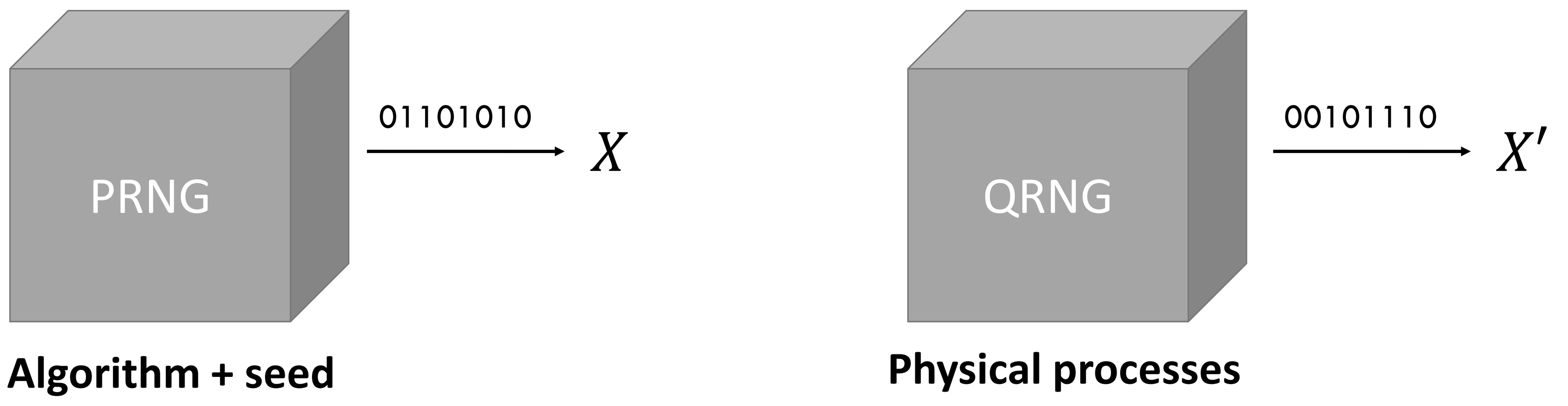
Certification of Random Number Generators using Machine Learning

Hong Jie Ng[1], Raymond Ho[1], Syed Assad[2], Ping Koy Lam[2], Omid Kavehei[3,4], Chao Wang[1], Nhan Duy Truong[3,4], Jing Yan Haw[1]

[1]Department of Electrical & Computer Engineering, National University of Singapore
[2]CQC2T, Department of Quantum Science, The Australian National University
[3]Australian Research Council Training Centre for Innovative BioEngineering, The University of Sydney
[4]NeuroSyd Lab, School of Biomedical Engineering, Faculty of Engineering, The University of Sydney

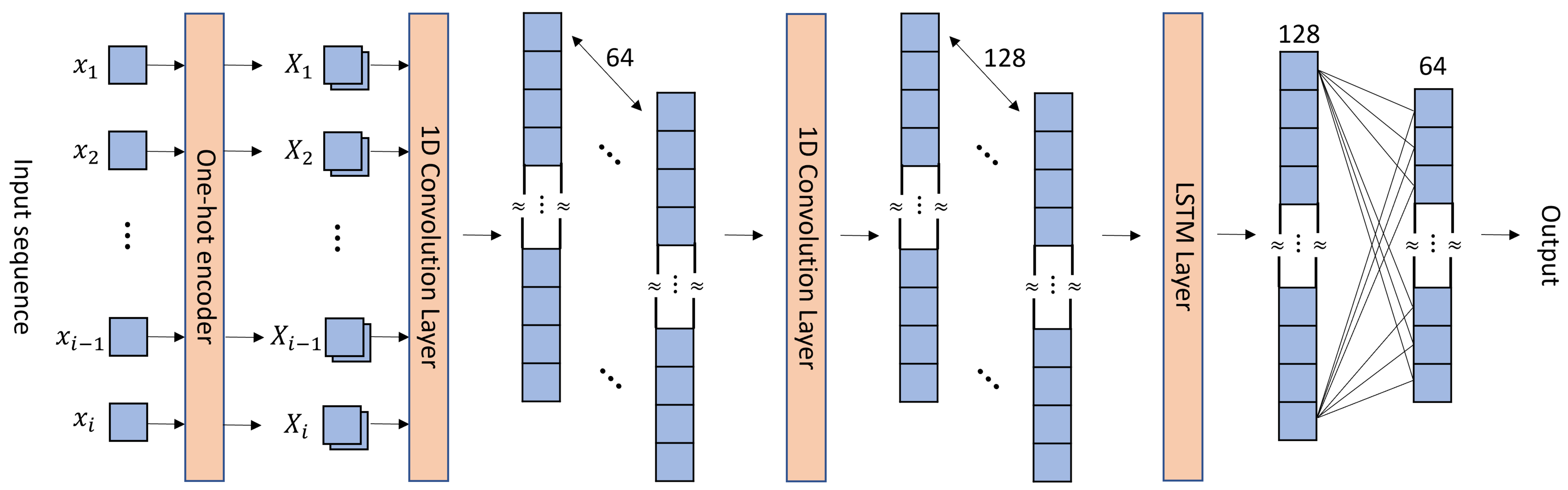
Introduction



Given X, the output from a pseudo-RNG (PRNG), and X', the output from a Quantum-RNG, is there a way to differentiate X from X' without the use of a fixed set of test statistics derived from the samples?

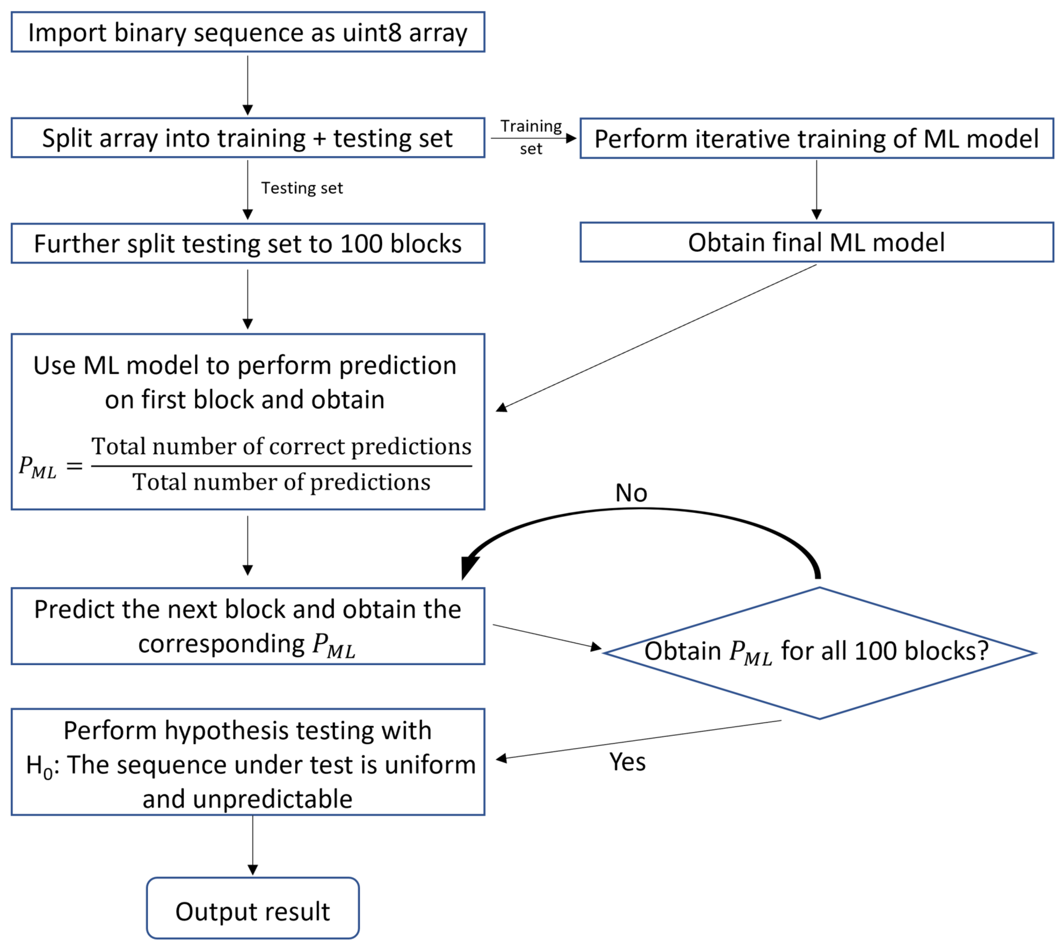
We investigate the use of **machine learning** (ML) as a potential tool for this purpose.

ML model [1]



Recurrent Convolutional Neural Network + Long-Short Term Memory (LSTM)

ML Flow



Testing of PRNGs

Linear Congruential Generator $x_{n+1} = ax_n + b \bmod P$

RNG Parameters	$P = 2^{24}$	$P = 2^{26}$	$P = 2^{28}$	$P = 2^{30}$
Result	Reject H_0	Reject H_0	Reject H_0	Pass

Inversive Congruential Generator $y_{n+1} = cy_n^{-1} + d \bmod P$

RNG Parameters	$P = 2^{17} - 1$	$P = 2^{19} - 1$	$P = 2^{23} - 595$	$P = 2^{24} - 75$
Result	Reject H_0	Reject H_0	Reject H_0	Pass

Mersenne Twister (MT19937)

RNG Parameters	Default
Result	Pass

Linear Feedback Shift Register

RNG Parameters	State size = 24	State size = 28	State size = 32
Result	Reject H_0	Pass	Pass

We observe that our ML model is able to pinpoint the deviations from randomness that is present in PRNGs for the cases where the period is relatively small.

Uniformity & Correlation Test*

Biased RNG $P(Z = 1) = 1/2 + \varepsilon$

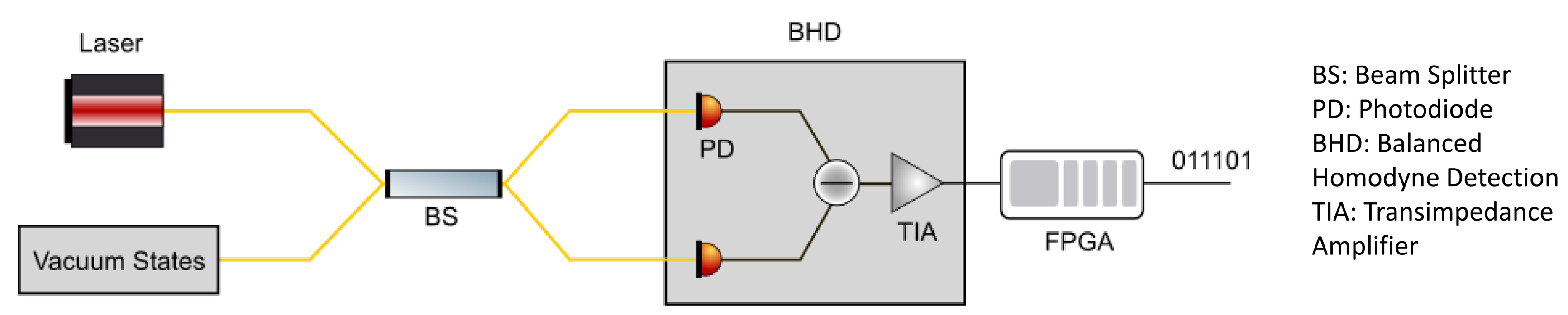
RNG Parameters	$\varepsilon \geq 0.01$	$\varepsilon = 0.0015$	$\varepsilon = 0.001$
Result	Reject H_0	Reject H_0	Pass

RNG with correlation r = Pearson r correlation

RNG Parameters	$r = 0.18323$	$r = 0.13739$	$r = 0.10970$
Result	Reject H_0	Reject H_0	Pass

* Generated from PRNGs

Testing of a QRNG



QRNG based on quantum vacuum states [2] α = Level of significance

RNG Parameters	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2.5\%$	$\alpha = 1\%$
Result	Pass	Pass	Pass	Pass

Conclusion

- In our experiment, we have used 960 Mbits of data for each RNG to train and test our ML model. With a PC setup with 32 GB of RAM and an Nvidia Quadpro P400 GPU processor, the total run time is around 3 hours (training 2 hours, testing 1 hour).
- Provided the training data is sufficiently large, our ML model is sensitive to imperfect randomness (deterministic sequence, bias & correlation).
- Compared to the NIST statistical test suite [3] and Dieharder [4], ML-based approach can evaluate the quality of the randomness using only a single model.

[1] N. D. Truong, J. Y. Haw, S. M. Assad, P. K. Lam, and O. Kavehei, 'Machine Learning Cryptanalysis of a Quantum Random Number Generator', *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 2, pp. 403–414, Feb. 2019.
[2] J. Y. Haw *et al.*, "Maximization of Extractable Randomness in a Quantum Random-Number Generator," *Phys. Rev. Applied*, vol. 3, no. 5, p. 054004, May 2015
[3] A. Rukhin *et al.*, 'NIST 800-22 A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications', Apr. 2010.
[4] R. G. Brown, D. Eddelbuettel, and D. Bauer, "dieharder: A random number test suite," URL https://webhome.phy.duke.edu/~rgb/General/dieharder.php. C program archive dieharder, version, vol. 2, no. 3, 2020

Contact: hongjie@u.nus.edu
Quantum Communications Lab Group Website: https://www.ccwlim.com/