

This section explains how the scaling constants h_w and h_f, used to approximate both the white FM and flicker FM noise magnitude, are selected. Based on the obtained noise corner, the ERO TRNG sampling speed is determined as well. Table 1 lists the numerical values used in the remainder of this work.

4.1.1 White FM noise magnitude

Throughout this work, a single magnitude for the white FM noise is used: $h_w = 18.9$ fs. As provided by Table 9 in [PV24], this value is lower than most other estimates in the field and can therefore be considered as a conservative estimate.

There exists a one-to-one relation between the white FM noise component worst-case Shannon entropy and the accumulation time. Figure 8 depicts this relation for the selected white FM noise magnitude (solid red) and other magnitudes (dashed gray). Higher noise magnitudes give a higher entropy value at a given accumulation time.

4.1.2 Flicker FM noise magnitude

This work considers three different magnitudes for the flicker FM noise component. At the higher end of the spectrum, there is the magnitude as measured by [PV24], $h_f = 9.48 \times 10^{-9}$. At the lower end, the noise corner derived from the measurements in [HTBF14]

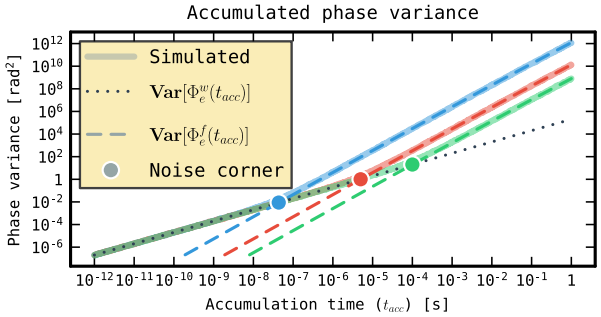

Figure 9: Oscillator phase variance versus accumulation time for the three different flicker FM noise magnitudes: $\{6.22 \times 10^{-12}, 1.04 \times 10^{-10}, 9.48 \times 10^{-9}\}$, represented by distinct colors, used in this work.

lead to a magnitude $h_f = 6.22 \times 10^{-12}$, when using the white noise strength estimate from Section 4.1.1. The third value is selected in between, $h_f = 1.04 \times 10^{-10}$, and approaches what has been reported by [LB15] and [FL14]. A frequency value $f_l = 1\,\text{mHz}$ is used for the lower frequency bound in Eq. (6).

4.1.3 White FM - flicker FM noise corner

The presence of both white FM and flicker FM noise gives rise to a noise corner. The noise corner represents a pair $\big(t_{cor}, \mathbf{Var}\big[\Phi(t_{cor})\big]\big)$, for which the accumulated oscillator white FM phase variance equals the accumulated flicker FM phase variance. From Eqs. (4) and (6), the corner accumulation time satisfies the relation: $t_{cor}\big(3 - 2\gamma - 2\ln(2\pi f_l t_{cor})\big) = \frac{h_w}{h_f}$.

Figure 9 depicts the accumulated oscillator phase variance versus accumulation time, for the three flicker FM noise magnitudes considered in this work. For accumulation times below the noise corner, the white FM noise component is dominant and the oscillator phase variance increases linearly. Above the noise corner, the flicker FM noise component dominates and the variance increases in a quadratic way. The dotted and dashed lines in Fig. 9 represent the theoretical phase variance from Eqs. (4) and (6) respectively. The simulation results, when using the Gaussian process model are shown as solid opaque curves. The noise corner accumulation time values obtained for the flicker FM magnitudes are provided in Table 1.

4.1.4 Sampling speed

Depending on the flicker FM magnitude, the accumulation time is selected as 25 %, 100 % and 400 % of the value of the noise corner. At 25 %, the white FM noise component will dominate and at 400 %, the flicker FM noise component dominates. Table 1 provides the accumulation times at interest and the corresponding worst-case Shannon entropy for the white FM component, also visible in Fig. 8, when the ERO TRNG is sampled at t_{acc} time intervals.

4.2 Entropy estimation

This section presents numerical results for the conditional ERO TRNG worst-case entropy, from implementing the theory outlined in Sections 3.2 and 3.4. The subsections are arranged in decreasing knowledge of the oscillator's state: in Section 4.2.1, we assume the observation of the complete oscillator phase, whereas in Section 4.2.2 only the produced TRNG output bits are assumed to be known.

4.2.1 Knowledge of the previous phase values

The worst-case Shannon entropy is evaluated when p previous sample phase values are known, $\mathbf{H_{worst}}[\vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}} \mid \vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o},\boldsymbol{p}} = \vec{\boldsymbol{\varphi}}]$. The entropy for the sixth bit from an ERO TRNG is calculated, given the knowledge of the previous p sample phases, for p ranging from zero up to five. The unobserved/observed excess phase vectors become

$$\vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}} = \left(\Phi_e^w(t_6), \Phi_e^f(t_6)\right)^\intercal,$$

$$\vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o},\boldsymbol{p}} = \left(\Phi_e^w(t_5), \Phi_e^w(t_4), \ldots, \Phi_e^w(t_{6-p}), \Phi_e^f(t_5), \Phi_e^f(t_4), \ldots, \Phi_e^f(t_{6-p})\right)^\intercal,$$

for $p \in \{0, 1, \ldots, 5\}$. When p equals zero, no phase information is known and the entropy becomes unconditioned.

Note that from Eq. (11), the conditional covariance matrix for $\vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}} \mid \vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o},\boldsymbol{p}} = \vec{\boldsymbol{\varphi}}$ is independent from the actual realized value of the previous sample phases, $\vec{\boldsymbol{\varphi}}$, and that the worst-case entropy is independent from a phase shift introduced by the conditional mean. The worst-case entropy, given the knowledge of p previously observed sample phases is therefore not influenced by the realized value itself: $\mathbf{H_{worst}}[\vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}} \mid \vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o},\boldsymbol{p}} = \vec{\boldsymbol{\varphi}}] = \mathbf{H_{worst}}[\vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}} \mid \vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o},\boldsymbol{p}}]$.

Knowledge of p previous sample phases Figures 10 to 12 provide the worst-case Shannon entropy and phase standard deviation for the sixth bit, given knowledge of p previous phase values, for flicker FM noise magnitudes 6.22×10^{-12}, 1.04×10^{-10} and 9.48×10^{-9} respectively. The entropy and phase standard deviation values are given for white FM and flicker FM noise separately, $\mathbf{H_{worst}}\left[\Phi_e^y(t_6) \mid \vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o},\boldsymbol{p}}\right]$ and $\sqrt{\mathbf{Var}\left[\Phi_e^y(t_6) \mid \vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o},\boldsymbol{p}}\right]}$, for $y \in \{w, f\}$, respectively.

As seen from these figures, the entropy reduces significantly when the previous sample phase is known, both for white FM and flicker FM noise. For white FM noise, the entropy remains constant for $p \geq 1$ and the phase variance equals the variance accumulated between the fifth and sixth sample: $\mathbf{Var}[\Phi_e^w(t_6) \mid \vec{\mathbf{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{o},\boldsymbol{p}}] = 4\boldsymbol{\pi}^2 f_n^2 h_w t_{acc}$. For flicker FM noise, the phase variance and therefore also the worst-case entropy keep reducing for increasing p, although the reduction is minor compared to the reduction for p from zero to one and reduces for higher p.

Knowledge of only the previous sample phase Elaborating on knowing only the phase of the previous sample ($p = 1$), Fig. 13 depicts the worst-case Shannon entropy for white FM and flicker FM noise separately, versus the accumulation time between the samples. Curves are given for three time instances: t_6, t_{1000} and $t_{1000000}$, the sixth (given in previous paragraph), thousandth and millionth bit, respectively. The worst-case white FM noise Shannon entropy shown in Fig. 13 is identical to the solid red curve from Fig. 8. As seen from this figure, given the knowledge of the previous sample's phase, the worst-case entropy for the flicker FM noise component is significantly higher, comparable or significantly lower than the worst-case entropy for the white FM noise component, when using the high, mid or low flicker FM noise magnitude estimate respectively from Table 1. Additionally, increasing from the sixth to the millionth sampled bit, reduces the knowledge gained over the current sample, when observing the previous sample's phase value.

4.2.2 Knowledge of the previous bit values

In this section, only the value for the previous p sampled bits instead of the full oscillator phase is assumed to be known. The entropy for the 300-th bit from an ERO TRNG is calculated, given the knowledge of the previous p sampled bits, for p ranging from zero up

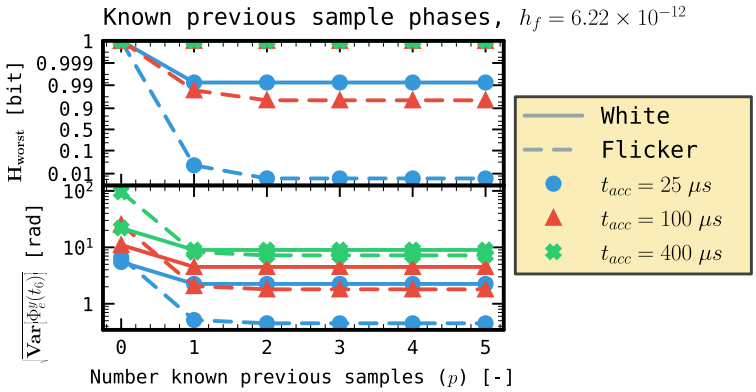

Figure 10: Worst-case Shannon entropy (top) and oscillator phase standard deviation (bottom), given the knowledge of p previous sample phase values, for $p \in \{0, 1, \ldots, 5\}$ and a flicker FM noise magnitude $h_f = 6.22 \times 10^{-12}$. Results are provided both for white FM and flicker FM noise and for accumulation lengths: $t_{acc} \in \{25.0\,\mu s, 100\,\mu s, 400\,\mu s\}$.

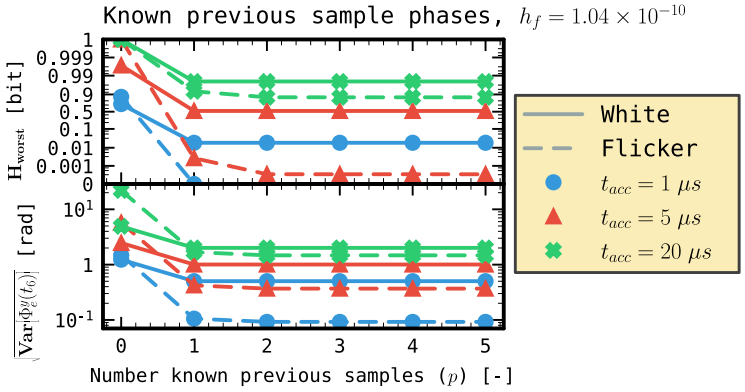

Figure 11: Worst-case Shannon entropy (top) and oscillator phase standard deviation (bottom), given the knowledge of p previous sample phase values, for $p \in \{0, 1, \ldots, 5\}$ and a flicker FM noise magnitude $h_f = 1.04 \times 10^{-10}$. Results are provided both for white FM and flicker FM noise and for accumulation lengths: $t_{acc} \in \{1.25\,\mu s, 5.00\,\mu s, 20.0\,\mu s\}$.

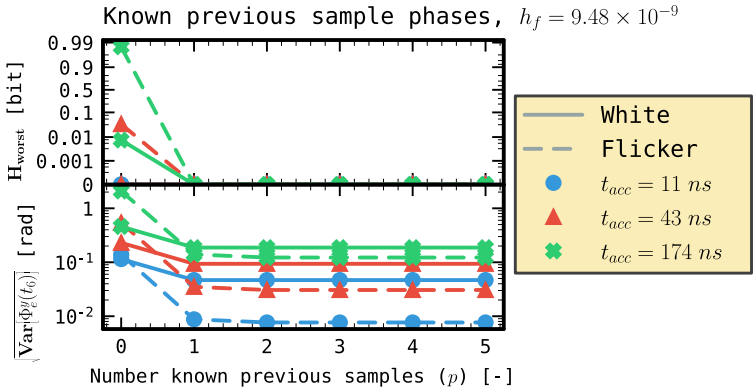

Figure 12: Worst-case Shannon entropy (top) and oscillator phase standard deviation (bottom), given the knowledge of p previous sample phase values, for $p \in \{0, 1, \ldots, 5\}$ and a flicker FM noise magnitude $h_f = 9.48 \times 10^{-9}$. Results are provided both for white FM and flicker FM noise and for accumulation lengths: $t_{acc} \in \{10.9\,ns, 43.4\,ns, 174\,ns\}$.

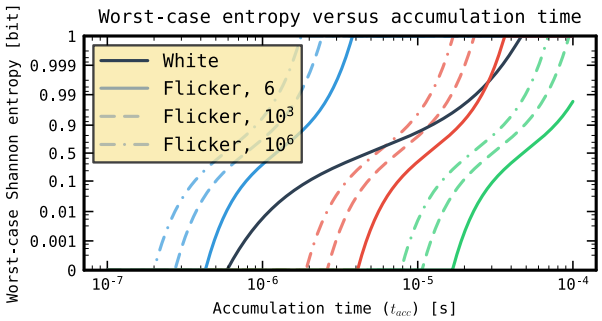

Figure 13: Worst-case flicker FM noise Shannon entropy, given the knowledge of the previous sample's phase value versus the accumulation time (t_{acc}) between two samples. Curves are plotted for three different flicker FM noise magnitudes: $h_f \in \{6.22 \times 10^{-12}, 1.04 \times 10^{-10}, 9.48 \times 10^{-9}\}$, a higher noise magnitude gives a higher entropy value. For each flicker FM noise magnitude, three curves corresponding to the sixth, thousandth and millionth bit are shown. The white FM noise entropy, from Fig. 8, is shown for reference.

to ten. When p equals zero, the entropy for the unconditioned distribution is given. The unobserved excess phase and observed bit vector become

$$\vec{\boldsymbol{\Phi}}_{\boldsymbol{e}}^{\boldsymbol{u}} = \left(\Phi_e^w(t_{300}), \Phi_e^f(t_{300})\right)^\intercal,$$

$$\boldsymbol{B}^{\vec{\boldsymbol{o}},\boldsymbol{p}} = (B_{299}, B_{298}, \ldots, B_{300-p})^\intercal.$$

Figures 14 to 16 show the worst-case Shannon entropy for the 300-th bit, given the knowledge of p previous sample bits, for flicker FM noise magnitudes 6.22×10^{-12}, 1.04×10^{-10} and 9.48×10^{-9} respectively. The entropy values are given for white FM and flicker FM noise separately, $\mathbf{H_{worst}}\left[\Phi_e^y(t_{300}) \mid \boldsymbol{B}^{\vec{\boldsymbol{o}},\boldsymbol{p}}\right]$, for $y \in \{w, f\}$, respectively.

Figures 14 and 15 show high worst-case Shannon entropy values for both white FM and flicker FM noise. The entropy reduces slightly with increasing number of known bits, as each bit reveals some amount of information on the current oscillator phase. For the higher sampling speeds in Fig. 16, the flicker FM worst-case Shannon entropy drastically reduces even when a single bit is known.

4.3 Summary of simulation findings

The simulation results reveal several key insights regarding the worst-case entropy generated by white FM and flicker FM noise in an ERO TRNG. Firstly, the worst-case entropy decreases more significantly when previous phase values are known for flicker FM noise compared to white FM noise. This could mainly be attributed to the dependencies in consecutive period lengths caused by flicker FM noise. Additionally, at higher sampling speeds, flicker FM noise may appear to contain more entropy than it actually does. This overestimation of entropy is only detected when previous samples are also considered. Lastly, the h_f value plays a crucial role in determining the ratio of entropy contribution between white and flicker noise, making it impossible to assert that flicker noise always provides a meaningful contribution to the output entropy.

5 Conclusion and further research directions

This work presented a method for modeling the excess phase process of a free-running oscillator. The time-domain model is based on the theory of Gaussian processes and is

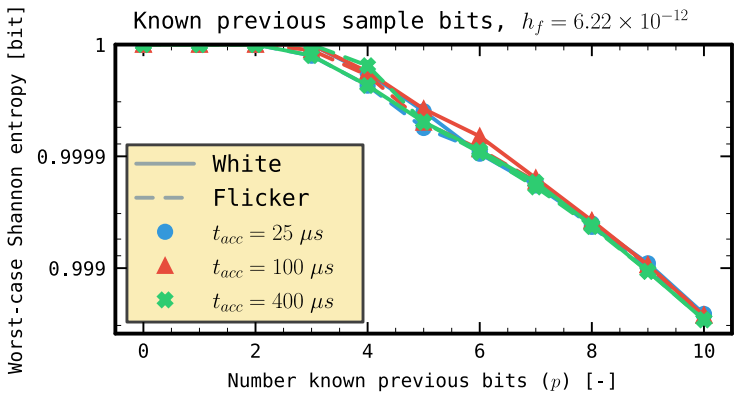

Figure 14: Worst-case Shannon entropy, given the knowledge of p previous sample bit values, for $p \in \{0, 1, \ldots, 10\}$ and a flicker FM noise magnitude: $h_f = 6.22 \times 10^{-12}$. Results are provided both for white FM and flicker FM noise and for accumulation lengths: $t_{acc} \in \{25.0\,\mu s, 100\,\mu s, 400\,\mu s\}$.

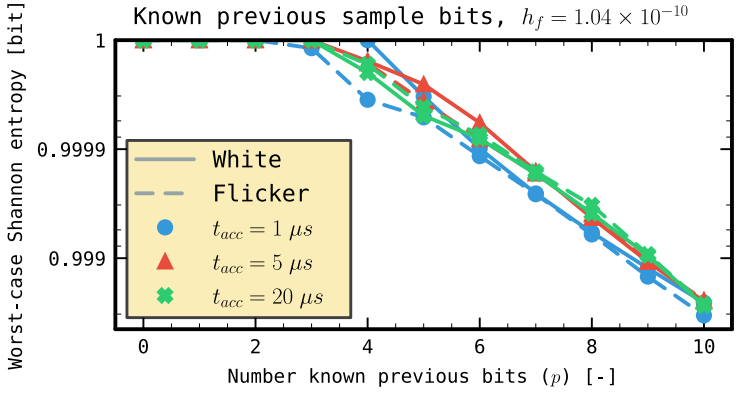

Figure 15: Worst-case Shannon entropy, given the knowledge of p previous sample bit values, for $p \in \{0, 1, \ldots, 10\}$ and a flicker FM noise magnitude: $h_f = 1.04 \times 10^{-10}$. Results are provided both for white FM and flicker FM noise and for accumulation lengths: $t_{acc} \in \{1.25\,\mu s, 5.00\,\mu s, 20.0\,\mu s\}$.

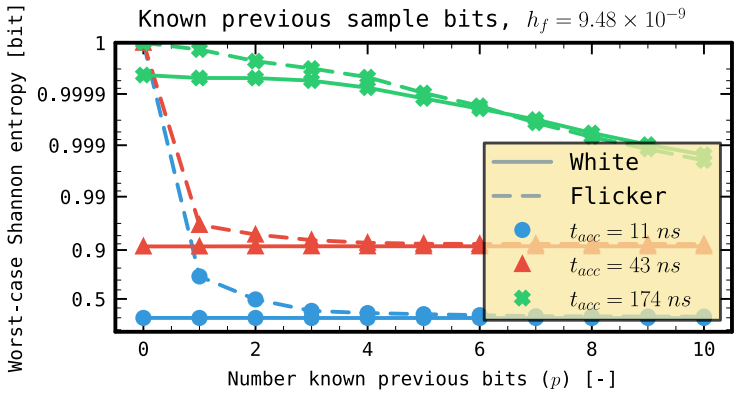

Figure 16: Worst-case Shannon entropy, given the knowledge of p previous sample bit values, for $p \in \{0, 1, \ldots, 10\}$ and a flicker FM noise magnitude: $h_f = 9.48 \times 10^{-9}$. Results are provided both for white FM and flicker FM noise and for accumulation lengths: $t_{acc} \in \{10.9\,ns, 43.4\,ns, 174\,ns\}$.

specifically tailored for use of estimating the entropy produced by a TRNG. The focus of this work was on the most prevalent noise types: white FM and flicker FM noise, but the proposed model could be applied to noise sources with other spectral shapes (e.g. random walk FM noise as described by [HAB81]) as well. For the two noise classes, the ACF was analytically derived from the shape of the oscillator's relative frequency deviation spectrum.

Using Bayes' theorem, the conditional ERO TRNG output bit distribution is analytically derived from the Gaussian process excess phase model. These distributions allow observing the change in phase PDF shape, when further knowledge on the TRNG state becomes available. Additionally, the entropy produced by the TRNG is derived from the obtained phase PDFs and the worst-case entropy concept was introduced to remove the deterministic influence of the phase offset on the derived entropy figure.

Finally, this work presents some exploratory simulation results for the proposed entropy model, using three different magnitudes for the flicker FM noise component, encountered in the literature. The results show that flicker FM noise can indeed in some cases be a valid source of TRNG entropy. However, due to the inherent long-lasting dependency, this noise should be harvested with great care. Given a low flicker FM noise magnitude, the authors conclude from Fig. 13 that the flicker FM noise only bears minimal entropy compared to white FM noise at practical sampling speeds. Especially as there is a wide range of flicker FM noise estimates available in literature, ranging from $h_f = 9.48 \times 10^{-9}$ in [PV24] down to $h_f = 6.22 \times 10^{-12}$ in [HTBF14], more experimental evidence on potentially a wider range of platforms should become available before flicker FM noise could be widely accepted as a reliable source of TRNG entropy.

Besides from working on a more profound experimental validation of the flicker FM noise magnitude, the authors believe further research should be focused on applying the Gaussian process model on a more extended set of TRNG architectures, e.g. situations where multiple oscillators are used. Additionally, studying the stopping time, when a specified phase level is reached by the random process, is necessary to determine the distribution for the oscillator's period length, which in turn enables to augment existing TRNG stochastic models with the existence of flicker FM noise.

Acknowledgments

This work was supported by CyberSecurity Research Flanders with reference number VR20192203 and by the European Commission through the Horizon 2020 research and innovation program Belfort ERC Advanced Grant 101020005 695305.

References

[BCF+24] Licinius Benea, Mikael Carmona, Viktor Fischer, Florian Pebay-Peyroula, and Romain Wacquez. Impact of the Flicker Noise on the Ring Oscillator-based TRNGs. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2024(2):870–889, March 2024.

[BLMT11] Mathieu Baudet, David Lubicz, Julien Micolod, and André Tassiaux. On the Security of Oscillator-Based Random Number Generators. *Journal of Cryptology*, 24(2):398–425, April 2011.

[FL14] Viktor Fischer and David Lubicz. Embedded Evaluation of Randomness in Oscillator Based Elementary TRNG. In *Cryptographic Hardware and Embedded Systems — CHES 2014*, pages 527–543, 2014.

[GRN+91] G. Ghibaudo, O. Roux, Ch. Nguyen-Duc, F. Balestra, and J. Brini. Improved Analysis of Low Frequency Noise in Field-Effect MOS Transistors. *Physica Status Solidi (a)*, 124(2):571–581, April 1991.

[HAB81] D.A. Howe, D.U. Allan, and J.A. Barnes. Properties of Signal Sources and Measurement Methods. In *Annual Frequency Control Symposium*, pages 669–716, 1981.

[HFBN15] Patrick Haddad, Viktor Fischer, Florent Bernard, and Jean Nicolai. A Physical Approach for Stochastic Modeling of TERO-Based TRNG. In *Cryptographic Hardware and Embedded Systems – CHES 2015*, pages 357–372, 2015.

[HTBF14] Patrick Haddad, Yannick Teglia, Florent Bernard, and Viktor Fischer. On the assumption of mutual independence of jitter realizations in P-TRNG stochastic models. In *Design, Automation & Test in Europe Conference & Exhibition – DATE 2014*, pages 1–6, 2014.

[ISO19] ISO/IEC JTC 1/SC 27. Test and analysis methods for random bit generators within ISO/IEC 19790 and ISO/IEC 15408, October 2019.

[KG04] Paul Kohlbrenner and Kris Gaj. An embedded true random number generator for FPGAs. In *International Symposium on Field Programmable Gate Arrays – FPGA 2004*, pages 71–78, February 2004.

[LB15] David Lubicz and Nathalie Bochard. Towards an Oscillator Based TRNG with a Certified Entropy Rate. *IEEE Transactions on Computers*, 64(4):1191–1200, April 2015.

[LF24] David Lubicz and Viktor Fischer. Entropy Computation for Oscillator-based Physical Random Number Generators. *Journal of Cryptology*, 37(2):13, April 2024.

[PS22] Matthias Peter and Werner Schindler. A proposal for: Functionality classes for random number generators, September 2022.

[PV22] Adriaan Peetermans and Ingrid Verbauwhede. An energy and area efficient, all digital entropy source compatible with modern standards based on jitter pipelining. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2022(4):88–109, August 2022.

[PV24] Adriaan Peetermans and Ingrid Verbauwhede. Characterization of Oscillator Phase Noise Arising From Multiple Sources for ASIC True Random Number Generation. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 71(3):1144–1157, January 2024.

[SNR] Random sampling (numpy.random). https://numpy.org/doc/1.14/reference/routines.random.html. Accessed on July 11, 2024.

[TBK+18] Meltem Sönmez Turan, Elaine Barker, John Kelsey, Kerry A McKay, Mary L Baish, and Mike Boyle. Recommendation for the entropy sources used for random bit generation. Technical Report NIST SP 800-90b, National Institute of Standards and Technology, Gaithersburg, MD, January 2018.

[VD10] Michal Varchola and Milos Drutarovsky. New High Entropy Element for FPGA Based True Random Number Generators. In *Cryptographic Hardware and Embedded Systems – CHES 2010*, pages 351–365, 2010.

[YRG+18] Bohan Yang, Vladimir Rožic, Miloš Grujic, Nele Mentens, and Ingrid Verbauwhede. ES-TRNG: A High-throughput, Low-area True Random Number Generator based on Edge Sampling. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2018:267–292, August 2018.

A Combination of multiple noise sources

In terms of the relative phase acceleration, $\forall t \in \mathbb{R}_{\geq 0} : A(t) = \frac{\mathrm{d}}{\mathrm{d}t}Y(t)$, the relation from Eq. (3) becomes $S_A(f) = (2\pi f)^2 S_Y(f) = \sum_{\alpha=-2}^{2}(2\pi f)^2 S_{Y^\alpha}(f) = \sum_{\alpha=-2}^{2} S_{A^\alpha}(f)$. The relative phase acceleration is assumed stationary in [PV24], therefore $S_A(f) = \mathcal{F}\big\{R_A(\tau)\big\}$. Combine this by using the linearity of the Fourier Transform (FT):

$$S_A(f) = \sum_{\alpha=-2}^{2} S_{A^\alpha}(f) = \sum_{\alpha=-2}^{2} \mathcal{F}\big\{R_{A^\alpha}(\tau)\big\} = \mathcal{F}\Big\{\sum_{\alpha=-2}^{2} R_{A^\alpha}(\tau)\Big\} = \mathcal{F}\big\{R_A(\tau)\big\},$$

therefore, we have $R_A(\tau) = \sum_{\alpha=-2}^{2} R_{A^\alpha}(\tau)$, with $R_{A^\alpha}(\tau) = \mathcal{F}^{-1}\big\{(2\pi f)^2 S_{Y^\alpha}(f)\big\}$. The relative phase acceleration ACF is similarly composed of a sum of independent contributions.

Satisfying this relation, we assume the relative phase acceleration equals $\forall t \in \mathbb{R}_{\geq 0} : A(t) = \sum_{\alpha=-2}^{2} A^\alpha(t)$, with $\forall (t_i, t_j) \in \mathbb{R}^2_{\geq 0} : R_{A^\alpha}(t_i, t_j) = \mathbf{E}\big[A^\alpha(t_i)A^\alpha(t_j)\big]$. Indeed, the ACF now equals $R_A(t_i, t_j) = \mathbf{E}\big[A(t_i)A(t_j)\big] = \mathbf{E}\Big[\sum_{\alpha_i=-2}^{2} A^{\alpha_i}(t_i) \sum_{\alpha_j=-2}^{2} A^{\alpha_j}(t_j)\Big] = \sum_{\alpha_i=-2}^{2}\sum_{\alpha_j=-2}^{2} \mathbf{E}\big[A^{\alpha_i}(t_i)A^{\alpha_j}(t_j)\big] = \sum_{\alpha=-2}^{2} \mathbf{E}\big[A^\alpha(t_i)A^\alpha(t_j)\big] = \sum_{\alpha=-2}^{2} R_{A^\alpha}(t_i, t_j)$, using $\forall (t_i, t_j) \in \mathbb{R}^2_{\geq 0}, \alpha_i \neq \alpha_j : \mathbf{E}\big[A^{\alpha_i}(t_i)A^{\alpha_j}(t_j)\big] = 0$, due to the mutual independence of the noise contributions.

We now define the individual excess phase noise contributions: $\Phi_e^\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}$ by $\Phi_e^\alpha(t) = 2\pi f_n \int_0^t \int_0^\theta A^\alpha(\nu)\mathrm{d}\nu\mathrm{d}\theta$, or equivalently by $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Phi_e^\alpha(t) = 2\pi f_n A^\alpha(t)$. The total excess phase then becomes

$$\begin{aligned}\Phi_e(t) &= 2\pi f_n \int_0^t \int_0^\theta A(\nu)\mathrm{d}\nu\mathrm{d}\theta = 2\pi f_n \int_0^t \int_0^\theta \sum_{\alpha=-2}^{2} A^\alpha(\nu)\mathrm{d}\nu\mathrm{d}\theta \\ &= \sum_{\alpha=-2}^{2} 2\pi f_n \int_0^t \int_0^\theta A^\alpha(\nu)\mathrm{d}\nu\mathrm{d}\theta = \sum_{\alpha=-2}^{2} \Phi_e^\alpha(t),\end{aligned}$$

with the excess phase noise contribution PSD equal to $S_{\Phi_e^\alpha}(f) = \frac{(2\pi f_n)^2}{(2\pi f)^4} S_{A^\alpha}(f) = \left(\frac{f_n}{f}\right)^2 S_{Y^\alpha}(f) = \left(\frac{f_n}{f}\right)^2 h_\alpha |f|^\alpha$. The total oscillator phase therefore equals a sum of independent noise contributions, added to a deterministic part, determined by the nominal frequency, f_n: $\Phi(t) = 2\pi f_n t + \phi_0 + \sum_{\alpha=-2}^{2} \Phi_e^\alpha(t)$.

B Flicker FM noise ACF

Changing the order of integration in Eq. (5) and working out the integral obtains, $\forall (t_i, t_j) \in \mathbb{R}^2_{\geq 0}$:

$$\begin{aligned}R_{\Phi_e}(t_i, t_j) &= 8\pi^2 f_n^2 h_f \int_{f_l}^{f_h} \frac{1}{f} \int_0^{t_i} \int_0^{t_j} \cos\big(2\pi f(\theta_j - \theta_i)\big)\mathrm{d}\theta_j\mathrm{d}\theta_i\mathrm{d}f \\ &= 4\pi f_n^2 h_f \int_{f_l}^{f_h} \frac{1}{f^2} \int_0^{t_i} \Big(\sin\big(2\pi f(t_j - \theta_i)\big) + \sin(2\pi f\theta_i)\Big)\mathrm{d}\theta_i\mathrm{d}f \\ &= 2 f_n^2 h_f \int_{f_l}^{f_h} \frac{1}{f^3} \Big(\cos\big(2\pi f(t_j - t_i)\big) - \cos(2\pi f t_j) - \cos(2\pi f t_i) + 1\Big)\mathrm{d}f.\end{aligned}$$

When $t_i \neq t_j$, $t_i \neq 0$ and $t_j \neq 0$, the integral becomes

$$\begin{aligned}
R_{\Phi_e}(t_i,t_j) = 2f_n^2 h_f \Bigg(& -\frac{\cos\big(2\pi f_h(t_j-t_i)\big)}{2f_h^2} + \pi(t_j-t_i)\frac{\sin\big(2\pi f_h(t_j-t_i)\big)}{f_h} \\
& - 2\pi^2(t_j-t_i)^2\,\mathrm{Ci}\big(2\pi f_h|t_j-t_i|\big) + \frac{\cos\big(2\pi f_l(t_j-t_i)\big)}{2f_l^2} \\
& - \pi(t_j-t_i)\frac{\sin\big(2\pi f_l(t_j-t_i)\big)}{f_l} + 2\pi^2(t_j-t_i)^2\,\mathrm{Ci}\big(2\pi f_l|t_j-t_i|\big) \\
& + \frac{\cos(2\pi f_h t_j)}{2f_h^2} - \pi t_j\frac{\sin(2\pi f_h t_j)}{f_h} + 2\pi^2 t_j^2\,\mathrm{Ci}(2\pi f_h t_j) \\
& - \frac{\cos(2\pi f_l t_j)}{2f_l^2} + \pi t_j\frac{\sin(2\pi f_l t_j)}{f_l} - 2\pi^2 t_j^2\,\mathrm{Ci}(2\pi f_l t_j) \\
& + \frac{\cos(2\pi f_h t_i)}{2f_h^2} - \pi t_i\frac{\sin(2\pi f_h t_i)}{f_h} + 2\pi^2 t_i^2\,\mathrm{Ci}(2\pi f_h t_i) \\
& - \frac{\cos(2\pi f_l t_i)}{2f_l^2} + \pi t_i\frac{\sin(2\pi f_l t_i)}{f_l} - 2\pi^2 t_i^2\,\mathrm{Ci}(2\pi f_l t_i) + \frac{1}{2f_l^2} - \frac{1}{2f_h^2}\Bigg),
\end{aligned} \tag{18}$$

with: $\mathrm{Ci}(\cdot)$, the cosine integral, defined as $\forall x \in \mathbb{R}_{>0} : \mathrm{Ci}(x) = -\int_x^\infty \frac{\cos(\theta)}{\theta}\mathrm{d}\theta$. When either $t_i = t_j$, $t_i = 0$ or $t_j = 0$, the corresponding term simplifies to

$$-\frac{\cos(2\pi f_x t_y)}{2f_x^2} + \pi t_y\frac{\sin(2\pi f_x t_y)}{f_x} - 2\pi^2 t_y^2\,\mathrm{Ci}(2\pi f_x t_y) = -\frac{1}{2f_x^2},$$

for $x \in \{l,h\}$ and $t_y \in \{t_i, t_j, |t_j - t_i|\}$.

Similar as in [PV24], the upper frequency bound, f_h, is assumed very large: $f_h \to \infty$. Using the property of the cosine integral: $\lim_{x\to\infty}\mathrm{Ci}(x) = 0$, Eq. (18) is simplified:

$$\begin{aligned}
R_{\Phi_e}(t_i,t_j) = 2f_n^2 h_f \Bigg(& \frac{\cos\big(2\pi f_l(t_j-t_i)\big)}{2f_l^2} - \pi(t_j-t_i)\frac{\sin\big(2\pi f_l(t_j-t_i)\big)}{f_l} \\
& + 2\pi^2(t_j-t_i)^2\,\mathrm{Ci}\big(2\pi f_l|t_j-t_i|\big) \\
& - \frac{\cos(2\pi f_l t_j)}{2f_l^2} + \pi t_j\frac{\sin(2\pi f_l t_j)}{f_l} - 2\pi^2 t_j^2\,\mathrm{Ci}(2\pi f_l t_j) \\
& - \frac{\cos(2\pi f_l t_i)}{2f_l^2} + \pi t_i\frac{\sin(2\pi f_l t_i)}{f_l} - 2\pi^2 t_i^2\,\mathrm{Ci}(2\pi f_l t_i) + \frac{1}{2f_l^2}\Bigg).
\end{aligned}$$

Reordering further to

$$\begin{aligned}
R_{\Phi_e}(t_i,t_j) = {} & 4\pi^2(t_j-t_i)^2 f_n^2 h_f \Bigg(-\frac{1}{2}\frac{\sin^2\big(\pi f_l(t_j-t_i)\big)}{\pi^2 f_l^2(t_j-t_i)^2} - \frac{\sin\big(2\pi f_l(t_j-t_i)\big)}{2\pi f_l(t_j-t_i)} \\
& + \mathrm{Ci}\big(2\pi f_l|t_j-t_i|\big) + \frac{1}{4\pi^2 f_l^2(t_j-t_i)^2}\Bigg) + \frac{1}{f_l^2} f_n^2 h_f \\
& + 4\pi^2 t_j^2 f_n^2 h_f \left(\frac{1}{2}\frac{\sin^2(\pi f_l t_j)}{\pi^2 f_l^2 t_j^2} + \frac{\sin(2\pi f_l t_j)}{2\pi f_l t_j} - \mathrm{Ci}(2\pi f_l t_j) - \frac{1}{4\pi^2 f_l^2 t_j^2}\right) \\
& + 4\pi^2 t_i^2 f_n^2 h_f \left(\frac{1}{2}\frac{\sin^2(\pi f_l t_i)}{\pi^2 f_l^2 t_i^2} + \frac{\sin(2\pi f_l t_i)}{2\pi f_l t_i} - \mathrm{Ci}(2\pi f_l t_i) - \frac{1}{4\pi^2 f_l^2 t_i^2}\right).
\end{aligned} \tag{19}$$

As in [PV24], it is assumed that the lower frequency limit, f_l, is much smaller than the inverse of the observed time, $t_y \in \{t_i, t_j, |t_j - t_i|\} : \forall t_y \in \mathbb{R}_{>0} : f_l \ll \frac{1}{t_y}$. Therefore $\pi f_l t_y \ll 1$. Using the property $\lim_{x\to 0}\frac{\sin(x)}{x} = 1$, and using the Taylor series for $\mathrm{Ci}(x)$

around $x = 0$: $\mathrm{Ci}(x) \approx \gamma + \ln(x) + \sum_{k=1}^{\infty} \frac{(-x^2)^k}{2k(2k)!}$, and γ representing the Euler-Mascheroni constant, $\gamma \approx 0.577$, Eq. (19) is further simplified for $t_i > 0$, $t_j > 0$ and $t_i \neq t_j$:

$$\begin{aligned} R_{\Phi_e}(t_i, t_j) = {} & 4\pi^2 (t_j - t_i)^2 f_n^2 h_f \left(-\frac{3}{2} + \gamma + \ln\big(2\pi f_l |t_j - t_i|\big) \right) \\ & + 4\pi^2 t_j^2 f_n^2 h_f \left(\frac{3}{2} - \gamma - \ln(2\pi f_l t_j) \right) + 4\pi^2 t_i^2 f_n^2 h_f \left(\frac{3}{2} - \gamma - \ln(2\pi f_l t_i) \right). \end{aligned} \tag{20}$$

When $t_y = 0$ for $t_y \in \{t_i, t_j, |t_j - t_i|\}$, the corresponding term in Eq. (20) reduces to zero: $4\pi^2 t_y^2 f_n^2 h_f \big(\frac{3}{2} - \gamma - \ln(2\pi f_l t_y) \big) = 0$. Equation (20) can then further be simplified to obtain Eq. (6).