

# Some notes on Interrogating Random Quantum Circuits

Luís Brandão and René Peralta

Cryptographic Technology Group  
National Institute of Standards and Technology

Presentation at NIST/Google meeting  
December 13, 2019 @ NIST Gaithersburg, USA

# Outline

1. Introduction
2. Exponential model
3. Distinguishability
4. Min-entropy estimation
5. Concluding remarks

# Outline

1. Introduction
2. Exponential model
3. Distinguishability
4. Min-entropy estimation
5. Concluding remarks

## Goals of the presentation:

- ▶ Convey our preliminary understanding of the certifiable-QRNG setting
- ▶ Discuss distinguishability / parametrization aspects
- ▶ Identify questions for subsequent followup / research directions (?)

# Outline 1

1. Introduction
2. Exponential model
3. Distinguishability
4. Min-entropy estimation
5. Concluding remarks

# The protocol at a high-level

## **Towards certified/certifiable randomness.**

1. The operator is given a freshly chosen random quantum circuit.
2. Soon after, the operator publishes many circuit output strings.
3. Client extracts randomness for use in applications.
4. Long after, a supercomputer outputs the “P-values” of the strings.
5. By analysis of the “P-values”, get a retroactive statistical assurance that a sufficiently large set of outputs were sampled from the quantum circuit.

# The protocol at a high-level

## **Towards certified/certifiable randomness.**

1. The operator is given a freshly chosen random quantum circuit.
2. Soon after, the operator publishes many circuit output strings.
3. Client extracts randomness for use in applications.
4. Long after, a supercomputer outputs the “P-values” of the strings.
5. By analysis of the “P-values”, get a retroactive statistical assurance that a sufficiently large set of outputs were sampled from the quantum circuit.

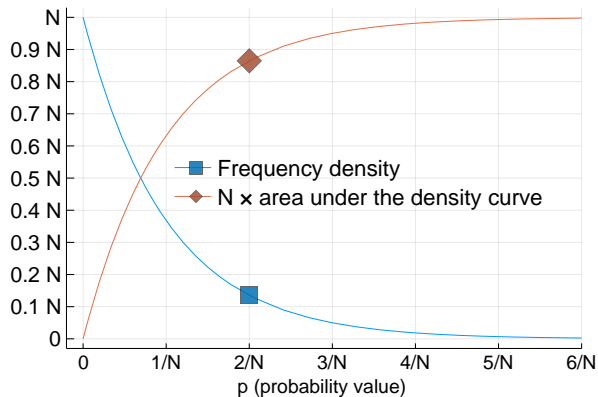
We want to look at suitable parameters for implementation of this protocol

# Outline 2

1. Introduction
2. Exponential model
3. Distinguishability
4. Min-entropy estimation
5. Concluding remarks

## Exponential model: frequency density

$f(p)$ : **Counting** the number of strings that, when sampling from a quantum random circuit, occur with each **probability** ( $p$ ).

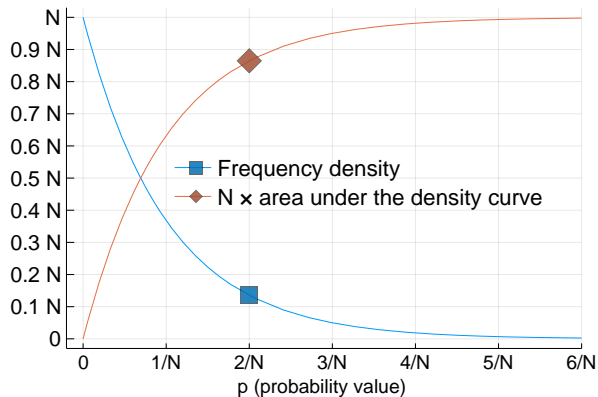


$$f(p) = N \cdot e^{-N \cdot p}$$



## Exponential model: frequency density

$f(p)$ : **Counting** the number of strings that, when sampling from a quantum random circuit, occur with each **probability** ( $p$ ).

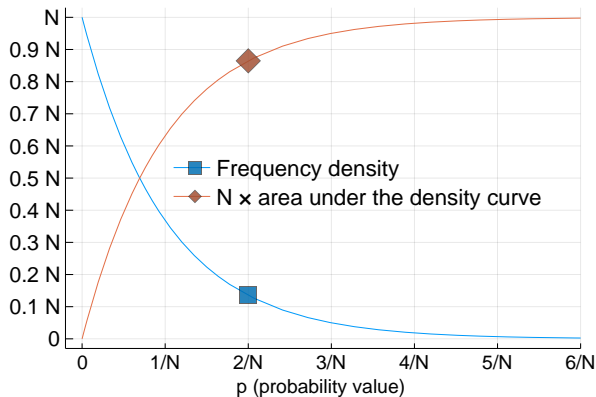


$$f(p) = N \cdot e^{-N \cdot p}$$

**Terminology:**  
we denote these  
particular  
probabilities ( $p$ )  
as “P-values”

## Exponential model: frequency density

$f(p)$ : **Counting** the number of strings that, when sampling from a quantum random circuit, occur with each **probability** ( $p$ ).



$$f(p) = N \cdot e^{-N \cdot p}$$

**Terminology:**  
we denote these particular probabilities ( $p$ ) as “P-values”

**Note:** the “frequency density” is a probability density (a continuous approximation) across the P-values, rather than across the strings.

## More on P-values

Once we obtain a freshly random quantum circuit  $C$ :

- ▶ Evaluating the circuit ( $s \leftarrow C$ ) is easy/fast with a quantum computer and super slow with a classical computer.
- ▶ There is a map  $P_{\text{val},C} : \{0, 1\}^n \rightarrow [0, 1[$ , where  $P_{\text{val},C}(s) = p$  means the string  $s$  has probability  $p$  of being output by an quantum-evaluation of  $C$
- ▶ Computing  $P_{\text{val}}(s)$  is very expensive for any  $s \in \{0, 1\}^n$
- ▶ A priori, without need to actually compute  $P_{\text{val},C}(\cdot)$ , the range  $\{P_{\text{val},C}(s) : s \in \{0, 1\}^n\}$  of P-values is assumed to be match the frequency characterization of function  $f = N \cdot e^{-N \cdot p}$ .

## More on P-values

Once we obtain a freshly random quantum circuit  $C$ :

- ▶ Evaluating the circuit ( $s \leftarrow C$ ) is easy/fast with a quantum computer and super slow with a classical computer.
- ▶ There is a map  $P_{\text{val},C} : \{0, 1\}^n \rightarrow [0, 1[$ , where  $P_{\text{val},C}(s) = p$  means the string  $s$  has probability  $p$  of being output by an quantum-evaluation of  $C$
- ▶ Computing  $P_{\text{val}}(s)$  is very expensive for any  $s \in \{0, 1\}^n$
- ▶ A priori, without need to actually compute  $P_{\text{val},C}(\cdot)$ , the range  $\{P_{\text{val},C}(s) : s \in \{0, 1\}^n\}$  of P-values is assumed to be match the frequency characterization of function  $f = N \cdot e^{-N \cdot p}$ .

This is a model — which this presentation simply assumes.

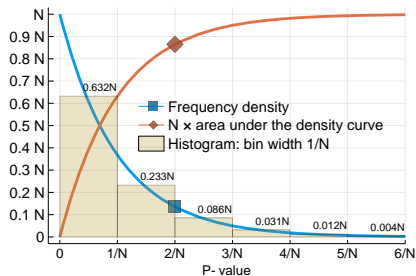
## Histogrammic perspective

**What is the probability that a sampled string has a P-value below  $x$ ?**

# Histogrammic perspective

What is the probability that a sampled string has a P-value below  $x$ ?

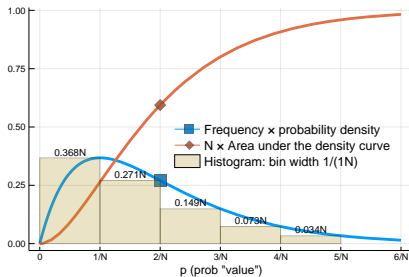
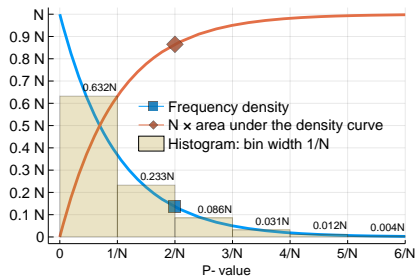
$x$	$1/N$	$2/N$	$3/N$	$4/N$
Upon uniform sampling	63%	95%	98%	99%



# Histogrammic perspective

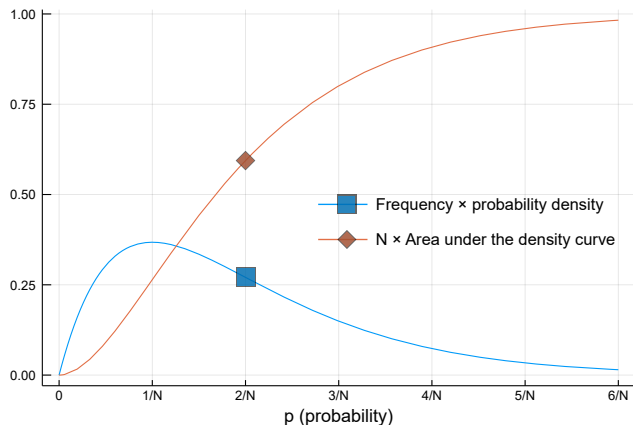
What is the probability that a sampled string has a P-value below  $x$ ?

$x$	$1/N$	$2/N$	$3/N$	$4/N$
Upon uniform sampling	63%	95%	98%	99%
Upon circuit evaluation	26%	59%	80%	91%



# Frequency times P-value

$f(p) \cdot p$ : Frequency times P-value as a function of P-value

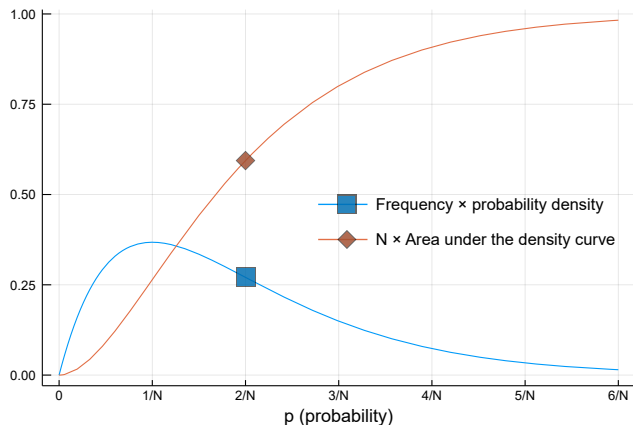


Useful to compute the probability with which each P-value occurs.



# Frequency times P-value

$f(p) \cdot p$ : Frequency times P-value as a function of P-value



$$E[X_Q] = 2/N$$

$$V[X_Q] = 2/N^2$$

Useful to compute the probability with which each P-value occurs.

# Fidelity

We are told that making a correct quantum evaluation is hard:

- ▶ Correct evaluation happens with probability  $\phi$
- ▶ Otherwise the output is uniform

Statistics for sum of P-value of  $m$  sampled strings:

Sampling type	Random variable $X_{*,m[*]}$	Expected value $E(X)$	Variance $V(X)$
<b>Uniform</b>	$X_{U,m}$	$m/N$	$m/N^2$
<b>Pure Quantum</b>	$X_{Q,m}$	$2 \cdot m/N$	$2 \cdot m/N^2$
<b>Q-Fidelity <math>\phi</math></b>	$X_{F,m,\phi}$	$(1 + \phi) \cdot m/N$	$(1 + \phi \cdot (2 - \phi)) \cdot m/N^2$

## Analyzing the empirical distribution of Q-values

We will want to compare obtained P-values vs. several distributions.

What kind of random variable  $X_{F,m,\phi}$  makes sense to analyze?

# Analyzing the empirical distribution of Q-values

We will want to compare obtained P-values vs. several distributions.

What kind of random variable  $X_{F,m,\phi}$  makes sense to analyze?

- ▶ Sum of obtained P-values
- ▶ Sum of the maximum  $k$  obtained P-values
- ▶ Kolmogorov-Smirnov of empirical distribution
- ▶ ...

## Analyzing the empirical distribution of Q-values

We will want to compare obtained P-values vs. several distributions.

What kind of random variable  $X_{F,m,\phi}$  makes sense to analyze?

- ▶ Sum of obtained P-values
- ▶ Sum of the maximum  $k$  obtained P-values
- ▶ Kolmogorov-Smirnov of empirical distribution
- ▶ ...

For simplicity we focus here on the “Sum of  $m$  obtained P-values”. Rationale:

- ▶ The  $E[X]$  is the *mean* times the *number of samples*
- ▶ We already know that  $\text{mean}_{\text{Honest}} > \text{mean}_{\text{uniform}}$
- ▶ Easy to approximate analytically (CLT), allowing faster simulations.

## Analyzing the empirical distribution of Q-values

We will want to compare obtained P-values vs. several distributions.

What kind of random variable  $X_{F,m,\phi}$  makes sense to analyze?

- ▶ Sum of obtained P-values
- ▶ Sum of the maximum  $k$  obtained P-values
- ▶ Kolmogorov-Smirnov of empirical distribution
- ▶ ...

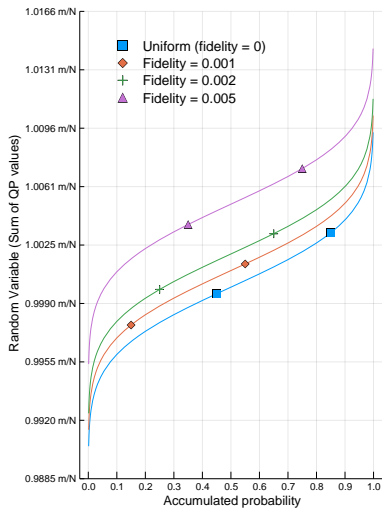
For simplicity we focus here on the “Sum of  $m$  obtained P-values”. Rationale:

- ▶ The  $E[X]$  is the *mean* times the *number of samples*
- ▶ We already know that  $\text{mean}_{\text{Honest}} > \text{mean}_{\text{uniform}}$
- ▶ Easy to approximate analytically (CLT), allowing faster simulations.

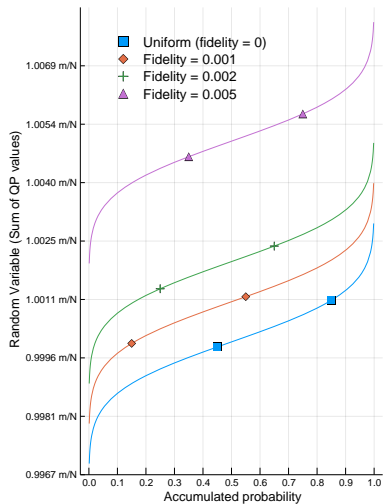
$$X_{F,m,\phi} \approx \mathcal{N} \left( \frac{(1+\phi) \cdot m}{N}, \frac{\sqrt{(1+\phi(2-\phi)) \cdot m}}{N} \right)$$

Curves for  $M = 10^5$  and  $M = 10^6$ 

Several string sampling experiments  
( $N=2^53$ ;  $M=10^5$ ;  $m/N=1.11022E-11$ )



Several string sampling experiments  
( $N=2^53$ ;  $M=10^6$ ;  $m/N=1.11022E-10$ )



# Outline 3

1. Introduction
2. Exponential model
3. Distinguishability
4. Min-entropy estimation
5. Concluding remarks



# Hypothesis testing

Some intro definitions:

- ▶ **False negative (FN):** reject when it is actually good (e.g., fid. 0.002)
- ▶ **False positive (FP):** accept when it is actually bad (e.g., uniform)

**Example:** If we have FN=20%, what do we get for FP?

It depends on the setup. In the last curves we had:

- ▶ If  $m = 10^5$ , then FP = 58.3%
- ▶ If  $m = 10^6$ , then FP = 12.4%

**Different FP:** We can formulate different definitions for FP, depending what we want to compare. For example, we can compare fidelity 0.002 (assumed honest) vs. 0.001 (the malicious case). This can be useful for entropy estimation. Then we would get

- ▶ If  $m = 10^5$ , then FP = 70.1%
- ▶ If  $m = 10^6$ , then FP = 12.4%

# What metrics for FN vs. FP?

Confusion matrix		Classification	
		Positive	Negative
Actual condition	Positive (Honest operator)	TP ratio	FN ratio
	Negative (Malicious operator)	FP ratio	TN ratio

accuracy =  $(TP + TN)/All$ ; precision =  $TP / (TP + FP)$ ; recall =  $TP / (TP + FN)$ ; ...

Is TN or TP more costly than the other? May depend on the application.

## What metrics for FN vs. FP?

Confusion matrix		Classification	
		Positive	Negative
Actual condition	Positive (Honest operator)	TP ratio	FN ratio
	Negative (Malicious operator)	FP ratio	TN ratio

accuracy = (TP + TN)/All; precision = TP / (TP + FP); recall = TP / (TP + FN); ...

Is TN or TP more costly than the other? May depend on the application.

- ▶ **Are FN's worse?** Can a FN, determined after the fact, impose rolling out / impugn some past legal procedure? E.g., assume the “randomness” was used to select a small sample of voting booths to recount votes in a tied election, leading to a tight win to one candidate. Will the procedure be contested if later the sample is rejected?
- ▶ **Are FP's worse?** A cryptographic application that hinges on fresh randomness for security. What if a completely deterministic (PRG) output is accepted, and the randomness provider is in cohorts with an adversary?

## Setting thresholds for FN and FP

- ▶ A la cryptographer: let  $\text{FN} = \text{FP} = 2^{-40}$  (common benchmark for “one-shot” security applications, e.g., cut-and-choose protocols)
- ▶ Different criteria for other applications (?)

## Setting thresholds for FN and FP

- ▶ A la cryptographer: let  $\text{FN} = \text{FP} = 2^{-40}$  (common benchmark for “one-shot” security applications, e.g., cut-and-choose protocols)
- ▶ Different criteria for other applications (?)

Let us look at some tables ...

# Table: Fixed FN ratios vs. FP ratios (using $\phi=0.002$ )

$M \in \{10^5, 10^6\}$ ,  $\phi = .002$ .

What is a “FP” depends on the comparison (e.g., consider “Uniform  $P_U$ ”)

$M$	$\phi$	FN ratio	Threshold	(Uniform)	(Fidelity)	(Fidelity)	(Fidelity)
		$p_\phi$	$T_{H,M,\phi}$	$p_U$	$p_{\phi/4}$	$p_{\phi/2}$	$p_{3\phi/4}$
$10^5$	0.002	$2^{-40}$	1.08765E-11	1.00000	1.00000	1.00000	1.00000
		$2^{-30}$	1.09130E-11	1.00000	1.00000	1.00000	1.00000
		$2^{-20}$	1.09569E-11	0.99998	0.99999	1.00000	1.00000
		0.001	1.10157E-11	0.99313	0.99561	0.99726	0.99833
		0.01	1.10426E-11	0.95530	0.96825	0.97793	0.98498
		0.1	1.10794E-11	0.74269	0.79085	0.83321	0.86956
		1/3	1.11093E-11	0.42040	0.48296	0.54587	0.60760
$10^6$	0.002	$2^{-40}$	1.10460E-10	1.00000	1.00000	1.00000	1.00000
		$2^{-30}$	1.10576E-10	0.99997	1.00000	1.00000	1.00000
		$2^{-20}$	1.10714E-10	0.99722	0.99946	0.99992	0.99999
		0.001	1.10901E-10	0.86355	0.94471	0.98188	0.99524
		0.01	1.10986E-10	0.62967	0.79689	0.90819	0.96624
		0.1	1.11102E-10	0.23703	0.41458	0.61173	0.78317
		1/3	1.11196E-10	0.05839	0.14279	0.28507	0.47277

Table: Fixed FN ratios vs. FP ratios (using  $\phi=0.005$ ) $M \in \{10^5, 10^6\}$ ,  $\phi = .005$ 

$M$	$\phi$	FN ratio $p_H$	Threshold $T_{H,M,\phi}$	(Uniform) $p_U$	(Fidelity) $p_{\phi/4}$	(Fidelity) $p_{\phi/2}$	(Fidelity) $p_{3\phi/4}$
$10^5$	0.005	$2^{-40}$	1.09091E-11	1.00000	1.00000	1.00000	1.00000
		$2^{-30}$	1.09457E-11	1.00000	1.00000	1.00000	1.00000
		$2^{-20}$	1.09897E-11	0.99933	0.99984	0.99997	0.99999
		0.001	1.10487E-11	0.93630	0.97240	0.98954	0.99654
		0.01	1.10757E-11	0.77541	0.87506	0.93865	0.97353
		0.1	1.11125E-11	0.38468	0.54060	0.69010	0.81308
		1/3	1.11425E-11	0.12543	0.22601	0.36062	0.51494
$10^6$	0.005	$2^{-40}$	1.10791E-10	0.98136	0.99956	1.00000	1.00000
		$2^{-30}$	1.10907E-10	0.85066	0.98888	0.99979	1.00000
		$2^{-20}$	1.11046E-10	0.41555	0.84976	0.98873	0.99979
		0.001	1.11233E-10	0.02909	0.25992	0.72711	0.96775
		0.01	1.11318E-10	0.00388	0.07922	0.43578	0.86079
		0.1	1.11434E-10	0.00010	0.00697	0.11332	0.51507
		1/3	1.11529E-10	0.00000	0.00046	0.01960	0.20780

## Table: Fixed FN ratios vs. FP ratios (higher fidelity)

$$M = 10^4, \phi \in \{.05, .1\}$$

$M$	$\phi$	FN ratio	Threshold	(Uniform)	(Fidelity)	(Fidelity)	(Fidelity)
		$p_H$	$T_{H,M,\phi}$	$p_U$	$p_{\phi/4}$	$p_{\phi/2}$	$p_{3\phi/4}$
$10^4$	0.05	$2^{-40}$	1.08376E-12	0.99142	0.99983	1.00000	1.00000
		$2^{-30}$	1.09584E-12	0.90243	0.99404	0.99989	1.00000
		$2^{-20}$	1.11034E-12	0.49593	0.88965	0.99246	0.99985
		0.001	1.12979E-12	0.03898	0.30630	0.76418	0.97245
		0.01	1.13868E-12	0.00519	0.09734	0.47553	0.87404
		0.1	1.15083E-12	0.00013	0.00870	0.12927	0.53560
		1/3	1.16072E-12	0.00000	0.00056	0.02275	0.22038
$10^4$	0.1	$2^{-40}$	1.13589E-12	0.01039	0.57286	0.99486	1.00000
		$2^{-30}$	1.14847E-12	0.00029	0.17824	0.93119	0.99992
		$2^{-20}$	1.16356E-12	0.00000	0.01225	0.57414	0.99413
		0.001	1.18382E-12	0.00000	0.00003	0.05998	0.79225
		0.01	1.19307E-12	0.00000	0.00000	0.00938	0.51407
		0.1	1.20572E-12	0.00000	0.00000	0.00029	0.15147
		1/3	1.21603E-12	0.00000	0.00000	0.00001	0.02886



## Other random variables

Once all P-values are assessed, what is the best strategy for confirmation?

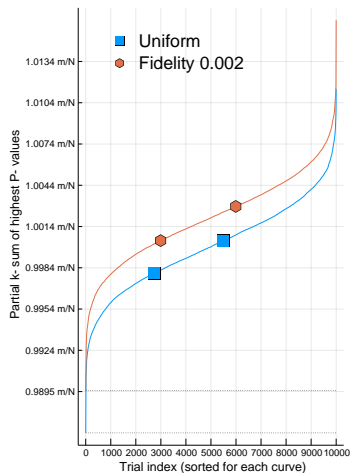
**Example:** Client has a “small” budget to verify P-values, e.g., 10% of them. How should they be chosen?

- ▶ Uniformly?
- ▶ the 10% highest?
- ▶ Sampling related to the  $f$  distribution?
- ▶ Something else?

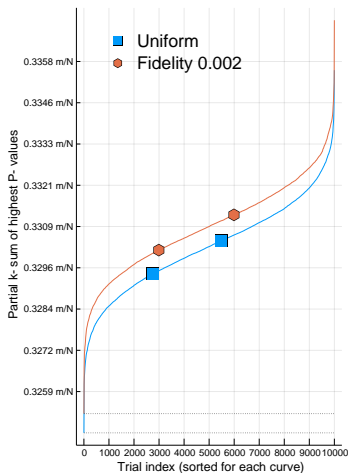
# Example: partial sum of the highest 10% P-values

Using  $M = 10^5$ , compare the cases  $k = 10^5$  vs.  $k = 10^4$

Several string sampling experiments  
( $N=2^{53}$ ;  $M=10^5$ ;  $k=10^5$ ;  $m/N=1.11022$ )



Several string sampling experiments  
( $N=2^{53}$ ;  $M=10^5$ ;  $k=10^4$ ;  $m/N=1.11022$ )



## Table: comparing some FP ratios for the same FN

### Example:

- ▶ Positive case: honest circuit evaluation with fidelity  $\phi = 0.002$ .
- ▶ Negative case: uniform string sampling.

$M$	$k$	$k/M$	(FN = 0.25) FP	(FN = 0.1) FP
$10^6$	$10^3$	.001	0.64	0.82
	$10^4$	.01	0.45	0.68
	$10^5$	.1	0.21	0.41
$10^5$	$10^3$	.01	0.69	0.86
	$10^4$	.1	0.59	0.79
	$10^5$	1	0.50	0.74

(Each curve based on simulation of  $10^4$  trials of partial-sums)

**Observation:** for fixed  $k$  and  $\phi$ , higher  $M$  leads to better results.

# Outline 4

1. Introduction
2. Exponential model
3. Distinguishability
- 4. Min-entropy estimation**
5. Concluding remarks

# Entropy needs / assumptions

**Assume a correct experiment execution with a honest operator:**

- ▶  $(n \text{ qubits, \# samples, fidelity } \phi) = (n, M, \phi) = (53, 10^5, 0.002)$
- ▶ Let  $H_Q$  be the entropy of a circuit generated string.
- ▶ Let  $q = M \cdot \phi$ , e.g.,  $(M, \phi) = (10^5, 0.002) \rightarrow q = 200$

Then entropy  $\approx (M - q) \cdot 2^n + q \cdot H_Q \approx 5 \times 10^6$  bits

# Entropy needs / assumptions

## Assume a correct experiment execution with a honest operator:

- ▶  $(n \text{ qubits, \# samples, fidelity } \phi) = (n, M, \phi) = (53, 10^5, 0.002)$
- ▶ Let  $H_Q$  be the entropy of a circuit generated string.
- ▶ Let  $q = M \cdot \phi$ , e.g.,  $(M, \phi) = (10^5, 0.002) \rightarrow q = 200$

Then entropy  $\approx (M - q) \cdot 2^n + q \cdot H_Q \approx 5 \times 10^6$  bits

- ▶ **Pre-sampling (sample size question):**

- ▶ **Post-sampling (min-entropy question):**

# Entropy needs / assumptions

**Assume a correct experiment execution with a honest operator:**

- ▶  $(n \text{ qubits, \# samples, fidelity } \phi) = (n, M, \phi) = (53, 10^5, 0.002)$
- ▶ Let  $H_Q$  be the entropy of a circuit generated string.
- ▶ Let  $q = M \cdot \phi$ , e.g.,  $(M, \phi) = (10^5, 0.002) \rightarrow q = 200$

Then entropy  $\approx (M - q) \cdot 2^n + q \cdot H_Q \approx 5 \times 10^6$  bits

- ▶ **Pre-sampling (sample size question):** Given FN ratio and FP ratio needed by my application, how many ( $M$ ) strings do I need to collect from a fidelity- $\phi$  experiment to get something useful (enable a high enough lower-bound on entropy)?
  
- ▶ **Post-sampling (min-entropy question):**

# Entropy needs / assumptions

## Assume a correct experiment execution with a honest operator:

- ▶  $(n \text{ qubits, \# samples, fidelity } \phi) = (n, M, \phi) = (53, 10^5, 0.002)$
- ▶ Let  $H_Q$  be the entropy of a circuit generated string.
- ▶ Let  $q = M \cdot \phi$ , e.g.,  $(M, \phi) = (10^5, 0.002) \rightarrow q = 200$

Then entropy  $\approx (M - q) \cdot 2^n + q \cdot H_Q \approx 5 \times 10^6$  bits

- ▶ **Pre-sampling (sample size question):** Given FN ratio and FP ratio needed by my application, how many ( $M$ ) strings do I need to collect from a fidelity- $\phi$  experiment to get something useful (enable a high enough lower-bound on entropy)?
- ▶ **Post-sampling (min-entropy question):** Given a list of P-values, measured for some set of strings,\* what is the highest min-entropy that we should estimate, under an adversarial scenario, with assurance  $p$ ?



# Conceivable attacks

## Setup:

- ▶ Quantum computer operator: advertises  $\phi$
- ▶ Client: chooses  $FP < \epsilon$ ,  $FN < \epsilon'$  (Negative means uniform).

# Conceivable attacks

## Setup:

- ▶ Quantum computer operator: advertises  $\phi$
- ▶ Client: chooses FP  $< \epsilon$ , FN  $< \epsilon'$  (Negative means uniform).

## Attack 0 (repeated strings):

- ▶ Select single string from circuit evaluation ( $E[X] = 2/N$ )
- ▶ Repeat the same string  $M$  times ... High probability of acceptance

# Conceivable attacks

## Setup:

- ▶ Quantum computer operator: advertises  $\phi$
- ▶ Client: chooses FP  $< \epsilon$ , FN  $< \epsilon'$  (Negative means uniform).

## Attack 0 (repeated strings):

- ▶ Select single string from circuit evaluation ( $E[X] = 2/N$ )
- ▶ Repeat the same string  $M$  times ... High probability of acceptance

Trivial fix: disallow repeated strings.

# Conceivable attacks

## Setup:

- ▶ Quantum computer operator: advertises  $\phi$
- ▶ Client: chooses FP  $< \epsilon$ , FN  $< \epsilon'$  (Negative means uniform).

## Attack 0 (repeated strings):

- ▶ Select single string from circuit evaluation ( $E[X] = 2/N$ )
- ▶ Repeat the same string  $M$  times ... High probability of acceptance

Trivial fix: disallow repeated strings.

## Attack 1 (full PRG generation):

- ▶ If FP is reasonable high (e.g., 0.1):
- ▶ Operator PRG-generates all  $M = 10^5$  strings and hopes to be lucky.

# Conceivable attacks

## Setup:

- ▶ Quantum computer operator: advertises  $\phi$
- ▶ Client: chooses FP  $< \epsilon$ , FN  $< \epsilon'$  (Negative means uniform).

## Attack 0 (repeated strings):

- ▶ Select single string from circuit evaluation ( $E[X] = 2/N$ )
- ▶ Repeat the same string  $M$  times ... High probability of acceptance

Trivial fix: disallow repeated strings.

## Attack 1 (full PRG generation):

- ▶ If FP is reasonable high (e.g., 0.1):
- ▶ Operator PRG-generates all  $M = 10^5$  strings and hopes to be lucky.

Conclusion: entropy = 0

# Conceivable attacks

## Setup:

- ▶ Quantum computer operator: advertises  $\phi$
- ▶ Client: chooses  $FP < \epsilon$ ,  $FN < \epsilon'$  (Negative means uniform).

## Attack 0 (repeated strings):

- ▶ Select single string from circuit evaluation ( $E[X] = 2/N$ )
- ▶ Repeat the same string  $M$  times ... High probability of acceptance

Trivial fix: disallow repeated strings.

## Attack 1 (full PRG generation):

- ▶ If FP is reasonable high (e.g., 0.1):
- ▶ Operator PRG-generates all  $M = 10^5$  strings and hopes to be lucky.

Conclusion: entropy = 0 ... but attack does not work if  $FP_U$  is very small

# Conceivable attacks

## Attack 2 (higher fidelity):

1. Operator has a fidelity 1 computer, but claims to only have fidelity .05.
2. PRG-compute  $M' = M \cdot (1 - \phi/2)$  strings (P-values distributed as  $X_{U,M'}$ )
3. Circuit-evaluate  $\phi/2$  strings

# Conceivable attacks

## Attack 2 (higher fidelity):

1. Operator has a fidelity 1 computer, but claims to only have fidelity .05.
2. PRG-compute  $M' = M \cdot (1 - \phi/2)$  strings (P-values distributed as  $X_{U,M'}$ )
3. Circuit-evaluate  $\phi/2$  strings

Conclusion: entropy =  $M \cdot \phi/2 \cdot H_Q$ , e.g.,  $10^5 \cdot 0.002/2 \cdot 52? = 5200$



# Conceivable attacks

## Attack 2 (higher fidelity):

1. Operator has a fidelity 1 computer, but claims to only have fidelity .05.
2. PRG-compute  $M' = M \cdot (1 - \phi/2)$  strings (P-values distributed as  $X_{U,M'}$ )
3. Circuit-evaluate  $\phi/2$  strings

Conclusion: entropy =  $M \cdot \phi/2 \cdot H_Q$ , e.g.,  $10^5 \cdot 0.002/2 \cdot 52? = 5200$

## Attack 3 (use lower fidelity):

- ▶ Change the FP — another Negative condition (Uniform  $\rightarrow$  half fidelity)
- ▶ Example:  $(\phi, FN) = (0.05, 0.1) \Rightarrow FP_U = 0.0013$ , but  $FP_{\phi/2} = 0.129 \approx 1/8$
- ▶ Attackers try their luck ( $\approx 1/8$  chance of winning) using half entropy.

# Conceivable attacks

## Conceivable attacks

### Attack 4 (post-sampling choice — in complement to attacks 2 and 3):

1. Operator PRG-generates  $M - q$  strings (0 entropy), e.g., with  $q = 100$
2. With fidelity 1, privately evaluate circuit about  $2^{25} \cdot q$  times
3. Choose  $q$  strings whose first 25 bits are zero after some transformation

## Conceivable attacks

### Attack 4 (post-sampling choice — in complement to attacks 2 and 3):

1. Operator PRG-generates  $M - q$  strings (0 entropy), e.g., with  $q = 100$
2. With fidelity 1, privately evaluate circuit about  $2^{25} \cdot q$  times
3. Choose  $q$  strings whose first 25 bits are zero after some transformation

Entropy:  $\approx q \cdot (H_Q - 25) \approx 100 \cdot 27 \approx 2700$

(more subtleties are needed, e.g., PR order of strings ...)

### To-do:

- ▶ Play with concrete parameters, get concrete results.
- ▶ Application appropriate parameters
- ▶ If you trust PRGS, why would you need thousands of bits?

# Outline 5

1. Introduction
2. Exponential model
3. Distinguishability
4. Min-entropy estimation
5. Concluding remarks

## Some questions worth exploring:

- ▶ Suitable (FN,FP) threshold for conceivable applications?
- ▶ Verification budget of P-Values for the user? (and oracle budget)
- ▶ What are the best statistics to measure? Full-sum, partial-sum, KS, ...?
- ▶ Application motivation: when are more than 512 random bits actually needed at once?
- ▶ Security proofs
- ▶ Research problem: (efficiently-verifiable) probabilistic checkable proofs (PCPs) for this problem

## Some questions worth exploring:

- ▶ Suitable (FN,FP) threshold for conceivable applications?
- ▶ Verification budget of P-Values for the user? (and oracle budget)
- ▶ What are the best statistics to measure? Full-sum, partial-sum, KS, ...?
- ▶ Application motivation: when are more than 512 random bits actually needed at once?
- ▶ Security proofs
- ▶ Research problem: (efficiently-verifiable) probabilistic checkable proofs (PCPs) for this problem

Overall this field has interesting challenges

Engaging in this has a potential to foster the understanding of applications of quantum randomness.

# A major caveat

There is a major caveat in our analysis!





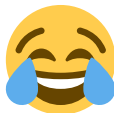
## A major caveat

There is a major caveat in our analysis!



Our simulations used classical randomness!

Would we get better results with quantum randomness?



- ▶ NISTIR 8213: <https://doi.org/10.6028/NIST.IR.8213-draft>
- ▶ Beacon project: <https://csrc.nist.gov/Projects/Interoperable-Randomness-Beacons>

# Thank you

- ▶ NISTIR 8213: <https://doi.org/10.6028/NIST.IR.8213-draft>
- ▶ Beacon project: <https://csrc.nist.gov/Projects/Interoperable-Randomness-Beacons>

## Some notes on Interrogating Random Quantum Circuits

luis.brandao@nist.gov; rene.peralta@nist.gov

Presentation at NIST/Google meeting  
December 13, 2019 @ NIST Gaithersburg, USA

**Disclaimer.** Opinions expressed in this presentation are from the author(s) and are not to be construed as official or as views of the U.S. Department of Commerce. The identification of any commercial product or trade names in this presentation does not imply endorsement or recommendation by NIST, nor is it intended to imply that the material or equipment identified are necessarily the best available for the purpose.

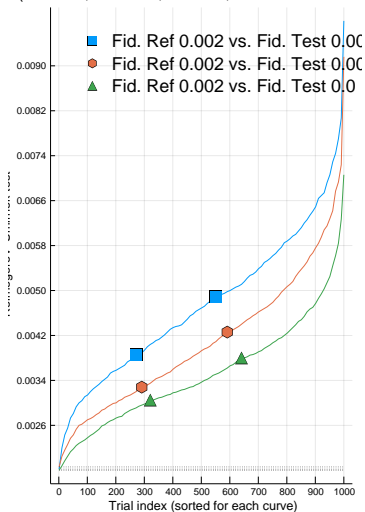
**Disclaimer.** Some external-source images and cliparts were included/adapted in this presentation with the expectation of such use constituting licensed and/or fair use.

# Using Kolmogorov-Smirnov

This slide and the next are tentative.

Results obtained this morning ... requires further sanity check.

Several string sampling experiments  
 ( $N=2^{53}$ ;  $M=10^5$ ;  $k=10^3$ ;  $m/N=1.00000E+$ )



## Table: Fixed FN ratios vs. FP ratios (higher fidelity)

$$M = 10^4, \phi \in \{.05, .1\}$$

$M$	$\phi$	FN ratio $p_H$	Threshold $T_{H,M,\phi}$	(Uniform) $p_U$	(Fidelity) $p_{\phi/2}$
$10^5$	0.002	$2^{-20}$	1.85008E-03	0.99600	1.00000
		0.001	1.92992E-03	0.98800	0.99700
		0.01	2.22990E-03	0.94600	0.97600
		0.1	3.16900E-03	0.62100	0.75400
		2/3	5.28000E-03	0.05000	0.16900

# List of slides

1. Promoting Public Randomness ...
2. Outline
3. Outline 1
4. The protocol at a high-level
5. Outline 2
6. Exponential model: frequency density
7. More on P-values
8. Histogramic perspective
9. Frequency times P-value
10. Fidelity
11. Analyzing the empirical distribution of Q-values
12. Curves for  $M = 10^5$  and  $M = 10^6$
13. Outline 3
14. Hypothesis testing
15. What metrics for FN vs. FP?
16. Setting thresholds for FN and FP
17. Table: Fixed FN ratios vs. FP ratios (using  $\phi=0.002$ )
18. Table: Fixed FN ratios vs. FP ratios (using  $\phi=0.005$ )
19. Table: Fixed FN ratios vs. FP ratios (higher fidelity)
20. Other random variables
21. Example: partial sum of the highest 10% P-values
22. Table: comparing some FP ratios for the same FN
23. Outline 4
24. Entropy needs / assumptions
25. Conceivable attacks
26. Conceivable attacks
27. Conceivable attacks
28. Outline 5
29. Some questions worth deepening:
30. A major caveat
31. Thank you
32. Using Kolmogorov-Smirnov
33. Table: Fixed FN ratios vs. FP ratios (higher fidelity)
34. List of slides