1 Introduction

This is a companion appendix as well as a more detailed description of the post-processing method used in our conference paper titled Probability integral based post-processing for photonic quantum random number generators, submittted to K&I'25. The goal of the presented post-processing method is to produce a close to uniformly distributed output sequence from physical measurement samples of quantum random number generators (QRNGs), following some other than uniform, but well characterized distribution.

2 General working principle of the post-processing method

2.1 Ideal case

Consider the following transformation based on the continuous probability integral transform [1] for processing measurement samples of physical random number generators:

Let D be a discrete random variable with possible d_i values from the set $\{\, d_i : i = 0, 1, 2, \dots \,\}$ with $p_i = \Pr(D = d_i)$ probabilities. Define a bijective $t(\cdot)$ function that assigns a c_i coordinate value to each possible d_i value, such that

$$c_i = t(d_i) = \sum_{j=0}^{i} \Pr(D = d_j) = \sum_{j=0}^{i} p_j. \tag{1}$$

This way, the $C = t(D)$ transformed discrete random variable has possible $c_i \in [0, 1]$ values from the set $\{\, c_i : i = 0, 1, 2, \dots \,\}$ with probabilities $p_i = \Pr(C = c_i) = \Pr(D = d_i)$. For notational convenience in later sections also set the edge case $c_{-1} \triangleq 0$. Note, that c_i form a monotonically increasing sequence in i on $[0, 1]$.

Then for $y \in [0, 1]$,

$$\Pr(C < y) = \sum_{\forall c_j : c_j \leq y} \Pr(C = c_j) = \sum_{j=0}^{N_y} \Pr(C = c_j) = \sum_{j=0}^{N_y} p_j = c_{N_y}, \tag{2}$$

where $N_y = \arg\max_i \{c_i \mid c_i \leq y\}$ is the biggest index of c_i for which $c_i \leq y$ still holds. Notice that for a random variable U uniformly distributed on $[0, 1]$, $\Pr(U < y) = y$; therefore, (2) approximates the behavior of an ideal uniform distribution with an error of $0 \leq y - c_{N_y} < c_{N_y+1} - c_{N_y}$.

Furthermore, the approximation error is further upper bounded by the maximum of $(c_{i+1} - c_i)$, giving

$$\begin{aligned} \Pr(U < y) - \Pr(C < y) = y - c_{N_y} < c_{N_y+1} - c_{N_y} \\ \leq \max_i (c_{i+1} - c_i) = \max_i \left(\sum_{j=0}^{i+1} p_j - \sum_{j=0}^{i} p_j \right) = \max_i p_i \end{aligned} \tag{3}$$

Similarly,

$$\Pr(x \leq C < y) = \Pr(C < y) - \Pr(C < x) = \sum_{j=0}^{N_y} p_j - \sum_{j=0}^{N_x} p_j = c_{N_y} - c_{N_x} \quad (4)$$

where $N_x = \arg\max_i \{c_i \mid c_i \leq x\}$ is the biggest index of c_i for which $c_i \leq x$ still holds, and

$$\begin{aligned} |\Pr(x \leq U < y) - \Pr(x \leq C < y)| &= \left|y - c_{N_y} - (x - c_{N_x})\right| \\ &< \max(c_{N_y+1} - c_{N_y}, c_{N_x+1} - c_{N_x}) \leq \max_i p_i. \end{aligned} \quad (5)$$

The previous equation can also be rewritten using the $H_\infty(D)$ min-entropy giving:

$$|\Pr(x \leq U < y) - \Pr(x \leq C < y)| < \max_i p_i = 2^{-H_\infty(D)} \quad (6)$$

2.2 Generating close to uniform output sequences

Typically, the expected output distribution of random number generators is uniform. Given a D input distribution over $\{0,1\}^{n_e}$ with 2^{n_e} possible values, the goal of post-processing is to create an R output result distribution on the $\{\, r_0, r_1 \ldots r_l \ldots r_{2^{m_e}-1} \,\}, l \in \mathbb{N}, 0 \leq l < 2^{m_e}$, sample space over $\{0,1\}^{m_e}$ that is statistically ϵ-close to the U uniform distribution over $\{0,1\}^{m_e}$ with 2^{m_e} possible values.

If the D distribution is known, the previously presented idea can be used to transform an input distribution into one approximating the uniform distribution with an upper bound on approximation error using the following scheme:

1. Using the $t(\cdot)$ function presented in in Sec. **??** assign c_i values on $[0,1]$ to all possible d_i values, thus creating a $C = t(D)$ transformed discrete random variable.

2. Assign r_l output values to all c_i such that if $\frac{l}{2^{m_e}} < c_i \leq \frac{l+1}{2^{m_e}}$ then r_l is assigned.

According to (4) and (5), the following holds for each of the r_l output values:

$$\Pr(R = r_l) = \Pr\left(\frac{l}{2^{m_e}} \leq C < \frac{l+1}{2^{m_e}}\right), \quad (7)$$

therefore,

$$d(r_l) = \left|\Pr\left(\frac{l}{2^{m_e}} \leq U < \frac{l+1}{2^{m_e}}\right) - \Pr(R = r_l)\right| < \max_i p_i, \quad (8)$$

meaning that the R distribution of the r_l output values approximate a uniform one over $\{0,1\}^m$ with an upper bound on approximation error given by (8).

2.2.1 Upper bound on bit bias

Note, that this upper bound on approximation error can also be used to give an upper bound on the bias of a single bit in the created m_e bit long sequence. Let $\mathbf{B} = \{\, B_0, B_1, \ldots, B_{m_e-1} \,\}$ be the bits in the m_e bit long sequence with $\mathbf{x} = \{\, x_0, x_1, \ldots, x_{m_e-1} \,\}$ bit values and A be the event of $\{B_0 = x_0, \ldots, B_{i-1} = x_{i-1}, B_{i+1} = x_{i+1}, \ldots, B_{m_e-1} = x_{m_e-1}\}$, where all $\mathbf{x}$ values for the $\mathbf{B}$ bits in the sequence are known except for the ith bit. Then, for any ith bit in the sequence:

$$\begin{aligned} \Pr(B_i = x_i \mid A) &= \frac{\Pr(\mathbf{B} = \mathbf{x})}{\Pr(A)} \\ &= \frac{\Pr(\mathbf{B} = \mathbf{x})}{\Pr(A, B_i = x_i) + \Pr(A, B_i = \overline{x_i})}. \end{aligned} \tag{9}$$

Looking at a worst-case scenario, this gives

$$\begin{aligned} &\left|\Pr(B_i = x_i) - \frac{1}{2}\right| \le \left|\frac{\frac{1}{2^{m_e}} + \max|d(r_l)|}{\frac{1}{2^{m_e}} - \max|d(r_l)| + \frac{1}{2^{m_e}} + \max|d(r_l)|} - \frac{1}{2}\right| \\ &= \left|\frac{\max|d(r_l)|}{2 \cdot \frac{1}{2^{m_e}}}\right| = \left|2^{m_\mathrm{e}-1} \max|d(r_l)|\right| < 2^{m_\mathrm{e}-1} \max_i p_i \end{aligned} \tag{10}$$

for an upper bound on individual bit bias.

2.2.2 Statistical distance from the uniform distribution

The statistical distance (or the total variation distance) between two discrete probability distributions X and Y is given by

$$\Delta(X, Y) = \frac{1}{2} \sum_v |\Pr(X = v) - \Pr(Y = v)|. \tag{11}$$

An upper bound for the statistical distance of the output of the previously presented scheme from the goal uniform distribution can then be written as:

$$\begin{aligned} \Delta(U, R) = \frac{1}{2} \sum_l &\left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \le U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| \\ &\le 2^{m_\mathrm{e}-1} \max|d(r_l)| < 2^{m_\mathrm{e}-1} \max_i p_i = 2^{-(H_\infty(D) - m_\mathrm{e} + 1)} \end{aligned} \tag{12}$$

according to the upper bound on approximation error presented in (8). Notice that the end result of (10) and (12) is the same, meaning that this expression upper bounds both individual bit bias and statistical distance. Adopting the common ϵ notation for the upper bound of statistical distance, the criterion for the output distribution to be ϵ-close to the expected uniform distribution in terms of m_e, ϵ, and $H_\infty(D)$ is then:

$$\log_2 \epsilon < m_\mathrm{e} - H_\infty(D) - 1. \tag{13}$$

2.3 Bit generation scheme using the joint distribution of multiple independent identically distributed measurement samples

In the following let us consider as the input distribution the joint probability distribution of multiple i.i.d. samples taken from the same distributon. Then, the possible $\underline{d}$ values of the input $\underline{D}$ joint distribution are $\underline{q} = (q_0, q_1, \ldots, q_{Q-1})$ Q-tuples, where the elements are the possible outcome combinations of the Q individual samples, with $\underline{p}_i = \Pr(\underline{D} = \underline{d}_i) = \Pr(D_0 = q_{0,i}, \ldots, D_{Q-1} = q_{Q-1,i}) = \prod_{\vartheta=0}^{Q-1} p_{q_{\vartheta,i}}$. Also, let us denote the probabilities of individual samples with $p_{q_\vartheta} = \Pr(D_\vartheta = q_\vartheta)$.

Let $t_Q(\cdot)$ be the bijective function that determines the i index assignment of the possible $\underline{d}_i$ Q-tuples meaning that $i = t_Q(\underline{q}) = t_Q((q_{0,i}, q_{1,i}, \ldots, q_{Q-1,i}))$ and $\underline{q} = (q_{0,i}, q_{1,i}, \ldots, q_{Q-1,i}) = t_Q^{-1}(i)$. According to the previous section,

$$\underline{c}_i = t(\underline{d}_i) = \sum_{j=0}^{i} \Pr\big(\underline{D} = \underline{d}_j\big) = \sum_{j=0}^{i} \underline{p}_j = \sum_{j=0}^{i} \prod_{\vartheta=0}^{Q-1} p_{q_{\vartheta,j}}, \tag{14}$$

where $q_{\vartheta,j}$ are the respective elements of the $q = t_Q^{-1}(j)$ tuples and $p_{q_{\vartheta,j}}$ are the associated probabilities. According to (8)

$$\begin{aligned} d(r_l) = \left|\Pr\!\left(\frac{l}{2^{m_\mathrm{e}}} \le U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| &< \max(c_{N_y+1} - c_{N_y}, c_{N_x+1} - c_{N_x}) \\ &\le \max_i \underline{p}_i = \prod_{\vartheta=0}^{Q-1} \max_i p_{q_{\vartheta,i}} = \Big(\max_i p_i\Big)^Q, \end{aligned} \tag{15}$$

and the upper bound for the statistical distance is:

$$\begin{aligned} \Delta(U, R) &= \frac{1}{2} \sum_l \left|\Pr\!\left(\frac{l}{2^{m_\mathrm{e}}} \le U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| \\ &< 2^{m_\mathrm{e}-1} \max_i \underline{p}_i = 2^{-(H_\infty(\underline{D}) - m_\mathrm{e} + 1)} = 2^{-(QH_\infty(D) - m_\mathrm{e} + 1)}, \end{aligned} \tag{16}$$

giving

$$\log_2 \epsilon < m_\mathrm{e} - H_\infty(\underline{D}) - 1 = m_\mathrm{e} - QH_\infty(D) - 1, \tag{17}$$

where $H_\infty(D)$ is the min-entropy of an individual sample and $H_\infty(\underline{D})$ is the min-entropy of the joint distribution of multiple samples, respectively. Notice that by increasing Q, ϵ decreases, allowing for creating output distributions that are arbitrarily close to the uniform distribution by using the joint distribution of sufficiently many individual measurement samples as input for the scheme.

3 Output characteristics in the presence of general estimation errors

3.1 General distributions

For the approximation scheme to work correctly, the distribution of D must be known. However, in most situations, fully characterizing D may not be practically feasible. Therefore, it is worth investigating the case where the probability of the d_i values is not exactly known but with some e_i error, giving $\Pr(D = d_i) = p_i = \hat{p}_i + e_i$, where $\hat{p}_i$ represents our best estimate of the true p_i probability.

Assuming that the $t(\cdot)$ coordinate assignment function assigns c_i coordinate values to d_i values according to the estimated $\hat{p}_i$ probabilities, the assignment rule presented in (1) changes to

$$c_i = \sum_{j=0}^{i} \hat{p}_j = \sum_{j=0}^{i} \left(p_j - e_j \right), \tag{18}$$

and

$$\Pr(C < y) = \sum_{j=0}^{N_y} p_j = c_{N_y} + \sum_{j=0}^{N_y} e_j. \tag{19}$$

The approximation error according to (3) is then

$$\begin{aligned} \Pr(U < y) - \Pr(C < y) &= y - \left(c_{N_y} + \sum_{j=0}^{N_y} e_j \right) \\ &< c_{N_y+1} - c_{N_y} + \sum_{j=0}^{N_y} e_j \le \max_i \hat{p}_i + \sum_{j=0}^{N_y} e_j. \end{aligned} \tag{20}$$

Similarly,

$$\begin{aligned} \Pr(x \le C < y) &= \Pr(C < y) - \Pr(C < x) \\ &= c_{N_y} + \sum_{j=0}^{N_y} e_j - \left(c_{N_x} + \sum_{j=0}^{N_x} e_j \right) = c_{N_y} - c_{N_x} + \sum_{j=N_x+1}^{N_y} e_j, \end{aligned} \tag{21}$$

and

$$\begin{aligned} &|\Pr(x \le U < y) - \Pr(x \le C < y)| \\ &= \left| y - \left(c_{N_y} + \sum_{j=0}^{N_y} e_j \right) - \left(x - \left(c_{N_x} + \sum_{j=0}^{N_x} e_j \right) \right) \right| \\ &< \left| c_{N_y+1} - c_{N_y} - (c_{N_x+1} - c_{N_x}) + \sum_{j=N_x+1}^{N_y} e_j \right| \le \max_i \hat{p}_i + \left| \sum_{j=N_x+1}^{N_y} e_j \right|. \end{aligned} \tag{22}$$

Naturally, this can also be written in terms of the estimated $H_\infty(\hat{D}) = -\log_2 \max_i \hat{p}_i$ min-entropy giving:

$$|\Pr(x \le U < y) - \Pr(x \le C < y)| < 2^{-H_\infty(\hat{D})} + \left| \sum_{j=N_x+1}^{N_y} e_j \right| \tag{23}$$

3.1.1 Output bit generation

Similarly to the ideal case, the error of output bit assignment can be rewritten with errors taken into account as:

$$\begin{aligned} d(r_l) &= \left| \Pr\left(\frac{l}{2^{m_\mathrm{e}}} \le U < \frac{l+1}{2^{m_\mathrm{e}}} \right) - \Pr(R = r_l) \right| \\ &< \left| c_{N_y+1} - c_{N_y} - (c_{N_x+1} - c_{N_x}) + \sum_{j=N_x+1}^{N_y} e_j \right| \le \max_i \hat{p}_i + \left| \sum_{j=N_x+1}^{N_y} e_j \right| \end{aligned} \tag{24}$$

with the choice of $x = \frac{l}{2^{m_\mathrm{e}}}$ and $y = \frac{l+1}{2^{m_\mathrm{e}}}$ for the rest of the section. We can also give the following upper bound for output bit bias:

$$\left| \Pr(B_i = x_i) - \frac{1}{2} \right| \le \left| 2^{m_\mathrm{e}-1} \max |d(r_l)| \right| < 2^{m_\mathrm{e}-1} \max_{i,l} \left(\hat{p}_i + \left| \sum_{j=N_x+1}^{N_y} e_j \right| \right). \tag{25}$$

The statistical distance from the expected uniform output is then

$$\begin{aligned} \Delta(U, R) &= \frac{1}{2} \sum_l \left| \Pr\left(\frac{l}{2^{m_\mathrm{e}}} \le U < \frac{l+1}{2^{m_\mathrm{e}}} \right) - \Pr(R = r_l) \right| \\ &< \frac{1}{2} \sum_l \left(\max_i \hat{p}_i + \left| \sum_{j=N_x+1}^{N_y} e_j \right| \right) = 2^{m_\mathrm{e}-1} \max_i \hat{p}_i + \frac{1}{2} \sum_l \left| \sum_{j=N_x+1}^{N_y} e_j \right| \\ &\le 2^{m_\mathrm{e}-1} \max_i \hat{p}_i + 2^{m_\mathrm{e}-1} \max_l \left| \sum_{j=N_x+1}^{N_y} e_j \right| = 2^{-(H_\infty(\hat{D}) - m_\mathrm{e} + 1)} + 2^{m_\mathrm{e}-1} \max_l \left| \sum_{j=N_x+1}^{N_y} e_j \right|. \end{aligned} \tag{26}$$

Alternatively, the sum of the individual $|e_i|$ errors can also be used for an upper bound:

$$\begin{aligned}
&\Delta(U,R) = \frac{1}{2}\sum_l \left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \leq U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| \\
&< \frac{1}{2}\sum_l \left(\max_i \hat{p}_i + \left|\sum_{j=N_x+1}^{N_y} e_j\right|\right) = 2^{m_\mathrm{e}-1}\max_i \hat{p}_i + \frac{1}{2}\sum_l \left|\sum_{j=N_x+1}^{N_y} e_j\right| \\
&\leq 2^{m_\mathrm{e}-1}\max_i \hat{p}_i + \frac{1}{2}\sum_i |e_i| = 2^{-(H_\infty(\hat{D})-m_\mathrm{e}+1)} + \frac{1}{2}\sum_i |e_i| \\
&\leq 2^{-(H_\infty(\hat{D})-m_\mathrm{e}+1)} + \frac{1}{2}\sum_i \left|\max_i e_i\right|.
\end{aligned} \tag{27}$$

Results of (24) and (26) show that the quality (distance from the expected uniform) of the output R distribution is heavily influenced by the accuracy of estimation of the input D distribution. In particular by the $\left|\sum_{j=N_x+1}^{N_y} e_j\right|$ accumulated errors for each (N_x, N_y) interval with $x = \frac{l}{2^{m_\mathrm{e}}}, y = \frac{l+1}{2^{m_\mathrm{e}}}$ corresponding to the r_l output values.

3.2 Effect of estimation errors when using the joint distribution of multiple samples

Similarly to Sec. 3.1 let us investigate the effect of estimation errors when using the joint distribution of multiple samples utilizing the notations introduced in Sec. 2.3. The probabilities of the individual samples are then $p_{q_\vartheta} = \Pr(D_\vartheta = q_\vartheta) = \hat{p}_{q_\vartheta} + e_{q_\vartheta}$, with $\hat{p}_{q_\vartheta}$ estimated probability and e_{q_ϑ} estimation error for a particular individual sample in the joint distribution. Using the $\underline{\hat{p}}_i = \prod_{\vartheta=0}^{Q-1} \hat{p}_{q_{\vartheta,i}}$ estimate for the joint probability, the actual $\underline{p}_i$ joint probabilities are then:

$$\underline{p}_i = \Pr(\underline{D} = \underline{d}_i) = \underline{\hat{p}}_i + \underline{e}_i = \prod_{\vartheta=0}^{Q-1} p_{q_{\vartheta,i}} = \prod_{\vartheta=0}^{Q-1} \left(\hat{p}_{q_{\vartheta,i}} + e_{q_{\vartheta,i}}\right) \tag{28}$$

and

$$\underline{e}_i = \underline{p}_i - \underline{\hat{p}}_i = \prod_{\vartheta=0}^{Q-1} \left(\hat{p}_{q_{\vartheta,i}} + e_{q_{\vartheta,i}}\right) - \prod_{\vartheta=0}^{Q-1} \hat{p}_{q_{\vartheta,i}}, \tag{29}$$

with

$$\underline{c}_i = \sum_{j=0}^{i} \underline{\hat{p}}_j = \sum_{j=0}^{i} \left(\underline{p}_j - \underline{e}_j\right). \tag{30}$$

The approximation errors and statistical distance can also be calculated by substituting the relevant $\underline{p}_i, \underline{\hat{p}}_i$ and $\underline{e}_i$ into results of Sec. 3.1, eventually giving:

$$\begin{aligned}&\left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \le U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| \\ &< \left|c_{N_y+1} - c_{N_y} - (c_{N_x+1} - c_{N_x}) + \sum_{j=N_x+1}^{N_y} \underline{e}_j\right| \le \max_i \underline{\hat{p}}_i + \left|\sum_{j=N_x+1}^{N_y} \underline{e}_j\right|\end{aligned} \tag{31}$$

and

$$\begin{aligned}&\Delta(U,R) = \frac{1}{2}\sum_l \left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \le U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| \\ &< \frac{1}{2}\sum_l \left(\max_i \underline{\hat{p}}_i + \left|\sum_{j=N_x+1}^{N_y} \underline{e}_j\right|\right) = 2^{m_\mathrm{e}-1} \max_i \underline{\hat{p}}_i + \frac{1}{2}\sum_l \left|\sum_{j=N_x+1}^{N_y} \underline{e}_j\right| \\ &\le 2^{m_\mathrm{e}-1} \max_i \underline{\hat{p}}_i + 2^{m_\mathrm{e}-1} \max_l \left|\sum_{j=N_x+1}^{N_y} \underline{e}_j\right| = 2^{-(QH_\infty(\hat{D})-m_\mathrm{e}+1)} + 2^{m_\mathrm{e}-1} \max_l \left|\sum_{j=N_x+1}^{N_y} \underline{e}_j\right|\end{aligned} \tag{32}$$

or

$$\begin{aligned}&\Delta(U,R) = \frac{1}{2}\sum_l \left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \le U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| \\ &< \frac{1}{2}\sum_l \left(\max_i \underline{\hat{p}}_i + \left|\sum_{j=N_x+1}^{N_y} \underline{e}_j\right|\right) = 2^{m_\mathrm{e}-1} \max_i \underline{\hat{p}}_i + \frac{1}{2}\sum_l \left|\sum_{j=N_x+1}^{N_y} \underline{e}_j\right| \\ &\le 2^{m_\mathrm{e}-1} \max_i \underline{\hat{p}}_i + \frac{1}{2}\sum_i |\underline{e}_i| = 2^{-(QH_\infty(\hat{D})-m_\mathrm{e}+1)} + \frac{1}{2}\sum_i |\underline{e}_i| \\ &\le 2^{-(QH_\infty(\hat{D})-m_\mathrm{e}+1)} + \frac{1}{2}\sum_i \left|\max_i \underline{e}_i\right|.\end{aligned} \tag{33}$$

Let us investigate the behavior of the $\sum_i |\underline{e}_i|$ error term. For this, start with the simplified general case of the joint distribution of two independent (but not necessarily identically distributed) samples. Then, $\Pr(D_1 = q_1) = p_{q_1} = \hat{p}_{q_1} + e_{q_1}$, where p_{q_1} is the actual probability of the first sample being q_1 and $\hat{p}_{q_1}$ is our best estimate for p_{q_1} with e_{q_1} error. Similarly, for the second sample $\Pr(D_2 = q_2) = p_{q_2} = \hat{p}_{q_2} + e_{q_2}$, where p_{q_2} is the actual probability and $\hat{p}_{q_2}$ is the estimate with e_{q_2} error. Then the joint distribution of the two samples is:

$$\begin{aligned}\Pr(D_1 = q_1, D_2 = q_2) &= \underline{p}_{(q_1,q_2)} = p_{q_1}p_{q_2} = (\hat{p}_{q_1} + e_{q_1})(\hat{p}_{q_2} + e_{q_2}) \\ &= \hat{p}_{q_1}\hat{p}_{q_2} + \hat{p}_{q_1}e_{q_2} + \hat{p}_{q_2}e_{q_1} + e_{q_1}e_{q_2} = \underline{\hat{p}}_{(q_1,q_2)} + \underline{e}_{(q_1,q_2)}\end{aligned} \tag{34}$$

Knowing that, $\underline{\hat{p}}_{(q_1,q_2)} = \hat{p}_{q_1}\hat{p}_{q_2}$,

$$\underline{e}_{(q_1,q_2)} = \underline{p}_{(q_1,q_2)} - \underline{\hat{p}}_{(q_1,q_2)} = \hat{p}_{q_1}e_{q_2} + \hat{p}_{q_2}e_{q_1} + e_{q_1}e_{q_2}. \tag{35}$$

The $\sum_{q_1,q_2} \underline{e}_{(q_1,q_2)}$ error term for this joint distribution is then:

$$\begin{aligned}\sum_{q_1,q_2} \underline{e}_{(q_1,q_2)} &= \sum_{q_1}\sum_{q_2} \underline{e}_{(q_1,q_2)} = \sum_{q_1}\sum_{q_2} \left(\hat{p}_{q_1} e_{q_2} + \hat{p}_{q_2} e_{q_1} + e_{q_1} e_{q_2}\right) \\ &= \sum_{q_1} \hat{p}_{q_1} \sum_{q_2} e_{q_2} + \sum_{q_1} e_{q_1} \sum_{q_2} \hat{p}_{q_2} + \sum_{q_1} e_{q_1} \sum_{q_2} e_{q_2} \\ &= \sum_{q_1} e_{q_1} + \sum_{q_2} e_{q_2} + \sum_{q_1} e_{q_1} \sum_{q_2} e_{q_2}.\end{aligned} \tag{36}$$

Similarly,

$$\begin{aligned}\sum_{q_1,q_2} \left|\underline{e}_{(q_1,q_2)}\right| &= \sum_{q_1}\sum_{q_2} \left(\hat{p}_{q_1} |e_{q_2}| + \hat{p}_{q_2} |e_{q_1}| + |e_{q_1}||e_{q_2}|\right) \\ &= \sum_{q_1} |e_{q_1}| + \sum_{q_2} |e_{q_2}| + \sum_{q_1} |e_{q_1}| \sum_{q_2} |e_{q_2}|.\end{aligned} \tag{37}$$

Using the results of (36) and (37), the calculation of these error terms for the joint distribution of multiple samples is also possible, by using a step-by-step approach[1]. Importantly, these results also show that $\sum_{q_1,q_2} \left|\underline{e}_{(q_1,q_2)}\right| \geq \sum_{q_1} |e_{q_1}|$ and $\sum_{q_1,q_2} \left|\underline{e}_{(q_1,q_2)}\right| \geq \sum_{q_2} |e_{q_2}|$. This means that by using the joint distribution of multiple samples as input, the error terms present in the upper bound for ϵ increase, and thus, ϵ cannot be made arbitrarily small.

4 Using the scheme with photon time-of-arrival based QRNGs

4.1 Ideal case

In the practical scenario of a QRNG based on photon arrival times, the input D distribution is the distribution of the measured discretized time differences between photon detection events, which are supposed to be exponentially distributed in an ideal scenario. Therefore,

$$p_i = p_n = \Pr(D = d_i = n) = \mathrm{e}^{-n\lambda\tau}\left(1 - \mathrm{e}^{-\lambda\tau}\right), \tag{38}$$

allowing the convenient choice of $d_i = n, i = n$ (an indexing scheme where the nth possible measurement outcome is the one with the value of n) since both the i index of d_i and our possible n measurement outcomes are non-negative integers. The $t(\cdot)$ assignment function of c_n is then:

$$t(d_i) = t(n) = c_n = \sum_{j=0}^{n} p_j = \sum_{j=0}^{n} \mathrm{e}^{-j\lambda\tau}\left(1 - \mathrm{e}^{-\lambda\tau}\right) = 1 - \mathrm{e}^{-(n+1)\lambda\tau}. \tag{39}$$

[1] For example, to calculate results for the joint distribution of three samples, one can first calculate the quantities corresponding to the joint distribution of two samples, and then use it paired with the quantities corresponding to a single sample, to get results for the case of three samples.

The post-processing scheme can then be defined following the steps presented in Sec. 2.2. The upper bound of the scheme's approximation error according to (8) is then:

$$\begin{aligned}\left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \leq U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| &< \max(c_{N_y+1} - c_{N_y}, c_{N_x+1} - c_{N_x}) \\ &\leq \max_n p_n = \left(1 - \mathrm{e}^{-\lambda\tau}\right).\end{aligned} \tag{40}$$

The value of N_y and N_x can also be calculated using the $t^{-1}(\cdot)$ inverse function, giving

$$N_y \leq -\frac{\ln(1-y)}{\lambda\tau} - 1 \Rightarrow N_y = \left\lfloor -\frac{\ln(1-y)}{\lambda\tau} - 1 \right\rfloor \tag{41}$$

where $\lfloor\cdot\rfloor$ is the floor function (or greatest integer function). The value for N_x can be calculated similarly. The $c_{N_y+1} - c_{N_y}$ quantity in (40) is then:

$$c_{N_y+1} - c_{N_y} = 1 - \mathrm{e}^{-(N_y+2)\lambda\tau} - \left(1 - \mathrm{e}^{-(N_y+1)\lambda\tau}\right) = \mathrm{e}^{-\left\lfloor -\frac{\ln(1-y)}{\lambda\tau}\right\rfloor\lambda\tau}\left(1 - \mathrm{e}^{-\lambda\tau}\right). \tag{42}$$

Notice, that (42) is decreasing in N_y and therefore

$$\begin{aligned}&\max(c_{N_y+1} - c_{N_y}, c_{N_x+1} - c_{N_x}) = c_{N_x+1} - c_{N_x} \\ &= \mathrm{e}^{-\left\lfloor -\frac{\ln(1-x)}{\lambda\tau}\right\rfloor\lambda\tau}\left(1 - \mathrm{e}^{-\lambda\tau}\right) = \mathrm{e}^{\left\lceil \frac{\ln(1-x)}{\lambda\tau}\right\rceil\lambda\tau}\left(1 - \mathrm{e}^{-\lambda\tau}\right).\end{aligned} \tag{43}$$

The upper bound for statistical distance according to (12) is

$$\begin{aligned}&\Delta(U,R) < \frac{1}{2}\sum_l \mathrm{e}^{\left\lceil \ln\left(1-\frac{l}{2^{m_\mathrm{e}}}\right)\frac{1}{\lambda\tau}\right\rceil\lambda\tau}\left(1 - \mathrm{e}^{-\lambda\tau}\right) \\ &\leq 2^{m_\mathrm{e}-1}\max_n p_n = 2^{m_\mathrm{e}-1}\cdot\left(1 - \mathrm{e}^{-\lambda\tau}\right) = 2^{-(H_\infty(D) - m_\mathrm{e} + 1)},\end{aligned} \tag{44}$$

with (13) taking the form of

$$\log_2 \epsilon < m_\mathrm{e} - H_\infty(D) - 1 = m_\mathrm{e} - \log_2\left(1 - \mathrm{e}^{-\lambda\tau}\right) - 1. \tag{45}$$

These quantities can also be calculated for the case of using the joint distribution of multiple measurement results, giving:

$$\begin{aligned}\underline{p}_i &= \Pr(\underline{D} = \underline{d}_i) = \Pr(D_0 = q_{0,i}, D_1 = q_{1,i}, \ldots, D_{Q-1} = q_{Q-1,i}) \\ &= \prod_{\vartheta=0}^{Q-1} p_{q_\vartheta} = \left(1 - \mathrm{e}^{-\lambda\tau}\right)^Q \prod_{\vartheta=0}^{Q-1} \mathrm{e}^{-q_{\vartheta,i}\lambda\tau}.\end{aligned} \tag{46}$$

The c_i values are:

$$c_i = \sum_{j=0}^{i} \Pr\left(\underline{D} = \underline{d}_j\right) = \sum_{j=0}^{i} \underline{p}_j = \sum_{j=0}^{i}\left(\left(1 - \mathrm{e}^{-\lambda\tau}\right)^Q \prod_{\vartheta=0}^{Q-1} \mathrm{e}^{-q_{\vartheta,j}\lambda\tau}\right), \tag{47}$$

and the upper bound of approximation error is

$$\left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \leq U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| < \max(c_{N_y+1} - c_{N_y}, c_{N_x+1} - c_{N_x}) \\ \leq \max_i \underline{p}_i = \left(1 - \mathrm{e}^{-\lambda\tau}\right)^Q . \tag{48}$$

Similarly to (41) and (42) N_y and N_x can be calculated:

$$N_y \leq t^{-1}(y) \Rightarrow N_y = \left\lfloor t^{-1}(y) \right\rfloor , \tag{49}$$

as well as,

$$c_{N_y+1} - c_{N_y} = \sum_{j=0}^{N_y+1} \underline{p}_j - \sum_{j=0}^{N_y} \underline{p}_j = \underline{p}_{N_y+1} = \left(1 - \mathrm{e}^{-\lambda\tau}\right)^Q \prod_{\vartheta=0}^{Q-1} \mathrm{e}^{-q_{\vartheta,N_y+1}\lambda\tau}. \tag{50}$$

Finally, the upper bound for the statistical distance is:

$$\begin{aligned}\Delta(U,R) &= \frac{1}{2}\sum_l \left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \leq U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| \\ &< \frac{1}{2}\sum_l \left|\max(c_{N_y+1} - c_{N_y}, c_{N_y+1} - c_{N_y})\right| = \frac{1}{2}\sum_l \left|\max(\underline{p}_{N_y+1}, \underline{p}_{N_x+1})\right| \\ &\leq 2^{m_\mathrm{e}-1} \max_i \underline{p}_i = 2^{m_\mathrm{e}-1}\left(1 - \mathrm{e}^{-\lambda\tau}\right)^Q = 2^{-(H_\infty(\underline{D}) - m_\mathrm{e}+1)} = 2^{-(QH_\infty(D) - m_\mathrm{e}+1)}.\end{aligned} \tag{51}$$

4.2 Effect of estimation errors of the input photon rate

An error in the estimation of the input rate causes e_i estimation errors mostly with the same sign, making it a major contributor to the accumulated $\left|\sum_{j=N_x+1}^{N_y} e_j\right|$ errors both in the case of using the distribution of single detection events or the joint distribution of multiple detections as input for the post-processing scheme.

4.2.1 Using the distribution of single measurement samples

First, consider the case of estimating the actual $\lambda = \hat{\lambda} + \lambda_e$ input rate with $\hat{\lambda}$ estimation value and λ_e estimation error. Then,

$$p_n = \Pr(D = d_i = n) = \mathrm{e}^{-n(\hat{\lambda}+\lambda_e)\tau}\left(1 - \mathrm{e}^{-(\hat{\lambda}+\lambda_e)\tau}\right), \tag{52}$$

$$\hat{p_n} = \mathrm{e}^{-n\hat{\lambda}\tau}\left(1 - \mathrm{e}^{-\hat{\lambda}\tau}\right), \tag{53}$$

while

$$c_n = \sum_{j=0}^{n} \hat{p}_j = \sum_{j=0}^{n} (p_j - e_j) = 1 - \mathrm{e}^{-(n+1)\hat{\lambda}\tau}. \tag{54}$$

and the e_n errors can of course be calculated as $e_n = p_n - \hat{p_n}$. Using these results, the approximation errors can be calculated similarly to Sec. 3.1:

$$\begin{aligned}\Pr(C < y) &= \sum_{j=0}^{N_y} p_j = 1 - \mathrm{e}^{-(N_y+1)(\hat{\lambda}+\lambda_e)\tau} \\ &= c_{N_y} + \mathrm{e}^{-(N_y+1)\hat{\lambda}\tau}\left(1 - \mathrm{e}^{-(N_y+1)\lambda_e\tau}\right) = c_{N_y} + \sum_{j=0}^{N_y} e_j.\end{aligned} \tag{55}$$

Using (21), (22) and (24),

$$\begin{aligned}&\left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \leq U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| < \max(c_{N_y+1} - c_{N_y}, c_{N_x+1} - c_{N_x}) + \left|\sum_{j=N_x+1}^{N_y} e_j\right| \\ &= \mathrm{e}^{\left\lceil \ln\left(1-\frac{l}{2^{m_\mathrm{e}}}\right)\frac{1}{\lambda\tau}\right\rceil \lambda\tau}\left(1 - \mathrm{e}^{-\lambda\tau}\right) + \left|\sum_{j=N_x+1}^{N_y} e_j\right| \leq \max_i \hat{p}_i + \left|\sum_{j=N_x+1}^{N_y} e_j\right| \\ &= 1 - \mathrm{e}^{-\hat{\lambda}\tau} + \left|\left(\sum_{j=0}^{N_y} p_j - \sum_{j=0}^{N_x+1} p_j\right) - \left(\sum_{j=0}^{N_y} \hat{p}_j - \sum_{j=0}^{N_x+1} \hat{p}_j\right)\right| \\ &= 1 - \mathrm{e}^{-\hat{\lambda}\tau} + \left|\mathrm{e}^{-(N_y+1)\hat{\lambda}\tau} - \mathrm{e}^{-(N_y+1)(\hat{\lambda}+\lambda_e)\tau} + \mathrm{e}^{-(N_x+2)(\hat{\lambda}+\lambda_e)\tau} - \mathrm{e}^{-(N_x+2)\hat{\lambda}\tau}\right|,\end{aligned} \tag{56}$$

where $N_x = \left\lfloor -\frac{\ln\left(1-\frac{l}{2^{m_\mathrm{e}}}\right)}{\lambda\tau} - 1 \right\rfloor$ and $N_y = \left\lfloor -\frac{\ln\left(1-\frac{l+1}{2^{m_\mathrm{e}}}\right)}{\lambda\tau} - 1 \right\rfloor$.

Total magnitude of errors Knowing that the

$$e_n = p_n - \hat{p}_n = \mathrm{e}^{-n(\hat{\lambda}+\lambda_e)\tau}\left(1 - \mathrm{e}^{-(\hat{\lambda}+\lambda_e)\tau}\right) - \mathrm{e}^{-n\hat{\lambda}\tau}\left(1 - \mathrm{e}^{-\hat{\lambda}\tau}\right) = 0 \tag{57}$$

equation has only a singular solution for n in the form of $n_{e=0} = -\frac{\ln\left(\frac{1-\mathrm{e}^{-\hat{\lambda}\tau}}{1-\mathrm{e}^{-(\hat{\lambda}+\lambda_e)\tau}}\right)}{\lambda_e\tau}$, it can be seen that e_n values for $n < n_{e=0}$ have the same sign as λ_e, while e_n values for $n > n_{e=0}$ have a sign opposite of λ_e. Additionally considering that $\sum_n e_n = \sum_n p_n - \sum_n \hat{p}_n = 1 - 1 = 0$, means that

$$\sum_{n=0}^{\lfloor n_{e=0} \rfloor} e_n = - \sum_{n=\lceil n_{e=0} \rceil}^{\infty} e_n, \tag{58}$$

and therefore,

$$\sum_{n} |e_n| = 2\cdot \left| \sum_{n=0}^{\lfloor n_{e=0} \rfloor} e_n \right| = 2\cdot \left| \left(1 - \mathrm{e}^{-(\lfloor n_{e=0} \rfloor+1)(\hat{\lambda}+\lambda_e)\tau}\right) - \left(1 - \mathrm{e}^{-(\lfloor n_{e=0} \rfloor+1)\hat{\lambda}\tau}\right) \right|. \tag{59}$$

Statistical distance The statistical distance from the expected uniform output according to (26) can be written as:

$$\begin{aligned}
\Delta(U,R) &= \frac{1}{2}\sum_{l} \left| \Pr\left(\frac{l}{2^{m_\mathrm{e}}} \leq U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l) \right| \\
&< \frac{1}{2}\sum_{l} \left(\mathrm{e}^{\left\lceil \ln\left(1-\frac{l}{2^{m_\mathrm{e}}}\right)\frac{1}{\lambda\tau}\right\rceil \lambda\tau} \left(1 - \mathrm{e}^{-\lambda\tau}\right) + \left| \sum_{j=N_x+1}^{N_y} e_j \right| \right) \\
&\leq \frac{1}{2}\sum_{l} \mathrm{e}^{\left\lceil \ln\left(1-\frac{l}{2^{m_\mathrm{e}}}\right)\frac{1}{\lambda\tau}\right\rceil \lambda\tau} \left(1 - \mathrm{e}^{-\lambda\tau}\right) + \frac{1}{2}\sum_{i} |e_i| \\
&\leq \frac{1}{2}\sum_{l} \max_{i} \hat{p}_i + \frac{1}{2}\sum_{i} |e_i| = 2^{m_\mathrm{e}-1}\left(1 - \mathrm{e}^{-\hat{\lambda}\tau}\right) + \frac{1}{2}\sum_{i} |e_i|.
\end{aligned} \tag{60}$$

4.2.2 Using the joint distribution of multiple photon arrival time differences

The notable quantities in this case of using the joint distribution of multiple measurement samples are: According to (52) and (53),

$$p_{q_\vartheta} = \Pr(D = q_\vartheta) = \mathrm{e}^{-q_\vartheta(\hat{\lambda}+\lambda_e)\tau}\left(1 - \mathrm{e}^{-(\hat{\lambda}+\lambda_e)\tau}\right), \tag{61}$$

$$\hat{p}_{q_\vartheta} = \mathrm{e}^{-q_\vartheta\hat{\lambda}\tau}\left(1 - \mathrm{e}^{-\hat{\lambda}\tau}\right), \tag{62}$$

giving

$$\underline{p}_i = \prod_{\vartheta=0}^{Q-1} p_{q_\vartheta} = \left(1 - \mathrm{e}^{-(\hat{\lambda}+\lambda_e)\tau}\right)^{Q} \prod_{\vartheta=0}^{Q-1} \mathrm{e}^{-q_\vartheta(\hat{\lambda}+\lambda_e)\tau}, \tag{63}$$

$$\underline{\hat{p}}_i = \prod_{\vartheta=0}^{Q-1} \hat{p}_{q_\vartheta} = \left(1 - \mathrm{e}^{-\hat{\lambda}\tau}\right)^{Q} \prod_{\vartheta=0}^{Q-1} \mathrm{e}^{-q_\vartheta\hat{\lambda}\tau}, \tag{64}$$

$$\underline{e}_i = \underline{p}_i - \underline{\hat{p}}_i, \tag{65}$$

and

$$\underline{c}_i = t(\underline{d}_i) = \sum_{j=0}^{i} \underline{\hat{p}}_j = \sum_{j=0}^{i} \left(\left(1 - \mathrm{e}^{-\hat{\lambda}\tau}\right)^{Q} \prod_{\vartheta=0}^{Q-1} \mathrm{e}^{-q_{\vartheta,j}\hat{\lambda}\tau} \right) = \sum_{j=0}^{i} \left(\underline{p}_j - \underline{e}_j\right). \tag{66}$$

The upper bound of approximation error is the same as in (31):

$$\begin{aligned} &\left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \leq U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| \\ &< \left| c_{N_y+1} - c_{N_y} - (c_{N_x+1} - c_{N_x}) + \sum_{j=N_x+1}^{N_y} \underline{e}_j \right| \leq \max_i \hat{\underline{p}}_i + \left| \sum_{j=N_x+1}^{N_y} \underline{e}_j \right|, \end{aligned} \tag{67}$$

while N_y and N_x can be calculated similarly to (49) and (50) with the only difference of using the estimated probabilities in $t^{-1}(\cdot)$, giving

$$N_y \leq t^{-1}(y) \Rightarrow N_y = \left\lfloor t^{-1}(y) \right\rfloor, \tag{68}$$

as well as,

$$c_{N_y+1} - c_{N_y} = \sum_{j=0}^{N_y+1} \hat{\underline{p}}_j - \sum_{j=0}^{N_y} \hat{\underline{p}}_j = \hat{\underline{p}}_{N_y+1} = \left(1 - \mathrm{e}^{-\hat{\lambda}\tau}\right)^Q \prod_{\vartheta=0}^{Q-1} \mathrm{e}^{-q_{\vartheta, N_y+1}\hat{\lambda}\tau}. \tag{69}$$

The upper bound for the statistical distance is:

$$\begin{aligned} \Delta(U,R) &= \frac{1}{2}\sum_l \left|\Pr\left(\frac{l}{2^{m_\mathrm{e}}} \leq U < \frac{l+1}{2^{m_\mathrm{e}}}\right) - \Pr(R = r_l)\right| \\ &< \frac{1}{2}\sum_l \left|\max(\underline{p}_{N_y+1}, \underline{p}_{N_x+1})\right| + \frac{1}{2}\sum_l \left| \sum_{j=N_x+1}^{N_y} \underline{e}_j \right| \\ &= \frac{1}{2}\sum_l \left|\max(\underline{p}_{N_y+1}, \underline{p}_{N_x+1})\right| + \frac{1}{2}\sum_i |\underline{e}_i| \\ &\leq 2^{m_\mathrm{e}-1} \max_i \hat{\underline{p}}_i + \frac{1}{2}\sum_i |\underline{e}_i| = 2^{m_\mathrm{e}-1}\left(1 - \mathrm{e}^{-\lambda\hat{\tau}}\right)^Q + \frac{1}{2}\sum_i |\underline{e}_i| \end{aligned} \tag{70}$$

Note that according to (37) the $\sum_i |\underline{e}_i|$ error term grows when using the joint distribution of multiple samples compared to the single sample case. Due to this, using the joint distribution of more samples is only beneficial up to a certain point, from where the magnitude of the $\Delta(U,R)$ statistical distance becomes dominated by the error term.

4.3 A practical bit generation scheme

In practice it may be beneficial to limit the possible number of measurement outcomes. Limiting the number of possible d_i input values also means a limited number of $i = t_Q(\underline{q})$ indexes. Consider the following modified distribution with N_O finite number of outcomes for the individual measurement samples:

$$\begin{aligned} &p_i = \\ &\begin{cases} \mathrm{e}^{-i\lambda\tau}\left(1 - \mathrm{e}^{-\lambda\tau}\right) = p_n & \text{if } n < N_O - 1, \\ 1 - \sum_{j=0}^{N_O-2} p_j = \mathrm{e}^{-\lambda(N_O-1)\tau} = p_{N_O-1} & \text{if } n \geq N_O - 1. \end{cases} \end{aligned} \tag{71}$$

Notice, that in the case of $\hat{\lambda}$ estimated rate with λ_{e} rate estimation error, the error terms corresponding to the modified part of $n \geq N_O - 1$ and the original exponential $n < N_O - 1$ part have the same sign if $N_O \geq n_{e=0} = -\frac{\ln\left(\frac{1-\mathrm{e}^{-\hat{\lambda}\tau}}{1-\mathrm{e}^{-(\hat{\lambda}+\lambda_e)\tau}}\right)}{\lambda_e \tau}$, therefore (59) holds if this criterion for N_O is satisfied.

For i index assignment, use the following rule:

$$i = t_Q(\overline{q}) = N_O^0 q_0 + N_O^1 q_1 + \ldots + N_O^{Q-1} q_{Q-1} = \sum_{\vartheta=0}^{Q-1} N_O^{\vartheta} q_{\vartheta}. \tag{72}$$

Then, the individual q_ϑ outcomes can be calculated from i as

$$q_\vartheta = \left\lfloor \left(i \mod N_O^{\vartheta+1}\right) / N_O^{\vartheta} \right\rfloor. \tag{73}$$

Notice that the

$$c_i = \sum_{j=0}^{i} \underline{p}_j = \sum_{j=0}^{i} \prod_{\vartheta=0}^{Q-1} p_{q_{\vartheta,j}} \tag{74}$$

values can be separated into smaller sums using the presented index assignment rules:

$$\begin{aligned} c_i = \sum_{j=0}^{i} \underline{p}_j = & \sum_{j=0}^{t_Q((N_O-1,N_O-1,\ldots,N_O-1,q_{Q-1}-1))} + \sum_{j=t_Q((0,0,\ldots,0,q_{Q-1}))}^{t_Q((N_O-1,N_O-1,\ldots,q_{Q-2}-1,q_{Q-1}))} \\ + \ldots + & \sum_{j=t_Q((0,0,\ldots,q_{Q-2},q_{Q-1}))}^{t_Q((N_O-1,q_1-1,\ldots,q_{Q-2}-1,q_{Q-1}))} + \sum_{j=t_Q((0,q_1,\ldots,q_{Q-2},q_{Q-1}))}^{t_Q((q_0-1,q_1,\ldots,q_{Q-2}-1,q_{Q-1}))} + \underline{p}_{t_Q((q_0,q_1,\ldots,q_{Q-2},q_{Q-1}))}. \end{aligned} \tag{75}$$

Using $\underline{p}_j = \prod_{\vartheta=0}^{Q-1} p_{q_{\vartheta,j}}$ and the fact that $\sum_{j=0}^{N_O-1} p_j = 1$, we can rewrite (75) as:

$$\begin{aligned} c_i = \sum_{j=0}^{i} \underline{p}_j = & \sum_{j=0}^{q_{Q-1}-1} p_j + p_{q_{Q-1}} \sum_{j=0}^{q_{Q-2}-1} p_j + \ldots + \prod_{k=2}^{Q-1} p_{q_k} \sum_{j=0}^{q_1-1} p_j + \prod_{k=1}^{Q-1} p_{q_k} \sum_{j=0}^{q_0-1} p_j + \prod_{k=0}^{Q-1} p_{q_k} \\ = & \sum_{\vartheta=0}^{Q-1} \left(\prod_{k=\vartheta+1}^{Q-1} p_{q_k} \sum_{j=0}^{q_\vartheta-1} p_j \right) + \prod_{k=0}^{Q-1} p_{q_k}. \end{aligned} \tag{76}$$

For a more convenient expression for calculating c_i.

Additionally, the $\sum_{j=0}^{q_\vartheta-1} p_j$ term can be further simplified as:

$$\sum_{j=0}^{q_\vartheta-1} p_j = \begin{cases} 1 - \mathrm{e}^{-q_\vartheta \lambda \tau} & \text{if } q_\vartheta < N_O - 1, \\ 1 & \text{if } q_\vartheta \geq N_O - 1. \end{cases} \tag{77}$$

A practical scheme for bit generation can then be constructed according to the steps of Sec. 2.2. Notice that since the target output is an m_{e}-bit long binary string, after calculating c_i, the r_l output value assignment can be realized by simply taking the most significant m_e bits of the binary representation of c_i.

Note that the previous results for upper bounds on estimation error and statistical distance hold for this practical bit generation scheme as long as $p_0 = (1 - \mathrm{e}^{-\lambda\tau}) \geq \mathrm{e}^{-\lambda(N_O-1)\tau} = p_{N_O-1}$, giving the criterion of $N_O \geq 1 - \ln\bigl(1 - \mathrm{e}^{-\lambda\tau}\bigr)/\lambda\tau$.

References

[1] F. N. David and N. L. Johnson, "The probability integral transformation when parameters are estimated from the sample," *Biometrika*, vol. 35, no. 1/2, p. 182, May 1948. [Online]. Available: http://dx.doi.org/10.2307/2332638