Université du Québec
Institut National de la Recherche Scientifique
Énergie, Matériaux, Télécommunications

# A Scalable Design for Photonic Quantum Random Number Generation

Par

Shashwath Shankar Bharadwaj

Mémoire ou thèse présenté(e) pour l’obtention du grade de

Maître ès Sciences (M.Sc.)

en sciences de l'énergie et des matériaux

**Jury d’évaluation**

| | |
|---|---|
| Président du jury et examinateur interne | Long Le<br>Institut National de la Recherche Scientifique |
| Examinateur externe | Li Qian<br>University of Toronto |
| Directeur de recherche | Roberto Morandotti<br>Institut National de la Recherche Scientifique |

# ACKNOWLEDGEMENTS

Conducting cutting-edge experimental physics research has been a personal dream of mine since my high school days. Over the last three years, I have had the opportunity to fulfil this dream through my MSc studies, while working on a subject that holds my fascination both at a scientific and a philosophical level. In reconciling my experiences, I feel grateful for the contributions and influences of many individuals, without which undertaking this journey would not have been possible.

First and foremost, I would like to thank my supervisor Prof. Roberto Morandotti for giving me an opportunity to conduct my research in his lab at INRS. Without his support, both in the form of academic resources and financial help, the career transition from computer science to quantum physics would have been an arduous challenge. I am especially grateful for his encouragement to pursue new and challenging research ideas and also undertake leadership activities through the student chapters, which has been an incredibly fulfilling experience for me.

I thank Mitacs for enabling my move from India to Canada, first through the Globalink Research Internship and later through the Globalink Graduate fellowship. The academic experience and the financial support I received have been life changing.

I would like to acknowledge the support from my colleagues at the Nonlinear Photonics Group, in the form of technical guidance, moral support and with writing papers for conferences, journals as well as patents. My interactions with them have ensured that I gained a well-rounded perspective of life as a researcher. In particular, I thank Prof. James van Howe for his immense and extensive support at both a professional and personal level. His timely guidance ensured that I could scale the learning curve associated with the thesis project and his advice continues to shape my scientific thinking. I thank Dr. Simone Atzeni, for his meticulous help especially with the experimental side of the project, including the setting up of the single photon source. His guidance in the lab sped up the completion of this project by many months. I would like to thank Piotr Roztocki for numerous discussions and help at all stages of this project, right from its inception to the drafting of various papers. His technical insights on conducting quantum experiments and analyzing data helped me develop an intuition for quantum phenomena. I thank Renuka Narayanan for her help with developing the Minimum Entropy equations presented in this thesis, collecting, and analyzing experimental data including the implementation of postprocessing algorithms. I am grateful for the inputs from Prof. William J Munro in completing and submitting

the journal manuscript resulting from this work. I thank Nicolas Perron for his help with drafting the French summary of this thesis.

Special thanks to Robin Helsten, Dr. Mario Chemnitz, Dr. Stefania Sciara, Giacomo Balistreri, Agnes George, Benjamin Maclellan, Bennet Fischer, Young Zhang and Yoann Jestin for their valuable inputs and assistance at various stages of my MSc. My experience at INRS has been memorable thanks to these interactions.

I thank Dr. Matthew Posner, Meg All, Dirk Fabian from SPIE for providing numerous leadership opportunities through the student chapter programme.

Finally, I would like to thank my family and friends, both in India and Canada, who have been a pillar of support during three very tumultuous years of my life. Moving continents, switching careers, and starting from scratch has not been easy and I would not have been able to accomplish this without their continuous encouragement and support. Specifically, I thank my parents Shankar Billemane and Namitha Shankar, my girlfriend Vasuretha Chandar, and my Canadian family Robyn and Aso Urum, for always being present when I needed their help. I am blessed to receive their unconditional love and support.

# RÉSUMÉ

Le caractère aléatoire est l'une des propriétés fondamentales de la nature. Il est largement répandu dans de nombreux aspects de la vie quotidienne. Depuis l'Antiquité, plusieurs méthodes de génération de nombres aléatoires (représentant la source de l'indéterminisme) ont été explorées, prenons par exemple le premier générateur de nombres aléatoires, c'est-à-dire un dé, qui a été créé il y a des milliers d'années. Au cours des dernières années, des générateurs plus sophistiqués ont été développés en raison des exigences complexes des applications modernes telles qu'en cryptographie, en communications et en simulation. Poussés par les progrès des industries des semi-conducteurs et de l'informatique, de nombreux dispositifs physiques ont été créés au cours des 70 dernières années. Il est maintenant possible de produire des nombres aléatoires pouvant correspondre aux vitesses de fonctionnement des ordinateurs. Les générateurs aléatoires classiques sont largement utilisés dans les applications scientifiques et technologiques de pointe, ainsi que dans des domaines tels que le sport et la loterie. Bien que les performances de ces dispositifs classiques soient suffisantes pour les besoins actuels, l'émergence rapide des technologies quantiques menace sérieusement leurs fonctionnalités ainsi que l'intégrité des protocoles utilisés. En exploitant un paradigme informatique fondamentalement différent, les systèmes quantiques peuvent nuire au fonctionnement de nombreux protocoles machines et logiciels tels que les générateurs de nombres aléatoires classiques, l'algorithme Rivest-Shamir-Adleman, les méthodes cryptographiques, etc. Ces protocoles sont d'une grande importance dans le secteur des banques, des finances et de la défense. Afin de faire face à cette menace, il serait possible de concevoir et de déployer des générateurs de nombres aléatoires quantiques (QRNG) véritablement indéterministes. Ces systèmes quantiques ont été étudiés comme générateurs de nombres aléatoires pendant de nombreuses décennies. Alors que les études initiales utilisaient des systèmes radioactifs afin d'explorer théoriquement le caractère aléatoire, des études récentes se sont concentrées sur les plateformes optiques et photoniques en raison de leurs nombreux avantages. Elles offrent, en outre, de multiples degrés de liberté, une facilité d'intégration et d'utilisation en communications, etc. De récentes études telles que « Quantum Supremacy » et « Quantum Advantage » effectuées par des sociétés comme Google ont mis en valeur la puissance potentielle des machines quantiques, signalant le besoin urgent des QRNGs, dont la sécurité a déjà été démontrée, et qui peuvent être utilisés et déployés dans diverses applications pratiques.

Les recherches actuelles ont tenté d'optimiser deux paramètres de performance, à savoir le débit (vitesse de l'appareil) et la qualité du caractère aléatoire réel (sécurité de l'appareil). Par

conséquent, les méthodes actuelles peuvent être classées en trois catégories, soit a) les configurations d'appareils de confiance qui obtiennent un débit binaire rapide, mais une qualité aléatoire médiocre, b) les configurations dites autotestées qui obtiennent une bonne qualité, mais des débits extrêmement lents et c) les appareils semi-autotestés qui maintiennent un compromis entre ces deux paramètres. Par conséquent, une recherche active est en cours pour réaliser un générateur idéal pouvant fournir à la fois un débit binaire et un degré de confiance élevé. Cependant, les méthodes actuelles pour améliorer les QRNGs nécessitent généralement une modification complète du système ainsi que le remplacement de composants couteux. Ces problèmes découlent des deux stratégies principalement utilisées lors de la conception. D'une part, il est possible d'augmenter la dimensionnalité de la mesure, afin d'améliorer le nombre de bits aléatoires possibles par photon. D'autre part, il est aussi possible d'utiliser des composants clés ayant une meilleure performance tels que des sources ayant une probabilité plus faible de générer des états multiphotons ou des détecteurs ayant une meilleure résolution et des temps morts plus courts. Toutefois, cette dernière option est difficile à réaliser en raison des limitations des sciences des matériaux, des techniques de fabrication et de la compatibilité des différents composants. De plus, étant donné que ces designs sont très différents dans leur conception, il est difficile de les comparer et donc de déterminer lequel offre de meilleures performances dans une application spécifique.

Afin de résoudre ces problèmes, à travers les travaux menés dans cette thèse, nous introduisons une entropie d'information minimale par bit comme paramètre de conception important pour les QRNGs. Cette mesure sert de métrique universelle à travers la conception de QRNGs pour déterminer la quantité maximale et pratique pour laquelle le caractère aléatoire d'un appareil donné peut être utilisé pour générer des bits. Nous implémentons ensuite un QRNG de type temporel pour lequel l'efficacité de la génération de bits aléatoires peut être augmentée en plaçant en cascade un nombre croissant de diviseurs de faisceau fibrés, le tout en augmentant de manière minimale la complexité de la configuration. En utilisant comme signal un laser atténué et une source de photons intriqués, nous montrons expérimentalement que le caractère aléatoire peut être mis à l'échelle, conformément à la théorie. En fonction de la qualité des sources et des détecteurs, nous montrons que le caractère aléatoire peut être augmenté de 3% à 30%. De plus, dans notre montage, le laser atténué montre une entropie minimale par bit beaucoup plus grande par rapport à notre source de photons intriqués. Cela suggère qu'il existe un avantage potentiel à utiliser des sources laser plus simples et disponibles sur le marché afin de générer des nombres aléatoires élevés dans des QRNGs de confiance de type temporel. De plus, puisque les performances de ce design sont compatibles avec différentes sources utilisées, cette méthode

pourrait être adaptée en tant que mécanisme universel de mise à l'échelle des paramètres de n'importe quel QRNGs. Compte tenu de la simplicité, de l'interopérabilité de sa conception et du cout minimal nécessaire à l'optimisation de l'appareil, ce schéma et ses itérations futures pourraient accélérer le développement des générateurs de nombres aléatoires photoniques pour les applications commerciales.

**Mots clés:** Génération de nombres aléatoires; Optique quantique; Information quantique; Photonique quantique

# ABSTRACT

Randomness is one of the most fundamental properties of nature which is widely prevalent in many aspects of day-to-day life. Methods to produce random numbers (representing the source of indeterminism) have been explored since ancient times, as evidenced by the first random number generators (i.e.,dice), which were created thousands of years ago. In recent times, more sophisticated generators have been developed due to the complex requirements of modern applications like cryptography, communications, and simulations. Driven by advances in the semiconductor and computing industries, many physical devices have been created over the last 70 years, which can produce random numbers at a rate that can match the operating speeds of electronic computers. As such, electronic generators are widely used in state-of-the-art scientific and technological applications as well as in less critical fields like sports and lottery. While the performance of these classical devices are sufficient for current requirements, with the rapid advancement in the state of emerging quantum technologies, there exists a serious threat to their functionalities as well as the integrity of protocols that employ them. By exploiting a fundamentally different paradigm of computing, quantum systems can undermine the working of many critical hardware and software protocols such as classical random number generators, the Rivest-Shamir-Adleman algorithm, cryptographic methods etc. These are extensively used in sensitive sectors like banking, finance and defence. To address this threat, truly indeterministic Quantum Random Number Generators (QRNGs) can be designed and deployed. Quantum systems have been studied for use as random number generators for many decades. While the initial studies used radioactive systems for a theoretical exploration of randomness, more recent realizations have focused on optics and photonics platforms owing to their numerous advantages like multiple degrees of freedom, ease of integration, use in communications etc. Recent demonstrations of “Quantum Supremacy” and “Quantum Advantage” by companies like Google, which showcase the potential strength of even short-term quantum machines, signal the urgent need for demonstrably secure, and practically usable QRNGs that can be deployed in various applications.

Current research in this direction optimizes for two performance parameters, namely the bitrate (speed of the device) and certification of genuine randomness (security of the device). Consequently, current methods can be classified into three categories, i.e. a) trusted-device configurations which achieve a fast bitrate but poor certification, b) self-testing configurations which achieve good certification but extremely slow rates and c) semi self-testing devices which maintain a trade-off between the two parameters. Hence, active research is being undertaken to realize an ideal generator that can provide both a high bitrate and a high degree of trust. However, current methods to improve QRNGs typically require a complete re-design of the system along with costly component upgrades. This is primarily because of design strategies which focus on increasing the measurement dimensionality (in order to augment the possible random bits per photon) and/or using key components with better performance (such as sources with lower likelihood of multiphoton generation, and detectors with higher resolution and lower dead times), which are difficult to accomplish due to challenges in material sciences, fabrication techniques and compatibility of different components. Furthermore, since these schemes are widely different in their designs, it is tough to compare them and thus determine which is truly best for use in a specific application.

To address these issues, through the work carried out in this thesis, we introduce minimum information entropy per bit as an important design parameter for photonic QRNGs. This serves as a universal metric across QRNG designs to determine the maximum and practical amount for which a given device's randomness can be used to generate bits. We then implement a temporal mode QRNG for which the efficiency of random bit generation can be scaled-up by cascading an increasing number of fiber-based beam splitters, while minimally increasing setup complexity. Using an attenuated laser signal and a source of entangled photons, we experimentally show that the randomness can be scaled, consistently with theory. Depending on the quality of the sources and detectors, we show that randomness can be increased anywhere from 3%-30%. Furthermore, in our implementation, the attenuated laser shows much larger minimum entropy per bit compared to our entangled photon pair source. This suggests a potential advantage in using simpler, off-the-shelf laser sources for high random number generation rates in trusted-device temporal mode QRNGs. Additonally, since the performance of this scheme is consistent with different sources, this method can be adapted as a universal mechanism to scale the parameters of any QRNG. Given the simplicity and interoperability of its design and minimal cost for improving device performances, this scheme and its extensions could accerelate the development of photonic random number generators for commercial applications.

**Keywords:** Random Number Generation; Quantum Optics; Quantum Information; Quantum Photonics

# SOMMAIRE RÉCAPITULATIF : DESIGN ÉVOLUTIF DE GÉNÉRATION DE NOMBRES ALÉATOIRES QUANTIQUES PHOTONIQUES

Les nombres aléatoires sont une ressource centrale pour des applications allant de la cryptographie aux simulations statistiques [1]. La plupart des générateurs de nombres aléatoires (Random Number Generator - RNG) actuels fonctionnent selon les lois de la physique classique. Ils ont, par conséquent, une nature déterministe, ce qui limite leur capacité à modéliser correctement des phénomènes vraiment aléatoires. Pour surmonter cette limitation, il est de mise d'exploiter l'indéterminisme inhérent aux systèmes quantiques pour générer des nombres véritablement aléatoires [2]. Actuellement, la plupart des générateurs de nombres aléatoires quantiques de pointe (Quantum Random Number Generator - QRNG) [3,4] sont des systèmes photoniques, ce qui leur permet de bénéficier d'une bande passante élevée, d'un faible encombrement et d'une grande simplicité expérimentale. De plus, ils ont la robustesse des technologies optiques offerte avec les sources laser ultrarapides, les composants fibrés et les plateformes microoptiques sur puce. La plupart des recherches actuelles sur les QRNGs se sont concentrées sur l'optimisation de deux critères de performance soit le débit de génération de bits aléatoires et l'assurance que le débit binaire est vraiment aléatoire (la certifiabilité). Les réalisations actuelles maintiennent un compromis entre ces deux quantités de telle sorte que les QRNGs dits de confiance (qui utilisent des sources et des composants bien caractérisés et non malveillants) atteignent un débit rapide au détriment de la nature aléatoire [5,6], tandis que les QRNGs autotestés (qui ne font aucune hypothèse sur les paramètres de configuration ou sur la confiance dans l'appareil) offrent un bon caractère aléatoire, mais de mauvais débits [7-10].

Nous considérons les QRNGs de confiance comme étant l'option la plus prometteuse afin d'obtenir le débit binaire (« bit rate ») aléatoire nécessaire aux applications commerciales. Dans ces dispositifs, un débit binaire aléatoire est généralement obtenu au moyen de deux stratégies soit en optimisant la source ou le détecteur [8,9], en augmentant la dimensionnalité de l'état (le nombre de niveaux/modes qu'un photon peut occuper lors de la détection) [10,11] ou les deux.

Cette première stratégie est importante pour les appareils dits de confiance, car ils fonctionnent sous l'hypothèse que les composants ne sont pas vulnérables aux attaques malveillantes. Par conséquent, la nature de la source garantit son caractère aléatoire. Cela signifie que les seules raisons pour lesquelles un QRNG de confiance serait incapable d'atteindre la limite maximale de son caractère aléatoire (c'est-à-dire traduire tous ses états de photons en bits aléatoires) sont

dues aux imperfections du système telles que le bruit, la génération d'états multiphotons, la résolution finie du détecteur, le faible contraste des franges d'interférence de nature quantique, etc. La deuxième stratégie, quant à elle, est critique, car la dimensionnalité de l'état quantifie l'étendue possible du caractère aléatoire dans le QRNG, et ce, sans les imperfections du monde réel. Ainsi, le débit binaire aléatoire des dispositifs peut être augmenté en pratique en réduisant les imperfections d'un système, en augmentant le nombre total de bits récupérables ou les deux. Cependant, en pratique, plusieurs inconvénients découlent de l'utilisation de ces méthodes pour augmenter le débit binaire des QRNGs. D'un côté, un débit binaire qui dépend entièrement des paramètres de la source ou du détecteur a une flexibilité limitée. De plus, à mesure que les sources et les détecteurs se dégradent (ou toute autre modification des autres paramètres de fonctionnement), le débit binaire évolue également. L'amélioration de la performance de tels dispositifs nécessite la révision de composants couteux, par exemple le remplacement complet de la source ou du détecteur. D'un autre côté, l'amélioration du débit binaire grâce à la mise à l'échelle de la dimensionnalité du système s'accompagne généralement d'une complexité expérimentale et de couts accrus. Par exemple, des travaux récents de Grafe *et coll.* [11] démontrent un mécanisme potentiel de mise à l'échelle réalisé en augmentant le nombre de modes spatiaux dans un système de QRNG basé sur les guides d'ondes. Cette technique nécessite toutefois huit détecteurs de photons uniques pour chacun des chemins possibles.

Dans ce travail, nous démontrons expérimentalement une méthode de mise à l'échelle flexible du débit binaire pour un QRNG de confiance basée sur les modes temporels, et ce sans changer la source ou le détecteur. De plus, celle-ci n'induit qu'une augmentation minimale de la complexité du système. La motivation derrière notre conception expérimentale vient d'un paramètre que nous étudions, soit l'entropie d'information minimale par bit, $H_\eta$ (détaillée à la section théorie), qui représente l'efficacité avec laquelle un QRNG peut, en pratique, utiliser son caractère aléatoire. Par exemple, dans une configuration avec $H_\eta$= 0,5, seule la moitié des détections de photons est considérée comme vraiment aléatoire. Comme ce paramètre est directement proportionnel au débit binaire aléatoire dans les systèmes QRNG de confiance, il s'agit d'un paramètre particulièrement intéressant pour la conception et l'optimisation des systèmes. Notre implémentation fonctionne à l'aide de séparateurs de faisceau fibrés et de lignes à délais fibrées en cascade, afin d'augmenter la dimensionnalité de l'espace d'état des temps d'arrivée de photon et ainsi augmenter $H_\eta$. Avec cette technique, le nombre de canaux détecteurs reste le même, même si on augmente physiquement le nombre de détection d'états de photon possibles à l'aide du multiplexage temporel. En ajoutant simplement des séparateurs de faisceau et des lignes à délai fibrées, nous pouvons réduire l'impact des imperfections de la source et du détecteur. Cette méthode permet d'améliorer la génération de bits aléatoires ou l'utilisation des dispositifs de qualité inférieure pour obtenir des performances comparables. Dans cette thèse, nous montrons une mise à l'échelle du caractère aléatoire, conforme à la théorie, pour deux sources de photons uniques différentes : 1) un laser atténué et 2) des photons intriqués temps-énergie continus générés par une conversion paramétrique descendante spontanée (SPDC).

Nos sources spécifiques démontrent une amélioration de l'entropie minimale par bit entre 3 et 20 % à mesure que le nombre de séparateurs de faisceau passe de 0 à 4. Cependant, notre méthode est très prometteuse pour des QRNGs dits économiques qui utilisent des sources et des détecteurs de qualité inférieure et pour lesquels l'amélioration du caractère aléatoire peut atteindre 29 %. De plus, nos expériences indiquent que pour des implémentations avec des appareils de confiance, les sources de paires de photons annonceurs ne semblent pas offrir d'avantages inhérents par rapport aux sources de laser atténuées concernant la vitesse de génération aléatoire de bit. Cela indique que les lasers disponibles sur le marché peuvent être suffisants pour répondre aux conditions expérimentales de la génération quantique de nombres aléatoires. Cependant, les paires de photons annonceurs peuvent être utilisées pour la régénération et l'investigation simultanées de l'espace d'état des temps d'arrivée de photon en utilisant une seule mesure au lieu d'utiliser deux mesures séparées.

## **Approche théorique**

Dans les QRNGs temporels basés sur le temps d'arrivée des photons [12], la quantité d'éléments aléatoires dans la distribution brute est mesurée par l'entropie minimale :

$$H_{min} = log_2(N) \tag{1}$$

en unités de bits, où $N$ est le nombre total de tranches de temps dans la fenêtre d'observation. Pour un QRNG générique, $N$ fait référence au nombre d'états possibles qu'un photon peut occuper lors d'un évènement de détection donné. Dans le scénario idéal, tous les intervalles de temps ont une probabilité d'occupation égale et donc $P_i$ = 1/ $N$. Dans un système où la distribution de probabilité des états n'est pas uniforme, on utilise $H_{min} = log_2\left(\frac{1}{P_{max}}\right)$. Indépendamment de la source ou du mécanisme de détection utilisé, le nombre de bits qui peuvent être obtenus en pratique à partir d'une telle implémentation de dispositif de confiance est

$$H_{min} = log_2(N) - H_{dev} \tag{2}$$

où $H_{dev}$ correspond aux imperfections de la source ou du détecteur soit, par exemple, le comptage de multiphotons, la variation du moment de détection, les temps morts, la résolution temporelle limitée, etc., ce qui dégrade la capacité de génération aléatoire de bits. En divisant $H_{min}$ par le nombre de bits que le système pourrait idéalement produire, nous introduisons

$$H_\eta = 1 - \frac{H_{dev}}{log_2(N)} \tag{3}$$

qui a une valeur maximale de 1 lorsque $H_{min}$ = 0. En d'autres termes, dans un système parfaitement aléatoire, il n'y a pas d'imperfection expérimentale et donc toutes les détections de photons peuvent être considérées comme réellement aléatoires. Quand $H_{dev}$ est différent de zéro, $H_\eta$ peut être optimisé en réduisant $H_{dev}$ ou en augmentant *N*. Bien que ce ne soit pas les objectifs principaux de leurs travaux, Xu *et coll.* [13] montrent un taux de génération de bits aléatoires impressionnant dans les QRNG de confiance en améliorant $H_\eta$. Les efforts pour améliorer $H_\eta$ en diminuant $H_{dev}$ nécessitent une optimisation de la source et du détecteur ou une révision complète du design. Toutefois, les méthodes qui améliorent $H_\eta$ en augmentant la dimensionnalité *N*, s'accompagnent soit d'une complexité expérimentale supplémentaire, soit d'un cout accru. Dans cette thèse, nous caractérisons $H_{min}$ et $H_{dev}$ sur un nouveau QRNG pour deux sources et montrons comment ils affectent $H_\eta$. De plus, nous expliquons comment de petits changements apportés à la configuration expérimentale peuvent évoluer en optimisant le terme $H_{dev}/log_2(N)$.

La figure 1 montre le schéma expérimental nécessaire afin de réaliser notre technique de mise à l'échelle d'entropie. Nous considérons que notre système est composé de sources et de détecteurs imparfaits (c'est-à-dire avec une émission multiphotonique, une résolution de détecteur finie, une efficacité imparfaite et une présence de bruit à la fois dans le fond et dans l'obscurité). Au cœur de notre montage se trouve un ensemble de diviseurs de faisceau fibrés en cascade pour lesquels la source est divisée en deux chemins par le premier et par la suite,

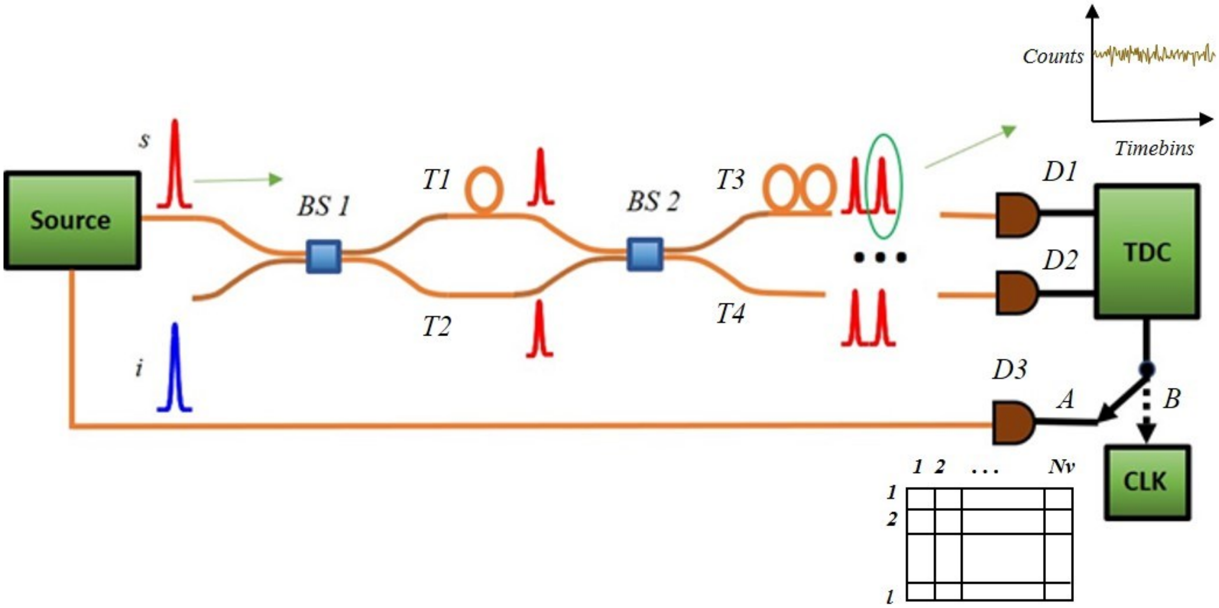


**Fig. 1 :** Schéma expérimental de la configuration du séparateur de faisceau de fibre en cascade pour la mise à l'échelle de l'entropie. BS - Séparateur de faisceau, TDC - Convertisseur numérique temporel, CLK - Horloge, T1, T2, T3, T4 - Retards temporels introduits par différentes longueurs de fibre, s, i - Paire de photons enchevêtrés signal-idler, D1, D2, D3 – Détecteurs de Photon unique

recombinée dans l'entrée suivante. Le tout est répété à nouveau jusqu'au dernier diviseur de faisceau. Les chemins optiques possibles sont constitués de différentes longueurs de fibre tels que montrés à la figure 1. Compte tenu du temps de génération $T_{gen}$ pour chaque photon, on obtient $l = 2^n$ états temporels dans lesquels le photon peut être détecté (où *n* est le nombre de séparateurs de faisceau). Par exemple, si nous considérons le cas où il n'y a que deux séparateurs de faisceau (illustré à la figure 1), le schéma aléatoire introduit quatre états temporels que chaque photon peut occuper lors de son arrivée aux détecteurs, correspondant aux retards $T_{gen}$ + T1 + T3, $T_{gen}$ + T1 + T4, $T_{gen}$ + T2 + T3 et $T_{gen}$ + T2 + T4, où T1, T2, T3 et T4 correspondent aux retards temporels introduits par les chemins de fibre illustrés sur la figure 1.

Étant donné que ces états temporels correspondent aux composants physiques du système, ils sont appelé des « intervalles de temps physiques ». Si $T_{gen}$ est mesuré en utilisant un photon annonciateur (« heralding photon ») de la paire de photons signal-idler d'une source corrélée, ces intervalles de temps physiques peuvent être mesurés. Cela correspond à la position de commutation A sur la figure 1, dans laquelle le photon de signal entre dans les diviseurs de faisceau au même moment où l'idler est émis afin d'être utilisé comme déclencheur « d'horloge ». Cette horloge permet ensuite de mesurer le temps de retard des photons (voir les données de la figure 3).

Cependant, en l’absence d’un annonceur, les intervalles de temps physique ne peuvent être mesurés en raison de l’incertitude sur le temps de génération $T_{gen}$ qui augmente à cause des caractéristiques temporelles de la source (telles que l’émission spontanée ou l’émission continue du laser). Comme la génération de photons est un processus spontané, $T_{gen}$ varie de manière aléatoire par rapport à tout signal d’horloge dans le référentiel du laboratoire. Sans synchronisation avec le temps de génération, le détecteur affiche un chevauchement des intervalles de temps physiques introduits par notre système. Cela résulte en une distribution équiprobable et uniforme des temps d’arrivée des photons par rapport au référentiel du laboratoire. Cette façon d’utiliser une horloge arbitraire dans le laboratoire (position de commutation B sur la figure 1) pour mesurer les arrivées aléatoires de photons dans le référentiel de l’idler a été utilisée avec succès dans un certain nombre de QRNGs de type temporel à grande dimension [14,15] pour encoder des bits aléatoires. Puisque cet encodage représente des bits dans des positions qui ne sont pas déterminées par les composants physiques dans la configuration, nous les appelons des « intervalles de temps virtuels », dont le nombre total est $N_v$. On peut noter que les paramètres de mesure représentés par les positions A et B de l’interrupteur sur la figure 1 correspondent à la mesure de l’incertitude/du caractère aléatoire dans deux degrés de liberté différents. Alors que la mesure à la position A de l’interrupteur permet la caractérisation de l’incertitude sur le chemin traversé par les photons détectés, la mesure à la position B de l’interrupteur permet la caractérisation de l’incertitude sur le temps de génération des photons. L’utilisation de deux degrés de liberté distincts dans un système de multiplexage temporel est une caractéristique unique de notre montage expérimental.

La configuration que nous proposons combine des intervalles de temps virtuels et physiques pour mettre à l’échelle la dimensionnalité et par conséquent $H_\eta$. Les temps d’arrivée des photons après la cascade de séparateurs de faisceau sont interprétés comme une matrice où les deux vecteurs de base sont les temps de génération et les positions physiques dans une fenêtre de temps (voir l’encadré de la figure 1). L’encodage d’une fenêtre de temps virtuelle conventionnelle, d’une seule dimension physique, représente un espace d’états d’une matrice ligne. Notre façon de faire, en revanche, fournit un espace d’états égaux à l’aire de la matrice (qui est dans notre cas $l$ = 2-16 intervalles de temps physiques et $N_v$= 100 intervalles de temps virtuels, équivalents à $l$ x $N_v$ = 200 – 1,600 états temporels). De plus, notre méthode présente un avantage important autre que d’augmenter simplement les états qu’un photon peut occuper. En utilisant des intervalles de temps physiques, nous pouvons mettre à l’échelle l’entropie globale par photon sans changer les propriétés de la source ou du détecteur d'origine. Cela n'est pas possible dans les QRNGs à créneaux temporels virtuels conventionnels, où l'on ne peut augmenter le nombre d'états qu'en allongeant la fenêtre de temps de détection ou en améliorant la résolution du détecteur [16].

Les effets de la configuration de séparateur de faisceau en cascade sur les temps d'arrivée des photons sont également illustrés sur la figure 2. Cet exemple représente le cas ou n'y a que 2 séparateurs de faisceau. Comme le montre cette figure, les deux séparateurs de faisceau donnent $2^2$ = 4 trajets au niveau des détecteurs. Chaque photon avec un $T_{gen}$ distinct peut traverser chacun de ces chemins avec une probabilité égale. En outre, les photons avec des valeurs différentes peuvent occuper le même intervalle de temps (par rapport à une horloge externe) dans la séquence de bits résultante, car ils ont des retards différents correspondant au chemin qu'ils auraient traversé. Par conséquent, le nombre total de façons dont une position de bit donnée peut être remplie augmente (de 1 à 4) avec l'ajout de séparateurs de faisceau. Plus généralement, une détection dans un intervalle de temps, qui a pour référence une horloge externe opérant à une fréquence donnée (ce qui crée $N_v$ intervalle temporel dans la fenêtre d'observation), peut signifier l'arrivée de $2^n$ photons distincts dus aux $n$ séparateurs de faisceaux. Ici, les incertitudes spatiales sont combinées avec les incertitudes temporelles de la

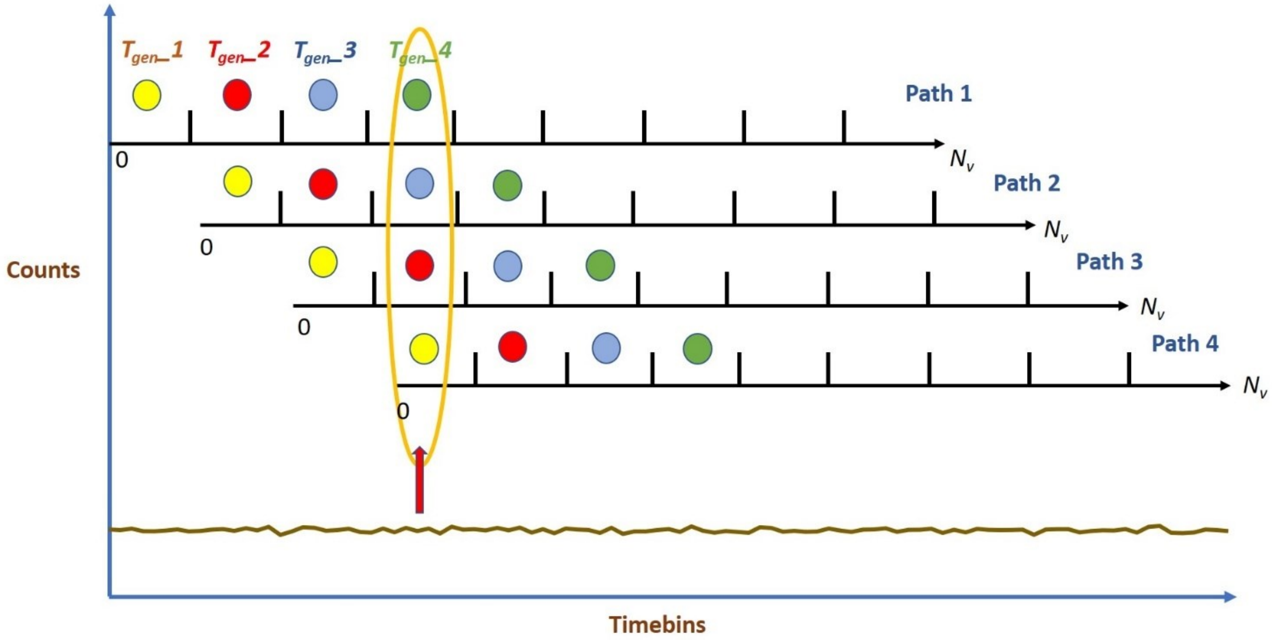


**Fig 2 :** Illustration de la mise à l'échelle aléatoire à l'aide de deux séparateurs de faisceau dans la configuration. Les chemins 1, 2, 3 et 4 sont créés au niveau des détecteurs qui sont temporellement distincts les uns des autres. Les photons avec des temps de génération différents peuvent arriver à la même position de bit dans la séquence finale, augmentant ainsi le caractère aléatoire par bit.

génération/émission des photons provenant de la sources (ce qui en découle une variation aléatoire de $T_{gen}$ vis-à-vis de l'horloge du laboratoire). Par conséquent, ce système génère un plus grand nombre d'états aléatoires que les QRNGs temporels traditionnels qui n'utilisent qu'un seul degré de liberté pour l'extraction du caractère aléatoire soit l'arrivée d'un photon. Ainsi les $N_v$ intervalles temporels définis par l'horloge externe et les $2^n$ modes spatiaux provenant des séparateurs de faisceau en cascade, et ce pour l'ensemble des positions possibles dans les intervalles temporels, résulte en une séquence binaire ayant $N_v$*$2^n$ possibilités. Cette quantité représente la dimensionnalité de l'espace d'état des photons détectés. Dans le cas où $N_v = 5$ et $n = 1$, une séquence binaire de 10000 pourrait être mesurée si un photon est détecté soit au détecteur du canal D1 ou D2 dans le premier intervalle temporel. Sans séparateur de faisceau, la séquence binaire serait dépendante seulement d'une détection unique (correspondant à un $T_{gen}$ unique). L'introduction d'un séparateur de faisceau multiplie donc le nombre possible d'états de photons et augmente ainsi le caractère aléatoire du système.

Pour démontrer l'efficacité de notre méthode et que la mise à l'échelle aléatoire est indépendante de la source et du détecteur, nous dérivons l'entropie minimale attendue par bit pour le laser CW atténué et pour les sources de paires de photons intriqués.

### a. Laser CW atténué

La lumière émise par le laser atténué a un faible nombre moyen de photons $\bar{n} < 0{,}1$ par période d'observation. Elle suit une distribution de poisson et peut être traitée comme une source de photons unique. L'entropie minimum associée à la détection d'un tel photon peut s'écrire comme suit [16] pour $\bar{n} < 0{,}1$ :

$$H_{min} = log_2(N_v) + log_2\left(1 - e^{-\lambda T \gamma}\right) - log_2(\lambda T \gamma) \qquad (4)$$

où $N_v$ est le nombre de tranches de temps virtuelles dans le schéma d'encodage, $\lambda$ caractérise l'intensité du laser, $\gamma$ est l'efficacité du détecteur et $T$ est la durée d'observation telle que déterminée par le temps mort du détecteur. Les termes $log_2\left(1 - e^{-\lambda T \gamma}\right) - log_2(\lambda T \gamma)$ sont obtenus après la correction des comptages de multiphotons, de la gigue de synchronisation du détecteur, etc., qui contribuent tous à tort au caractère aléatoire des données brutes collectées - comme expliqué dans [16].

Cependant, puisque la dimensionnalité de l'espace d'état est $N_v * 2^n$ au lieu de $N_v$ comme dans les QRNGs traditionnels, cette équation devient (voir l'Annexe pour les étapes détaillées) :

$$H_{min} = log_2(N_v) + n \ +log_2\left(1 - e^{-\lambda T\gamma}\right) \ - log_2(\lambda T\gamma) \qquad (5)$$

L'entropie additionnelle résultant de la section du séparateur de faisceau en cascade peut être surveillée à travers l'entropie minimum par bit, tel que montré dans l'équation suivante, en utilisant les valeurs expérimentales de $N_v$ et de $n$:

$$H_\eta \ = \ 1 - \frac{H_{dev}}{log_2(N_v)+n} \qquad (6)$$

avec

$$H_{dev} = log_2\left(1 - e^{-\lambda T\gamma}\right) - \ log_2(\lambda T\gamma) \qquad (7)$$

### b. Paires de photons intriquée

Un état biphoton $N_v$–dimensionnel peut être obtenu par différents processus spontanés dans des milieux non linéaires tels que le mélange spontané à quatre ondes (SFWM) dans des résonateurs microanneau [16,17] ou par SPDC dans les guides d'ondes de type PPLN [18]. Cet état peut être caractérisé comme:

$$|\psi\rangle = \ \sum_{i=1}^{N_v} |i\rangle_A \otimes \ |i\rangle_B \qquad (8)$$

où $|i\rangle$ représente un photon unique à un intervalle de temps discrétisé $i$ [13]. Il a été démontré que le minimum d'entropie lisse (qui prend en considération le bruit de l'environnement) d'un état intriqué temporellement de haute dimensionnalité, lorsque mesuré dans deux bases mutuellement non biaisées orthogonales de POVMs (Positive Operator-Valued Measures) (telles que les intervalles temporels virtuels et physiques correspondant aux positions A et B du montage expérimental de la figure 1) peut être liée en utilisant le principe d'incertitude comme dans l'équation ci-dessous [19,20]:

$$H_{min} \geq -log_2 c - H_{max} \qquad (9)$$

où $c$ est la superposition maximale (voir éq. 10) entre deux mesures projectives mutuellement non biaisées venant de l'incompatibilité du POVMs sur la base des intervalles de temps physique et virtuel. L'entropie maximale $H_{max}$ est l'entropie de Renyi de l'ordre ½, ce qui réduit l'importance des mesures ayant peu d'occurrences [21]. Cette quantité peut être modélisée de telle sorte qu'elle suit certaines sources d'inexactitude durant une mesure telles que des fluctuations statistiques, une variation de la visibilité d'un photon unique, etc. ce qui a été montré dans [13]. Plus encore, $H_{max}$ peux être approximée de façon à être proportionnelle à $N_v$ qui est le paramètre le plus significatif [22]. Ici, les paramètres $H_{min}$ et $H_{max}$ correspondent respectivement aux mesures de l'entropie des intervalles de temps virtuel et physique et forment une base orthogonale mutuellement non biaisée. Puisque pour chaque valeur de $i$ il existe $2^n$ états temporels dans lesquels le photon peut être mesuré au niveau du détecteur,

$$c = \frac{1}{N_v * 2^n} \qquad (10)$$

En outre, l'entropie minimale par bit peut également être décrite par l'équation (6) en remplaçant $H_{dev}$ par $H_{max}$

## Détails et résultats expérimentaux

La source atténuée qui est utilisée est un laser CW (NetTest Tunics Plus) centré à 1547,6 nm avec une largeur de raie de 100 MHz. La puissance du laser a été atténuée au moyen de deux atténuateurs optiques variables, pour atteindre un flux de photons de 25 kHz au niveau du détecteur. Cela correspond à un nombre moyen de photons, $\bar{n}$ <0,1, sur la fenêtre temporelle de 100 ns que nous avons choisie pour mesurer les temps d'arrivée des photons. La source SPDC utilise le même laser comme pompe pour un système de guide d'ondes au niobate de lithium à polarisation périodique (PPLN) à deux étages. Dans cette configuration, le premier PPLN est utilisé pour la génération de secondes harmoniques (Second Harmonic Generation - SHG) pour convertir la lumière de 1547,6 nm à 773,8 nm. Ensuite, cette paire de photons est utilisée pour pomper le deuxième guide d'ondes PPLN afin d'engendrer un effet SPDC dans le but de générer des paires de photons intriqués (pour lesquels des photons dégénérés sont centrés à 1547,6 nm).

Les deux PPLN sont des dispositifs commerciaux (Srico 2000), dont les paramètres sont similaires à ceux de la référence [23]. Des filtres passe-bandes à efficacité élevée ont été utilisés pour bloquer la lumière résiduelle de la pompe non convertie respectivement par SHG ou SPDC. Bien que la bande passante totale du SPDC soit supérieure à 4 THz (trop grande pour être mesurée par notre équipement de laboratoire), nous utilisons un filtre de télécommunications programmable (Finisar 4000A WaveShaper) pour sélectionner 25 GHz de la bande passante des photons signal et idler SPDC non dégénérés. En pompant le premier PPLN à 1 mW, nous produisons 100 $\mu$W de SHG, ce qui nous donne un taux de photons SPDC détectés de 13 kHz pour les canaux de signal et idler. Des détecteurs de photons uniques à nanofils supraconducteurs (Quantum Opus One) ont été utilisés pour toutes les expériences. Cet instrument a un temps mort de 80 ns, une efficacité de 85 % à 1550 nm et une résolution (gigue) d'environ 400 ps. Les évènements de détection de photons ont été enregistrés par un convertisseur temps-numérique (Picoquant Hydraharp 400), qui a été utilisé pour donner la différence de temps d'arrivée des photons entre un déclenchement et un évènement de détection. Comme décrit dans la section théorie, nous utilisons deux déclencheurs différents : le photon idler détecté (position A du commutateur de la figure 1) et une horloge de laboratoire arbitraire (position B du commutateur de la figure 1), soit un train d'impulsions à 10 MHz à partir d'un générateur de fonctions arbitraires (Tektronix AFG 3251).

Notre système consiste en 1 à 4 coupleurs de fibre 50/50 à maintien de polarisation (AFW PFC-15-2-50-BB) connectés comme indiqué sur la figure 1. Les coupleurs utilisés dans notre expérience avaient une perte d'insertion moyenne d'environ 1 dB chacun (y compris la perte due aux manchons d'accouplement et aux fibres à délais) et un rapport de division de 50/50 ayant une erreur d'au plus 2 %. Il convient de noter que des écarts plus importants dans les rapports de division seraient néfastes à la génération de bits aléatoires, car ils entraineraient une réduction de l'entropie minimale par bit en raison d'une diminution des états de photons possibles.

Afin de générer des intervalles de temps physiques séparés, nous avons mis des fibres à délais sur toutes les sorties des BS allant de 0,20 à 4,0 m. Notez que la longueur des fibres ne représente pas la différence de chemin optique entre les bras des interféromètres. Elles sont plutôt incluses pour assurer que les intervalles de temps physiques mesurés sont distincts et ne se chevauchent pas. Le chevauchement des intervalles de temps mènerait à une réduction du nombre total d'états distincts de photon et du nombre de bits par détection de photon pouvant être extrait, ce qui n'est pas souhaitable pour un QRNG. De plus, la différence de chemin minimum entre les bras des interféromètres est maintenue à 2m afin d'éviter l'interférence de premier ordre provenant du laser et de la source PPLN, qui nécessitent une différence de chemin minimale de 1 m. La figure 3 montre l'histogramme des temps d'arrivée des photons en utilisant l'idler comme déclencheur. De plus, le taux de coïncidence accidentelle (CAR) du système varie de 200 à 1200 en fonction du nombre de séparateurs de faisceau (et donc des pertes) dans le système. Dans une implémentation idéale, où il n'y aurait aucune perte, la distribution de mesure de photon dans chaque intervalle de temps serait approximativement la même. Cependant, la distribution de mesure de photon observée montre une non-uniformité qui peut être due à une déviation du rapport réflexion/transmission (par rapport à l'idéal de 50/50) des séparateurs de faisceau en succession et aux différentes pertes introduites par l'ajout de séparateurs de faisceau, de fibre à délai et de connecteurs supplémentaires dans le montage expérimental. L'impact du coefficient d'asymétrie dans la distribution de mesure de photon peut être estimé à partir de la valeur expérimentale de l'entropie minimum. Dans le cas d'une déviation significative par rapport au comportement idéal, le nombre de bits utilisables serait inférieur au nombre de séparateurs de faisceau (soit $log_2(2^n) = n$ ) du système, étant donné qu'il y aurait moins d'états de photon distinct. Cependant, pour la distribution observée, la valeur expérimentale du minimum d'entropie était près de la valeur théorique attendue, indiquant que l'asymétrie observée dans la

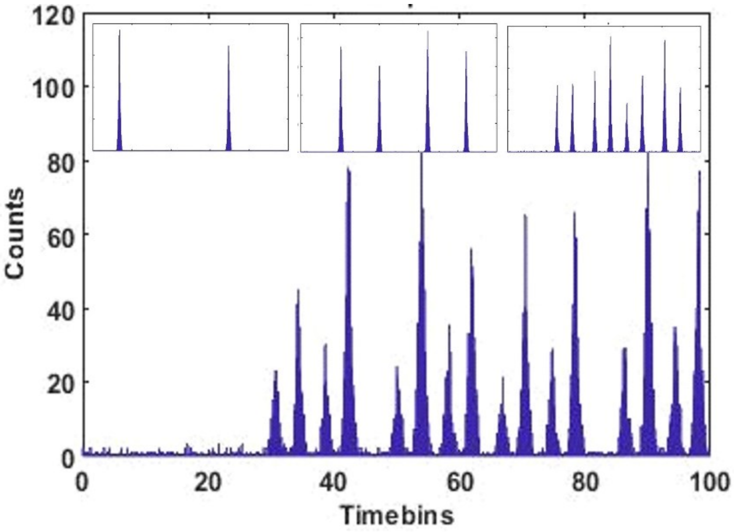


**Fig 3 : Histogramme des temps d'arrivée des photons uniques (cas photon heralded).** L'histogramme des arrivées de photons de signal (intervalles de temps physiques) utilisant l'idler comme déclencheur d'horloge pour les séparateurs de faisceau 1,2,3 et 4 (encadré: gauche - droite)

distribution de mesure de photon n'impacte pas significativement le caractère aléatoire pouvant être extrait.

La figure 4 montre la distribution des temps d'arrivée des photons lorsque l'horloge est en position B sur la figure 1. Cela correspond au mode actif du QRNG pour générer des bits aléatoires. Ici, nous choisissons une fenêtre d'observation de 100 ns (~ temps mort du détecteur) et une largeur d'intervalle de temps de 1 ns, ce qui donne un $N_v$ = 100. Nous choisissons délibérément cet intervalle de sorte qu'il soit d'au moins 2 inférieures à notre résolution réelle afin de simuler une situation expérimentale où la résolution réelle d'un détecteur serait de 1,0 ns. Finalement, $N_v$ est un paramètre « libre » limité par le temps mort du détecteur (qui définit la plus grande fenêtre d'observation) et par la résolution du détecteur (qui définit le plus petit intervalle de temps). Les distributions des deux sources sont presque uniformes comme prévu. En effet, la distribution uniforme est obtenue par une superposition des intervalles de temps créés par l'arrangement des séparateurs de faisceau (tels que montrés en figure 3) à diffèrent temps de génération $T_{gen}$. La

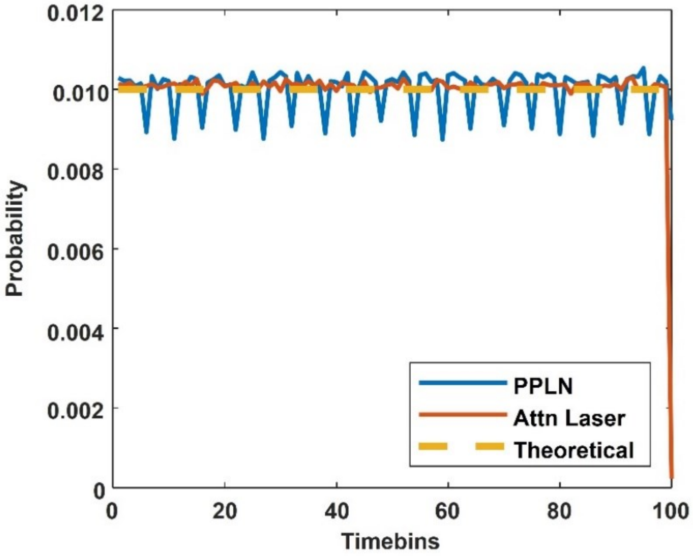


**Fig 4 : Histogramme des temps d'arrivée des photons (non heralded).** Intervalles de temps virtuels) au niveau des détecteurs pour les photons émis respectivement par une source de guide d'ondes PPLN intriqué et un laser atténué, en utilisant une horloge de laboratoire de 10 MHz comme déclencheur. Les données ont été collectées en 3 minutes.

Figure 5 illustre la genèse de cette distribution uniforme en utilisant les données de la figure 3 dans le cas utilisant un séparateur de faisceau et en variant la valeur de $T_{gen}$.

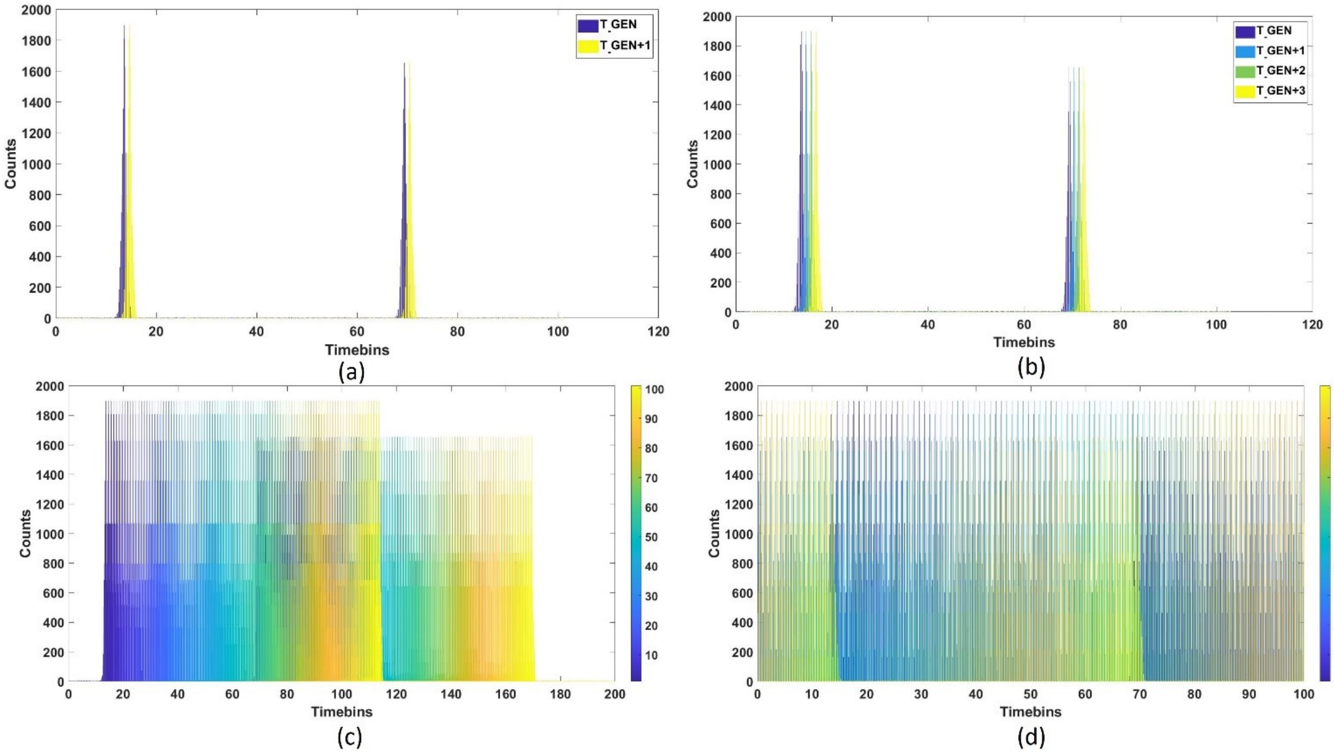

(a)
(b)
(c)
(d)

**Fig 5. Illustration de la création d'une distribution uniforme aléatoire dû aux variations de $T_{gen}$ et des positions physiques des intervalles de temps.** (a) Intervalle de temps physique en fonction de $T_{gen}$ variant dans un intervalle de 1 ns. (b) Intervalles de temps physiques en fonction de $T_{gen}$ variant dans un intervalle de 3 ns. (c) Translation linéaire et superposition des intervalles de temps physiques en fonction de $T_{gen}$ à travers 100 intervalles de temps physiques. (d) Distribution uniforme obtenue en mesurant des intervalles de temps physiques avec une horloge de référence à 10 MHz et en variant $T_{gen}$ à travers 100 intervalles de temps.

Avec un seul séparateur de faisceau dans le système, deux intervalles de temps physiques et distincts sont créés pour un $T_{gen}$ donné. Cependant, lorsque $T_{gen}$ varie en fonction de l'horloge de référence, les positions des intervalles de temps physique générés varient aussi. Cet effet est montré sur les figures 5 (a) et 5 (b), où $T_{gen}$ varie respectivement de 1 ns et 3 ns. Les intervalles de temps physique créés subissent une translation sur l'axe des intervalles de temps, correspondant aux variations de $T_{gen}$.

Comme $N_v$ = 100 et la longueur de chaque intervalle de temps est de 1 ns dans cette expérience, les valeurs mesurées pour $T_{gen}$ sont de 1 ns et 100 ns en incréments de 1 ns (égaux à la durée de l'intervalle de temps). La variation de la position temporelle des dimensions temporelles physique correspondant aux variations de $T_{gen}$ pendant un intervalle de 100 ns est montrée sur la figure 5 (c). Nous mesurons une superposition proche de 70 ns causé par un évènement simultané du premier intervalle de mesure de temps physique d'un $T_{gen}$ et d'un deuxième intervalle de mesure de temps d'intervalle d'un autre $T_{gen}$. Comme cette distribution est sur un intervalle de temps linéaire, alors la translation des positions des intervalles de temps physiques est aussi linéaire.

Cependant, la mesure d'un intervalle de temps relatif à une horloge externe avec un taux de répétition donné (10 MHz, utilisé pour la mesure du mode B de la figure 1) engendre un cadre de mesure avec une longueur de durée fixe. Les intervalles de temps se produisant en dehors du cadre de mesure (100 ns) sont vus dans les phases de mesures suivantes. Ceci engendre une distribution uniforme sur 100 intervalles de temps pour n'importe quel cadre de mesure, tel que montré sur la figure 5 (d). Ici, les intervalles de temps physique correspondant à plusieurs valeurs de $T_{gen}$ sont mesurées dans ce cadre référentiel, avec aussi des intervalles de temps qui se superposent et se produisent en dehors de la durée de mesure du cadre précédent; ceci engendre une probabilité inégale de détecter un photon à un de ces 100 intervalles de temps donné. Cette figure illustre donc la génération d'une distribution uniforme aléatoire (telle que celle sur la figure 4) dans la base temporelle dû à des valeurs changeantes et la création d'intervalles de temps physiques, qui sont les deux sources aléatoires utilisées dans cette expérience.

La figure 6 montre l'entropie minimale par bit $H_{\eta}$ mesurée expérimentalement, en fonction du nombre de diviseurs de faisceau pour les deux sources, ainsi que les courbes théoriques pour d'autres sources potentielles et détecteurs avec des quantités variables d'imperfections. Pour déterminer $H_{\eta}$, il existe deux grandeurs importantes: $H_{dev}$ et le nombre total de bits aléatoires potentiels, $log_2(N_v) + n$, générés par cette configuration du QRNG. On mesure expérimentalement la dernière en utilisant les données de la figure 4, en calculant l'entropie à partir des données brutes et en utilisant sa définition empirique comme étant $-log_2$(Max $P_i$) (le nombre maximal de photons comptés dans tous les intervalles temporels confondus) divisé par le nombre total de photons comptés dans cette fenêtre d'observation. De même, les données de la figure 3 sont utilisées pour déterminer expérimentalement le nombre d'états physiques, $l$. Par exemple, dans la configuration avec les quatre séparateurs de faisceau, la valeur théorique des bits devrait être $\log_2(100) + 4 = 10,64$ bits. Nos données de la figure 3 et de la figure 4 quant à eux montrent que le nombre expérimental de bits est de 10,47$\pm$ 0.039. $H_{dev}$ est déterminé à partir de la caractérisation des sources et des détecteurs (à partir des mesures de l'efficacité et de la résolution des détecteurs, de la génération de multiphotons, etc., voir section théorie) similaires aux références [13,14]. Pour notre source laser atténuée, nous avons calculé le $H_{dev}$ moyen à 0,56 et 3,55 pour la source guide d'ondes PPLN.

On peut voir sur la figure 6, que la mise à l'échelle expérimentale de $H_{\eta}$ correspond étroitement aux prévisions théoriques. En général, on s'attend à ce que l'augmentation du nombre de chemins physiques, l, améliore l'efficacité du caractère aléatoire. La valeur de $H_{\eta}$ du laser atténué qui est utilisé dans cette expérience est d'une valeur élevée de 91,52% et atteint 94,8% avec l'ajout de séparateurs de faisceau. Pour ce laser, $H_{dev}$ est si faible que l'augmentation du nombre de modes a peu d'effets sur le caractère aléatoire, car le système permet déjà l'utilisation de l'ensemble du caractère aléatoire pour la génération de bits.

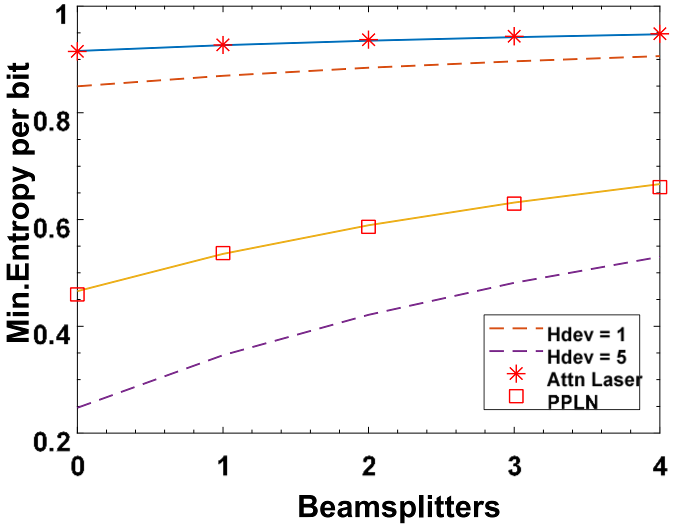


**Fig 6 :** Entropie min par bit en fonction du nombre de séparateurs de faisceau. Les points de données montrent des mesures expérimentales, les lignes pleines montrent des prédictions théoriques pour les valeurs Hdev mesurées et les courbes en pointillé montrent des prédictions théoriques pour différentes non-idéalités de source / détecteur.

Cependant, l'augmentation des pertes due à l'inclusion de plus de diviseurs de faisceau fibrés affecte le flux de photons détecté, ce qui empêche l'augmentation arbitraire du débit binaire de ce système. Pour la source intriquée de type guide d'ondes PPLN, $H_{\eta}$ augmente beaucoup plus, c'est-à-dire de 45,97 % à 66,09 %, lorsque le nombre de séparateurs de faisceau augmente de 0 à 4. La figure 6 montre que l'augmentation du nombre de séparateurs de faisceau est significative pour des valeurs plus élevées de $H_{dev}$. Par exemple, lorsque $H_{dev}$ = 5, ce qui peut être obtenu en changeant les paramètres de la source et du détecteur tels que *N* (paramètre le plus important) est dans l'intervalle de 100 à ~ 500 (par exemple, en utilisant une source PPLN appropriée), l'augmentation de $H_\eta$ peut atteindre jusqu'à 29 %. Cela indique que notre schéma est particulièrement utile dans les cas où les sources ou les détecteurs ont un nombre d'imperfections accrues, ce qui le rend encore plus intéressant dans un contexte commercial. Il est à noter que la valeur expérimentale de $H_\eta$ change de façon négligeable lorsque le temps d'intégration varie de 30 s à 60 minutes. Cela indique qu'un comportement asymptotique est obtenu pour des temps de collecte de données relativement courts. Ainsi, les effets de taille finie ne sont pas dominants dans nos données, c'est-à-dire qu'un changement significatif des valeurs de $H_\eta$ n'est pas susceptible d'être observé avec l'accumulation de données. Enfin, nous notons que nos données de la figure 6 montrent que le laser atténué utilisé dans l'expérience surpasse les paramètres d'aléa de la source de photons intriqués. Cela indique que les sources de photons intriqués n'offrent pas un avantage inhérent à une configuration de QRNG en mode temporel.

Nous utilisons également les données de la figure 6 pour calculer la mise à l'échelle du débit binaire aléatoire à partir de la relation suivante:

$$\begin{aligned} \text{Bitrate} &= \text{Total Minimum Entropie * Flux de photons} \\ &= H_\eta * (log_2(N_v) + n) * Détecté\ taux\ Photon \qquad (11) \end{aligned}$$

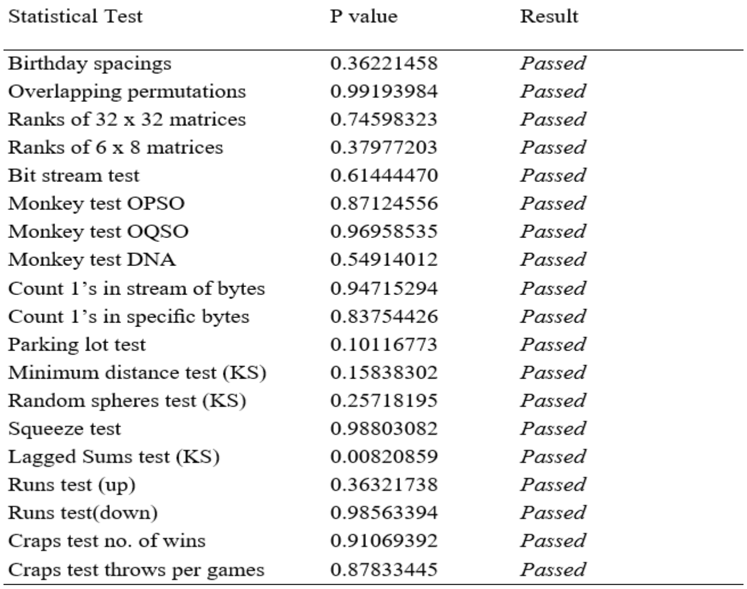
Les sources utilisées ont atteint un débit binaire aléatoire de 50 kb/s à 150 kb/s. Cependant, une valeur plus élevée du débit binaire pourrait être obtenue en utilisant différentes combinaisons de source et de détecteur plus rapides. On peut voir que tous les paramètres de l'équation 11 varient avec le nombre de séparateurs de faisceau dans le montage. Par conséquent, le débit binaire augmente ou diminue selon le paramètre dominant. Par exemple, lorsque les pertes sont faibles (~ 0,5 - 1 dB), les deux premiers paramètres de l'équation 11 augmentent beaucoup plus rapidement que la baisse du flux de photons. Par conséquent, dans de tels cas, il y a une augmentation nette du débit binaire avec l'ajout d'un séparateur de faisceau. Cependant, à mesure que les pertes deviennent plus importantes (~ 2-5 dB), la baisse du flux de photons devient le contribuant principal au débit binaire aléatoire. Il est donc nuisible d'ajouter plus de séparateurs de faisceau à ce montage dans ce cas précis. Sur la figure 4, la courbe pour le laser atténué atteint un maximum d'entropie min par bit de 0,948. Cependant, avec l'ajout du quatrième séparateur de faisceau, il n'y a qu'une légère augmentation, soit d'environ 0,5 %, alors que le flux de photons continue de diminuer. De même, la courbe des guides d'ondes PPLN augmente d'environ 5 % lorsque le nombre de séparateurs de faisceau passe de 2 à 3 avec une diminution identique du flux de photons. Cependant, comme le nombre d'états fait également augmenter le terme $(log_2(N_v) + n)$, on observe une hausse du débit de 3 % dans cette région. Ainsi, pour nos paramètres de source et de détecteur, le fonctionnement optimal du QRNG nécessite entre 2 et 3 séparateurs de faisceau. Cet effet est évidemment plus important lorsque la qualité de la source est inférieure, car, dans ce cas, une augmentation plus importante de $H_\eta$ est observée avec l'ajout de chaque séparateur de faisceau.

Pour obtenir les séquences de bits finales, un extracteur de type Toeplitz Hashing [24] est mis en œuvre afin de séparer le bruit dans les canaux du caractère aléatoire provenant de la source et des composantes. L'extracteur récupère une séquence binaire aléatoire de longueur $m$ à partir d'une séquence binaire brute de longueur $n$ en la multipliant par une matrice de Toeplitz de dimensions $n \times m$. Dans notre implémentation, nous avons choisi une grande valeur de $n$ (= 4096) pour restreindre les effets de taille finie en concaténant plusieurs séquences de bits brutes obtenues à partir du convertisseur numérique temporel (Time Digital Converter – TDC) (chacune de longueur 100). De plus, pour chacun des montages, une grande valeur correspondante de $m$ a été déterminé expérimentalement ($m \geq n * H_\eta,$). Les séquences de bits résultantes sont transmises dans une suite de tests Diehard qui détermine la qualité du caractère aléatoire statistique. Au total, 96 Mbit de données collectées à partir de différentes configurations (avec 0, 1, 2, 3 et 4 séparateurs de faisceau) ont été testés et l'ensemble des tests ont été réussis. À titre d'exemple, les résultats de l'un de ces tests sont présentés à la figure 7.

On peut noter que notre générateur produit un nombre aléatoire en raison de sa dimensionnalité accrue, ce qui le distingue des autres réalisations où un bit aléatoire est généré. Ainsi, pour vérifier le caractère aléatoire statistique de nos données, les tests Diehard sont plus pertinents, car ils sondent le contexte d'occurrence des séquences plutôt que les propriétés de la séquence elle-même (comme dans le cas des tests NIST). En d'autres termes, pour notre implémentation, il est plus pertinent de tester les corrélations à longue portée entre différentes chaines de bits plutôt que la répétition de valeurs de bits individuelles dans une chaine donnée. En outre, la réussite de ces tests assure que les nombres testés satisfont aux exigences statistiques d'un véritable hasard. Cela ne certifie cependant pas que les processus physiques impliqués dans leur génération sont vraiment aléatoires, car aucune information sur la source ou sur les composants n'est utilisée dans la conception de ces tests. De plus, des attaques sophistiquées peuvent être conçues pour générer des séquences de bits prédéterminées qui passent les tests statistiques

[25]. Par conséquent, l'adéquation de cette suite de tests ainsi que la pertinence de chaque test individuel doivent être déterminées avant d’attester de leur véracité dans une implémentation spécifique. De même, selon la mise en œuvre, une méthode de traitement de données simple ou plus complexe peut être employée.

Ma et *al*, fournissent une comparaison robuste des techniques d'extraction populaires et des débits binaires qui en résultent pour les QRNG [24].



| Statistical Test | P value | Result |
|---|---|---|
| Birthday spacings | 0.36221458 | *Passed* |
| Overlapping permutations | 0.99193984 | *Passed* |
| Ranks of 32 x 32 matrices | 0.74598323 | *Passed* |
| Ranks of 6 x 8 matrices | 0.37977203 | *Passed* |
| Bit stream test | 0.61444470 | *Passed* |
| Monkey test OPSO | 0.87124556 | *Passed* |
| Monkey test OQSO | 0.96958535 | *Passed* |
| Monkey test DNA | 0.54914012 | *Passed* |
| Count 1’s in stream of bytes | 0.94715294 | *Passed* |
| Count 1’s in specific bytes | 0.83754426 | *Passed* |
| Parking lot test | 0.10116773 | *Passed* |
| Minimum distance test (KS) | 0.15838302 | *Passed* |
| Random spheres test (KS) | 0.25718195 | *Passed* |
| Squeeze test | 0.98803082 | *Passed* |
| Lagged Sums test (KS) | 0.00820859 | *Passed* |
| Runs test (up) | 0.36321738 | *Passed* |
| Runs test(down) | 0.98563394 | *Passed* |
| Craps test no. of wins | 0.91069392 | *Passed* |
| Craps test throws per games | 0.87833445 | *Passed* |

**Fig 7 : Résultats de la suite de tests Diehard pour le cas de 4 séparateurs de faisceau**. On peut voir que tous les tests sont réussis et des résultats similaires ont été obtenus pour tous les cas avec des tailles différentes de la matrice Toeplitz, correspondant à la valeur mesurée expérimentalement de $H_{\eta}$.

Alors que la méthode décrite ici concerne les implémentations de dispositifs de confiance, le schéma peut être utilisé plus universellement, comme indiqué par les données de deux sources différentes, conduisant à la réalisation de QRNGs hybrides semi-autotestés. Les conclusions de ce travail, son impact et les travaux potentiels qui en découlent sont discutés dans les sections suivantes.

## Conclusions

Nous avons étudié l'entropie d'information minimale d'un photon par bit (qui est une estimation du caractère aléatoire total dans le générateur) en tant que paramètre de conception important des QRNGs photoniques. Nous avons ensuite réalisé expérimentalement un nouveau QRNG de confiance pour lequel le minimum d'entropie par bit peut être réglé et mis à l'échelle indépendamment de la source et du mécanisme de détection, avec des changements minimes au dispositif. Contrairement aux méthodes précédentes, notre QRNG augmente la flexibilité des états qu'un photon peut occuper en combinant les intervalles de temps virtuels utilisés dans un dispositif classique avec des intervalles de temps physiques (plusieurs chemins qu'un photon peut emprunter), à travers les diviseurs de faisceau en cascade. En connectant ou déconnectant les séparateurs de faisceau, nous sommes en mesure de régler l'entropie minimale par bit et d'optimiser la génération de bits aléatoires d'une manière simple et pratique. La robustesse de ce mécanisme de réglage est montrée en comparant la génération de bits aléatoires à partir de deux sources de photons uniques et par les simulations des imperfections dans la configuration qui contribuent à l'écart par rapport au comportement aléatoire réel. Nous montrons également que pour les QRNGs de dispositifs de confiance, il existe des régimes ou des modes opérationnels dans lesquels une simple source laser atténuée surpasse les sources de paires de photons annoncées en ce qui concerne le taux de génération de bits aléatoires, offrant des avantages à la fois en termes de vitesse, de praticité et de coût dans les dispositifs commerciaux, bien que dans certaines configurations (comme les mesures T2) les sources annoncées permettent une investigation simultanée de l'espace d'état entier des arrivées de photons.

## Travaux futurs

Bien que démontré pour les QRNGs de confiance, le mécanisme de mise à l'échelle introduit ici est prometteur pour le développement de QRNGs indépendants de l'appareil (Device Independant QRNGs – DIQRNGs) dits semi-autotestés hybrides qui offrent des degrés de confiance plus élevés via la certification apportée par la mécanique quantique. Comme indiqué par les données récoltées par deux sources de photons uniques très différentes, les caractéristiques de mise à l'échelle ne dépendent pas strictement de la source pour obtenir un véritable comportement aléatoire. À cette fin, cette méthode pourrait agir comme un mécanisme de mise à l'échelle universelle et peu couteuse pour améliorer les paramètres de caractère aléatoire dans n'importe quelle configuration de QRNG. De même, comme indiqué par les estimations d'entropie min et les tests statistiques, les performances de l'appareil ne dépendent pas strictement du nombre de séparateurs de faisceau utilisés. En d'autres termes, les performances de l'appareil sont indépendantes de sa conception. Par conséquent, cette façon de faire peut être utilisée pour implémenter les QRNGs semi-autotestés dans des configurations où la source et le dispositif de mesure sont indépendants [2]. De plus, alors que les deux ensembles de mesures utilisés dans l'expérience fournissent une estimation indirecte de l'espace d'état en caractérisant les deux vecteurs de base individuels, une reconstruction simultanée de l'espace d'état peut être mise en œuvre en utilisant le mode T2 de l'HydraHarp avec une source de paires de photons annoncés.

En outre, les caractéristiques de mise à l'échelle ont pu être observées grâce à l'étude de l'entropie min par bit d'un photon en tant que paramètre clé de la conception autonome. Cela montre que si le nombre total de bits obtenus à partir d'un dispositif n'augmente pas nécessairement, le nombre de bits utilisables par détection de photon peut être augmenté en raison d'une plus grande efficacité de la conversion aléatoire. Bien que ce paramètre ait déjà été mesuré expérimentalement, il n'a pas été étudié en tant que paramètre de conception indépendant. De plus, avec des designs de multiplexage plus complexes, d'autres métriques de conception similaires pourraient émerger. Cela permettrait d'obtenir de meilleures caractéristiques de mise à l'échelle. Ainsi, la poursuite des tests sur des designs similaires pourrait aboutir à des démonstrations expérimentales de protocoles d'expansion aléatoire dans des QRNGs bien caractérisés.

Enfin, avec les progrès rapides de l'intelligence artificielle et des algorithmes d'apprentissage automatique [26], les ordinateurs sont devenus extrêmement puissants pour déterminer des paramètres expérimentaux. Cette avancée peut être mise à profit pour identifier certaines relations qui ne peuvent être définies purement physiquement ou mathématiquement. L'analyse potentielle des données des QRNGs via les systèmes d'apprentissage automatique pourrait conduire à leur analyse comparative définitive. Ainsi, il serait possible d'aboutir à une mesure de performance minimale requise bien définie pour que différents générateurs soient utilisés dans diverses applications. Une telle analyse comparative est présentement difficile à réaliser simplement parce qu'il n'y a pas suffisamment d'informations sur la manifestation macroscopique du hasard. Par conséquent, une combinaison de la description physique, mathématique et statistique du caractère aléatoire associée à des modèles d'apprentissage automatique pourrait avoir un impact fondamental sur la recherche du caractère aléatoire. Plus précisément, en utilisant la méthode introduite dans cette thèse, il serait possible de comparer les QRNGs ayant des paramètres de performance et leur contrepartie améliorée par l'apprentissage automatique.

Les résultats d'une telle enquête pourraient indiquer si certains systèmes quantiques sont encore sensibles aux attaques des systèmes d'intelligence artificielle ou si le hasard de la mécanique quantique est vraiment indéterministe dans toutes les configurations. Les résultats récents de Truong et *al* sur la cryptanalyse par apprentissage automatique avec QRNGs sont très prometteurs pour une prochaine enquête sur ce domaine [27].

Ainsi, le travail présenté dans cette thèse procure le contexte et les motivations nécessaires pour poursuivre les recherches sur la génération de nombres aléatoires quantiques en photoniques.

## Références

[1] N. Metropolis and S. Ulam, Journal of the American Statistical Association **44**, 335 (1949).

[2] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, npj Quantum Information **2**, 16021 (2016).

[3] F. Raffaelli, P. Sibson, J. E. Kennard, D. H. Mahler, M. G. Thompson, and J. C. F. Matthews, Opt. Express **26**, 19730 (2018).

[4] M. Imran, V. Sorianello, F. Fresi, L. Potì, and M. Romagnoli, in *Optical Fiber Communication Conference (OFC) 2020* (Optical Society of America, San Diego, California, 2020), p. M1D.5.

[5] C. Abellán, W. Amaya, M. Jofre, M. Curty, A. Acín, J. Capmany, V. Pruneri, and M. W. Mitchell, Opt. Express **22**, 1645 (2014).

[6] Z. L. Yuan, M. Lucamarini, J. F. Dynes, B. Fröhlich, A. Plews, and A. J. Shields, Applied Physics Letters **104**, 261112 (2014).

[7] Y. Liu *et al.*, Physical Review Letters **120**, 010503 (2018).

[8] Y. Liu *et al.*, Nature **562**, 548 (2018).

[9] P. Bierhorst *et al.*, Nature **556**, 223 (2018).

[10] Y. Zhang *et al.*, Physical Review Letters **124**, 010505 (2020).

[11] M. Gräfe *et al.*, Nature Photonics **8**, 791 (2014).

[12] M. A. Wayne, E. R. Jeffrey, G. M. Akselrod, and P. G. Kwiat, Journal of Modern Optics **56**, 516 (2009).

[13] F. Xu, J. H. Shapiro, and F. N. C. Wong, Optica **3**, 1266 (2016).

[14] Y.-Q. Nie, H.-F. Zhang, Z. Zhang, J. Wang, X. Ma, J. Zhang, and J.-W. Pan, Applied Physics Letters **104**, 051110 (2014).

[15] M. A. Wayne and P. G. Kwiat, Opt. Express **18**, 9351 (2010).

[16] M. Kues *et al.*, Nature **546**, 622 (2017).

[17] M. Kues, C. Reimer, J. M. Lukens, W. J. Munro, A. M. Weiner, D. J. Moss, and R. Morandotti, Nature Photonics **13**, 170 (2019).

[18] S. Tanzilli, W. Tittel, H. De Riedmatten, H. Zbinden, P. Baldi, M. DeMicheli, D. B. Ostrowsky, and N. Gisin, The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics **18**, 155 (2002).

[19] M. Tomamichel and R. Renner, Physical Review Letters **106**, 110506 (2011).

[20] G. Vallone, D. G. Marangon, M. Tomasin, and P. Villoresi, Physical Review A **90**, 052327 (2014).

[21] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, Reviews of Modern Physics **89**, 015002 (2017).

[22] F. Furrer, T. Franz, M. Berta, A. Leverrier, V. B. Scholz, M. Tomamichel, and R. F. Werner, Physical Review Letters **109**, 100502 (2012).

[23] L. Oesterling *et al.*, Journal of Modern Optics **62**, 1722 (2015).

[24] X. Ma, F. Xu, H. Xu, X. Tan, B. Qi, and H.-K. Lo, Physical Review A **87**, 062327 (2013).

[25] A. Sarkar and C. M. Chandrashekar, Scientific Reports **9**, 12323 (2019).

[26] E. Trentin, F. Schwenker, N. El Gayar, and H. M. Abbas, Neural Processing Letters **48**, 643 (2018).

[27] N. D. Truong, J. Y. Haw, S. M. Assad, P. K. Lam, and O. Kavehei, IEEE Transactions on Information Forensics and Security **14**, 403 (2018).

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF EQUATIONS

# LIST OF SYMBOLS AND ABBREVIATIONS

| | |
|---|---|
| PRNG | Pseudo Random Number Generator |
| QRNG | Quantum Random Number Generator |
| RNG | Random Number Generator |
| SFWM | Spontaneous Four Wave Mixing |
| SPDC | Spontaneous Parametric Down Conversion |
| TRNG | True Random Number Generator |
| $H_{min}$ | Minimum Information Entropy or Minimum Entropy |
| $H_{\eta}$ | Minimum Entropy per bit |
| $H_{dev}$ | Entropy of non-idealities |
| $H_{max}$ | Maximum Entropy or Max Entropy |
| $\lambda_m$ | Decay constant of radioactive material |
| $\lambda$ | Laser intensity |
| $T$ | Detector deadtime |
| $N$ | Dimensionality of measurement basis |
| $N_v$ | Virtual time bins |
| $l$ | Spatial modes corresponding to the beam splitters |
| $\bar{n}$ | Average photon number |
| $T_{gen}$ | Photon generation time |
| $T1, T2, T3, T4$ | Temporal delays |

# 1 INTRODUCTION

## 1.1 Importance of random number generators

Events with random outcomes are not considered favourable in daily life. For example, systems in the cockpit of an aircraft should have fixed, known responses for specific inputs by the pilot. Performing financial transactions should reflect definite amounts in our bank accounts and not random sums of money. Even in generic, less critical settings such as a living room or an office space, randomness in the form of unorganized cupboards or haphazard seating arrangements is conventionally frowned upon. Despite a natural human preference for determinism and order, a world without randomness is simply unimaginable. In fact, randomness can be considered an essential natural resource that is central to many aspects of daily life. As Hayes [1] points out, there is an entire industry built around the availability and usability of randomness for various applications. Consider computer networks. When two nodes try to communicate with each other as part of an information exchange protocol, they may reach a "deadlock" where both nodes are waiting for the other to respond to their previous message. To break such a deadlock, each node is typically programmed to pick a random number between 1 and *n* (in units of time) after which it resumes operation. Computing has a whole class of algorithms called "randomised algorithms" [2] which find application in different contexts such as multiprocessing in operating systems [3], Very Large Scale Integration (VLSI) [4], signal processing [5] and data mining [6]. In addition, simulation methods like the Monte Carlo system (which is commonly used in physics) require high quality random numbers [7,8] to correctly model complex phenomena like molecular reactions [9,10] or changes in climatic conditions [11,12]. Moreover, cryptographic protocols such as the Rivest-Shamir-Adleman (RSA) use random numbers [13-17] to ensure privacy and security on the internet. Modeling techniques, such as those used to predict the spread and impact of the SARS-CoV-2 virus [18,19], require random sampling [20,21] of patients and test cases to aid governments and healthcare organizations develop contingency plans. Even recreational activities such as the Roulette, BINGO, the lottery or deciding who bats first in a cricket game are made possible due to randomness.

Given its extensive occurrence in natural phenomena and usage in human activities, randomness has been studied as a scientific and mathematical concept over many centuries [22]. While the initial investigations (dating back to several thousand years) were more philosophical in nature [23], recent works have been focused on random number generation techniques for practical uses, especially in computing applications [24,25]. Section 1.2 provides a more detailed history of random numbers and their study. State-of-the-art Random Number Generators (RNGs) that are used in most commercial applications today, are either complex physical systems such as Complimentary Metal Oxide Semiconductor (CMOS) integrated semiconductor devices [26,27], or mathematical models with certain statistical properties [28]. Both these approaches for random number generation are described in detail in section 1.4. While the performance parameters of these methods (such as fast rate of production and producing mutually independent numbers) meet application requirements, they are known to be pseudo-random, i.e. the exhibited behaviour appears random enough for the specific application but is not truly random from a physical or mathematical standpoint. This is because the physical devices are classical systems which are fundamentally deterministic, while the algorithms are periodic by definition. This means that they cannot be used to accurately represent truly random phenomena. Nonetheless, the impact of this pseudo-random behaviour can be minimized (e.g. by defining an exceptionally large seed length in the algorithmic approach), so that the performance of protocols that employ these methods is not compromised. This is perfectly illustrated in the implementation of the RSA algorithm, where the public keys generated are so large that not even the most powerful supercomputers today can factor them in a realistic amount of time [29-31].

However, it has been long known that quantum mechanical systems can improve over the performance of classical systems owing to the inherent indeterminism, and phenomena such as superposition and entanglement [32,33]. Since the nature of quantum information [34] is fundamentally different from classical information, it has captured the imagination of physicists and computer scientists alike since the latter half of the 20th century. As such, many advantages of quantum technologies over their classical alternatives have already been demonstrated theoretically. Specifically, Peter Shor developed a landmark algorithm in 1994 which can factor large prime numbers in polynomial timescales compared to exponential timescales in classical systems [35]. Similarly, employing Grover's search algorithm [36] lead to an order of magnitude reduction in the querying time of database searches. More recently, there has been a rapid improvement in the capabilities of quantum hardware, thanks to renewed interest from various universities, governments and industries. As silicon technologies approach the limit of their computing power [37,38], there is a concerted effort across the world to find alternatives that can allow us to make progress in areas like data privacy, cryptography and information processing at the same rate as the last few decades. This is reflected, not only in ambitious projects being undertaken at major universities but also by government policy initiatives such as the US National Quantum Initiative [39], European Quantum Flagship Initiative [40], and Quantum Canada [41], which have collectively invested billions of dollars in commercializing quantum research. In addition, leading technological companies such as IBM and Google already have different teams working on areas ranging from the quantum internet to quantum chemistry. As a direct result of these efforts, initial prototypes of Noisy and Intermediate Scale Quantum (NISQ) computers (IBM Q-series, D Wave) are already available for commercial use via the cloud [42].

As these technologies mature and become widely available, they will pose a very serious threat to state-of-the-art classical systems [43]. The RSA algorithm for example, which is central to cryptographic systems used in banking, data privacy on the internet etc., may be broken by a quantum computer [44]. This necessitates the development of a robust “quantum-proof” encryption and communications strategy ahead of time, so that alternatives can quicky replace compromised systems. Thankfully, quantum mechanics which gives rise to this threat, also provides a solution. The insufficiency of current systems against quantum-enhanced attacks is due to the fact that the RNGs being used in their implementation are pseudo – random [45]. If in a cryptographic method, there was no way to factor a public key of any length in practical timescales, such an encryption algorithm could be trusted to be truly secure. In fact, the limitations of RNG performances due to their pseudo random behaviour and their impacts have been well researched, which has led to the development of genuinely random True RNGs or TRNGs [46,47]. Specifically, quantum mechanical systems can exploit the inherent indeterminism to generate numbers in unpredictable patterns, giving rise to Quantum Random Number Generators (QRNGs) [48]. Chapter 2 discusses the field of quantum random number generation in detail covering QRNG designs, implementations, performance parameters, their advantages, and drawbacks. Since research in this direction began as early as the 1960’s [49], state-of-the-art systems [50,51] already employ advanced photonics techniques, mathematical calculations and manufacturing processes as a result of the vast literature and knowledge developed over the decades. Consequently, there exist commercially available prototypes of QRNGs such as IDQuantique’s Quantis which can be purchased as a USB device for interfacing with regular laptops and desktops [52]. Furthermore, the use of such QRNGs in full fledged Quantum Key Distribution (QKD) and quantum communications protocols is also well researched, especially in recent years [53,54]. However, some concerns such as scalability of design, reliability and certifiability of randomness are yet to be fully addressed. Due to the lack of a unified solution, QRNG usage in mainstream technological applications remains limited and has only recently seen progress with Samsung releasing the first class of smartphones with integrated QRNG technology in May, 2020 [55].

In addition to being a cybersecurity solution, the study of QRNGs and surrounding principles is pertinent to investigations in several areas of natural sciences. For example, QRNG experiments have proved to be a useful tool in the study of nonlocality and experimental demonstration of loophole free Bell tests [56]. In biological systems, the study of molecular processes like RNA folding, protein structures, ion exchanges etc. can be greatly enhanced by the genuine randomness in QRNGs [57,58]. Similarly, the accuracy of prediction models used in atmospheric sciences, which are crucial to contingency planning in case of natural disasters, can be boosted using genuine RNGs [59]. Since the techniques impacted by using QRNGs in place of classical RNGs (such as simulation, numerical and statistical methods) are fundamental to the study of several areas of science, QRNG research can have a strong impact across a wide range of subjects.

In order to address difficulties in QRNG characterization and design, this thesis introduces a new parameter, the minimum entropy per bit, which can be used to optimize and scale-up randomness in photonic QRNGs. The manifestation of this parameter in experimental setups is demonstrated by realizing a simple, fiber-integrated, photonic QRNG whose bitrate can be increased while minimally affecting the design complexity. To provide the necessary background and context, Chapter 1 covers the history of randomness studies and the scientific methods that have been used to understand its properties over centuries. It also provides the mathematical and physical background required to study randomness generation schemes discussed in the next chapters. Chapter 2 provides an in-depth overview of quantum random number generation. It describes the different types of QRNGs – optical, non-optical etc. and the broad design considerations, advantages and drawbacks of each. It specifically covers state-of-the-art photonic QRNGs and their classification into categories like trusted-device, self-testing, and semi self-testing, along with the important performance parameters. Chapter 3 discusses a solution to the problem of scaling QRNG architectures, which is identified through the literature review in chapters 1 and 2. The novel scaling mechanism introduced here can enhance the performance of commercial QRNGs without a significant increase in cost and complexity of the device. This scheme uses simple off-the-shelf telecommunication components like 50/50 beam splitters and temporal delays to achieve an increase in both speed and security of the device in a controllable way. The observed performance improvement is especially high for non-ideal source/detector parameters and environmental conditions. In addition, this scheme is compatible with different embodiments (polarization, frequency, radioactive decay etc.), integration platforms and source – detector combinations, which makes it well-suited for commercial prototyping. The post-processing methods required to verify genuine random behaviour, experimental results and their usefulness in different applications are also discussed. Finally, the thesis concludes with Chapter 4 providing a summary of quantum random number generation and an overview of the work carried out as part of this Master’s thesis along with its potential impact on present-day random number generation research. It also discusses future works that can arise from an extension of the presented scheme, both for fundamental physics experiments and commercialization efforts.

## 1.2 History of random number generation

Despite the heavy reliance of many modern-day applications on random numbers, their usefulness and methods of production are not recent discoveries. In fact, the earliest methods utilized to generate randomness made use of primitive devices such as coins and dice, which were first used in ancient Rome, India, China and the Middle East between 4000 and 6000 years ago. Dice excavated from archaeological sites in these locations reveal imprints on each face, indicating the different outcomes, in place of numeric representations. As Pierre L'Ecuyer notes [60], this indicates that randomness and its uses were understood even before a formal notation was developed for representing numbers!

It is interesting to note that in addition to physical devices that produced random outcomes, there also existed philosophical debates about randomness and determinism in the context of the nature of the universe and their consequences to human actions, dating back to the pre-Socratic era. Leucippus and Democritus believed that the world was fundamentally deterministic and certain events appear to be random either due to a gap in our understanding or an inability to comprehend the underlying mechanisms [61,62]. On the contrary, Epicurus believed that there does in fact exist true indeterminism in nature, which makes aspects of human life such as free will possible [63]. Over centuries, both these schools of thought were developed to give rise to Epistemic and Ontic interpretations of randomness, with Epistemic philosophers believing the former while the Ontic philosophers argued the latter [64]. This philosophical context is relevant even to more recent studies of randomness. For example, initial scientific investigations into the quantitative study of seemingly random human behaviour (which started between the 15th and 16th centuries) were motivated by similar socio-political reasons. 'Social numbers', such as mortality rates and birth counts, were formulated in the late 1600s to help guide political policy and their distributions by region and gender were studied to derive insights into human social behaviour. A crucial result by Laplace in 1781 – where he showed that the Gaussian curve, primarily used for error calculation in scientific experiments, can also correctly model social statistics – marked the convergence of sociological studies with experimental science. Apart from giving rise to probabilistic models in classical physics (such as the distribution of velocities in Maxwell's kinetic theory), investigations of randomness in later centuries also helped shape the debate around consequences of a quantum world - giving rise to theories like the many worlds interpretation of quantum mechanics [65,66]. This debate in fact saw participation from some highly noted quantum physicists such as Einstein, Schrodinger, Bohr and others who famously debated the origin, nature and interpretations of quantum randomness [67]. In fact, the question of defining true randomness in quantitative, measurable terms remains a subject of interest in quantum research even today.

Despite a historical interest in randomness, methods to produce random numbers saw significant improvements (beyond the use of primitive devices such as dice and coins) only in the late 20th century – coinciding and being largely driven by advances in the computing industry. As late as 1890, Francis Galton, a noted statistician who worked on preliminary methods of randomness expansion, remarked: “As an instrument for selecting at random, I have found nothing superior to dice” [68]. Even in the late 1930’s, published tables containing digits randomly sampled from sources like census reports were being used for statistical experiments [69,70]. In a landmark contribution in 1938, Kendall and Babington-Smith not only questioned the veracity of such tables and their usefulness in experiments, but also designed the first machine to produce random numbers, made up of a rotating cardboard disk with divisions [71]. They argued that any number of a given size appearing in a table of random digits has, in theory, the same probability of occurrence as any other and thus, no two tables can be compared for quality of randomness. They further proposed the first set of statistical tests for verification of randomness by defining sub-sequences with specific properties (such as number of successive 1’s and 0’s) that had to be met by the complete sequence for it to be considered random. This work is particularly significant as some of these statistical tests are still in use today, for example in the NIST test suite, to verify true random behaviour. Extending their work, in 1955, the RAND corporation published the first book with a million random digits generated in a completely automated fashion using physical equipment [72]. This table was also made available in the form of punched cards for use by emerging electronic computing equipment in other scientific programs and experiments, such as in FORTRAN.

As the nascent computing industry grew between the 1940s and 1950s, it was no longer sufficient to read random numbers from tables or external storage devices. Additionally, the size of the main memory in these early electronic computers was too small to store large tables or punched cards. This imposed a limitation on the kind of tasks that stored random numbers could be used for. For example, physicists who wanted to simulate large systems with complex variables and state descriptions could not use these punched cards to run Monte Carlo simulations. To overcome these, two broad strategies were employed in the early 1950’s which continue to form the basis of random number generation techniques today: 1) Design and create a fast physical device capable of producing random numbers on demand or 2) Design a mathematical algorithm that outputs numbers which look random under a given set of operating conditions. For example, a module 10 operation can result in random numbers in the interval (0,9). These techniques are discussed in greater detail in section 1.3.

## 1.3 Definition of randomness

Despite being studied for centuries, providing a single, complete definition of randomness remains a complex scientific challenge [73,74]. This is because a random process involves both physical phenomena and has mathematical properties. While physical and mathematical requirements of randomness can be independently defined, it is harder to formulate a definition which encapsulates both sets of requirements. For example, a bit sequence can be considered mathematically random if it satisfies certain statistical properties [75]. However, from a purely statistical study, no insights can be derived on the physical processes that result in the bit sequence. Similarly, if a process produces uncorrelated events which follow a Poisson distribution, it can be considered truly random from a physics standpoint [76-78]. However, defining correlations can be a challenge as this requires mathematical analysis of observed phenomena. Furthermore, defining genuine randomness looks like a paradoxical problem: if a process is truly indeterministic, there should not be any measurable parameters which indicate this property. Hence, as discussed in more detail in Chapter 2 in the context of self-testing QRNGs, certifying that a process in truly random remains a difficult scientific problem to solve. Even so, it is especially important to develop a working definition of randomness, given the extensive use of random numbers in several technologies. This section provides some mathematical notions of randomness, which can help evaluate whether a given set of data is random enough for a specific application. Although these definitions may not be mathematically complete, they provide insights into the statistical requirements for the usability of bit sequences for various applications.

As Kolmogorov *et al* defined in their work in 1988 [79], a truly random sequence should have the following properties:

1. Being typical:

   Sequences such as 0010110100 or 11010001010 are considered typical whereas a sequence like 00000000000 is considered atypical for studies of randomness. In other words, if a sequence appears to have special properties (like only 0's or only 1's), it is considered atypical. It is important to note however, that all the three bit sequences mentioned above have the same probability of occurrence. Hence, it is incorrect to conclude that the third bit sequence is atypical because it is less likely to occur. Rather, the property of being typical is the property of belonging to any reasonable majority. On choosing a sequence at random, we can have the justifiable expectation that the sequence is typical [79].

2. Being chaotic:

   A truly random sequence is chaotic in that it has no simple law governing the alternation of terms throughout its length. In other words, every bit position within a sequence of a given length can be filled in any number of ways as allowed by the bases of the observed system.

3. Stability of frequencies:

   The frequencies of 0's and 1's in the bit sequences as a proportion of the total string length must converge to ½ if they are equally likely to occur. Further, this property must be true for any properly chosen substring of the bit sequences as well. Bit strings that posses this property are called stochastic sequences.

In applying these definitions for evaluation of randomness, the following are to be noted:

- These conditions on random sequences are not mathematically complete, i.e. there exist other requirements for verification of genuine randomness. However, they serve as good indicators for initial estimates of randomness. If a sequence fails to demonstrate these properties, there is a high likelihood that it will also fail to meet other conditions, although this is not certain.
- These tests assume a uniform Bernoulli distribution of bits. However, as Knuth *et al* have shown [80], RNGs can be designed to produce any distribution of numbers which can be converted to a uniform distribution for statistical analysis.
- These definitions provide only an information theoretical understanding of randomness and provide no insights on the physical processes that may have resulted in the bit sequences. Hence, the suitability of these definitions for a particular implementation of an RNG is to be determined on a case-by-case basis. In other words, it cannot be concluded that a bit sequence is not random if it fails to conform with any of the above definitions. Conversely, it cannot be guaranteed that meeting these definitions implies true randomness.
- These definitions only look for properties within a bit sequence. However, in practice, it may be more meaningful to consider the appearance context of a given bit string for studies of randomness. For example, the number 3 has the bit representation '11'. Looking only at the bit sequence, it would be meaningless to conclude that 3 is a random

number. It would be more relevant to consider the other bit strings that are obtained from the same process that produces the number 3. Hence tests of true randomness should not only look the statistical properties of bit sequences, but also the context of their occurrences.

A more comprehensive description of statistical requirements of random sequences can be found in the NIST and Diehard instruction manuals [75,81], which are briefly discussed in Chapter 3.

## 1.4 Types of random number generators

Since the inception of sophisticated random number generation techniques in the mid-1900's, many different types of RNGs have been implemented both for theoretical studies and for practical applications as outlined in section 1.1. These devices can be grouped into two main categories based on their design strategies, namely, algorithmic RNGs and physical RNGs. State-of-the-art classical systems employ either of these methods depending on the specific needs of the application. Each of these classes can be further categorized based on metrics such as source of randomness, performance parameters, and ease of use. This section details the physical and mathematical concepts used in the design and implementation of classical RNGs, their advantages, drawbacks and suitability for practical application. In discussing the fundamental limitations of classical systems for random number generation, it also explains the advantages of indeterministic quantum systems, which are covered in detail in Chapter 2.

### 1.4.1 Algorithmic Random Number Generators

Most RNGs used in commercial applications today are algorithmic RNGs in which mathematical models act as the source of randomness. For example, when the method rand(100) is executed on Matlab, the output random numbers are not produced by a physical device inside the computer, but are generated via mathematical manipulations. The most popular algorithmic RNG is the Lehmer RNG, named after D.H Lehmer who introduced this method in 1951 [82].

Lehmer RNGs are special cases of Linear Congruential Generators (LCGs) which yield pseudo random sequences according to the relation:

$$X_{n+1} = (aX_n + c)mod(m) \qquad (1.1)$$

Where $X_i$ is the $i^{\text{th}}$ digit of the sequence,

$\boldsymbol{m} > 0$ is the modulus,

$m > \boldsymbol{a} > 0$ is the multiplier,

$m$ > $\boldsymbol{c}$ ≥ 0 is the increment and

$m$ > $\boldsymbol{X_0}$ ≥ 0 is the seed or the start value.

The integer constants $m, a, c$ and $X_0$ collectively define the generator. An LCG with $c = 0$ is called a Lehmer RNG or a Multiplicative Congruential Generator (MCG). From equation 1.1, it is evident that there exists a period after which a certain integer value repeats in this generator. This is a concern as in an ideal RNG, there should be no correlations between any two generated numbers. However, by choosing appropriate values for the characteristic parameters, a long period can be defined such that there is no repetition of numbers within that window. Hence, the quality of random numbers generated by Lehmer RNGs or indeed LCGs is extremely sensitive to the choice of values for the parameters $m, a, c$ and $X_0$. In fact, it has been shown that a poor combination of these values can lead to questionable results in experiments that employ such RNGs [83,84]. Fig. 1.1 illustrates this concept by showing the working of three different generators. From the first two cases (top and middle generators), it can be seen that a change in seed value (from 1 to 3) results in a change in the repetition length (from 5 to 1). Different combinations of parameter values hence lead to different performances in such RNGs.

Since periodicity and determinism is built into the definition of this method, it is obvious that algorithmic RNGs can only produce pseudo randomness. However, the distribution of numbers obtained through these methods already satisfy statistical requirements of randomness. Hence, it may be advantageous to use these models in applications where a fast rate of production is more important than a high quality of randomness since PRNG techniques are typically faster than TRNG methods. In addition, these results are reproducible and hence this method is useful when the set of generated random numbers needs to be the same, in order to compare different systems that use them (e.g, comparing simulation models).

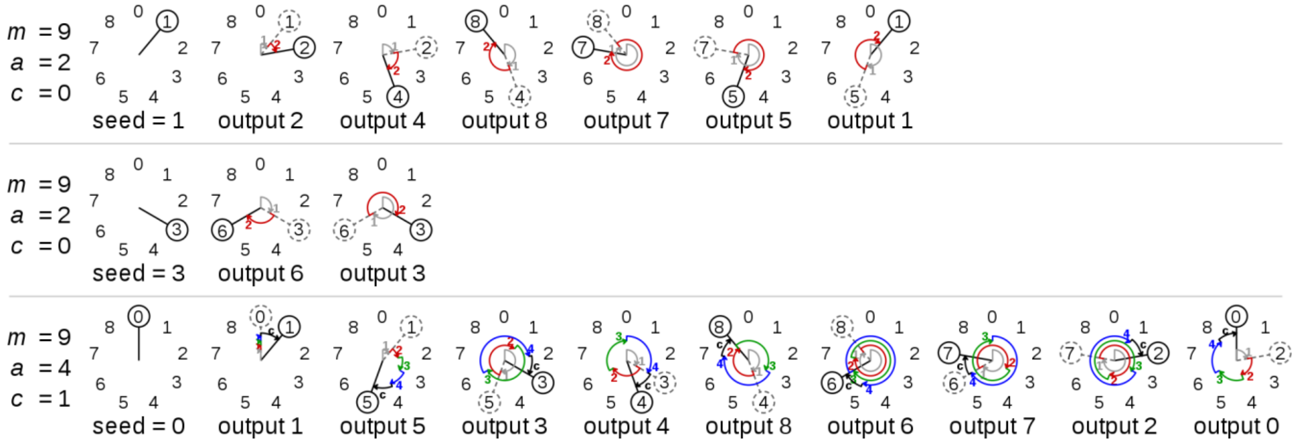


**Figure 1.1 Evolution of three LCGs with different values of $m, a, c$ and $X_0$.** The first two generators produce different period lengths (6 and 2 respectively) for different seed values despite same *m,a, and c* values. The last generator produces a longer period of length 9.
By Cmglee - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=38637545

### 1.4.2 Physical Random Number Generators

As explained in section 1.1, the need to develop physical systems capable of producing random numbers on-the-fly arose out of the shortcomings associated to reading pre-generated numbers from punched cards and memories. Early implementations of such devices used electronic noise as the source of randomness combined with a sampling scheme that produced numbers at regular intervals [85]. These schemes typically monitored a signal periodically and output bits whose values depended on some signal parameters. For example, a famous device named ERNIE (Electronic Random Number Indicator Equipment) that could produce about 50 bits of information per second was used to determine the winning numbers in the British Lottery in 1957 [86]. In this machine, current was produced by applying a high voltage at each end of a glass tube filled with neon gas. Collisions between electrons and atoms of neon made the current noisy, which was then amplified and finally sampled to produce random digits [60]. Several thousand numbers at each draw were produced during the lottery using this method. Since then, thousands of prototypes for physical generators have been proposed, making use of phenomena such as the photoelectric effect, thermionic emission and radioactive decay as sources of randomness [87]. In addition to producing numbers on demand, a key advantage of physical generators is that they can produce genuine randomness by employing phenomena that exhibit intrinsic indeterminism. However, the early implementations were all classical systems that are fundamentally deterministic. This meant that the numbers being generated were still pseudo random, even though they passed all the statistical tests such as uniformity and regular frequencies. In other words, although there seemed to be no apparent pattern in the generated numbers, the working of the device could be exactly understood using mathematical and physical descriptions of the processes involved in their production. This information could then be used to control the output of the device. Additionally, physical imperfections and environmental factors such as noise introduced slight biases or correlations in the generated sequences of numbers. Hence, von Neumann once famously remarked “Anyone who attempts to generate random numbers by deterministic means is, of course, living in a state of sin”. [87]

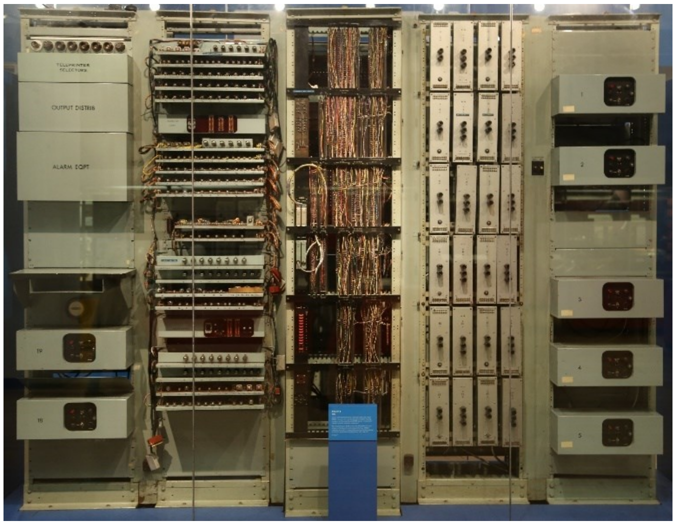


**Figure 1.2 ERNIE 1, used to generate the premium bond lottery numbers.** This device was used till 1972 after which it underwent several modifications. The most recent version ERNIE 4 is in use since 2004. By Geni – Photo by user:geni, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=20243440

Since both algorithmic methods and physical RNGs (classical systems) exhibit deterministic behaviour, they belong to the class of pseudo random RNGs or PRNGs. However, unpredictability is a key requirement in many applications like cryptography and communications which has pushed recent efforts towards the design of systems that can produce genuine randomness. Two strategies have been widely employed to obtain truly random numbers:

1. Using post-processing methods to reduce the bias and imperfections introduced by classical systems. For example, a bitwise XOR operation can be performed on two blocks of bits by aligning their Most Significant Bits (MSBs). This ensures that correlations in corresponding bit positions are eliminated [88]. Similarly, a modulo *b* operation can be performed to reduce biases as shown by Horton and Smith [87]. More sophisticated techniques make use of hashing extractors such as von Neumann’s extractor, Trevisan

extractor, Toeplitz matrix etc. which have all been studied and implemented [89] in combination with different PRNG schemes. One commonly used postprocessing method (the Toeplitz Matrix) is discussed in Chapter 3. It is regular practice to use a PRNG in combination with any of these methods for practical applications.

2. Using indeterministic natural processes as sources of randomness. This approach is particularly promising since it reduces the post processing requirements. This also means that such devices can be directly deployed in industrial infrastructures as standalone systems. Hence, several true RNG or TRNG schemes have been implemented using Johnson's noise [90], laser phase noise [91], chaotic systems [92] etc. For example, Intel implemented an RNG using Johnson's noise in a limited series of computer processors which used the amplified thermal noise of a resistor in combination with a latch-based sampling mechanism to produce random numbers [93]. Similarly, Free Running Oscillators (FROs) [94] in which electronic oscillations are created by applying the output of an inverter circuit to its input (thus creating indeterminism at the final output) have been used in 3rd and 4th generation FGPA, CPLD and ASIC hardware for cryptography, as well as in VIA C3 processors [87,95].

While electronic, logic-based, physical RNGs can produce acceptable performances (in combination with post-processing methods, as needed by present-day applications), they have some limitations which makes them susceptible to failure. For instance, Intel's RNGs released in 2011 [96] which comprise of two Yin-Yang circuits as sources of randomness are extremely sensitive to the speed and strength of inverters. In FRO based implementations, multiple oscillators on the same chip are subject to phased interlocking [90] which leaves them vulnerable to attacks from external electromagnetic radiation. These limitations not only require additional processing stages in RNG architectures for satisfactory performances, but also necessitate the development of alternate schemes that do not pose similar security or performance challenges. On the other hand, quantum mechanical systems exhibit truly indeterministic behaviour without the need for extraneous physical or mathematical manipulations [97]. Hence, they are a natural choice for the implementation of TRNGs. Quantum RNGs or QRNGs are hence a special case of TRNGs. The first QRNGs were implemented using radioactive systems in the 1960s [98]. While the initial implementations were mostly academic studies, there has recently been a renewed interest in commercial QRNG prototypes due to increased investment in quantum technologies. As a result, numerous architectures have been explored to create devices which can produce fast and verifiable randomness. QRNGs are of particular interest, as they are not only important standalone components, but also play a crucial part in facilitating other quantum technologies and protocols such as QKD and multi-node quantum communication networks [99,100]. Furthermore, attempts to implement device-independent QRNGs [101], i.e, devices which can produce certifiable, genuine randomness irrespective of device imperfections and environmental conditions, has led to landmark demonstrations in fundamental sciences [102-104]. Chapter 2 covers in detail the popular implementation techniques of QRNGs, their advantages, shortcomings, and usability in daily applications.

# 2 TYPES OF QRNGs

## 2.1 Classification of QRNGs

Physical QRNGs can be grouped into categories based on many factors such as embodiment (optical, electronic, radioactive decay etc.), degrees of freedom (polarization, arrival time, spatial mode, spectral mode, etc.) and potential uses (academic or commercial). While each of these classes have been extensively studied, current QRNG research is primarily focused on photonic implementations, as they benefit from the high bandwidth, low footprint, experimental simplicity and robustness typical of optical technologies, such as ultrafast laser sources, fiber-integrated components and micro-optical chip-based platforms. As a result, physical QRNGs can be broadly categorized as optical and non-optical implementations for the purposes of academic classification. The reviews of Stipcevic and Koc (2014), Xiongfeng Ma *et al* (2016), and Herrero-Collantes and Garcia-Escartin (2017) exhaustively discuss the overarching design considerations for various QRNG types, their working principles, advantages, and limitations. This chapter outlines the main ideas discussed therein with a specific focus on photonic implementations and the current state-of-the-art optical QRNGs. The usability of each of these generator types for different applications is also reviewed.

## 2.2 Non-optical systems

### 2.2.1 Radioactive Decay

The first physical QRNGs were based on radioactive decay [105-107] as this is an easily accessible natural phenomenon (for e.g. through Geiger Muller tubes) which is intrinsically random. Many design principles used in the development of state-of-the-art optical QRNGs are borrowed from radioactive systems.

Most QRNGs based on radioactivity use the detection of $\beta$ - radiation as their source of randomness [88]. A detection event produces a pulse which is then translated into a digit. In a sample of radioactive material, the probability of decay follows an exponential random distribution of the form:

$$P(t) = \lambda_m e^{-\lambda_m t} \tag{2.1}$$

where $\lambda_m$ is the decay constant of the material.

Assuming that these devices operate at timescales much smaller than the half-life of the sample and under fixed experimental conditions (e.g. position of the sample, state of the gas in the Geiger

Muller tubes), the decay events have been shown to follow a Poisson distribution in a fixed interval of time [88]. Further, the time intervals between detected pulses also follow an exponential random distribution and are known to be independent [108] (i.e., they are uncorrelated events). Hence, from the physical characterisation of detection events, we expect that their temporal properties can be used to design a TRNG (as explained in section 1.2).

To encode the detection events as random numbers, a digital counter is used in combination with either a fast or a slow clock. A fast clock is defined as a signal with frequency $v > \lambda$ ( $\lambda$ is the mean rate of detection corresponding to $\lambda_m$ in the exponential random distribution) while a slow clock is defined so that $v < \lambda$ . The counter increments its value each time a signal is received and can be reset to start from 0. The timing diagram in Fig. 2.1 illustrates one possible encoding mechanism employing a slow clock as the reset signal and detection events as the increment signals to the counter, resulting in numbers 3, 2 and 0 as the outputs.

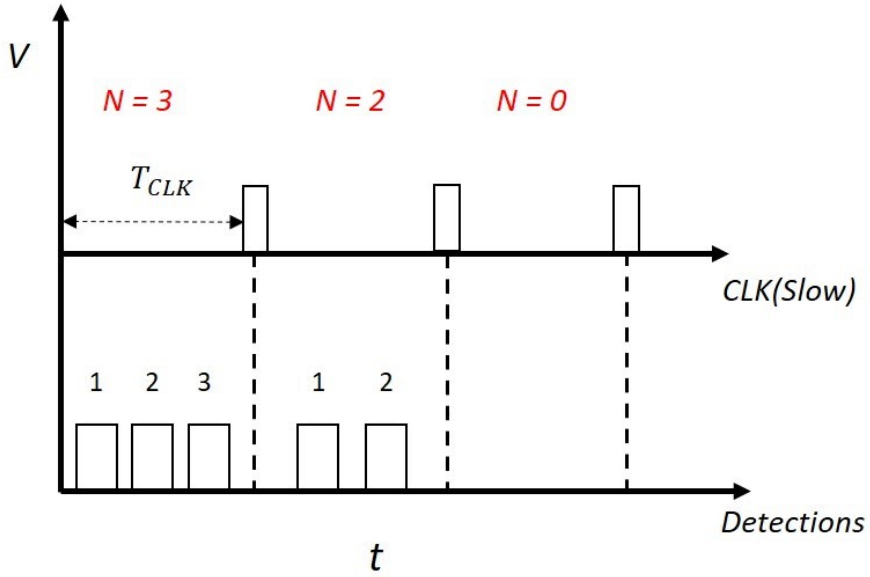


**Figure 2.1 Sampling scheme of a radioactive QRNG with a slow clock.** The pulses at the bottom correspond to detection of *β* radiation, which increase the value of the counter. A slow clock indicates the read-out operation, and the counter is reset to 0 for the next measurement window. In each window, the resulting number corresponds to the number of particle detections.

In this scheme, a counter increments its value each time a particle is detected. The value of the counter is monitored at regular intervals (as defined by the period of the slow clock), which corresponds to the number of detection events in that time window. Once this is read-out, the counter is reset to zero and the next measurement begins. This method has been employed in the scheme of Schmidt *et al* [109]. The converse of this method, as shown in Fig. 2.2, can also be used to encode detection events. In this scheme, a fast clock is used to increment the value of the counter and a particle detection event signals the read-out, resulting in numbers 2, 2 and 3 as outputs. This technique has been used in the works of Ishida and Ikeda [105] and Vincent *et al* [107].

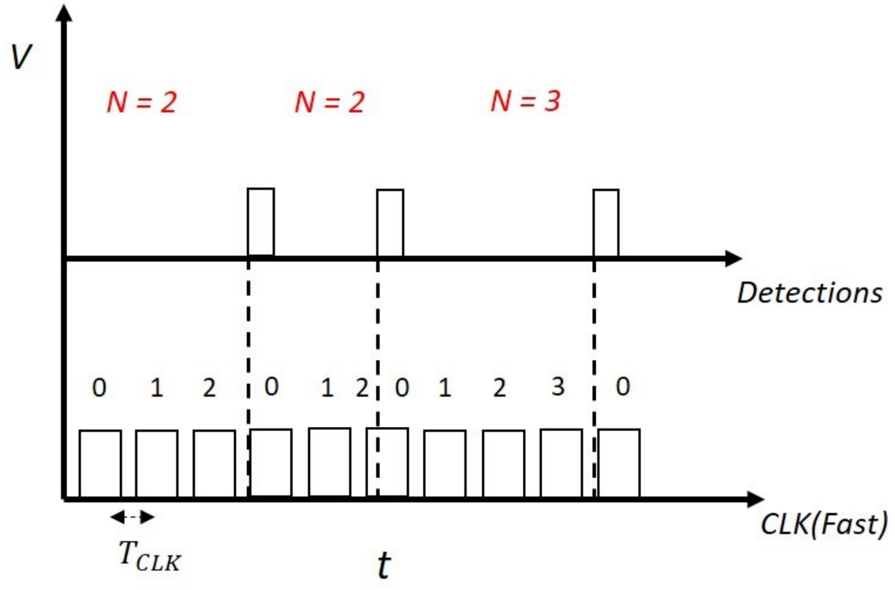


**Figure 2.2 Sampling scheme of a radioactive QRNG with a fast clock.** The pulses at the bottom correspond to a fast clock which signals an increment in counter value. A particle detection event (shown at the top) signals the read-out and reset of the counter. The resulting digit in each window is one less than the number of clock pulses.

Although these initial prototypes have undergone several modifications, the underlying techniques still remain popular. For example, the web-based server “Hotbits” (deployed in 1996) uses the pairwise time difference between arrival events in a radioactive system to produce random numbers [110]. If the difference in arrival times of the first pair is greater than that of the second pair, bit ‘0’ is recorded. On the contrary, if the difference in arrival times of the first pair is smaller than that of the second pair, bit ‘1’ is recorded. In subsequent models, Geiger counters have been replaced by semiconductor devices like PIN photodiodes [111]. These devices are safer and more convenient to use, as they do not require the same high voltages as Geiger detectors, although the strength of their output signals is much lower. Other proposals, like that of Duggirala *et al* [112], attempt to convert the exponential random distribution into a uniform distribution (as required in many RNG applications), and implement a uniform RNG with the use of electronic RC circuits. While radioactive systems provide a method to study truly random phenomena which can be translated into bits of information, these models have some severe limitations which prevent their usage in practical applications. The biggest limitation is the need for a highly radioactive source. Most demonstrated QRNGs use materials such as Cobalt-60, Strontium 90, and Nickel 63, which require additional safety equipment for proper deployment

[88]. This means that their integration with computing systems is not straightforward. Next, the components used in these setups such as Geiger tubes or semiconductor detectors degrade over time and with extensive use [88]. Hence, these devices require routine maintenance, and the outputs must be constantly monitored for correctness. While the impact of component degradation on the randomness of the output bits may be minimal in the short term, their quality may deteriorate significantly with prolonged use. Furthermore, due to biases in the physical processes, apparatus as well as the sampling scheme, the obtained bits require postprocessing to satisfy the statistical conditions of true randomness. Finally, even well-calibrated radioactive QRNGs have very low rates of production (few hundred kilobits per second) as defined by the decay rate of the sample and the recovery time of the detectors. Hence, their use in modern computing applications is fairly limited.

### 2.2.2 Electronic Systems

Noise in electronic systems is another source of randomness that is easily accessible for RNG design. Various types and sources of noise such as Johnson's noise [90], Zener noise [113,114], laser phase noise [91], and chaos noise [92] can be utilized in combination with commonly used circuit elements like resistors, capacitors, amplifiers, and op-amps to design efficient TRNGs. Fig. 2.3 shows the conceptual design of an electronic QRNG.

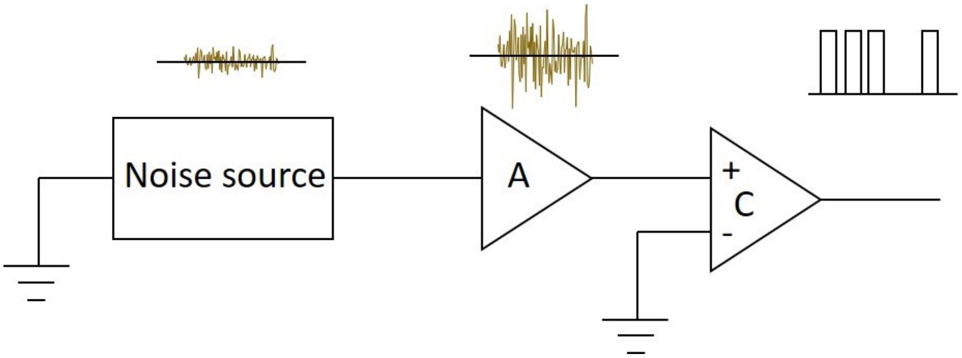


**Figure 2.3 Conceptual representation of an electronic RNG.** The design of electronic RNGs broadly consists of a source of randomness, an optional amplifying stage (A) and a comparator (C) which checks the amplified signal against a standard reference to encode bits

Voltage fluctuations in a circuit element caused by effects such as reverse bias in Zener diodes [114] or inverse base-emitter breakdown in bipolar transistors [115,116] are used as the sources of randomness. For instance, Johnson's noise which arises from the random thermal motion of quantized charges results in random voltages at the ends of any resistive material [90]. Similarly, in semiconductor Zener diodes, the tunnelling of charge carriers across quantum barriers of a particular height and width appear as voltage peaks [88]. This results in “pink noise” which is ideally suited for randomness applications. The voltage signals are then amplified to the desired level and compared with a threshold. If the signal crosses the threshold, bit ‘1’ can be obtained and bit ‘0’ is otherwise generated (or vice versa). In another encoding mechanism, instead of directly sampling the voltage signals, pulses are generated when the voltage crosses the threshold. In this scheme, the time difference between successive pulses follows a Poisson distribution and hence, the techniques explained in section 2.2.1 can be employed to obtain random bits.

While noise based electronic RNGs are readily accessible, they suffer from various limitations. The biggest of these is the presence of short-term and long-term correlations which results in a bias towards one of the bit positions. In RNGs employing Johnson noise for example, the long range carrier correlations imply that the voltages developed across a resistor are also correlated and hence not completely random [117]. Similarly, Zener diodes suffer from memory effects [87] which means that an instantaneous voltage developed across the barrier is dependent on some past voltage values and therefore not completely indeterministic. Even in RNGs employing other sources of noise, the randomness cannot be well characterized, measured, or controlled during the fabrication process. Further, in schemes employing an amplification stage, the nonlinearity and the gain bandwidth of the amplifier can cause deviations from ideal random behaviour [87]. Finally, the threshold (defined in the comparator for the encoding of bits) needs to be fine tuned to reduce the bias between the bit positions for 0s and 1s. However, due to issues with the stability of the device and the time needed for fine tuning, this has proven to be a difficult engineering problem to solve [87].

Apart from noise based RNGs, other electronic systems have also been explored for the implementation of TRNGs. One commonly employed method is to use a logical inverter with its output fed back into the input, resulting in the “Free Running Oscillator” (FRO) configuration [94]. In an ideal scenario, this scheme produces true indeterminism as the output will always be the logical inverse of the input and vice versa. However, in practice, the circuit behaves as an oscillator due to the finite propagation delay of the Boolean NOT operation. In an FRO based RNG, the output of a fast FRO (with high frequency of oscillations) is sampled by a slow FRO to produce random digits upon comparison with a threshold. Fig. 2.4 shows a schematic of VIA C3 Padlock RNGs [95] implemented using four FROs:

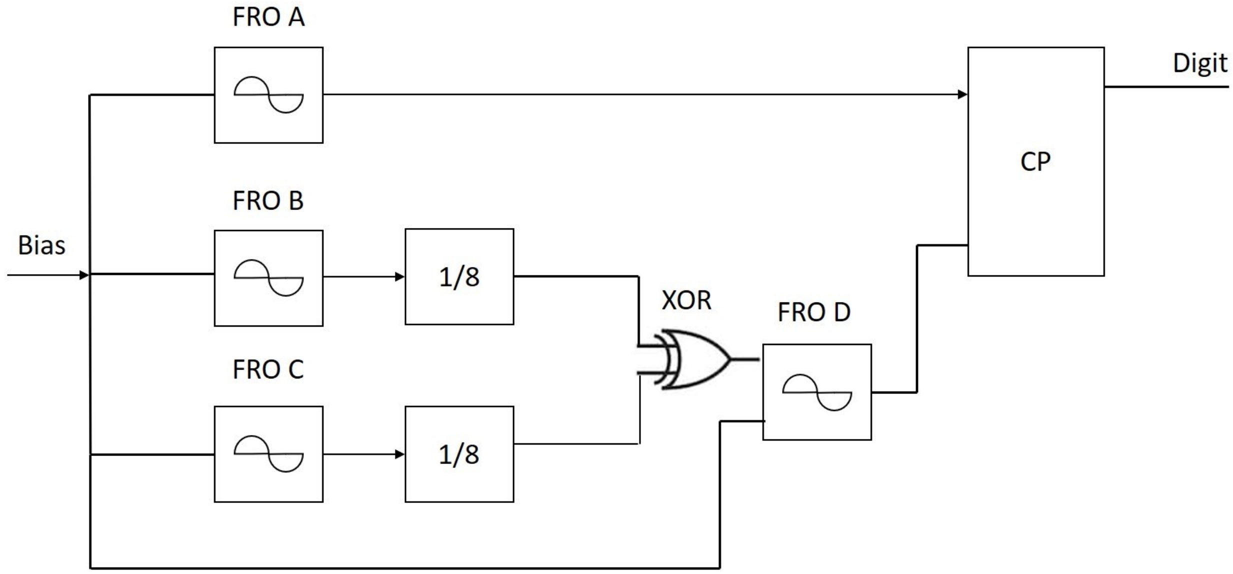


**Figure 2.4 Working principle of VIA C3 Padlock RNGs.** The generator samples FRO A (fast) using FRO D (slow), which is driven by FROs B and C, and outputs a digit using a reference threshold in the comparator CP. The outputs of FROs B and C are slowed down by 1/8 dividers and XOR-ed, which then serves as an input to FRO D. The sampling action of FRO D on FRO A is analogous to the sudden stoppage of a fast rotating wheel, which yields a random orientation each time. The final digit is produced at the output of the comparator which can be further processed using algorithms.

The working principle of this scheme is analogous to the sudden stoppage of a fast-rotating wheel. Since the wheel can be stopped in any position, its observed state at any given point of time is expected to be random [87]. In such schemes, it is important to maintain a relative phase jitter between the fast and slow FROs to prevent repetitive binary patterns or pseudo random behaviour. A key feature of these circuits is that they can "lock-in" or stabilize at a voltage level between 0 and 1 resulting in minor amplitude oscillations which are incapable of driving further logical circuitry. This is due to the finite gain of the circuit and can be overcome by intentionally adding some reactances in the feedback loop (the same effect can also be provided by stray reactances in the circuit) [87]. However, the oscillations would remain extremely sensitive to variations in electrical power and temperature. Further, individual FROs with different input noise levels (expected to have random phase walk-offs) tend to get synchronized when on the same chip, due to the high gain of NOT gates. Such gain picks up any nearby interference (thereby causing the phased-interlocking effect [118,119]). Interlocked rings thus share nearly the same phase which leads to pseudo random behaviour. In addition, this effect leaves FRO based designs susceptible to attacks from external electromagnetic radiations [87].

Despite these limitations, FROs are used in many commercial RNG prototypes thanks to their easy accessibility (through integrated circuits etc.) and well characterized system parameters [95].

Apart from these embodiments, other non-optical systems such as Majorana fermions [120], spin noise [121] and trapped ion systems [122] have also been studied for the implementation of QRNGs. While some of these realizations are particularly relevant for fundamental investigations of randomness and quantum mechanics, their use in practical applications is severely limited due to complexity of their setups.

## 2.3 Optical systems

Optics and photonics systems are amenable to the implementation of QRNG hardware for a multitude of reasons. Firstly, there exist many degrees of freedom such as arrival time [123], path [124], polarization [125], and frequency [126] which can all be used independently or in combination (for example in hyper-entangled states [127]) as sources of randomness. With the rapid advancement in integrated optics and silicon photonics capabilities in recent years [128], optical components like on-chip lasers [129], integrated sources of single photons [130,131], directional couplers and beam splitters [132,133] have become widely accessible and interoperable with off-the-shelf telecommunication components such as wave shapers and spectral filters. Further, the proven superior performance of optical technologies for communications (as compared to electronic systems) necessitates the use of components that can interface with fiber optic networks even for classical communication protocols. Thus, photonic systems are of specific interest for the realization of future technological infrastructures such as multi-node quantum networks (leading to the so-called quantum internet), quantum cryptography, and quantum key distribution [134]. As QRNGs are crucial to the success of each of these areas, it is of particular interest to develop fast, efficient, and trustworthy photonic QRNGs which can be readily interfaced with other systems.

Since the 1980's, numerous techniques have emerged for the design and optimization of optical QRNGs utilizing parameters such as quadrature measurements of electromagnetic radiation [135], phase noise [136] etc. The system architecture of a generic photonic QRNG is shown in Fig. 2.5:

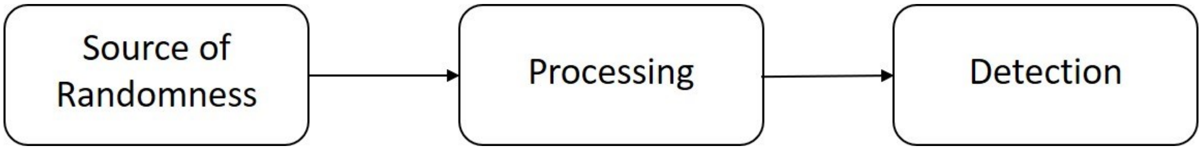


**Figure 2.5 System architecture of a generic photonic QRNG.** The design of a photonic QRNG generally consists of a photonic source of genuine randomness which is processed and measured to encode bits. Each of these stages can further comprise of additional functionalities like postprocessing and randomness extraction (upon detection).

It comprises of a photonic source of genuine randomness, a processing stage which translates the available randomness into bit positions (made up of optical components like beam splitters, temporal delay lines etc.) and a detection mechanism. The raw light detected can be further processed using randomness extraction algorithms, unbiasing operations, etc., to obtain genuinely random numbers or bit sequences. Depending on the design of the device, the processing stage can be grouped with either the source or the detector for the purposes of device characterization.

In most state-of-the-art realizations, the source is an emitter of single photons. For example, the photon number distribution of an attenuated laser follows a Poisson distribution. The probability of finding '$k$' photons in a given interval of time ($T$) can be determined using the equation:

$$P(k) = \frac{e^{-\lambda T}(\lambda T)^k}{k!} \tag{2.2}$$

where $\lambda$ characterizes the laser intensity (counts/s) and $\lambda T$ is its mean photon number. When the number of photons is sufficiently low (such that $\lambda T$ < 0.1, e.g. for attenuated light), the photon state follows sub-Poissonian statistics which is a clear signature of the quantum nature of light [78]. In pulsed laser systems, this criterion is equivalent to having one photon per pulse and hence, such sources can be treated as emitters of single photons. As a result, the properties of attenuated lasers (most commonly, the distribution of arrival times) can be used for the implementation of QRNGs, analogous to radioactive systems with photons replacing decaying particles.

Alternatively, the excitation of certain nonlinear materials like lithium niobate [137] or silicon nitride [138] can result in the generation of entangled photon pairs which can be used to implement interesting QRNG schemes. For instance, either photon of the signal-idler pair produced from parametric processes [139] can be used individually (as single photons) or in combination. In the latter case, the breaking of coherent superposition of photon states by a projection measurement results in an intrinsically random observation at the detectors [124]. The working of a typical QRNG based on this principle in explained in the next section. Sources of entangled photons are particularly interesting as they also facilitate the investigation of fundamental quantum mechanical properties like non-locality, teleportation, and causality due to the availability of superposition and entanglement, in addition to probing randomness.

The design of the processing and detection stages depends on the parameter exhibiting random behaviour in the source. While many degrees of freedom of a photon can be used to design the RNG, the most popular techniques use either its spatial or temporal properties. These techniques are discussed in detail below.

### 2.3.1 Spatial Mode QRNGs

The path traveled by a single photon from the source to the detectors is an easily accessible parameter which can be encoded to derive bits of information. If a photon has equal probabilities of traversing all possible paths, then the observed path information or the spatial mode will be intrinsically random. Fig. 2.6 shows the schematic of a spatial mode QRNG.

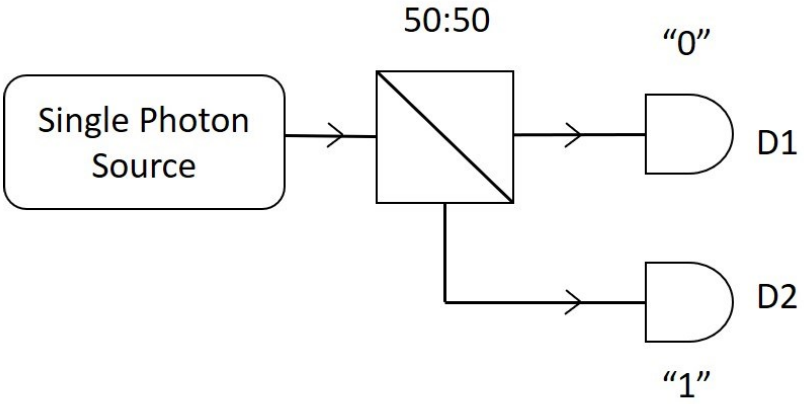


**Figure 2.6 Schematic of a spatial mode QRNG.** Action of the 50:50 beam splitter on a single photon state results in equal probabilities of photon detection at detectors D1 and D2, which are encoded as bits '0' and '1' respectively.

This scheme was first introduced by Rarity *et al* in 1994 [124] and has since become extremely popular thanks to its simplicity and reproducibility. It consists of a 50:50 beam splitter which divides the input classical light so that 50% of the input power is available at each output port. The action of this beam splitter on a single photon state can be understood as follows:

Let the two paths resulting from the beam splitter be represented by states $|1\rangle_1|0\rangle_2$ and $|0\rangle_1|1\rangle_2$, where the first state represents a photon detection at detector D1 and no photon detection at detector D2 while the second state represents a photon detection at D2 and absence of detection at D1. When a single photon impinges onto the balanced beam splitter, it can traverse either the first or the second optical path, giving rise to the following superposition state at the output of the device:

$$\frac{|1\rangle_1|0\rangle_2 + |0\rangle_1|1\rangle_2}{\sqrt{2}} \tag{2.3}$$

The probability of measuring this state at either detector (thus resulting in the breaking of superposition) is $\frac{1}{2}$. This effect can also be understood using the particle representation for single photons. Every photon has equal probability of taking either path at the output of the beam splitter (since it cannot be further split) and hence, the probability of detection at each detector is 0.5. In other words, the output from this system will have clicks produced by either D1 or D2 (but not simultaneously) with equal frequency. If a detection at D1 is encoded as "0" and at D2 as "1", the clicks can be translated into a sequence of bits in which the value of each bit position will be unpredictable due to the inherent indeterminism of the system. A bias towards either bit value due to factors such as imbalance in splitting ratio, accidental counts or background noise can be overcome using simple post-processing methods like a bitwise XOR operation (as explained in Chapter 1). This technique can be extended to other implementations beyond spatial modes. For instance, QRNGs based on polarization [140] and surface plasmons [141] have been realized using the same design principle. In addition, this scheme can have a variety of embodiments such as integrated chips, fiber-based systems, or free space setups. Due to the conceptual and experimental simplicity of this method, it has been widely used with modifications for both standalone QRNG implementations and as a part of other applications. However, it may be noted that the performance of such QRNGs is severely limited. The speed of this device is directly dependent on the speed of the source and detectors. Since each photon results in only a bit of information, a fast source/detector combination with high photon flux is required to obtain long bit sequences at a fast rate. As ultrafast single photon sources and detectors cannot be easily deployed with other systems (due to requirements such as the use mode-locked lasers, superconducting detectors etc.), this method is rarely used for practical QRNG prototypes without significant modifications.

A popular technique to overcome this limitation involves increasing the amount of information that can be extracted from a detection event. This can be achieved by increasing the dimensionality of spatial modes of the photon. An example of such a scheme is shown in Fig. 2.7.

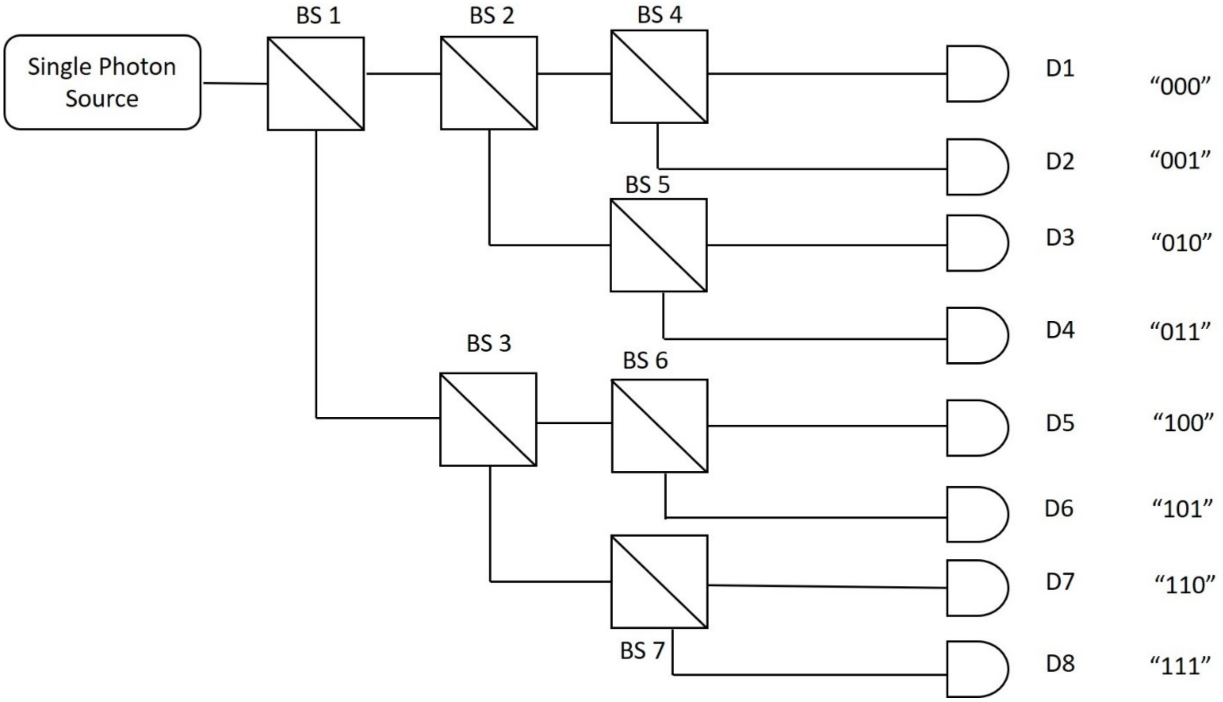


**Figure 2.7 Schematic of a higher-dimensional spatial mode QRNG.** Multiple 50:50 beam splitters can be used in combination to increase the number of possible paths a photon can undertake. A detection event at each detector corresponds to a unique path taken by a photon from the source to the detector. Hence, the possible paths can be encoded as a sequence of bits.

In this setup, an array of 50:50 beam splitters connected end-to-end increases the number of spatial modes of the photon by successively introducing new paths that it can take until detection. Since a photon can traverse each path with equal probability (due to the 50/50 splitting), clicks at detectors D1-D8 should follow a truly random, uniform distribution. As the clicks in each detector channel correspond to a unique spatial mode, they can be encoded as shown in the figure to derive bit sequences. Note that in the setup shown in the figure, each photon results in 3 bits vs a single bit in the previous realization. More generally, in such schemes, if $N$ is the number of beam splitters used, then the number of detectors required is $N$ + 1 and the total number of resulting bits is given by $log_2N$. A similar scheme was realized on-chip in the experiment of Grafe *et al* in 2014 [142].

Although this method increases the information retrieved per photon detection, it is evident that the number of required single photon detectors increases linearly with the number of beam splitters used. Given the associated cost and complexity of deploying a large number of single photon detector channels, this method is impractical for implementation at a large scale.

### 2.3.2 Temporal Mode QRNGs

Another method to increase the total information contained in a photon detection is to encode the arrival time of single photons as bits. QRNGs based on photon arrival time have been widely studied and implemented [48]. Many design principles used in these RNGs are derived from radioactive systems, with single photons replacing particle decay. The working principle of a typical time of arrival generator is shown in Fig. 2.8:

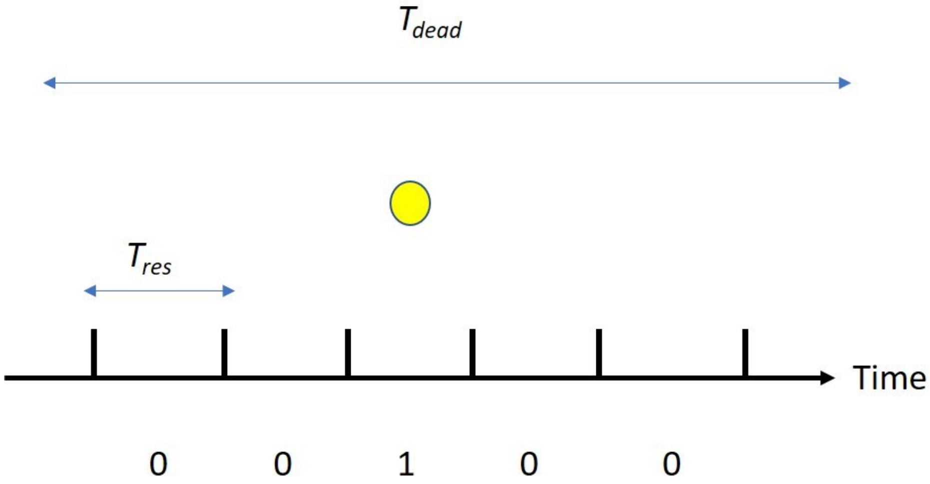


**Figure 2.8 Working principle of a temporal-mode QRNG.** Temporal observation windows are defined as allowed by the detector dead time ($\boldsymbol{T_{dead}}$), which are further divided into timebins as per the detector resolution ($\boldsymbol{T_{res}}$). Photon detection at a timebin is encoded as 1 and its absence is encoded as 0's.

In temporal mode QRNGs, photon arrivals are observed in specific time intervals, the width of which are determined by a combination of source and detector parameters. The width is chosen so that exactly one photon detection is possible in each interval - typically the same as the detector deadtime. This ensures that erroneous counts due to factors such as multiphoton generation, accidental counts, and noise are not registered as separate events. Each interval is then divided into smaller "timebins" as allowed by the detector resolution and the single photon coherence time. Random bits are obtained by encoding the detection of a photon in a timebin as '1' and its absence as '0', as shown in Fig. 2.8. Since the source exhibits true indeterminism, each timebin has the same probability of occupancy and hence the distribution of the bit positions corresponding to '1' will be uniform across multiple detections. In other words, each time interval will see a photon detection in a different timebin and thus, an accumulation of these events gives rise to a uniform distribution at the detectors, as expected in an ideal RNG.

To illustrate, an observation window of width 100 ns can be defined using a detector with a deadtime of the same duration. If its resolution is 1 ns, 100 timebins can be obtained in each interval. It follows that a photon detection in this measurement basis results in $\boldsymbol{log_2 100}$ = 6.64 bits of information. In general, if $\boldsymbol{T_{dead}}$ is the detector deadtime and $\boldsymbol{T_{res}}$ is its resolution, then $\boldsymbol{N = \frac{T_{dead}}{T_{res}}}$ timebins can be obtained per observation window, which yields $\boldsymbol{log_2 N}$ bits.

Note that in such schemes, it is important to choose sources whose parameters are compatible with those of the measurement basis. For instance, if the coherence time of a single photon is larger than the width of a timebin, a detection event would span multiple bit positions instead of just one as shown in Fig. 2.8. This would result in ambiguity (due to multiple bits representing the same event) and a consequent loss of information, as the number of bit positions required to represent each event increases, thereby reducing the total number of extractable bits per photon. Similarly, the brightness of the source (counts/s) that can be used is limited by the detector deadtime. A brighter source does not necessarily result in a faster device as the events occurring within the detector deadtime will not be recorded. Thus, optimization of source and detector parameters is necessary to achieve a high-performance temporal mode QRNG.

Time of arrival generators have become increasingly popular for the realization of fast and compact QRNGs owing to the simplicity of their designs and increased dimensionality of information encoding. As a result, numerous schemes have been implemented to study their performance parameters and suitability for various applications. For example, Nie *et al* [123] have implemented an extremely simple and fast QRNG with only an attenuated laser source and an external reference clock to encode arrival times as random bits. In contrast, Xu *et al* [143] also realized a high bitrate QRNG using a source of entangled photons and dynamically monitoring the visibility of quantum interference fringes as a measure of randomness in the system. Since temporal mode QRNGs can be realized in vastly different configurations, the needs of each application or study determine the exact design of the RNG. Further, since multiple time references can be used to encode bits (similar to the fast and slow clock techniques discussed for radioactive systems in section 2.2.1), such schemes are suitable candidates for the implementation of multiplexing schemes which can achieve very high performances. An example of such a scheme is studied in detail in Chapter 3.

## 2.4 Performance parameters of QRNGs

As discussed in the previous sections, QRNGs can be realized using a multitude of techniques. While the initial radioactive systems were more suitable for a theoretical study of quantum randomness, photonic implementations are best suited for commercial applications. Since the 1950s, thousands of QRNG prototypes have been developed and patented for various uses [60]. In order to compare these prototypes and determine the best scheme for a given application (or to develop a new one), it is necessary to quantify the performance of QRNGs using measurable quantities. State-of-the-art QRNG research focuses on optimizing two performance parameters, namely, bitrate and certification of genuine randomness. The bitrate of a device is equivalent to its speed and can be measured using the number of bits produced per unit time. For instance, a generator that can produce numbers at the rate of 10 Mbps is better than one which operates at 100 Kbps, considering only their speeds. In photonic implementations, bitrate is directly dependent on the deadtime of the detectors, brightness of the source, extractable bits per photon etc. This parameter is of specific relevance for commercial applications where the fast production of a large number of bits is required for use in other dependent protocols.

The second parameter i.e. certification of genuine randomness, is a direct measure of the security of the device. It addresses the question of whether the bits generated by a QRNG can be trusted to be truly random. Since present day generators already achieve statistical randomness, this parameter is crucial toward distinguishing the advantages of quantum systems over their classical counterparts. Further, it provides a measure of how immune the device is to external attacks which can compromise the functionality of the generator as well as other components with which it is interfaced. Unlike the bitrate, certifiability cannot be measured directly from the data produced by an RNG. Rather, it is determined through the verification of quantum mechanical criteria such as a loophole free violation of Bell tests, high visibility of quantum interference fringes and so on. For instance, if in a QRNG the setup used to encode bits can also be used to obtain quantum interference fringes with a high visibility (typically greater than 71% [144]), such a method can be trusted to be truly random. In devices where quantum mechanical properties cannot be directly measured (for example, in an attenuated laser with sub-Poissonian statistics), a detailed device characterization in combination with an estimation of the minimum information entropy of the device can be used to determine trust..

QRNGs can also be classified and studied based on their performance parameters. As discussed by Ma *et al.* [48] in their review of QRNGs, the following classes can be defined based on the usability of a generator for practical applications:

### 2.4.1 Self-testing QRNGs

Self-testing QRNG configurations provide the highest degree of trust in the randomness of the produced bits. The methods used to realize these devices lead to the development of the so-called Device-Independent QRNGs (DIQRNGs) [145] in which randomness does not depend on the components used in the setup or the design of the generator. In other words, these devices do not make any assumptions about component parameters, their deviation from ideal behaviour, presence or absence of an adversary etc., in order to generate truly random bits. The outputs from these QRNGs are guaranteed to be indeterministic, irrespective of design parameters, via the satisfaction of a number of complex quantum mechanical criteria, typically including the loophole-free violation of Bell inequalities. Self-testing and Device-Independent QRNGs are especially relevant for applications where security takes precedence over all other considerations and also for fundamental investigations into the nature of randomness.

Ideally, every QRNG should be self-testing so that the produced bits can be directly used in the required applications. However, due to the experimental complexity of these setups, their bitrates are extremely low (of the order of a few hundred bits per second), which makes them unusable in practical applications [48,88]. In addition, since there is no defined threshold for the physical verification of randomness (similar to the NIST and Diehard test suites for a statistical verification [75]), it is unclear if every application indeed requires a robust certification as performed in these realizations. To address this limitation, many recent works have sought to increase the rate of production in self-testing devices and this remains an active area of research with some promising results emerging only recently from the works of Rusca *et al* [146,147].

### 2.4.2 Trusted-device QRNGs

While self-testing and device-independent QRNGs achieve a very high level of trust, the performance of practical devices is often evaluated using other parameters such as speed, compactness, and ease of use. As a result, fast QRNG schemes with verifiable quantum behaviour have been developed for use in commercial applications, which are termed as trusted-device QRNGs. Such devices may not be completely “quantum-proof”, i.e. their designs may not be robust enough to withstand sophisticated quantum attacks. However, they exhibit genuine random behaviour and possess distinct quantum mechanical properties which can be verified through measurements such as coincidence-to-accidental ratios, correlation counts, quantum interference etc. Thus, trusted-devices are an effective intermediate between classical devices which are deterministic and self-testing QRNGs which are mostly experimental.

Since trusted-devices are mainly developed for practical purposes, factors such as the size of the device, its interoperability with other electronic or photonic systems, and shelf life (due to the degradation of components) are important considerations in addition to having a fast bitrate. As a result, performing complex quantum mechanical tests on their outputs would not be feasible to certify the generated bits as being truly random. Hence, security of these devices is assured through a detailed device characterization along with a quantitative estimate of the randomness in the obtained bits. Note that this is a crucial design principle used in trusted-devices. It is assumed that the components in the setup are non-malicious i.e., while they can deviate from ideal behaviour, no component behaves as an adversary, deliberately detracting from random behaviour. The device characterization includes a complete description of the generated photon states, their evolution caused by the processing stage and detection settings that yield the final bits. This information is then used to estimate the amount of genuine randomness in the system which can be quantified using the minimum information entropy or simply the minimum entropy. Chapter 3 describes the technique to derive mathematical expressions for the minimum entropy of QRNGs using source, detector, and environment parameters. In trusted-devices, the minimum entropy can be used in combination with other parameters as a figure of merit of the QRNG. For a well characterized quantum device, higher minimum entropy implies more randomness in the generated bits. Note that the estimation of minimum entropy is distinct from the certification of genuine randomness even though an investigation of source parameters is required for both. While minimum entropy quantifies the randomness, certification guarantees that its origin is truly indeterministic.

As discussed above, current QRNG research maintains a trade-off between speed and security. While trusted-devices achieve high speeds, they are not immune to adversarial attacks whereas self-testing QRNGs achieve high security, but operate at extremely slow rates. Hence, an appropriate implementation strategy can be chosen to meet the needs of a specific application. An ideal QRNG should have both high security and high speed in order to meet the requirements of both commercial applications and academic studies. Even as active research is being undertaken to realize such systems, an intermediate QRNG type can be implemented to partially meet both requirements. Semi-self-testing devices are a class of QRNGs in which some parts of the setup are trusted while other parts are not. For instance, in a source-independent QRNG [48], the source is uncharacterized (similar to self-testing protocols) while the measurement settings are well-defined. By rotating through a set of different, well-defined measurement bases, the indeterminism in the source can be translated into random bits. Conversely, a measurement-independent QRNG can be defined in which the source is well characterized but the measurement components are untrusted. In such devices, a specific measurement basis is determined through the use of auxiliary photon states from the source. This information about the measurement device can be then used to encode bits. It follows that in semi-self-testing QRNGs some quantum mechanical criteria for genuine randomness are met by either the source or the measuring components and a detailed characterization of the rest of the device provides the complete information necessary to verify genuine random behaviour.

It may be noted that irrespective of the design methodology being used, all approaches for the implementation of QRNGs involve a complete re-design of the device which results in a fixed set of performance parameters for each setup. Although this strategy could yield increasingly improved QRNGs, it typically involves replacing the source, the detectors, the processing elements or all of them together. This is an expensive and experimentally complex undertaking as interoperability and backward compatibility between various state-of-the-art components is not easily achievable. Therefore, it would be tremendously useful, especially in commercial applications, to realize a QRNG in which the trade-off between these parameters can be tuned. If a given device can be used with minor modifications to provide either a faster bitrate or higher security in different contexts, this would reduce the overhead involved in re-designing and re-implementing a QRNG when the application requirements change. One possible design solution to address this problem was developed through this Master's project which has been described in detail in Chapter 3.

# 3 SCALING RANDOMNESS PARAMETERS IN PHOTONIC TRUSTED-DEVICE QUANTUM RANDOM NUMBER GENERATORS

This chapter describes the design of a novel randomness scaling mechanism developed as a part of this Master's thesis. The motivation for this design, experimental realization, advantages and limitations are explained in detail below.

As explained in Chapter 2, commercial applications require a high random bit rate and hence, trusted-device QRNGs are considered the most promising practical implementation. In these devices, a high random bit rate is typically achieved by means of two strategies: source/detector optimization [8,9], and/or increasing the state dimensionality (the number of levels/modes a photon can occupy at detection) [10,11].

The first is important for trusted-devices as they operate under the assumption that the components are not vulnerable to malicious attacks and therefore genuine randomness is guaranteed due to the nature of the source. This means that the only reason a trusted-device QRNG may be unable to produce its maximum achievable randomness (that is, translate all photon states into random bits) is because of system nonidealities such as noise, multi-photon state generation, finite detector resolution, and imperfect quantum visibilities. The second strategy is critical as the state dimensionality quantifies the potential total randomness in the QRNG *without* real-world nonidealities. In this way, the random bit rate in practical devices can be increased by reducing system nonidealities and/or by increasing the total number of possible extractable bits. However, there are several practical drawbacks to increasing trusted-device bitrates using these methods. A QRNG bit rate that depends critically on source or detector parameters has limited flexibility. As sources and detectors degrade (or other operational parameters change) so does the bit rate. Improving the performance of such devices eventually requires the overhaul of costly components, e.g., the complete replacement of the source or detector. On the other hand, improving bit rate through dimensionality scaling typically comes with increased experimental complexity and cost. For example, recent work from Grafe *et al*. [142] shows a potential scaling mechanism accomplished by increasing the number of spatial modes in a waveguide implementation, but at the price of using eight single photon detectors for eight different possible photon paths.

In this chapter, we experimentally demonstrate a method for flexible, bitrate scaling in a temporal mode trusted-device QRNG without changing the source or the detector, and with minimal increase of system complexity. The motivation for our experimental design comes from a parameter we investigate, the minimum information entropy per bit, $H_\eta$ (detailed in section 3.1 ), which signifies the efficiency with which a QRNG can practically use its randomness. For example, in an implementation with $H_\eta$ = 0.5, only half of all photon detections can be trusted to be truly random. As this parameter is directly proportional to the random bitrate in trusted-device implementations, it is a particularly worthwhile parameter for QRNG design and optimization. Our specific implementation works by cascading fiber-based beam splitters and fiber-based delays in order to increase the dimensionality of state space of photon arrivals, and therefore $H_\eta$. In this technique, the number of detector channels remains the same even though we physically increase the possible detected photon states through a time multiplexing scheme. By simply adding beam splitters and fiber delays, we can reduce the impact of source and/or detector nonidealities to either improve random bit generation or, alternatively, allow the use of lower quality devices for comparable performances. Here we show randomness scaling, consistent with theory, for two different sources of single photons: 1) an attenuated laser and 2) continuous time-energy entangled photons generated by spontaneous parametric down-conversion (SPDC).

For our specific sources, we show an improvement in minimum entropy per bit between 3-20% as the number of beam splitters is increased from 0 to 4. However, our method shows great promise for low-cost, practical QRNGs that use affordable, lower quality sources and detectors and for which randomness improvement can be as large as 30%. Furthermore, our experiments suggest that for trusted-device implementations, there exist regimes or operational settings where sources of heralded photon pairs may not offer any inherent advantages over attenuated laser sources with regards to random bit generation rate, indicating that off-the-shelf lasers can be sufficient for practical quantum random number generation requirements. At the same time, in certain configurations (like T2 measurements) heralded photon pairs can be potentially used for a simultaneous recreation and investigation of the state space of photon arrivals using a single measurement instead of using two separate measurements as described in this chapter.

## 3.1 Theoretical approach

In temporal QRNGs based on the photon arrival time [148], the amount of genuine randomness in the raw distribution can be quantified using the minimum entropy given by:

$$H_{min} = log_2(N) \tag{3.1}$$

in units of bits, where $N$ is the total number of time bins in the observation window. For a generic QRNG, $N$ refers to the number of possible states a photon can occupy at a given detection event. In the ideal scenario, all time bins have an equal probability of occupancy and hence $P_i$ = 1/ $N$. In a system where the probability distribution of the states is not uniform, one uses $H_{min} = log_2\left(\frac{1}{P_{max}}\right)$, where $P_{max}$ is the maximum probability of a photon occupying a time bin. Regardless of the source or detection mechanism used, the number of bits that can be obtained practically from such a trusted-device implementation is

$$H_{min} = log_2(N) - H_{dev} \qquad (3.2)$$

where $H_{dev}$ corresponds to source/detector nonidealities such as multi-photon counts, detection jitter, dead times, limited temporal resolution, etc., which degrade the capability for random bit generation. In dividing $H_{min}$ by the number of bits the system could ideally produce, we introduce the minimum entropy per bit as

$$H_\eta = 1 - \frac{H_{dev}}{log_2(N)} \qquad (3.3)$$

which has a maximum value of 1 when $H_{dev}$ = 0. In other words, in a perfectly random system, there are no experimental nonidealities and thus all photon detections can be trusted to be truly random. When $H_{dev}$ is non-zero, $H_\eta$ can be optimized by either reducing $H_{dev}$ or increasing *N*. Though not the main objectives of their works, Xu *et al* [143] recently show impressive random bit generation rate in trusted-device QRNGs by improving $H_\eta$. Efforts to improve $H_\eta$ by decreasing $H_{dev}$ require source and detector optimization or a complete design overhaul. On the other hand, methods that improve $H_\eta$ by increasing the dimensionality *N*, come with either additional experimental complexity or increased cost. Here we characterize $H_{min}$ and $H_{dev}$ on a novel QRNG for two sources and show how they affect $H_\eta$. Further, we explain how small changes to the experimental setup can scale $H_\eta$ by optimizing the term $H_{dev}/log_2(N)$.

Fig. 3.1 shows the proposed experimental setup of our entropy scaling technique. Our theory assumes non-ideal sources and detectors (i.e., with multiphoton emission, finite detector resolution, imperfect efficiency, and the presence of noise in the form of both background and dark counts). The heart of our randomness scheme is a set of cascaded, fiber-based beam splitters in which the source is divided into two paths by the first and recombined into the two inputs of the next (and repeated again until the last beam splitter) with different amounts of fiber length between their outputs. Given a generation time $T_{gen}$ for each photon, we obtain $l =$

$2^n$ temporal states in which the photon can be detected (where *n* is the number of beam splitters). For example, in an implementation with two beam splitters (shown in Fig. 3.1), the randomness scheme introduces four temporal states that every photon can occupy at the detectors, corresponding to delays $T_{gen}$ + T1 + T3, $T_{gen}$+ T1 + T4, $T_{gen}$+ T2 + T3 and $T_{gen}$+ T2 + T4, where T1, T2, T3, and T4 correspond to the temporal delays introduced by the fiber paths shown in Fig. 3.1. Since these temporal states correspond to physical components in the system, we refer to them as "physical time bins." If $T_{gen}$ is used as a reference by using a heralding photon from the signal-idler photon pair of a correlated source, these physical time bins can be characterized. This corresponds to switch position A in Fig. 3.1, in which the signal photon from the source is input into the beam splitters and the simultaneously emitted idler is used as the "clock" trigger to measure the photon delay time (see also the data in Fig. 3.3).

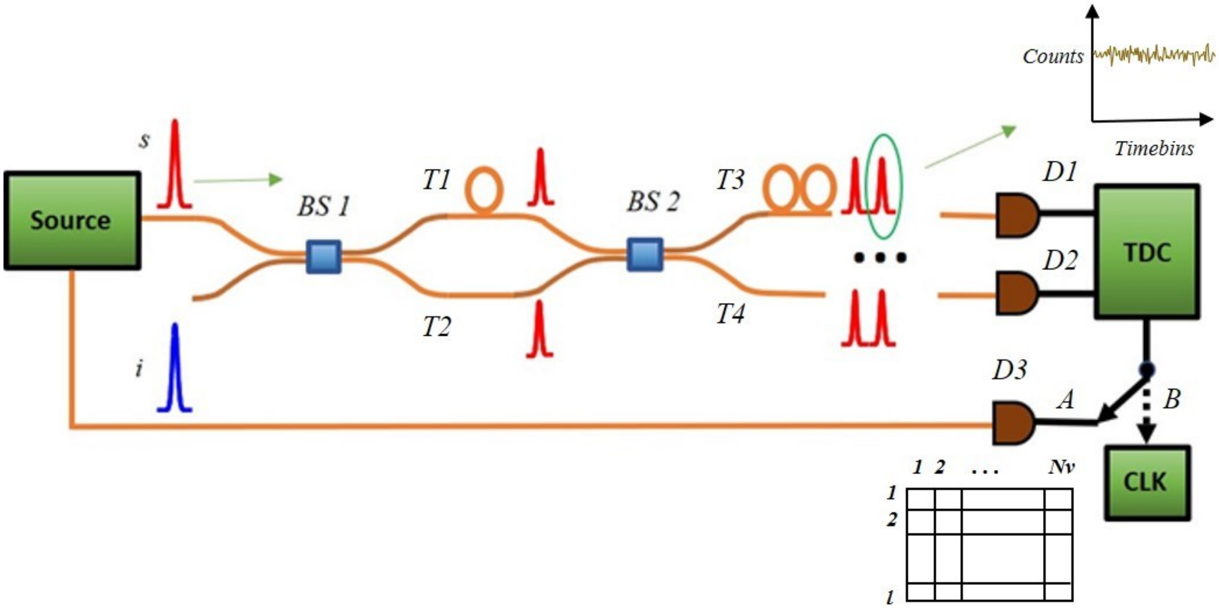


**Figure 3.1 Experimental scheme of the cascaded fiber beam splitter setup for entropy scaling**. *BS – Beam Splitter, TDC – Time Digital Converter, CLK – Clock, T1,T2,T3,T4 – Temporal delays introduced by different fiber lengths, s, i – entangled signal-idler photon pair, D1,D2,D3 – Single Photon Detectors*

However, in the absence of a herald, the physical time bins cannot be measured due to uncertainty in the generation time $T_{gen}$ which arises because of source temporal characteristics (such as spontaneous emission or continuous wave emission from the laser). Since photon generation is a spontaneous process, $T_{gen}$ varies randomly with respect to any clock signal in the lab frame. Without synchronization to the generation time, the detector displays an overlap of the physical time-bins introduced by the randomness scheme, resulting in a uniform distribution of equally probable random arrival times of photons with respect to the lab clock. In fact, this scheme of using an arbitrary lab clock (corresponding to switch position B in Fig. 3.1) to measure random photon arrival times within the clock window has been used successfully in a number of high-dimension temporal mode QRNGs [123,149]. Since this encoding represents bits in positions that are not determined by physical components in the setup, we refer to them as “virtual time-bins”, the total number of which is $N_v$. It may be noted that measurement settings represented by switch positions A and B in Fig. 3.1 correspond to the measurement of uncertainty/randomness in two different degrees of freedom. While measurements for switch position A allows the characterization of uncertainty in the path traversed by the detected photons, measurements for switch position B allow the characterization of the uncertainty in the generation time of the photons. The utilization of two distinct degrees of freedom in a time multiplexing scheme is a unique feature of our experimental setup.

In our proposed setup, we combine virtual and physical time-bins to scale the dimensionality and therefore $H_\eta$. The net effect of the cascaded beam splitter setup on photon arrival times can be summarized as a matrix where the two basis vectors are generation times and physical time-bin positions, see Fig. 3.1 (inset). Conventional virtual time-bin encoding with one physical dimension represents a state space of only one row of the matrix. Our scheme, on the other hand, provides a state space equal to the *area* of the matrix (which in our demonstration is $l$ = 2 - 16 physical time-bins and $N_v$=100 virtual time-bins, equivalent to $l$ x $N_v$ = 200 - 1,600 temporal states). However, there is an important advantage to our method besides simply increasing the states a photon can occupy. By using physical time bins, we can scale the overall entropy per photon without changing the source or detector properties of the setup. This is not possible in conventional virtual time-bin QRNGs, where one can only increase the number of states by lengthening the detection observation time window or improving detector resolution [15].

The net effect of the cascaded beam splitter setup on the photon arrival times is also illustrated in Fig. 3.2, using the case of 2 beam splitters as an example. As seen in the figure, the two beam splitters result in $2^2$ = 4 paths at the detectors. Every photon with a distinct $T_{gen}$ can traverse each of these paths with equal probability. Further, photons with different $T_{gen}$ values can occupy the same timebin (with respect to an external clock) in the resulting bit sequence, as they experience different delays corresponding to the path they traverse. Hence, the total number of ways in which a given bit position can be filled increases (from 1 to 4) with the addition of beam splitters. More generally, with respect to an external clock of a given repetition rate (giving rise to $N_v$ timebin positions within the window of observation), a detection in each timebin can be the result of $2^n$ distinct photon events due to $n$ beam spiltters in the setup. Since the uncertainty in the spatial mode exists in addition to the temporal uncertainty of photon generation/emission from the source (which results in a random variation of $T_{gen}$ with respect to the lab clock), this scheme generates more randomness in comparison to traditional temporal QRNGs which use only one degree of freedom for randomness extraction, namely, the arrival time of single photons. Thus, with $N_v$ timebin positions defined by the external clock, and $2^n$ spatial modes defined by their cascaded beam splitter setup for each of the $N_v$ positions, the total number of ways where a bit sequence can be generated is ($N_v$*$2^n$), which is the dimensionality of the state-space of photon arrivals. To illustrate this further, consider the case of $N_v = 5$ and $n = 1$. In this configuration, the bit sequence 10000 can be detected as a result of a photon arrival at either detector channel D1 or channel D2 in the first timebin position, whereas in the absence of a beam splitter, the bit sequence would be the result of an individual detection event (corresponding to a unique $T_{gen}$). Thus, the introduction of a beam splitter leads to more photon states, thereby increasing the randomness of the setup.

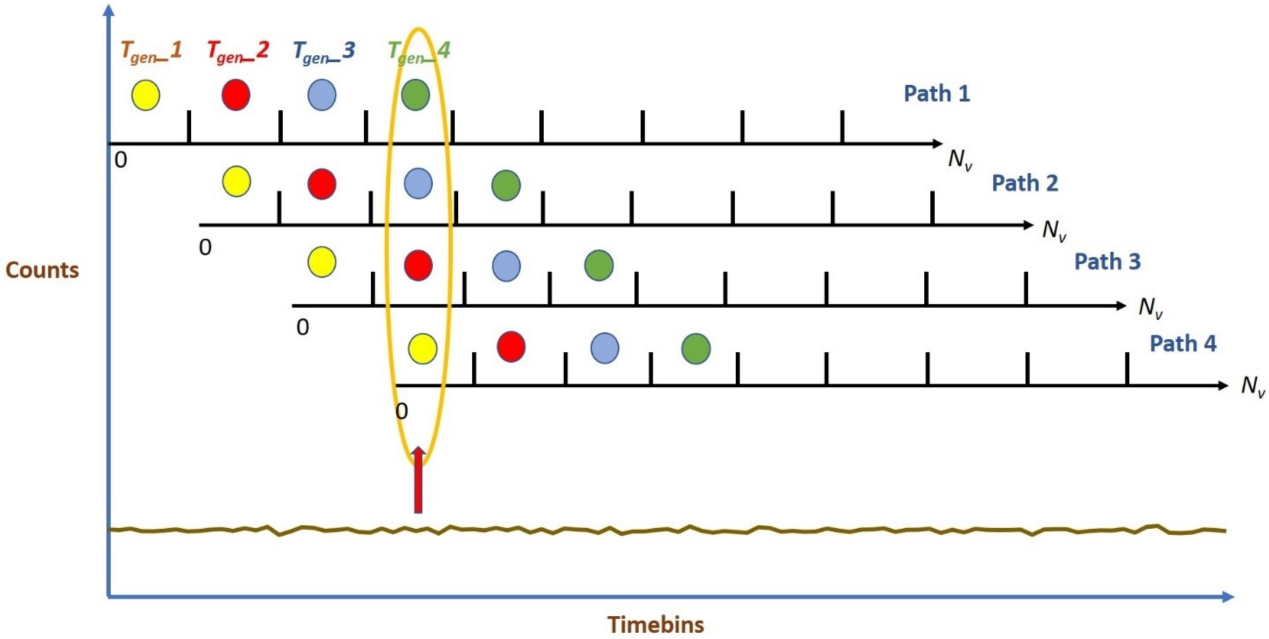


**Figure 3.2 Illustration of randomness scaling using two beam splitters in the setup**. Paths 1,2,3 and 4 which are temporally distinct from one another are created at the detectors. Photons with different generation times can arrive at the same bit position in the final sequence, thereby increasing the randomness per bit.

To demonstrate the effectiveness of our method and that randomness scaling is independent of the source and detector, we derive the expected minimum entropy per bit for both the attenuated CW laser and entangled photon pair sources.

### 3.1.1 Attenuated CW Laser

Light emitted from an attenuated laser with low average photon number, $\bar{n} < 0.1$ per observation time, follows a Poisson distribution and can be treated as a single photon source. The minimum entropy associated with the detection of such a photon can be written as follows [15]:

$$H_{min} = log_2(N_v) + log_2\left(1 - e^{-\lambda T\gamma}\right) - log_2(\lambda T\gamma) \qquad (3.4)$$

where $N_v$ is the number of virtual time bins in the encoding scheme, $\lambda$ characterizes the laser intensity, $\gamma$ is the efficiency of the detector and $T$ is the duration of observation as determined by the detector deadtime. The terms $log_2\left(1 - e^{-\lambda T\gamma}\right) - log_2(\lambda T\gamma)$ are obtained after correcting for multiphoton counts, timing jitter of the detector, etc., which contribute erroneously towards randomness in the collected raw data - as explained in [15].

However, since the dimensionality of the state space is $N_v * 2^n$ instead of $N_v$ as in traditional QRNGs, this equation becomes (see the Appendix for detailed steps):

$$H_{min} = log_2(N_v) + n + log_2\left(1 - e^{-\lambda T\gamma}\right) - log_2(\lambda T\gamma) \qquad (3.5)$$

The additional entropy resulting from the cascaded beam splitter section can be monitored through the minimum entropy per bit as follows, using the experimental values of $N_v$ and $n$:

$$H_\eta = 1 - \frac{H_{dev}}{log_2(N_v)+n} \qquad (3.6)$$

with

$$H_{dev} = log_2\left(1 - e^{-\lambda T\gamma}\right) - log_2(\lambda T\gamma) \qquad (3.7)$$

### 3.1.2 Entangled Photon Pairs

An $N_v$-dimensional biphoton state can be obtained through different spontaneous processes in nonlinear media such as Spontaneous Four Wave Mixing (SFWM) in microring resonators [126,150], SPDC in Periodically Poled Lithium Niobate (PPLN) waveguides [151]. This state can be characterized as:

$$|\psi\rangle = \sum_{i=1}^{N_v} |i\rangle_A \otimes |i\rangle_B \tag{3.8}$$

where $|i\rangle$ represents a single photon at a discretized time interval $i$ [143]. It has been shown that the smooth minimum entropy (accounting for environmental noise) of such a higher-dimensional timebin entangled state, when measured in two mutually unbiased bases with orthogonal Positive Operation Valued Measures (POVMs) (such as the virtual and physical timebins corresponding to switch positions A and B of the experimental setup in Fig. 3.1) can be bound as follows [152,153]:

$$H_{min} \geq -log_2 c - H_{max} \tag{3.9}$$

where $c$ is the maximum overlap (see eq. 3.10) between the two mutually unbiased projective measurements from the incompatible POVMs, onto the physical timebins and virtual timebins bases. The Maximum Entropy $H_{max}$ is the Renyi entropy of order ½ which gives more weight to events with small surprisal [154]. $H_{max}$ can be modeled so that it accounts for inaccuracies in the measurement, for example due to statistical fluctuations, variations in the visibility of the single photon source etc., as shown in [143]. Further, $H_{max}$ can be approximated such that it varies proportionally with $N_v$, which is the most significant parameter [155]. In our implementation, the parameters $H_{min}$ and $H_{max}$ correspond to the measured entropy of the virtual timebins and the physical timebins respectively, which form the two mutually unbiased, orthogonal bases. Since for each value of $i,$ there exist $2^n$ temporal states in which the photon can be measured at the detector,

$$c = \frac{1}{N_v * 2^n} \tag{3.10}$$

Thus, the minimum entropy per bit can also be described by eq (3.6), substituting $H_{dev}$ with $H_{max}$.

## 3.2 Experimental Details and Results

As an attenuated source, we used a CW laser and a pulsed laser (NetTest Tunics Plus) with emission centered at 1547.6 nm wavelength and with a linewidth of 100 MHz. The outputs of the lasers were attenuated by means of two stacked variable optical attenuators, leading to 25 kHz photon flux at the detector. This corresponds to an average photon number, $\bar{n}$ < 0.1, for the 100 ns temporal window we choose to measure photon arrival times. For the SPDC source, we use the same laser as a pump for a two-stage periodically poled lithium niobate (PPLN) waveguide system. In this setup, the first PPLN is used for second harmonic generation (SHG) [156] to convert the 1547.6 nm light to 773.8 nm, which is then used to pump the second PPLN waveguide for SPDC, so as to generate entangled photon pairs (for which degenerate photons are centered at 1547.6 nm). Both PPLNs are commercial devices (Srico 2000), the parameters of which are similar to those in [157]. Note that high-rejection bandpass filters were used to block residual pump light not converted by SHG or SPDC, respectively. Although the total SPDC bandwidth was > 4 THz (too large to be measured by our laboratory equipment) we use a telecom programmable filter (Finisar 4000A WaveShaper) to select 25 GHz each of signal and idler bandwidths of non-degenerate SPDC photons. By pumping the first PPLN at 1 mW we produce 100$\mu$W of SHG which gives us a detected SPDC photon rate of 13kHz for signal and idler channels. Superconducting nanowire single photon detectors (Quantum Opus One) were used for all experiments, featuring a deadtime of 80 ns, an efficiency of 85% at 1550 nm, and a resolution (jitter) of ~400 ps. Photon detection events were recorded by a time-to-digital converter (Picoquant Hydraharp 400), used to give the difference in photon arrival time between a trigger and a detection event. As described in section 3.1, we use two different triggers - the detected idler photon (Fig. 3.1 switch position A) and an arbitrary lab clock (Fig. 3.1 switch position B) which we choose to be a pulse train at 10 MHz from an arbitrary function generator (Tektronix AFG 3251).

Our randomness scheme consists of 1-4 polarization maintaining 50/50 fiber couplers (AFW PFC-15-2-50-BB) connected as shown in Fig. 3.1. The couplers used in our experiment had an average insertion loss (including loss due to mating sleeves and fiber delays) of ~1dB each and their splitting ratios deviated from 50/50 by at most 2%. Note that larger deviations in splitting ratios would be detrimental for random bit generation as it would cause a reduction in minimum entropy per bit due to a decrease in possible photon states.

In order to generate separate, nonoverlapping physical time-bins, we put fiber delays on all BS outputs ranging from 0.20 to 4.0 m. Note that these fiber lengths are not the path differences between the arms of the interferometers. Rather, they are included to ensure that the measured physical timebins are distinct and do not overlap. Overlapping timebins would lead to a reduction in the total number of distinct photon states and hence the number of extractable bits per photon detection, which is undesirable for a practical QRNG. Further, the minimum path difference between the arms of the interferometers is maintained at 2m to prevent first order interference from both the laser and the PPLN source, which require a minimum path difference of 1m.

Fig. 3.3 shows the histogram of photon arrival times using the idler as the trigger. Additionally, the coincidence-to-accidental rate (CAR) of the system ranges from 200 to 1200 depending on the number of beam splitters (and hence losses) in the system. In an ideal implementation with no losses in the different arms of the interferometers, the count distribution in each timebin would be approximately the same. However, the observed count distribution shows non-uniformity which can be attributed to deviation in splitting ratio from perfect 50/50 in successive beam splitters as well as to the different losses introduced by the inclusion of additional beam splitters, fiber delays and mating sleeves in the experimental setup. The impact of the skew in the count distribution can be estimated from the experimentally measured value of minimum entropy. In case of a significant deviation from an ideal behaviour, the number of usable bits would be less than the

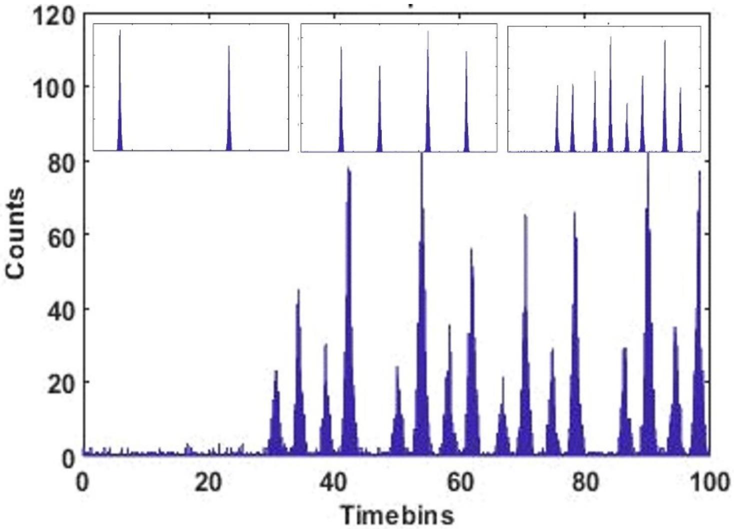


**Figure 3.3 Histogram of single photon arrival times (heralded case)**. The histogram of signal photon arrivals (physical time bins) using the idler as a clock trigger for the system with 4 beam splitters. The histograms when 1,2, and 3 beam splitters are employed are reported in the inset at the top (left – right)

number of beam splitters (or $log_2(2^n) = n$ ) in the setup, due to fewer distinct photon states. However, in the observed distribution, the experimentally measured value of minimum entropy closely matched the theoretically expected value, thus indicating that the observed skew in the count distribution does not significantly impact the extractable randomness.

Fig. 3.4 shows the distribution of photon arrival-times when the clock is in position B in Fig. 3.1, corresponding to the active mode of the QRNG for generating random bits. Here we choose the observation window to be 100 ns (~detector deadtime) and width of each time bin to be 1 ns, giving an $N_v$ = 100. Note that we purposely choose this width to be more than twice worse than our actual resolution to simulate practical systems where a detector resolution of 1.0 ns is a more realistic value. Ultimately $N_v$ is a 'free' parameter limited by the dead time of the detector (which sets the largest observation window) and detector resolution (which sets the smallest bin width). The distributions of both the sources are near-uniform as expected. Effectively, this uniform distribution is obtained due to an overlap of physical timebins created by the beamsplitter setup (as shown in Fig. 3.3) at different generation times $T_{gen}$. Fig. 3.5 illustrates the creation of this uniform distribution using the data from Fig. 3.3 for the case of one beamsplitter and by numerically varying the value of $T_{gen}$.

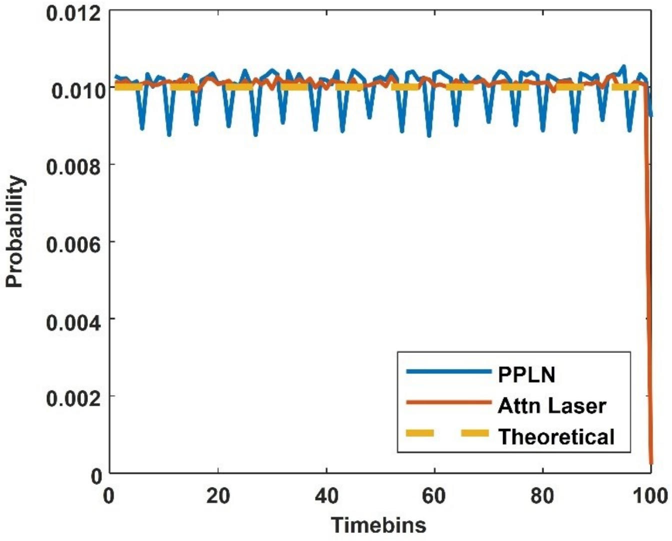


**Figure 3.4 Histogram of photon arrival times (unheralded)**. Virtual time bins at the detectors for photons emitted by an entangled PPLN waveguide source and an attenuated laser, respectively, using a 10 MHz lab clock as the trigger. Data were collected over 3 minutes.

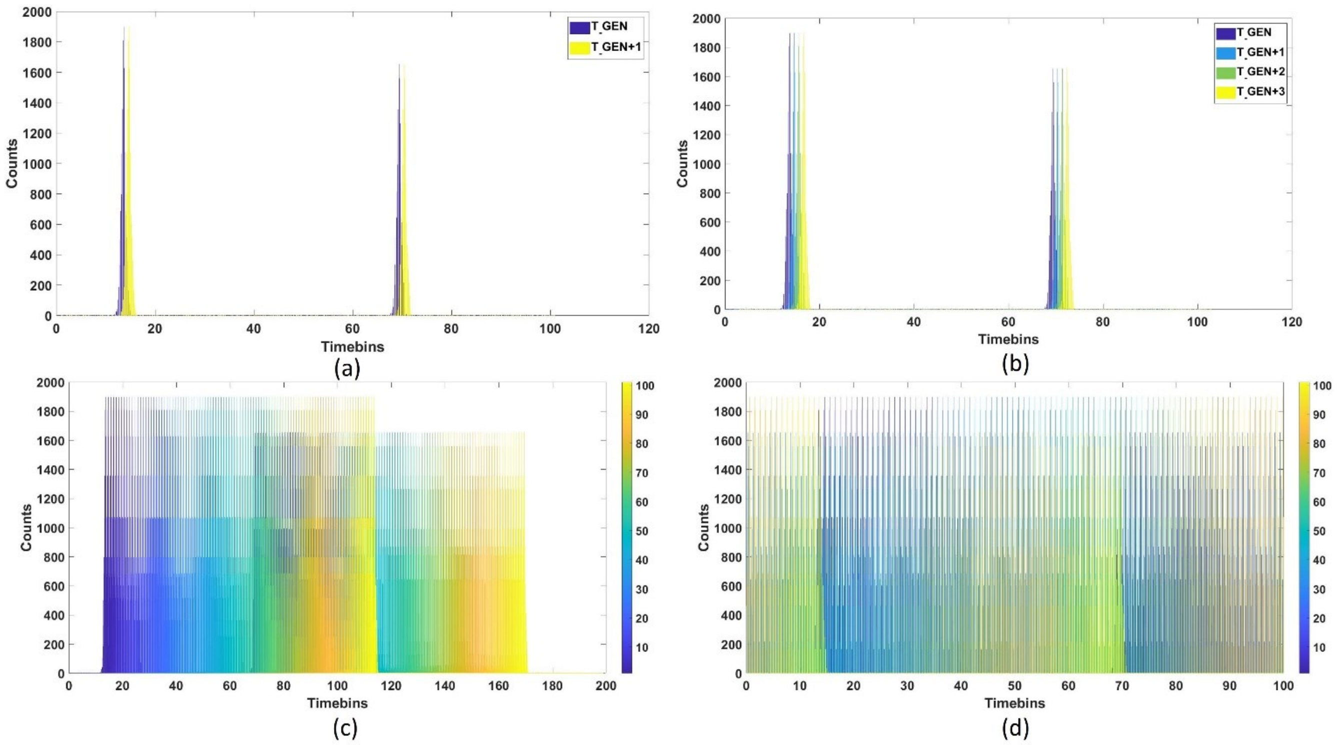


**Figure 3.5. Illustration of the creation of a uniform random distribution due to varying $T_{gen}$ and physical timebin positions.** (a) Physical timebin positions with $T_{gen}$ varying in a 1ns interval. (b) Physical timebins positions with $T_{gen}$ varying in a 3ns interval. (c) Linear shift and overlap of physical timebins positions with $T_{gen}$ varying across all 100 timebins. (d) Uniform distribution obtained by measuring physical timebin positions with respect to a 10MHz clock and $T_{gen}$ varying across all 100 timebins.

With a single beamsplitter in the setup, two distinct physical timebins are created for a given $T_{gen}$. However, as $T_{gen}$ varies with respect to a reference clock, the positions of the created physical timebins also change. This effect can be seen in Figs. 3.5 (a) and 3.5 (b) where $T_{gen}$ varies within intervals of 1 ns and 3 ns, respectively. The created physical timebins are shifted along the timebin axis corresponding to the variation in $T_{gen}$. Since $N_v$= 100 and the width of each timebin is 1 ns in the experiment, the observed values of $T_{gen}$ can change between 1 ns and 100 ns in increments of 1 ns (equal to the bin width). The variation in the temporal position of the physical timebins corresponding to varying $T_{gen}$ values across 100 ns is shown in Fig. 3.5 (c). It can be seen that some timebins start to overlap at ~70 ns due to a simultaneous occurrence of the first physical timebin from a certain $T_{gen}$ measurement and the second from a different one. Since this distribution is shown over a linear timescale, the shift in the positions of the physical timebins is also linear. However, measuring the timebin positions with respect to an external clock of a given repetition rate (10 MHz, used in measurement mode B of Fig. 3.1) results in measurement frames of a fixed width or measurement duration. Timebins occurring outside the frame width (100ns) are observed in subsequent measurement frames. Thus, a uniform distribution is obtained across the 100 timebins in each measurement frame as shown in Fig. 3.5 (d). Here, physical timebins corresponding to multiple values of $T_{gen}$ are observed within the frame, along with overlapping timebins which occur outside the measurement duration of the previous frame, resulting in equal probability of photon detection at any of the 100 timebins. Thus, this figure illustrates the generation of the ideal uniform random distribution (as in Fig. 3.4) in the temporal basis due to varying $T_{gen}$ values and the creation of physical timebins, which are the two sources of randomness used in the experiment.

Fig. 3.6 shows the experimentally measured minimum entropy per bit, $H_{\eta}$, as a function of the number of beam splitters for the two sources, as well as the theoretical curves for other potential sources and detectors with varying amounts of nonidealities. For the determination of $H_{\eta}$ there are two important quantities: $H_{dev}$, and the total number of potential random bits, $log_2(N_v) + n$, generated by this QRNG configuration. We experimentally measure the latter using the raw data from Fig 3.4., calculating the minimum entropy from the empirical definition $-log_2$(Max $P_i$) as the maximum photon counts in any of the 100 timebins divided by the total number of photon counts in that observation window. Similarly, the data in Fig. 3.3 are used to experimentally determine the number of physical states, $l$. For example, in the configuration with four beam splitters, the theoretical value of bits is expected to be $log_2(100)$ + 4 = 10.64 bits. Here our data from Fig. 3.3 and Fig. 3.4 show the experimental number of bits to be 10.47$\pm$ $0.039$. $H_{dev}$ is determined from the characterization of the sources and detectors (i.e. measurements of detector efficiency and resolution, multiphoton generation, etc., see section 3.1) similar to refs [123, 143] and subtracted from the minimum entropy of the raw data. For our attenuated laser source, we estimate the average $H_{dev}$ to be 0.56 and 3.55 for the PPLN waveguide source.

It can be seen from Fig. 3.6 that the experimental scaling of $H_{\eta}$ closely matches our theoretical predictions. In general, we expect that increasing the number of physical paths, $l$, should improve the efficiency of randomness. For the attenuated laser used in this experiment, $H_{\eta}$ already begins at a high value of 91.52% and reaches 94.8% with the addition of beam splitters. For this laser, $H_{dev}$ is so low that increasing the number of modes has a small effect on randomness since the system already allows almost all the randomness in the system to be used for bit generation. However, the increase in losses due to the inclusion of more fiber beam splitters affects the detected photon flux, which prevents an arbitrary increase in bitrate of this system. For the entangled PPLN waveguide source, $H_{\eta}$ increases much more dramatically, from 45.97% to 66.09%, as the number of beam splitters increases from 0 to 4.

In general, Fig. 3.6 shows that increasing the number of beam splitters is particularly significant with higher values of $H_{dev}$ . For example, when $H_{dev}$ = 5, which can be obtained by changing source and detector parameters such as *N* (most prominent parameter) from 100 to ~ 500 (e.g., by using a suitable PPLN source), the increase in $H_\eta$ is as high as 29%. This indicates that our scheme is particularly useful in cases with increased source/detector non-idealities, which makes it even more amenable to application in real world systems. Note that the experimental value of $H_\eta$ changes negligibly as the integration time varied from 30s to 60 minutes, which indicates that asymptotic behaviour is achieved for relatively short data collection times. Thus, finite size effects are not dominant in our data i.e., a significant change in the values of $H_\eta$ is not likely to be observed with the accumulation of more data. Finally, we note that our data in Fig. 3.6 show that the attenuated laser used for our experiment outperforms the source of heralded photons with regards to random bit generation rate , indicating that entangled photon sources do not offer an inherent advantage in such a temporal mode QRNG configuration.

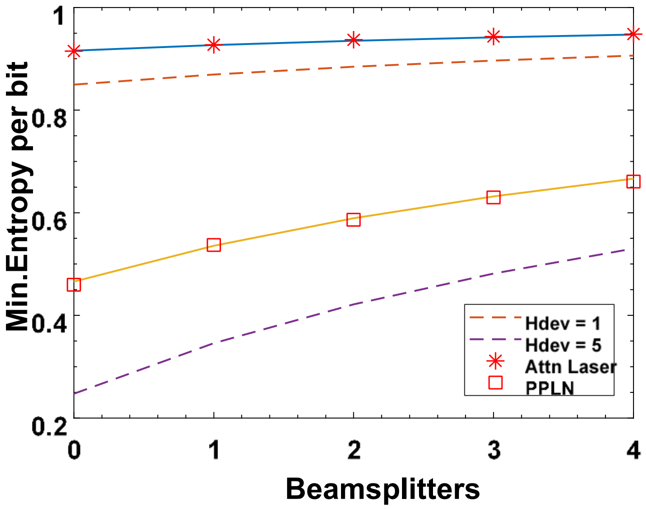


**Figure 3.6 Minimum Entropy per bit as a function of the number of beam splitters**. Data points show experimental measurements and dashed curves show theoretical predictions for different source/detector nonidealities.

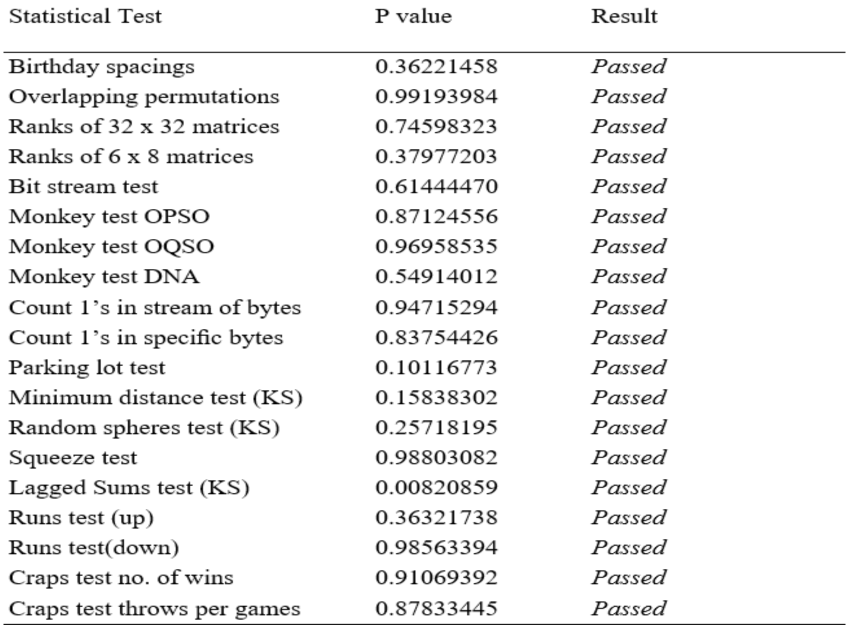
We also use the data in Fig. 3.6 to calculate the scaling in random bitrate from the relation:

$$\text{Bitrate} = \text{Total Minimum Entropy} * \text{Photon Flux}$$

$$= H_\eta * (log_2(N_v) + n) * \text{Detected Photon rate} \qquad (3.11)$$

With our sources, a random bitrate in the range of 50 kbps to 150 kbps was achieved. However, a higher value for bitrates can be obtained using faster source/detector combinations. It can be seen that all parameters in eq. 3.11 vary with the number of beam splitters in the setup. Consequently, the bitrate either increases or decreases, depending on which of these parameters has the dominant effect. For example, when the losses are low (~0.5 – 1dB), the first two parameters in eq. 3.11 increase much faster than the decline in the photon flux. Therefore, in such cases, there is a net increase in bitrate with the addition of a beam splitter. However, as losses become greater (~2- 5dB), the decline in photon flux becomes the dominant contributor to the random bit rate and hence it is indeed detrimental to add more beam splitters to the setup in this region of operation. In Fig. 3.6, the curve for attenuated laser achieves the maximum value for minimum entropy per bit of 0.948 when four beam splitters are used. However, with the addition of the fourth beam splitter, there is only a slight increase in $H_\eta$ of ~0.5%, whereas the photon flux continues to decrease. In contrast, the curve for PPLN waveguides sees an increase in $H_\eta$ of 6% when the number of beam splitters goes from 2 to 3 with an equal decrease in photon flux. However, since the number states also increases $((log_2(N_v)+n)$, we observe a net rise in bitrate of ~3% in this region. Thus, for our source/detector settings, the region between beam splitters 2 and 3 is optimal for QRNG operation. This effect is evidently more prominent when the quality of the source is lower, since a larger increase in $H_\eta$ is seen with every beam splitter in such case.

To obtain the final bit sequences, a Toeplitz Hashing Extractor [89] is implemented in order to separate noise in the channels from genuine randomness resulting from the source and the components. The extractor retrieves a random bit sequence of length $m$ from a raw bit sequence of length $n$ by multiplying it with a Toeplitz Matrix of dimensions $n \times m$. In our implementation, we choose a large value of $n$ (=4096) to restrict finite size effects, by concatenating multiple raw bit sequences obtained from the time-to-digital converter (each of length 100), and a corresponding large value of $m$ ( ≥ $n * H_\eta$, determined experimentally in each case). The resulting bit sequences are passed through the Diehard test suite for statistical randomness. In total, 96 Mbits of data collected from various configurations (with 0,1,2,3 and 4 beam splitters) were tested and all tests were passed in each case. The results of one of these test cases is shown in Fig. 3.7 as an example.



| Statistical Test | P value | Result |
|---|---|---|
| Birthday spacings | 0.36221458 | *Passed* |
| Overlapping permutations | 0.99193984 | *Passed* |
| Ranks of 32 x 32 matrices | 0.74598323 | *Passed* |
| Ranks of 6 x 8 matrices | 0.37977203 | *Passed* |
| Bit stream test | 0.61444470 | *Passed* |
| Monkey test OPSO | 0.87124556 | *Passed* |
| Monkey test OQSO | 0.96958535 | *Passed* |
| Monkey test DNA | 0.54914012 | *Passed* |
| Count 1's in stream of bytes | 0.94715294 | *Passed* |
| Count 1's in specific bytes | 0.83754426 | *Passed* |
| Parking lot test | 0.10116773 | *Passed* |
| Minimum distance test (KS) | 0.15838302 | *Passed* |
| Random spheres test (KS) | 0.25718195 | *Passed* |
| Squeeze test | 0.98803082 | *Passed* |
| Lagged Sums test (KS) | 0.00820859 | *Passed* |
| Runs test (up) | 0.36321738 | *Passed* |
| Runs test(down) | 0.98563394 | *Passed* |
| Craps test no. of wins | 0.91069392 | *Passed* |
| Craps test throws per games | 0.87833445 | *Passed* |

**Figure 3.7 Results of the Diehard test suite for the case of 4 beam splitters in the setup**. It can be seen that all the tests are passed and similar results were obtained for all cases with different sizes of the Toeplitz matrix, corresponding to the experimentally measured value of $H_\eta$.

It may be noted that our generator produces a random number owing to the increased dimensionality, which is distinct from other realizations where a random bit is generated. Hence, to verify the statistical randomness of our data, Diehard tests are more relevant as they probe the context of sequence occurrence rather than the properties of the sequence itself (as in the case of NIST tests [75]). In other words, for our implementation, it is more relevant to test the long-range correlations between different bit strings rather than repetition of individual bit values within a given string. Furthermore, the successful passing of these tests only verifies that the numbers being tested satisfy the statistical requirements of genuine randomness. This however, does not certify that the physical processes involved in their generation are truly random, since no information about the source or components is used in the design of these tests. In addition, sophisticated attacks can be engineered so that pre-determined bit sequences are generated which pass the statistical tests [158]. Therefore, the suitability of these test suites as well as the relevance of each individual test for a specific implementation needs to be determined before using them for verification. Similarly, depending on the implementation, a simple or more complex post-processing method may be employed. Ma *et al*, provide a robust comparison of popular extraction techniques and the resulting bitrates from them in QRNG implementations [89].

While the method described here pertains to trusted-device implementations, the scheme can be used more universally, as indicated by data from two different sources, leading to the realization of hybrid semi self-testing QRNGs. The impact of this work and the potential future directions that can emerge from it are discussed in Chapter 4.

# 4 CONCLUSIONS AND DISCUSSIONS

This thesis provides a broad overview of quantum random number generation with a specific focus on photonics implementations. As discussed in Chapter 1, techniques for random number generation have been studied for millennia – initially motivated by philosophical questions and more recently being driven by commercial requirements. The rapid upscale in the abilities of nascent quantum technologies in recent years, poses a significant threat to protocols like cryptography, communications, and simulations, which are critically reliant on classical RNGs today. Given their extensive use in day-to-day applications, developing quantum-proof QRNGs is crucial to the success of future quantum inspired technologies. The progress made in realizing physical QRNGs capable of producing random numbers at a fast rate was discussed in Chapter 2. Optics and photonics platforms are particularly well suited for these implementations since they offer a multitude of advantages ranging from the numerous degrees of freedom to the easy interfacing with existing telecommunications infrastructure. From the review of commonly used design principles and performance parameters of photonic QRNGs, the lack of scalability in prototypes (required especially for commercial applications) and the associated cost to achieve improved performances was identified as a pertinent unsolved problem.

The work carried out for this Master's thesis in order to address this issue was described in Chapter 3. In this work, we investigated minimum information entropy per bit of a photon (which is an estimate of the total randomness in the generator) as an important design parameter of photonic QRNGs. We then experimentally realized a novel trusted-device QRNG for which the minimum entropy per bit can be tuned/scaled independently of the source and the detection mechanism, with minimal changes to the device itself. Unlike previous methods, our QRNG gives flexibility to the states a photon can occupy by combining virtual time-bins used in a typical temporal-mode device with physical time-bins (multiple paths a photon can take) through many cascaded fiber-based beam splitters. By connecting/disconnecting beam splitters, we are able to tune the minimum entropy per bit and optimize random bit generation in a simple, practical fashion. The robustness of this tuning mechanism is shown by comparing random bit generation from two sources of single photons and simulating non-idealities in the setup which contribute to deviation from genuine random behavior. We also show that for trusted-device QRNGs there exist regimes or operational modes where a simple attenuated laser source outperforms sources of heralded photon pairs with regards to random bit generation rate, offering advantages in both speed, practicality, and cost in commercial devices, although in some configurations (like T2 measurements) heralded sources allow a simultaneous investigation of the entire state space of photon arrivals.

## 4.1 Future works

Although demonstrated here for trusted-devices, the introduced scaling mechanism shows promise for the development of hybrid semi self-testing QRNGs and device-independent QRNGs which offer higher degrees of trust via quantum mechanical certification. As indicated by data from two very different sources of single photons, the scaling characteristics do not strictly depend on the source for genuine random behaviour. To that end, our scheme could act as a universal, inexpensive scaling mechanism to improve randomness parameters in any QRNG configuration. Similarly, as indicated by the minimim entropy estimates and statistical tests, the performance of the device also does not strictly depend on the number of beam splitters used. In other words, the performance of the device is independent of its design. Hence, this scheme can be used to implement semi self-testing QRNGs in both the source-independent and measurement-device indepenedent configurations. Further, while the two sets of measurements used in the experiment provide an indirect estimate of the state space by characterizing the two individual basis vectors, a simultaneous reconstruction of the state space can be preferably implemented by using the T2 mode of the HydraHarp with a source of heralded photon pairs, and will be the subject of future work.

In addition, the scaling characteristics could be observed due to the investigation of minimum entropy per bit of a photon as a standalone design parameter. This shows that while the total number of bits obtained from a device may not increase, the number of usable bits per photon detection can be increased due to higher efficiency of randomness conversion. Although this parameter has been experimentally measured before, it has not been studied as an independent design parameter. In addition, with more complex multiplexing schemes, other similar design metrics could emerge which can achieve better scaling characterisitcs. Thus, the continued investigation of similar schemes could result in experimental demonstrations of randomness expansion protocols in well-characterized QRNGs. Access to mechanisms which allow tunable performances of these devices would be tremendously useful for the universal usage of quantum technologies.

With rapid advances in artifical intelligence and machine learning algorithms [159], computers have become enormously powerful in identifying patterns. This power can be leveraged to indentify correlations which may not be defined either purely physically or purely mathematically. Analysis of data from QRNGs with machine learning systems can lead to their definitive benchmarking, resulting in a well-defined minimum performance metric required for different generators to be usable in various applications. Such a benchmarking has been hard to accomplish simply because there is not enough information about the macroscopic manifestation of randomness. Hence, a combination of physical, mathematical, and statistical descriptions of randomness combined with machine learning models can fundamentally impact randomness research. Specifically, using the scheme introduced in this thesis, QRNGs with differing performance parameters can be tested against a machine learning enhanced adversary. The results of such an investigation could indicate whether some quantum systems are still susceptible to attacks from artificial intelligence systems or if indeed quantum mechanical randomness is truly indeterministic in all configurations. Recent results from Truong *et al* on machine learning cryptanalysis with QRNGs are very promising for future investigation in this direction [160].

# APPENDIX

## A.1 Expression for minimum entropy using an attenuated laser source

In trusted-device temporal mode QRNGs using a lab clock as a reference, the conditional probability of detecting a photon at a given timebin in the presence of $k$ photons (due to multiphoton events, detector jitter, noise, dark counts etc) is given by [123]:

$$P(\hat{n} = i|k) = \left(1 - \frac{i-1}{N_v * 2^n}\right)^k - \left(1 - \frac{i}{N_v * 2^n}\right)^k \tag{A.1}$$

where $i$ is a timebin position within the observation window and $N_v * 2^n$ is the dimensionality of the state space of photon arrivals.

As shown in Ref. [123], the randomness of the raw data can be evaluated using the expression for $P_i$ from the minimum entropy relation as:

$$H_{min} = -log_2(\max P_i) \tag{A.2}$$

It follows from eq (A.1) that $\max P_i$ occurs at $i = 1$. Further, the upper bound of $P_1$ occurs as $k$ tends to infinity. Thus,

$$P_1 = \frac{1}{1-e^{-\lambda T\eta}} \sum_{k=1}^{\infty} P(n = 1|k)P(k) \tag{A.3}$$

where $\frac{1}{1-e^{-\lambda T\eta}}$ is the normalizing factor for the Poisson distribution of the attenuated laser.

From eq (A.3) we find:

$$P_1 = \frac{1}{1-e^{-\lambda T\eta}} \sum_{k=1}^{\infty} \frac{(\lambda T\eta)^k e^{-\lambda T\eta}}{k!} \left[1 - \left(1 - \frac{1}{N_v * 2^n}\right)^k\right] \tag{A.4}$$

$$\leq \frac{1}{1-e^{-\lambda T\eta}} \sum_{k=1}^{\infty} \frac{(\lambda T\eta)^k e^{-\lambda T\eta}}{k!} * \frac{k}{N_v * 2^n} \tag{A.5}$$

$$= \frac{\lambda T\eta}{N_v * 2^n (1-e^{-\lambda T\eta})} \tag{A.6}$$

Thus, from eq (A.2), the lower bound of the minimum entropy can be evaluated as:

$$H_{min} \geq log_2(N_v) + p + log_2\left(1 - e^{-\lambda T\eta}\right) - log_2(\lambda T\eta) \tag{A.7}$$

## A.2 Expression for minimum entropy using a source of heralded photon pairs

The amount of genuine randomness is quantified through two mutually unbiased measurements corresponding to the physical and virtual timebins (switch positions A and B in Fig. 3.1). As such, the Entropic Uncertainty Principle (EUP) can be used to determine the number of extractable random bits as [154]:

$$H_{min}(x) + H_{max}(y) \geq log_2 \frac{1}{c} \qquad \text{(A.8)}$$

where, in our implementation, $x$ corresponds to the number of virtual timebins, $y$ corresponds to the number of physical timebins and the parameter $c = \frac{1}{N_v * 2^n}$ is the maximum overlap between the basis vectors of the two measurements settings (denoted by the 2D Matrix in the inset of Fig. 3.1). Further, it has been shown that the maximum entropy $H_{max}$ can be modeled as follows [143]:

$$H_{max} \leq log_2 \gamma(d_W^{L_1} + \lambda) \qquad \text{(A.9)}$$

where $\gamma(x),\ d_W^{L_1}$ and $\lambda$ can be defined as:

$$\gamma(x) = (x + \sqrt{1 + x^2})(\frac{x}{\sqrt{(1+x^2}-1})^x \qquad \text{(A.10)}$$

$$d_W^{L_1} \leq \frac{\sigma_{coh}\sigma_{cor}}{\delta \Delta T} \sqrt{\frac{16(1-V_0)}{\pi}} \qquad \text{(A.11)}$$

$$\lambda = N_v \sqrt{\frac{1}{n_T} ln \frac{1}{\frac{\epsilon_1}{4} - 2f(p, n_T)}} \qquad \text{(A.12)}$$

here $\sigma_{coh}$ is the single photon coherence time, $\sigma_{cor}$ is the biphoton correlation time, $\delta$ is the timebin duration, $\Delta T$ is the length of the observation window, $V_0$ is the chosen visibility of the source, $n_T$ is the total number of measurements in the window of observation, $\epsilon_1$ is the failure probability, $f(p, n_T) = \sqrt{2(1-(1-p)^{n_T}}$, and $p$ is the probability of occurrence of a photon event outside the window of observation.

It is evident from eq. (A.10), (A.11), (A.12) that the most significant parameter causing a variation in $H_{max}$ is $N_v$ and thus, this property can be used to study the variation of $H_\eta$ with the $N_v$.

# ASSOCIATED CONTRIBUTIONS

[1] **Bharadwaj S S**, van Howe J, Atzeni S, Roztocki P, Narayanan R, Osellame R, Azaña J, Munro J W, Morandotti R, 'A Scalable Design for Photonic Quantum Random Number Generators', *Conference on Lasers and Electro Optics (CLEO), May 2021*

[2] **Bharadwaj S S**, van Howe J, Atzeni S, Roztocki P, Narayanan R, Osellame R, Azaña J, Morandotti R, 'Entropy Scaling in trusted-device photonic quantum random number generators', *Advanced Photonics Congress, July 2020*

[3] **Bharadwaj S S**, van Howe J, Roztocki P, Jestin Y, Azaña J, Morandotti R, 'Method and system for extractable randomness scaling in quantum random number generators', *US Provisional Patent, 63051539, filed July 14 2020*

[4] **S.S. Bharadwaj**, J. van Howe, S. Atzeni, P. Roztocki, R. Narayanan, R. Osellame, J. Azaña, W. J. Munro, R. Morandotti, "Scaling randomness parameters in photonic trusted-device quantum random number generators", *(Manuscript under revision)*

[5] B. Fischer, P. Roztocki, C. Reimer, S. Sciara, Y. Zhang, M. Islam, L. Romero Cortés, **S.S. Bharadwaj**, D.J. Moss, J. Azaña, L. Caspani, M. Kues, R. Morandotti, "Nonlinear and quantum effects in integrated microcavities", in Nonlinear Meta Optics, Taylor & Francis (2020)

# 5 BIBLIOGRAPHY

[1] B. Hayes, American Scientist **89**, 300 (2001).

[2] R. Motwani and P. Raghavan, *Randomized algorithms* (Cambridge University Press, 1995).

[3] P. Joshi, C.-S. Park, K. Sen, and M. Naik, SIGPLAN Not. **44**, 110 (2009).

[4] K. Efe, Microprocessors and Microsystems **19**, 341 (1995).

[5] Y. Huang and D. P. Palomar, IEEE Transactions on Signal Processing **62**, 1093 (2014).

[6] O. Granichin, Z. Volkovich, and D. Toledano-Kitai, *Randomized Algorithms in Automatic Control and Data Mining* (Springer, 2015).

[7] A. M. Ferrenberg, D. P. Landau, and Y. J. Wong, Physical Review Letters **69**, 3382 (1992).

[8] M. Berblinger and C. Schlier, Computer Physics Communications **66**, 157 (1991).

[9] S. V. Ghaisas and A. Madhukar, Physical Review Letters **56**, 1066 (1986).

[10] M. Neurock, C. Libanati, A. Nigam, and M. T. Klein, Chemical Engineering Science **45**, 2083 (1990).

[11] M. New and M. Hulme, Integrated Assessment **1**, 203 (2000).

[12] U. Cubasch, B. D. Santer, A. Hellbach, G. Hegerl, H. Höck, E. Maier-Reimer, U. Mikolajewicz, A. Stössel, and R. Voss, Climate Dynamics **10**, 1 (1994).

[13] A. Juels and J. Guajardo, in *International Workshop on Public Key Cryptography* (Springer, 2002), pp. 357.

[14] W. J. Miller and N. G. Trbovich, *RSA Public-key Data Encryption System Having Large Random Prime Number Generating Microprocessor or the Like* (September 28, 1982), US Patent 4,351,982

[15] J. C. Schuldt and K. Shinagawa, in *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security* (2017), pp. 241.

[16] R. D. Silverman, *Fast Generation of Random, Strong RSA Primes* (Citeseer, 1997)

[17] S. A. Vanstone and R. J. Zuccherato, Journal of Cryptology **8**, 101 (1995).

[18] D. Ivanov, Transportation Research Part E: Logistics and Transportation Review **136**, 101922 (2020).

[19] T. Kaur, S. Sarkar, S. Chowdhury, S. K. Sinha, M. K. Jolly, and P. S. Dutta, medRxiv (2020).

[20] J. H. Stock, Data gaps and the policy response to the novel coronavirus, Report No. 0898-2937, 2020.

[21] J. Wang, Y. Liao, X. Wang, Y. Li, D. Jiang, J. He, S. Zhang, and J. Xia, Travel medicine and infectious disease, 101660 (2020).

[22] D. J. Bennett, *Randomness* (Harvard University Press, 2009).

[23] C. S. Calude and G. J. Chaitin, *Randomness and complexity: from Leibniz to Chaitin* (World Scientific, 2007).

[24] S. K. Monfared, O. Hajihassani, M. S. Kiarostami, S. M. Zanjani, D. Rahmati, and S. Gorgin, in *49th International Conference on Parallel Processing-ICPP: Workshops* (2020), pp. 1.

[25] A. Nisar, S. Dhull, B. K. Kaushik, and F. A. Khanday, in *Spintronics XIII* (International Society for Optics and Photonics, 2020), p. 114703X.

[26] G. Narasimman, J. Basu, P. Sethi, S. Krishnia, C. Yi, W. S. Lew, and A. Basu, IEEE Sensors Journal (2020).

[27] J. Nebhen, in *International conference on Modelling, Simulation and Intelligent Computing* (Springer, 2020), pp. 495.

[28] A. T. Fuller, The Computer Journal **19**, 173 (1976).

[29] D. R. Brown, IACR Cryptol. ePrint Arch. **2005**, 380 (2005).

[30] B. Ambedkar and S. Bedi, International Journal of Computer Science Issues (IJCSI) **8**, 242 (2011).

[31] K. Kurosawa, T. Ito, and M. Takeuchi, Cryptologia **12**, 225 (1988).

[32] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature **549**, 195 (2017).

[33] D. Ventura and T. Martinez, in *1998 IEEE International Joint Conference on Neural Networks Proceedings. IEEE World Congress on Computational Intelligence (Cat. No. 98CH36227)* (IEEE, 1998), pp. 509.

[34] M. A. Nielsen and I. Chuang, (American Association of Physics Teachers, 2002).

[35] P. W. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (1994), pp. 124.

[36] L. K. Grover, in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (1996), pp. 212.

[37] G. Shahidi, in *2019 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA)* (IEEE, 2019), pp. 1.

[38] T. N. Theis and H.-S. P. Wong, Computing in Science & Engineering **19**, 41 (2017).

[39] M. G. Raymer and C. Monroe, Quantum Science and Technology **4**, 020504 (2019).

[40] M. Riedel, M. Kovacs, P. Zoller, J. Mlynek, and T. Calarco, Quantum Science and Technology **4**, 020501 (2019).

[41] B. Sussman, P. Corkum, A. Blais, D. Cory, and A. Damascelli, Quantum Science and Technology **4**, 020503 (2019).

[42] D. Castelvecchi, Nature News **543**, 159 (2017).

[43] D. E. Denning, American Scientist **107**, 83 (2019).

[44] M. Mosca, IEEE Security & Privacy **16**, 38 (2018).

[45] M. G. Luby and M. Luby, *Pseudorandomness and cryptographic applications* (Princeton University Press, 1996), Vol. 1.

[46] F. Tehranipoor, W. Yan, and J. A. Chandy, in *2016 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)* (IEEE, 2016), pp. 79.

[47] S. Callegari and G. Setti, in *2007 IEEE International Symposium on Circuits and Systems* (IEEE, 2007), pp. 213.

[48] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, npj Quantum Information **2**, 16021 (2016).

[49] H. Isaksson, Teleteknik **3**, 25 (1959).

[50] F. Raffaelli, P. Sibson, J. E. Kennard, D. H. Mahler, M. G. Thompson, and J. C. F. Matthews, Opt. Express **26**, 19730 (2018).

[51] M. Imran, V. Sorianello, F. Fresi, L. Potì, and M. Romagnoli, in *Optical Fiber Communication Conference (OFC) 2020* (Optical Society of America, San Diego, California, 2020), p. M1D.5.

[52] I. Quantique and R. Extractor, Quantum number generator. idQuantique, Geneva, Switzerland (2001).

[53] T. K. Paraïso, I. De Marco, T. Roger, D. G. Marangon, J. F. Dynes, M. Lucamarini, Z. Yuan, and A. J. Shields, in *Optical Fiber Communication Conference* (Optical Society of America, 2019), p. Th1J. 4.

[54] A. Martin, B. Sanguinetti, C. C. W. Lim, R. Houlmann, and H. Zbinden, Journal of Lightwave Technology **33**, 2855 (2015).

[55] Samsung Surprise As World's First Smartphone With Quantum Technology Launches May 22, (Forbes) https://www.forbes.com/sites/daveywinder/2020/05/15/samsungs-surprising-new-5g-smartphone-is-worlds-first-with-quantum-technology/#6619bec930e0.

[56] Y. Zhang *et al.*, Physical Review Letters **124**, 010505 (2020).

[57] K. C. Wiese, A. Hendriks, A. Deschênes, and B. B. Youssef, IEEE Transactions on Nanobioscience **4**, 219 (2005).

[58] T. H. Click, A. Liu, and G. A. Kaminski, Journal of computational chemistry **32**, 513 (2011).

[59] H.-X. Yan, S.-S. Li, D.-L. Zhang, and S. Chen, Appl. Opt. **39**, 3023 (2000).

[60] P. L'Ecuyer, in *2017 Winter Simulation Conference (WSC)* (IEEE, 2017), pp. 202.

[61] H. Diels, *Die Fragmente der Vorsokratiker, griechisch und deutsch* (Рипол Классик, 1906), Vol. 1.

[62] K. Freeman, *Ancilla to Pre-Socratic Philosophers* (Harvard University Press, 1983).

[63] M. T. Cicero, *On the Nature of Gods* (Cambridge, MA: Harvard University Press, 1933).

[64] M. N. Bera, A. Acín, M. Kuś, M. W. Mitchell, and M. Lewenstein, Reports on Progress in Physics **80**, 124001 (2017).

[65] P. Ball, Nature **415**, 371 (2002).

[66] H. Everett III, Reviews of modern physics **29**, 454 (1957).

[67] P. Halpern, *Einstein's Dice and Schr dinger's Cat: How Two Great Minds Battled Quantum Randomness to Create a Unified Theory of Physics* (Basic Books, 2015).

[68] F. Galton, Nature **42**, 13 (1890).

[69] L. H. C. Tippett, *Random Sampling Numbers* (Cambridge University Press, 1927).

[70] R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research* (Oliver and Boyd Ltd, London, 1943).

[71] M. G. Kendall and B. B. Smith, Journal of the Royal Statistical Society **101**, 147 (1938).

[72] R. Corporation, *A million random digits with 100,000 normal deviates* (Free Press, 1955).

[73] K. Landsman, Foundations of Physics **50**, 61 (2020).

[74] K. Landsman, arXiv preprint arXiv:2003.03554 (2020).

[75] S.-J. Kim, K. Umeno, and A. Hasegawa, arXiv preprint nlin/0401040 (2004).

[76] J. Aitchison, Journal of the american statistical association **50**, 901 (1955).

[77] B. D. Ripley, Journal of the Royal Statistical Society: Series B (Methodological) **41**, 368 (1979).

[78] M. Fox, *Quantum optics: an introduction* (OUP Oxford, 2006), Vol. 15.

[79] A. N. Kolmogorov and V. A. Uspenskii, Theory of Probability & Its Applications **32**, 389 (1988).

[80] D. E. Knuth, *Art of computer programming, volume 2: Seminumerical algorithms* (Addison-Wesley Professional, 2014).

[81] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, and E. Barker, A statistical test suite for random and pseudorandom number generators for cryptographic applications, 2001.

[82] D. H. Lehmer, Annu. Comput. Lab. Harvard Univ. **26**, 141 (1951).

[83] W. H. Press, S. A. Teukolsky, B. P. Flannery, and W. T. Vetterling, *Numerical recipes in Fortran 77: volume 1, volume 1 of Fortran numerical recipes: the art of scientific computing* (Cambridge university press, 1992).

[84] P. D. Coddington, International Journal of Modern Physics C **7**, 295 (1996).

[85] J. Cobine and J. Curry, Proceedings of the IRE **35**, 875 (1947).

[86] W. Thomson, Journal of the Royal Statistical Society: Series A (General) **122**, 301 (1959).

[87] M. Stipčević and Ç. K. Koç, in *Open Problems in Mathematics and Computational Science* (Springer, 2014), pp. 275.

[88] M. Herrero-Collantes and J. C. Garcia-Escartin, Reviews of Modern Physics **89**, 015004 (2017).

[89] X. Ma, F. Xu, H. Xu, X. Tan, B. Qi, and H.-K. Lo, Physical Review A **87**, 062327 (2013).

[90] H. Nyquist, Physical Review **32**, 110 (1928).

[91] H. Guo, W. Tang, Y. Liu, and W. Wei, Physical Review E **81**, 051137 (2010).

[92] X. Li, A. B. Cohen, T. E. Murphy, and R. Roy, Opt. Lett. **36**, 1020 (2011).

[93] B. Jun and P. Kocher, Cryptography Research Inc. white paper **27**, 1 (1999).

[94] P. C. Mcleod Jr, (Google Patents, 1970).

[95] Talk notes: Hardware Startup Opportunities in Europe, (viatech.com) https://www.viatech.com/en/2016/02/hardware-startup-europe/.

[96] Behind Intel's New Random-Number Generator, https://spectrum.ieee.org/computing/hardware/behind-intels-new-randomnumber-generator (Accessed August 24 ).

[97] W. M. Dickson, *Synthese*, **107**, pp 55–82 (1996).

[98] J. Manelis, Electronics (US) **34** (1961).

[99] Y. Xu-Tao, X. Jin, and Z. Zai-Chen, Chinese Physics B **22**, 090311 (2013).

[100] Y. Nambu and A. Tomita, (Google Patents, 2003).

[101] Y. Liu *et al.*, Physical Review Letters **120**, 010503 (2018).

[102] B. Hensen *et al.*, Nature **526**, 682 (2015).

[103] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, Nature **557**, 400 (2018).

[104] J. Ahrens, P. Badziag, A. Cabello, and M. Bourennane, Nature Physics **8**, 592 (2012).

[105] M. Ishida, Annals of Institute of Statistical Mathematics **8**, 119 (1956).

[106] R. Maddocks, S. Matthews, E. Walker, and C. Vincent, Journal of Physics E: Scientific Instruments **5**, 542 (1972).

[107] C. Vincent, Journal of Physics E: Scientific Instruments **3**, 594 (1970).

[108] M. P. Silverman, W. Strange, C. R. Silverman, and T. C. Lipscombe, Physics Letters A **262**, 265 (1999).

[109] H. Schmidt, Journal of Applied Physics **41**, 462 (1970).

[110] J. Walker, Online: http://www. fourmilab. ch/hotbits (2001).

[111] A. Alkassar, T. Nicolay, and M. Rohe, in *International Conference on Computational Science and Its Applications* (Springer, 2005), pp. 634.

[112] R. Duggirala, A. Lal, and S. Radhakrishnan, *Radioisotope thin-film powered microsystems* (Springer Science & Business Media, 2010), Vol. 6.

[113] V. L. Pokrovsky and D. Sun, Physical Review B **76**, 024310 (2007).

[114] P. Somlo, Electronics Letters **11**, 290 (1975).

[115] J. Van Den Biesen, Solid-State Electronics **29**, 529 (1986).

[116] P.-F. Lu and C. Chuang, IEEE Transactions on Electron Devices **39**, 1902 (1992).

[117] C. Beenakker and M. Büttiker, Physical Review B **46**, 1889 (1992).

[118] R. Adler, Proceedings of the IRE **34**, 351 (1946).

[119] X. Zhang, B. J. Rizzi, and J. Kramer, IEEE Transactions on Microwave Theory and Techniques **44**, 2711 (1996).

[120] D.-L. Deng and L.-M. Duan, Physical Review A **88**, 012323 (2013).

[121] G. Katsoprinakis, M. Polis, A. Tavernarakis, A. Dellis, and I. Kominis, Physical Review A **77**, 054101 (2008).

[122] S. Pironio *et al.*, Nature **464**, 1021 (2010).

[123] Y.-Q. Nie, H.-F. Zhang, Z. Zhang, J. Wang, X. Ma, J. Zhang, and J.-W. Pan, Applied Physics Letters **104**, 051110 (2014).

[124] J. G. Rarity, P. Owens, and P. Tapster, Journal of Modern Optics **41**, 2435 (1994).

[125] X. Li, P. L. Voss, J. E. Sharping, and P. Kumar, Physical Review Letters **94**, 053601 (2005).

[126] M. Kues *et al.*, Nature **546**, 622 (2017).

[127] P. Roztocki *et al.*, in *European Quantum Electronics Conference* (Optical Society of America, 2019), p. ea_p_3.

[128] J. E. Bowers *et al.*, in *Next-Generation Optical Communication: Components, Sub-Systems, and Systems V* (International Society for Optics and Photonics, 2016), p. 977402.

[129] M. Feng *et al.*, IEEE Journal of Selected Topics in Quantum Electronics **24**, 1 (2018).

[130] A. Orieux, M. A. Versteegh, K. D. Jöns, and S. Ducci, Reports on Progress in Physics **80**, 076001 (2017).

[131] S. Tanzilli, A. Martin, F. Kaiser, M. P. De Micheli, O. Alibart, and D. B. Ostrowsky, Laser & Photonics Reviews **6**, 115 (2012).

[132] T. Lunghi *et al.*, Opt. Express **26**, 27058 (2018).

[133] W. Bogaerts, D. Pérez, J. Capmany, D. A. B. Miller, J. Poon, D. Englund, F. Morichetti, and A. Melloni, Nature **586**, 207 (2020).

[134] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Reviews of Modern Physics **74**, 145 (2002).

[135] T. Michel, J. Y. Haw, D. G. Marangon, O. Thearle, G. Vallone, P. Villoresi, P. K. Lam, and S. M. Assad, Physical Review Applied **12**, 034017 (2019).

[136] T. Tomaru, OSA Continuum **2**, 2554 (2019).

[137] H. Jin *et al.*, Physical review letters **113**, 103601 (2014).

[138] P. Imany, J. A. Jaramillo-Villegas, O. D. Odele, K. Han, D. E. Leaird, J. M. Lukens, P. Lougovski, M. Qi, and A. M. Weiner, Opt. Express **26**, 1825 (2018).

[139] M. C. Teich, B. E. Saleh, A. V. Sergienko, J. T. Fourkas, R. Wolleschensky, M. Kempe, and M. C. Booth, (Google Patents, 2006).

[140] P. Bronner, A. Strunz, C. Silberhorn, and J.-P. Meyn, European journal of physics **30**, 1189 (2009).

[141] J. T. Francis, X. Zhang, Ş. K. Özdemir, and M. Tame, Quantum Science and Technology **2**, 035004 (2017).

[142] M. Gräfe *et al.*, Nature Photonics **8**, 791 (2014).

[143] F. Xu, J. H. Shapiro, and F. N. C. Wong, Optica **3**, 1266 (2016).

[144] P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, Physical Review A **47**, R2472 (1993).

[145] Y. Liu *et al.*, Nature **562**, 548 (2018).

[146] D. Rusca, H. Tebyanian, A. Martin, and H. Zbinden, arXiv preprint arXiv:2004.08307 (2020).

[147] D. Rusca, T. van Himbeeck, A. Martin, J. B. Brask, W. Shi, S. Pironio, N. Brunner, and H. Zbinden, arXiv preprint arXiv:1904.04819 (2019).

[148] M. A. Wayne, E. R. Jeffrey, G. M. Akselrod, and P. G. Kwiat, Journal of Modern Optics **56**, 516 (2009).

[149] M. A. Wayne and P. G. Kwiat, Opt. Express **18**, 9351 (2010).

[150] M. Kues, C. Reimer, J. M. Lukens, W. J. Munro, A. M. Weiner, D. J. Moss, and R. Morandotti, Nature Photonics **13**, 170 (2019).

[151] S. Tanzilli, W. Tittel, H. De Riedmatten, H. Zbinden, P. Baldi, M. DeMicheli, D. B. Ostrowsky, and N. Gisin, The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics **18**, 155 (2002).

[152] M. Tomamichel and R. Renner, Physical Review Letters **106**, 110506 (2011).

[153] G. Vallone, D. G. Marangon, M. Tomasin, and P. Villoresi, Physical Review A **90**, 052327 (2014).

[154] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, Reviews of Modern Physics **89**, 015002 (2017).

[155] F. Furrer, T. Franz, M. Berta, A. Leverrier, V. B. Scholz, M. Tomamichel, and R. F. Werner, Physical Review Letters **109**, 100502 (2012).

[156] D. Kleinman, Physical Review **128**, 1761 (1962).

[157] L. Oesterling *et al.*, Journal of Modern Optics **62**, 1722 (2015).

[158] A. Sarkar and C. M. Chandrashekar, Scientific Reports **9**, 12323 (2019).

[159] E. Trentin, F. Schwenker, N. El Gayar, and H. M. Abbas, Neural Processing Letters **48**, 643 (2018).

[160] N. D. Truong, J. Y. Haw, S. M. Assad, P. K. Lam, and O. Kavehei, IEEE Transactions on Information Forensics and Security **14**, 403 (2018).