Quantum Science and Technology

PAPER

Post-selection in noisy Gaussian boson sampling: part is better than whole

Tian-Yu Yang, Yi-Xin Shen, Zhou-Kai Cao and Xiang-Bin Wang

1 State Key Laboratory of Low Dimensional Quantum Physics, Department of Physics, Tsinghua University, Beijing 100084, People's Republic of China
2 Jinan Institute of Quantum technology, SAICT, Jinan 250101, People's Republic of China
3 Shanghai Branch, CAS Center for Excellence and Synergetic Innovation Center in Quantum Information and Quantum Physics, University of Science and Technology of China, Shanghai 201315, People's Republic of China
4 International Quantum Academy, Shenzhen 518048, People's Republic of China
5 Frontier Science Center for Quantum Information, Beijing, People's Republic of China

* Author to whom any correspondence should be addressed.
E-mail: xbwang@mail.tsinghua.edu.cn

Keywords: quantum computation, boson sampling, error mitigation

Abstract

Gaussian boson sampling (GBS) is originally proposed to show quantum advantage with quantum linear optical elements. Recently, several experimental breakthroughs based on GBS pointing to quantum computing supremacy have been presented. However, due to technical limitations, the outcomes of GBS devices are influenced severely by photon loss. Here, we present a practical method to reduce the negative effect caused by photon loss. We first show with explicit formulas that a GBS process can be mapped to another GBS processes. Based on this result, we propose a post-selection method which discards low-quality data according to our criterion to improve the performance of the final computational results, say part is better than whole. As an example, we show that the post-selection method can turn a GBS experiment that would otherwise fail in a ‘non-classicality test’ into one that can pass that test. Besides improving the robustness of computation results of current GBS devices, this post-selection method may also benefit the further development of GBS-based quantum algorithms.

1. Introduction

The potential speedup [1–3] of quantum algorithms over their classical counterparts makes quantum computation a hot topic nowadays. In the past few years, people have witnessed the fast development of quantum computing technologies. One of the central issues these years is to demonstrate the quantum supremacy. Towards this end, several breakthroughs of quantum computing occurred in different physical systems, such as superconducting systems [4–7], trapped ions [8], and quantum optics [9–13], etc.

Notwithstanding these achievements, we are still in an era of noisy intermediate-scale quantum (NISQ) [2]. There are mainly two kinds of quantum algorithms that might show an advantage over their classical counterparts on NISQ devices. They are approximate optimization algorithms, e.g. variational quantum eigenvalue solver [14], quantum approximate optimization algorithm [15], and quantum sampling problems, e.g. random circuit sampling [4, 5, 16] and Gaussian boson sampling (GBS) [9–11, 17–21]. Due to the imperfections on NISQ devices, errors frequently occur while implementing those quantum algorithms. However, quantum error correction is very expensive in NISQ devices, as those devices only own tens of qubits and high-error quantum gates. Thus, it is important to implement error mitigation methods to improve the performance of NISQ devices.

GBS problem is a variant of boson sampling problem proposed by Aaronson and Arkhipov [18, 22]. GBS was originally designed to facilitate the implementation of the original boson sampling problem without changing the problem’s computational complexity. Besides showing evidence for quantum advantage, GBS
devices have been linked to several potential applications, such as quantum chemistry [23, 24], graph theory [25–27], quantum approximate optimization [28], and quantum machine learning [29].

In this work, we focus on error mitigation in the GBS problem. In a GBS device, one of the main error sources is photon loss. Recently, error mitigation of photon loss in GBS devices has been studied in [30]. Two methods based on extrapolation technique and probability estimation is presented. The aim of those methods is to improve the estimation of the sampling probabilities. For the first time, it shows that error mitigation methods are applicable for computing models based on quantum photonic devices. However, the methods in [30] do not work in the case of large photon loss and hence the expected quantum advantage of the result could be undermined. In addition, those methods require collecting a sufficient number of samples to obtain the exact probability distribution.

We first show with explicit formulas that a GBS process can be mapped to another GBS processes. Based on this result, we proposed our post-selection method enables us to conduct a GBS task with overall transmission rate $\eta'$ even when we only have a GBS device with overall transmission rate $\eta$. In comparison with methods in [30], our method is more robust to errors caused by photon loss. Moreover, our method does not need to collect the samples from the quantum computing devices for probability distribution of the output pattern, say, it generates high-quality samples directly.

As an example, we use a 'non-classicality test' [31] to test the performance of the post-selection method. This 'non-classicality test' is widely used in GBS experiments as a reference to check whether quantum advantage might exist [10, 11, 21]. Numerical results show that the post-selection method can turn a GBS experiment that would otherwise fail in the ‘non-classicality test’ into one that can pass that test. This example, i.e. the post-selection method can help the GBS experiment to surpass the ‘non-classicality test’, shows the value of the post-selection method. Besides, by improving the transmission rate, the post-selection method is also beneficial for increasing circuit depth, which is another issue in recent GBS experiments [32].

This article is organized as follows. In section 2, we give an overview of the GBS protocol and provide some background knowledge about quantum optics and the Gaussian state. In section 3, we prove several theorems that demonstrate the validity of the post-selection method and then explain the post-selection method in detail. In section 4, we give further analysis about the performance of the post-selection method. In section 4.1, we analyze the performance of the post-selection method under different experimental conditions and some numerical results are given. In section 4.2, we analyze the effect of other experimental imperfections for the post-selection process, including non-uniform loss, the limited photon number resolution (PNR) capability of the detector, and the dark count rate. In section 4.3, we take the ‘non-classicality test’ to test the performance of the post-selection method. In section 4.4, we compare the differences between the post-selection method and the existing error mitigation methods. Finally, a summary is presented in section 5.

2. GBS

The architecture of the GBS [19] is shown in figure 1. The first step is to inject an input quantum state consists of $K$ single-mode squeezed states (SMSSs) and $M$-K vacuum states into an $M$-mode passive linear optical network. PNR detectors then measure the output quantum state, and an output pattern $\vec{n} = n_1 n_2 \ldots n_M$ is generated, where $n_i$ is the detected photon number in the $i$th mode. The probability distribution of the output pattern $\vec{n} = n_1 n_2 \ldots n_M$ is

$$P_{\text{out}}(\vec{n}) = \frac{1}{n! \sqrt{\sigma_Q}} \text{Haf}(A_0)$$

$$\sigma_Q = \sigma + \frac{I_{2M}}{2}$$

where Haf is a computationally hard matrix function called hafnian and $\sigma$ is the covariance matrix of the output Gaussian state. The matrix $A_0$ is determined by the input state, the passive linear optical network and the output sample pattern [19].

An SMSS can be described by following equations [33–36]:

$$|r, \phi\rangle = \hat{S}(r, \phi)|0\rangle$$

$$= \frac{1}{\sqrt{\cosh(r)}} \exp\left( -\frac{1}{2} \exp(i\phi)(\hat{a}^\dagger)^2 \tanh r \right) |0\rangle$$

$$= \frac{1}{\sqrt{\cosh(r)}} \sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{n!} \left( -\frac{1}{2} \exp(i\phi) \tanh r \right)^n |2n\rangle,$$
Figure 1. Schematic setup of Gaussian boson sampling. First, $K$ single-mode squeezed states (SMSSs) are injected into a passive linear optical network with $M$ optical modes. Then the output quantum state is measured by photon number resolving detectors $D_1, D_2, \ldots, D_M$ in each mode and an output pattern $\bar{n} = n_1 n_2 \ldots n_M$ is generated.

where $\hat{S}(r, \phi)$ is the squeezing operator, $r$ is squeezing strength and $\phi$ is the phase. The photon number distribution of an SMSS is

$$p^{(r)}(n) = \begin{cases} \frac{1}{\cosh r} \frac{(\tanh r)^n}{(n/2)!} & n = 0, 2, 4, \ldots \\ 0 & n = 1, 3, 5, \ldots \end{cases},$$

which corresponds to the negative binomial distribution [37].

The probability distribution of the total photon number of the injected quantum state can be calculated through the convolution of probability distributions of SMSSs, i.e.

$$p^{(r)}(|\bar{n}|) = \sum_{n_i \text{ even for } i = 1, \ldots, K} \prod_{i=1}^{K} \left( \frac{1}{2} + \frac{n_i}{2} - 1 \right) \frac{(\tanh r_i)^{n_i}}{\cosh r_i},$$

where $r_i$ represents the squeezing strength of the SMSS in input mode $i$ and $|\bar{n}| = \sum_{i=1}^{K} n_i$. If $r_i = r$ where $i = 1, 2, \ldots, K$, the probability distribution of the total photon number is

$$p^{(r)}(N) = \left( \frac{N + K/2 - 1}{N/2} \right)^K \frac{1}{\cosh(r)} \tanh^N(r),$$

when $N$ is even, and

$$p^{(r)}(N) = 0,$$

when $N$ is odd.

Passive linear optical networks can preserve the total photon number between the input and output state. They can be used to describe the photon loss process [38–40]. Denote the transform operator corresponding to a passive linear optical network as $\hat{U}_0$ and $\hat{a}^T = (\hat{a}_1^\dagger, \hat{a}_2^\dagger, \ldots, \hat{a}_{2M}^\dagger)^T$ where $\hat{a}_i^\dagger$ is the creation operator. Let the first $M$ modes be actual modes, and the last $M$ modes be loss (environment) modes. Loss modes are initially in the vacuum state and will be traced out in the last step. When photon loss is uniform, we have

$$\hat{U}_0 \hat{a}^\dagger \hat{U}_0^\dagger = \Lambda_0 \hat{a}^\dagger,$$

where

$$\Lambda_0 = \begin{pmatrix} \frac{\sqrt{\eta}}{\sqrt{1 - \eta}} \\ \frac{\sqrt{1 - \eta}}{\sqrt{\eta}} \end{pmatrix},$$

and

$$\Lambda_0^{(K)} = \begin{pmatrix} \Lambda_M \\ 0 \end{pmatrix},$$

where $0 \leq \eta \leq 1$ is the transmission rate.

Denote the transform operator corresponding to the passive linear optical network in an ideal GBS as $\hat{U}_1$. We have

$$\hat{U}_1 \hat{a}^\dagger \hat{U}_1^\dagger = \Lambda_1 \hat{a}^\dagger,$$

where

$$\Lambda_1 = \begin{pmatrix} \Lambda_M & 0 \\ 0 & I_M \end{pmatrix},$$
where $\Lambda_M$ is a $M \times M$ unitary matrix and $I_M$ is an identity matrix with rank $M$. The output state of a lossy GBS experiment is thus

$$\hat{\rho}_{\text{out}} = \text{Tr}_M \left( \hat{U}_1 \hat{U}_0 (\hat{\rho}_{\text{in}} \otimes |0\rangle \langle 0|) \hat{U}_0^\dagger \hat{U}_1^\dagger \right).$$

(9)

A brief introduction of photon detectors can be found in [12]. There are mainly three types of errors in photon detectors, i.e. detection efficiency, PNR capability and dark count rate. The detection efficiency can be considered as a part of overall transmission rate. PNR capability indicates the maximum photon number that can be accurately resolved. The dark count rate indicates the probability of counting the cases where no photon reaches the detector.

3. Post-selection method

3.1. Mapping between different GBS experiments

We first derive a formula to calculate the probability distribution of GBS outputs in a different form from equation (1), which will be very useful in proving the validity of the post-selection method.

**Lemma 1.** Denote the input squeezing strengths of a GBS process as $\{ r_i \}$, the overall transmission rate as $\eta$. The probability distribution of the output pattern is equal to

$$P^{(r, \eta)}_{\text{out}} (\bar{n}) = \sum_{m,l} v_1 (\bar{m}|\bar{n}) v_1^* (\bar{l}|\bar{n}) \sqrt{P^{(r, \eta)}_{\text{in}} (\bar{m}) P^{(r, \eta)}_{\text{in}} (\bar{l})},$$

(10)

where $P^{(r, \eta)}_{\text{in}} (\bar{m})$ is the probability distribution of the input pattern undergoes a photon loss process with overall transmission rate $\eta$, $v_1 (\bar{m}|\bar{n}) = \langle \bar{m}| \hat{U}_1 |\bar{n} \rangle$, and $\hat{U}_1$ is the transform operator corresponding to the passive linear optical network.

The proof of lemma 1 is given in appendix A. Next, we give theorem 1. It considers the lossless case, i.e. $\eta = 1$. Theorem 1 shows that GBS experiments with different input squeezing strengths can be mapped to each other under certain conditions.

**Theorem 1.** Denote the probability distribution of the output pattern of GBS devices with input squeezing strengths $\{ r_i \}$ as $P^{(r_i)}_{\text{out}} (\bar{n})$, where $\{ r_i \} = \{ r_1, r_2, \ldots, r_M \}$ represents the squeezing strengths of input squeezed states, and $\bar{n} = n_1 \ldots n_M$ is the output pattern. We have

$$P^{(r, \eta)}_{\text{out}} (\bar{n}) = P^{(r)}_{\text{out}} (\bar{n}) \times \frac{P^{(r_i)}_{\text{in}} (\bar{m})}{P^{(r_i)}_{\text{out}} (\bar{m})},$$

(11)

provided that $\{ r_i' \}$ satisfies $\frac{\tanh r_i'}{\tanh r_i} = c$, for $i = 1, 2, \ldots, M$ and $c$ is a constant real number. Here $|\bar{n}| = \sum_{i=1}^{M} n_i$ is the total photon number, $n_i$ is even for all $i$ and $P^{(r_i)}_{\text{in}} (|\bar{n}|)$ is the probability of generating an output pattern with total photon number $|\bar{n}|$.

This result inspires us to find theorem 2. Theorem 2 shows that when there is uniform photon loss, GBS experiments with different input squeezing strengths can still be mapped to each other under certain conditions. This map changes the effective transmission rate in the sampling process.

**Theorem 2.** Denote the probability distribution of the output pattern of GBS devices with input squeezing strengths $\{ r_i \}$ and overall transmission rate $\eta$ as $P^{(r_i, \eta)}_{\text{out}} (\bar{n})$, where $\{ r_i \} = \{ r_1, r_2, \ldots, r_M \}$ represents the squeezing strengths of the input squeezed states, and $\bar{n} = n_1 \ldots n_M$ is the output pattern. We have

$$P^{(r_i, \eta')}_{\text{out}} (\bar{n}) = P^{(r_i, \eta)}_{\text{out}} (\bar{n}) \times \frac{P^{(r_i)}_{\text{out}} (|\bar{n}|)}{P^{(r_i)}_{\text{out}} (|\bar{n}|)} \left( \frac{\eta'}{\eta} \right)^{|\bar{n}|},$$

(12)

where $\eta' = 1 - (1 - \eta)c$, provided that $\{ r_i' \}$ satisfies $\frac{\tanh r_i'}{\tanh r_i} = c$, for $i = 1, 2, \ldots, M$ and $c$ is a constant real number. Here $|\bar{n}| = \sum_{i=1}^{M} n_i$ is the total photon number and $P^{(r_i)}_{\text{out}} (|\bar{n}|)$ is the probability of generating an output pattern with total photon number $|\bar{n}|$.

Proofs of theorems 1 and 2 are given in appendices B and C respectively. Notice that

$$\frac{P^{(r_i)}_{\text{out}} (|\bar{n}|)}{P^{(r_i)}_{\text{out}} (|\bar{n}|)} \left( \frac{\eta'}{\eta} \right)^{|\bar{n}|} = c^{-|\bar{n}|} \left( \frac{\eta'}{\eta} \right)^{|\bar{n}|}.$$
Theorem 2 shows that we can improve the transmission rate of the sampling process by a post-selection process. We will discuss this post-selection method in detail in the next subsection.

3.2. Post-selection method

We propose the following post-selection method to mitigate the effect of photon loss for GBS devices. For a sample obtained from a GBS experiment, we choose the following post-selection probability

\[ P_{\text{post}}(|n|) = \alpha^{s-|n|} \left( \frac{n'}{\eta} \right)^{|n|} \]

\[ \eta' = 1 - (1 - \eta) c \]

where
\[ c = \frac{\tanh r_i}{\tanh r_i^*} < 1, \]

to decide whether to keep the sample or not. Here a cut-off photon is needed. Denote the cut-off photon number as \( N_0 \). If the total photon number of a sample exceeds the cut-off photon number \( N_0 \), this sample will be discarded. This guarantees the correctness of the post-selection method as it ensures that the post-selection probability is always less than 1. In order to retain as many samples as possible in the post-selection process, the parameter \( \alpha \) in equation (14) is chosen to satisfy the following equation

\[ P_{\text{post}}(N_0) = 1. \] (15)

This makes samples with a total photon number \( N_0 \) to be retained with probability 1. The post-selection protocol is summarized in protocol 1.

According to theorem 2, apart from a cut-off and discarding a proportion of samples, the post-selection process maps a GBS experiment with input squeezing strength \( \{r_i\} \) and an overall transmission rate \( \eta \) to a GBS experiment with input squeezing strengths \( \{r_i'\} \) and an overall transmission rate \( \eta' \). If \( \eta = 1 \), according to theorem 1, the post-selection process maps a GBS experiment with input squeezing strength \( \{r_i\} \) to a GBS experiment with input squeezing strengths \( \{r_i'\} \). This means that our post-selection method can help to extend the application of GBS devices in the sense that the post-selection method makes it possible to conduct a GBS task with input squeezing strengths \( \{r_i'\} \) even when we only have a GBS device with input squeezing strength \( \{r_i\} \). If \( \eta \neq 1 \), by choosing \( c < 1 \), the effective transmission rate \( \eta' \) of the retained samples is improved. Below, we shall discuss the effect of the cut-off and the discard. As we will show, the effect of them can be small while an improvement in the effective transmission rate \( \eta' \) is obtained.

**Protocol 1. Post-selection method.**

**Step 1:** Choosing parameters \( c \) and \( N_0 \) depends on the experimental conditions.

**Step 2:** Inject input SMSSs with parameter set \( \{r_i\} \) into the passive linear optical network. The overall transmission rate is \( \eta \).

**Step 3:** Measure the output state and calculate the total photon number.

**Step 4:** According to the total photon number, we decide whether to retain or discard the specific outcome at that time based on post-selection probability given in equation (14).

**Step 5:** Repeat step 2–4 for many times we obtain the outcomes with probability distribution given by the input SMSSs with parameter set \( \{r_i'\} \) and overall transmission rate \( \eta' \).

Usually, the time cost for a classical computer to generate a specific sample increases exponentially with the total photon number of that specific sample [41]. Here we need to emphasize that the existence of a cut-off does not restrict the maximum photon number that can be obtained in the experiment. In fact, the cut-off photon number can be set as the maximum total photon number obtained in the experiment. According to equation (15), samples with a total photon number \( N_0 \) will be retained 100% of the time. So all samples with the maximum total photon number are retained in the post-selection process.

Moreover, to demonstrate the quantum advantage with GBS, a larger cut-off photon number does not necessarily mean better. In fact, hardness arguments of GBS nowadays are based on the conjunction of collision-free outcomes [18, 19, 42]. They assume that the number of modes is at least quadratic in the number of photons, i.e. \( N \sim O(M^2) \). So a choice of the cut-off photon number \( N_0 \) in order \( O(\sqrt{M}) \) might be enough.

Another concern of the post-selection method is about the yield. By yield, we mean the percentage of the samples that are retained after the post-selection process. The yield can be calculated using the following equation:
shows that the yield can be acceptable when an obvious improvement of effective
\[ y = \sum_{|n|=0}^{N_i} P_{\text{post}}(|n|) P_{\text{out}}^{r_i, \eta}(|n|). \] (16)

The above equation shows that the yield is determined by experimental conditions and the parameter \( c \) of the post-selection method. A low yield will increase the time needed for generating samples. This can be avoided by a proper choice of the parameter \( c \). By increasing \( c \), the yield can be improved. When \( c (c < 1) \) converges to 1, the yield \( y (y < 1) \) also converges to 1. In the same time, the effective transmission rate \( \eta' \) decreases to \( \eta \).

The sample generating speed of recent GBS experiments is about \( 10^4 \text{ s}^{-1} \) [9, 21]. The numerical results in sections 4.1 and 4.3 shows that the yield can be acceptable when an obvious improvement of effective transmission rate is obtained.

Although reducing the sample size, the post-selection process can mitigate the effect of photon loss. More specifically, our post-selection method allows us to conduct a GBS task with overall transmission rate \( \eta' \) even when we only have a GBS device with overall transmission rate \( \eta \). In the next section, we will show that sacrificing a proportion of samples could be worthwhile. An interesting example is given in section 4.3, where we show that the post-selection method can make GBS experiments that previously failed the ‘non-classicality test’ to pass the ‘non-classicality test’ [31].

4. Further analysis on performance of the post-selection method

4.1. Analysis and numerical results of performance

In this subsection, we analyze the performance of the post-selection method. The numerical simulations here are based on the analytical equations in section 3.2.

First, let us analyze how experimental conditions influence the performance of the post-selection method. According to equation (14), we find that

\[ \eta' = 1 - (1 - \eta) \frac{\tanh r_i}{\tanh r'_i} < 1 - (1 - \eta) \tanh r_i. \] (17)

This shows that the smaller the input squeezing strengths \( r_i \) and the transmission rate \( \eta \), the more the transmission rate could be increased by the post-selection method. Denote \( \eta'_{\text{max}} = 1 - (1 - \eta) \tanh r_i \) as the maximum effective transmission rate that can be obtained by the post-selection method. The relation between \( \eta'_{\text{max}}, r_i \), and \( \eta \) is shown in figure 2.

In the following, we evaluate the performance of the post-selection method by numerical simulations using parameters similar to two recent GBS experiments [11, 21]. As can be seen from figures 3(a) and (b), after post-selection, the effective transmission rate \( \eta' \) increases as the target squeezing strength \( r' \) increases while the yield is not too low comparing to the sample generating speed of those experiments.

The parameters in figure 3(a) correspond to those in the GBS experiment [21]. The squeezing strength of the input squeezed states is \( r = 1.1 \). The number of input squeezed states is 216. The overall transmission rate of the device is about 0.32. The maximum photon number of the measured sample is 219. When \( r' = 2.1 \), the effective transmission rate becomes \( \eta' \approx 0.44 \), which is about 38\% higher than the original transmission rate, while the yield is about \( 8.9 \times 10^{-7} \). The low yield is mainly due to the low probability of experimentally observing samples with total photon number \( N = 219 \). According to [21], the number of samples generated per second in the experiment is about \( 2 \times 10^4 \). So, if \( r' = 2.1 \) is chosen, the experimental device can averagely generate a sample in hundreds of seconds.

The parameters in figure 3(b) correspond to those in the GBS experiment [11]. The squeezing strength of the input squeezed states is \( r = 1.4 \). The number of input squeezed states is 50. The overall rate of the device is about 0.5. The maximum photon number of the measured sample is 113. When \( r' = 2.5 \), the effective transmission rate becomes \( \eta' \approx 0.55 \), which is about 10\% higher than the original transmission rate, while the yield is about 0.0617. The high yield reflects the high probability of observing a sample with total photon number \( N = 113 \) experimentally. Notice that the experiment [11] uses threshold detectors. So, the total photon number is not always equal to the total number of detector clicks. We will demonstrate in section 4.2 that the limited PNR capability does not negatively affect the performance of post-selection method.

4.2. Post-selection and other noise

In this subsection, we consider how other experimental imperfections affect the performance of the post-selection method, including non-uniform photon loss, the limited PNR capability of detectors and the dark count rate.

Although uniform loss is a good approximation for the current GBS experiments, non-uniform loss will inevitably occur in experiments [9–11, 21]. By splitting a general loss process into a uniform loss process
Figure 2. The largest effective transmission rate $\eta'_{\text{max}}$ for different transmission rates and squeezing strengths. Assuming $\eta_i = \eta$ for $i = 1, 2, \ldots, K$.

Figure 3. Performance of the post-selection method. $r$ is input squeezing strength, $r'$ is the target squeezing strength to be simulated. $K$ is the number of input single-mode squeezed states. $N_0$ is cut-off photon number. $\eta$ is the overall transmission rate. We use the results of Gaussian boson sampling devices with input squeezing strength $r$ to simulate that of $r'$ by the post-selection method of protocol 1. Red points correspond to $\eta'$ which is the effective transmission rate in final samples. Orange points correspond to yield which is the percentage of preserved outputs.
with transmission rate $\eta_u$ and a different non-uniform loss process, we show that the post-selection method can mitigate the effect of non-uniform loss process. The percentage of samples affected by the non-uniform loss part is decreased, while the transmission rate $\eta_u$ of the uniform loss part is improved. Proof of this can be found in appendix D. In addition, in most cases, the non-uniform loss in GBS experiments is largely attributed to the different detection efficiencies of photon detectors. We give a modified post-selection protocol to deal with the non-uniform loss due to the different detection efficiencies of photon detectors. A description of this can be found in appendix E.

We assumed in our previous analysis that the total photon number of detected samples can be precisely resolved in GBS experiments. However, even when the detectors’ PNR capability is limited, the post-selection method can still be beneficial for the GBS experiment. As we will show below, the post-selection method reduces the number of samples with wrong total photon numbers due to the limited PNR capability of detectors.

Let the real output pattern in a GBS experiment be $\bar{n}$ and the measured pattern be $\bar{m}$. This error event should be retained with probability $P_{\text{post}}(|\bar{m}|)$ in the post-selection process. But since the total photon number of the detected sample is $|\bar{m}|$, the actual probability that this error event is retained is $P_{\text{post}}(|\bar{m}|)$. Thus the probability of this error event occurring after post-selection is equivalent to $P_{\text{post}}(|\bar{m}|)$ times the original probability. According to equation (14), the smaller the total photon number of the sample, the smaller the probability of being retained. So we have $P_{\text{post}}(|\bar{m}|) < 1$. That is, after the post-selection process, the number of events exceeding the PNR capability of the detector is reduced.

Another imperfection in detectors is the inherent dark counts. A dark count leads to a larger total photon number in the measured sample. This causes an increase in the probability of that sample being retained by the post-selection process. The result of it is an increase in the effective dark count rate in the experiment.

Denote the dark count rate of a detector as $p_D$. If a dark count occurs in a detector, then the probability that the sample is retained is $P_{\text{post}}(N + 1)$, where $N$ is the actual total photon number of the sample. This corresponds to an increase in the effective dark count rate. Thus, the effective dark count rate becomes

$$P'_{D} = \frac{p_D P_{\text{post}}(N + 1)}{p_D P_{\text{post}}(N + 1) + (1 - p_D) P_{\text{post}}(N)}$$

$$= \frac{p_D \left(1/\cosh(1 - \eta)\right)}{p_D \left(1/\cosh(1 - \eta)\right) + 1 - p_D},$$

which is bounded by

$$P'_{D} \leq \frac{p_D \left(1 + 1/\tanh(1 - \eta)\right)}{p_D \left(1 + 1/\tanh(1 - \eta)\right) + 1 - p_D}. \quad (20)$$

So, the dark count rate is increased after the post-selection process. However, the increase of the dark count rate is bounded and is generally not very large. Moreover the dark count rates in experiments are usually very low (about $10^{-3}$ to $10^{-4}$). Thus this increase does not have a significant effect on the experimental results. As we will show in the examples of 4.3, the trade off of slightly increasing the dark count rate could be a worthwhile one.

# 4.3. Enhancing the robustness of GBS devices

In this subsection, we show that the post-selection method can enhance the robustness of recent GBS devices. Recent GBS experiments achieve great success, in the sense that those experiments achieved a large number of optical modes and detected samples with large total photon numbers (about 100 to 200) [9, 11, 21]. However, imperfections in those experiments have brought these quantum advantage results into question [31, 32, 43–45]. Classical methods can simulate the sampling results when the experimental errors exceed a certain range, e.g. [31, 32]. Therefore, it might be worthwhile to discard a proportion of samples in order to obtain more robust experimental results to meet the challenges from classical algorithms. In the following, we provide some examples as support for our view.

First, we will show that the post-selection method can make a GBS experiment that could have been effectively and approximately simulated by a classical algorithm incapable of being simulated by that classical algorithm. This so called ‘non-classicality test’ [31] is currently used by a variety of GBS experiments to test whether their experiments might achieve a quantum advantage result [10, 11, 21, 31]. The classical algorithm
used in this example can efficiently simulate the sampling process of a GBS experiment up to an error $\varepsilon$ when the following condition is satisfied:

$$\text{sech} \left( \frac{1}{2} \Theta \left[ \ln \left( \frac{1 - 2\eta \sqrt{D}}{\eta e^{-2r} + 1 - \eta} \right) \right] \right) > e^{-\varepsilon^2/4K},$$

where $\Theta$ is the ramp function $\Theta(x) = \max(x, 0)$, $r$ is the squeezing strength of the input squeezed states, $\eta$ is the overall transmission rate, $K$ is the total number of input squeezed states, and $q_D$ is the dark count rate of the photon detectors. When the inequality above has no solution for $\varepsilon \in [0, 1]$, the GBS experiment pass the ‘non-classicality test’. Denote $\varepsilon_0$ as the minimum value which satisfies the inequality given in equation (21), that is:

$$\text{sech} \left( \frac{1}{2} \Theta \left[ \ln \left( \frac{1 - 2\eta \sqrt{D}}{\eta e^{-2r} + 1 - \eta} \right) \right] \right) = e^{-\varepsilon_0^2/4K}. \tag{22}$$

When $\varepsilon_0$ is larger than 1, this classical simulation algorithm fails.

In figure 4, we give three examples to show that the post-selection method can make a GBS sampling process that could be approximately simulated by the classical algorithm unable to be simulated by it.

The blue lines show the error upper bound of the classical algorithm for simulating corresponding lossy GBS experiments. The squeezing strength of the input squeezed states is $r$. The dark count rate of the detectors is $q_D$. The input number of input SMSSs is $K$. The red lines show the error upper bound of the classical algorithm for simulating corresponding lossy GBS sampling experiments after a post-selection process. The target squeezing strength is $r'$. The vertical axis is the simulation error upper bound $\varepsilon_0$. The horizontal axis corresponds to the overall transmission rate $\eta$. Here we use the formula in section 4.2 to calculate the change of the dark count rate $q_D$. The effective dark count rate after the post-selection process is denoted as $q_D'$. The exponential decay of the transmission rate limits the increase of the circuit depth of GBS experiments. Post-selection method can help to mitigate the decrease in overall transmission rate when $\varepsilon_0$ exceeds 1. In this case the yield of the post-selection process is $y = 1.47 \times 10^{-5}$. When $\varepsilon_0$ is larger than 1, this classical simulation algorithm fails.

In figure 4(a), where $r = 1.38$, $r' = 1.5$, $K = 150$, $q_D = 0.0001$, if the overall transmission rate is $\eta = 0.105$, the sampling process can be approximately simulated by the classical algorithm (corresponding to the blue dots in the figure 4(a)). Set $N_0 = 117$, after the post-selection, the whole sampling process can no longer be simulated by the classical algorithm (corresponding to the red dots in figures), as the error upper bound $\varepsilon_0$ exceeds 1. In this case the yield of the post-selection process is $y = 1.47 \times 10^{-5}$. In figure 4(b), where $r = 1.6$, $r' = 1.8$, $K = 50$, $q_D = 0.0001$, if the overall transmission rate $\eta = 0.18$, this sampling process can be approximately simulated by the classical algorithm (corresponding to the blue dots in the figure 4(b)). Set $N_0 = 130$, after the post-selection, the whole sampling process can no longer be simulated by the classical algorithm (corresponding to the red dots in figures), as the error upper bound $\varepsilon_0$ exceeds 1. In this case the yield of this post-selection process is $y = 1.48 \times 10^{-5}$. Of course, there is no way to know whether one will discover new algorithms to enhance the simulation power of classical computers. However, the above examples show the value of using post-selection method to reduce photon loss for enhancing the robustness of GBS devices.

Moreover, the post-selection method is helpful for increasing the circuit depth of the GBS devices. In a recent article [32], a classical algorithm is designed to approximately simulate GBS experiments with shallow circuit depth. A major difficulty in increasing the circuit depth is that the transmission rate decreases exponentially with the number of layers in the circuit. If the transmission rate of one layer in the circuit is $\eta_0$, then the transmission rate of $D$ layer is $\eta = (\eta_0)^D$. The exponential decay of the transmission rate limits the circuit depth of the GBS experiments. Post-selection method can help to mitigate the decrease in transmission rate which is beneficial for increasing the circuit depth of GBS devices.

### 4.4. Comparison with the existing methods

Two error mitigation methods aim to mitigate the photon loss error of GBS devices are given in [30]. However, they are very different both in effectiveness and application scenarios from the post-selection method. The most significant difference is, according to examples and analysis given in [30], those methods are mainly applicable to the case of small photon loss and become unworkable when the effect of photon loss in GBS devices is severe. In contrast, according to our analysis in sections 3.2 and 4.1, the post-selection method can be used under conditions when photon loss rate is large.

Besides, the goal of the methods given in [30] is very different from the post-selection method. The methods in the article [30] require measuring the probability distribution of the samples first and then
Figure 4. Simulation errors of the classical algorithm given in [31] for different experimental parameters. The horizontal axis is the transmission rate $\eta$ of the real sampling process. The vertical axis is the error upper bound $\epsilon_0$ of the classical simulation algorithm. When the error upper bound $\epsilon_0$ on the vertical axis exceeds 1 (dashed line in the figure), the classical simulation algorithm fails. Parameter $r$ is the squeezing strength in a real setup and $r'$ is the target squeezing strength by our post selection. The blue lines show the corresponding simulation error upper bound of the classical algorithm for lossy Gaussian boson sampling experiments with $r$ and $\eta$. The red lines show the simulation error upper bound of the classical algorithm for corresponding lossy Gaussian boson sampling experiments with a post-selection by our method. Say, at any point $\eta$ in the red lines, the value of $\epsilon_0$ is calculated by target squeezing strength $r'$ and transmittance $\eta'$ mapped from $r$ and $\eta$ by our post selection method. $K$: number of input single-mode squeezed states, $q_D$: dark count rate, $q'_D$: effective dark count rate after the post selection process, $N_0$: cut-off photon number, $y$: yield.
obtaining the error-mitigated probability distribution. The post-selection method, on the other hand, does not require collecting and estimating the probability of the samples. In fact, the post-selection method works directly on the samples and the result of the post-selection process is the error-mitigated samples.

For GBS devices with a large number of optical modes, it is difficult to precisely measure the probability of generating a specific sample pattern. Because it requires collecting a huge number of samples. The solution given in [30] is to no longer do error mitigation for the generating probability of a specific sample, but for the generating probability of a class of samples (i.e. the coarse-grained probability). However, if error mitigation is done only for the coarse grained probability, it remains to be clarified what effect it has on the whole sampling process.

5. Summary

In this article, we proved that a lossy GBS process can be mapped to other lossy GBS processes under certain conditions. Based on this, we proposed a method to suppress the effect of photon loss for GBS devices. Our method takes post-selection of the results of real experiments so as to obtain high-quality sampling results of another sources. It enables us to conduct a GBS task with overall transmission rate $\eta'$ even when we only have a GBS device with overall transmission rate $\eta$.

This error mitigation method proposed in this paper is easy to be performed in nowadays’ GBS devices. Analytical and numerical results showed that the post-selection method has good performance in a wide range of experimental conditions. We also showed that the post-selection method can turn a GBS experiment that would otherwise fail in a ‘non-classicality test’ into one that can pass that test. This example shows the potential value of the post-selection method.

We also considered how other experimental imperfections affect the performance of the post-selection method, including non-uniform photon loss, the detector’s limited PNR capability and the dark count rate. The errors due to both the non-uniform loss and the limited PNR capability of the detector can be reduced by the post selection process, while the dark count rate is slightly increased by the post-selection process. Since the dark count rate is generally small, the increase of dark count rate has less impact on the GBS experimental results compared to the increase in the overall transmission rate.

Data availability statement

The data cannot be made publicly available upon publication because no suitable repository exists for hosting data in this field of study. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

We thank Ji-Qian Qin, Yun-Long Yu, Fan Yang and Zong-Wen Yu for valuable discussions. We acknowledge the financial support in part by National Natural Science Foundation of China Grant Nos. 11974204 and 12174215.

Appendix A. Proof of lemma 1

When there is no photon loss, the input state is a pure state

$$\hat{\rho}_m = \bigotimes_{i=1}^M (|r_i, \phi_i\rangle \langle r_i, \phi_i|).$$  \hspace{1cm} (A1)

For simplicity and without loss of generality, we can set the phase $\phi = \pi$. This is because the change in phases of the input squeezed states is equivalent to the change in phases of the passive linear optical network. The probability of the output pattern can be expressed as

$$P_{out}(\bar{m}) = \langle \bar{m}|\hat{U}\hat{\rho}_m\hat{U}^\dagger|\bar{m}\rangle$$
$$= \sum_{n,l} \langle \bar{m}|\hat{U}|\bar{n}\rangle \langle \bar{n}|\hat{\rho}_m|\bar{l}\rangle \langle \bar{l}|\hat{U}^\dagger|\bar{m}\rangle. $$  \hspace{1cm} (A2)

Denote $\sqrt{P_{in}(\bar{n})} = \langle \bar{n}|r, \phi \rangle = \sqrt{\langle \bar{n}|\hat{\rho}_m|\bar{n}\rangle}$. We get

$$P_{out}(\bar{m}) = \sum_{n,l} v(\bar{m}|\bar{n})v^*(\bar{m}|\bar{l})\sqrt{P_{in}(\bar{n})P_{in}(\bar{l})}, $$  \hspace{1cm} (A3)
where \( v(\bar{m} | \bar{n}) = \langle \bar{m} | \hat{U} | \bar{n} \rangle \). As passive linear optical networks preserve the photon number of quantum optical states, we have

\[
v(\bar{m} | \bar{n}) = 0, \tag{A4}
\]

when \( \sum_{i=0}^{M} m_i \neq \sum_{i=0}^{M} n_i \).

When there is photon loss, according to equation (9)

\[
P_{\text{out}}(\bar{m}) = \langle \bar{m} | \text{Tr}_M \left( \hat{U}_1 \hat{U}_0 (\hat{\rho}_m \otimes |\bar{0}\rangle \langle \bar{0}|) \hat{U}_1^\dagger \hat{U}_0^\dagger \right) | \bar{m} \rangle
\]

\[
= \sum_{n,l} v_1(\bar{m} | n) v_1^*(\bar{m} | l) \sum_{i,j} \langle n,i | \hat{U}_0 | j,0 \rangle
\]

\[
\times \langle \bar{i} | \hat{\rho}_m | \bar{j} \rangle \langle j,0 | \hat{U}_0^\dagger | \bar{l} \rangle, \tag{A5}
\]

where \( \hat{U}_1 \) and \( \hat{U}_0 \) are given in equations (7) and (8), \( v_1(\bar{m} | n) = \langle \bar{m} | \hat{U}_1 | n \rangle \) and \( v_1^*(\bar{m} | l) = \langle l | \hat{U}_1^\dagger | \bar{m} \rangle \).

According to equation (7), we have

\[
\hat{U}_0 | i,0 \rangle = \frac{1}{\sqrt{(i_1!)(i_2!)...(i_M!)}} \prod_{p=1}^{M} \left( \sqrt{\eta} a_{i_p}^\dagger + \sqrt{1-\eta} a_{i_p} \right)^{i_p} | 0,0 \rangle. \tag{A6}
\]

Then we have

\[
\langle n,k | \hat{U}_0 | i,0 \rangle = \frac{1}{\sqrt{(i_1!)(i_2!)...(i_M!)}} \prod_{p=1}^{M} \left( \sqrt{\eta} a_{i_p}^\dagger + \sqrt{1-\eta} a_{i_p} \right)^{i_p} | 0,0 \rangle.
\]

\[
\times \prod_{p=1}^{M} \beta_{i_p} \left( \eta \right)^{n_p} \left( \sqrt{1-\eta} \right)^{k_p} | 0,0 \rangle. \tag{A7}
\]

This equation is not zero only when

\[
k = i - \bar{n}. \tag{A8}
\]

So, we get

\[
\langle n,k | \hat{U}_0 | i,0 \rangle = \prod_{p=1}^{M} \sqrt{\frac{j_p}{n_p}} \left( \eta \right)^{n_p} \left( \sqrt{1-\eta} \right)^{k_p} | 0,0 \rangle. \tag{A9}
\]

Finally, bring equation (A9) back to equation (A5), we get

\[
P_{\text{out}}(\bar{m}) = \sum_{n,l} v_1(\bar{m} | n) v_1^*(\bar{m} | l) \sum_{i,j} \prod_{p=1}^{M} \prod_{q=1}^{M} \sqrt{P_{\text{in}}(i)P_{\text{in}}(j)}
\]

\[
\times \sqrt{\left( \frac{i_1}{n_1} \right) \left( \frac{i_2}{n_2} \right) ... \left( \frac{i_M}{n_M} \right)} \left( \eta \right)^{n_p} \left( \sqrt{1-\eta} \right)^{k_p} P_{\text{in}}(i), \tag{A10}
\]

where

\[
P_{\text{in}}^{(q)}(\bar{n}) = \sum_{i \geq n} \prod_{p=1}^{M} \frac{i_p}{n_p} \eta^{n_p} (1-\eta)^{i_p-n_p} P_{\text{in}}(i), \tag{A11}
\]

and

\[
\left( \frac{i}{n} \right) = \left( \frac{i_1}{n_1} \right) \left( \frac{i_2}{n_2} \right) ... \left( \frac{i_M}{n_M} \right). \tag{A12}
\]
Appendix B. Proof of theorem 1

According to equation (3), we have

\[
\frac{p_{\text{out}}^{(r)}(\bar{n})}{p_{\text{out}}^{(r')}((\bar{n})} = \frac{1}{\prod_{i=1}^{K} \cosh r_i} \sum_{j} v(\bar{n}|\bar{n}) v^*(\bar{l}|\bar{l}) \sqrt{p_{\text{in}}^{(r)}(\bar{n}) p_{\text{in}}^{(r')}((\bar{l})},
\]

where \(p_{\text{out}}^{(r)}(\bar{n})\) and \(p_{\text{out}}^{(r')}((\bar{n})\) are probability distributions of the output pattern with input squeezing strengths \(\{r_i\}\) and \(\{r'_i\}\) respectively, \(\bar{n} = n_1 n_2 \ldots n_K\) and \(\bar{m} = m_1 m_2 \ldots m_K\) are input patterns with the same total photon number \(N, n_i\) and \(m_i, (i = 1, \ldots, K)\) are photon numbers in each input mode. If we choose \(\{r'_i\}\) which satisfies \(\frac{\tanh r_i}{\tanh r'_i}\) = \(i\) for all \(i = 1, 2, \ldots, K\), we have

\[
\frac{p_{\text{in}}^{(r)}(\bar{n})}{p_{\text{in}}^{(r')}((\bar{n})} = \frac{1}{\prod_{i=1}^{K} \cosh r_i} \sum_{j} v(\bar{n}|\bar{n}) v^*(\bar{m}|\bar{m}) \sqrt{p_{\text{in}}^{(r)}(\bar{n}) p_{\text{in}}^{(r')}((\bar{m})},
\]

According to lemma 1, we find

\[
\frac{p_{\text{out}}^{(r)}(\bar{q})}{p_{\text{out}}^{(r')}((\bar{q})} = \frac{1}{\prod_{i=1}^{K} \cosh r_i} \sum_{j} v(\bar{q}|\bar{q}) v^*(\bar{l}|\bar{l}) \sqrt{p_{\text{in}}^{(r)}(\bar{q}) p_{\text{in}}^{(r')}((\bar{l})},
\]

where \(p_{\text{out}}^{(r)}(\bar{q})\) and \(p_{\text{out}}^{(r')}((\bar{q})\) are probability distributions of the output pattern with input squeezing strengths \(\{r_i\}\), \(\bar{q} = q_1 q_2 \ldots q_K\) and \(\bar{j} = j_1 j_2 \ldots j_K\) are output patterns with the same total photon number \(N, q_i\) and \(j_i\) \((i = 1, \ldots, K)\) are photon numbers in each output mode. Plugging equation (B2) into equation (B3) and we get

\[
\frac{p_{\text{out}}^{(r)}(\bar{q})}{p_{\text{out}}^{(r')}((\bar{q})} = \frac{1}{\prod_{i=1}^{K} \cosh r_i} \sum_{j} v(\bar{q}|\bar{q}) v^*(\bar{j}|\bar{j}) \sqrt{p_{\text{in}}^{(r)}(\bar{q}) p_{\text{in}}^{(r')}((\bar{j})},
\]

Taking summation over \(\bar{q}\), we obtain

\[
\frac{p_{\text{out}}^{(r)}(\bar{n})}{p_{\text{out}}^{(r')}((\bar{n})} = \frac{1}{\prod_{i=1}^{K} \cosh r_i} \sum_{j} v(\bar{n}|\bar{n}) v^*(\bar{m}|\bar{m}) \sqrt{p_{\text{in}}^{(r)}(\bar{n}) p_{\text{in}}^{(r')}((\bar{m})},
\]

which completes the proof.

Appendix C. Proof of theorem 2

According to lemma 1, we have

\[
P_{\text{out}}^{(r, \eta)}(\bar{n}) = \sum_{\bar{m}, \bar{f}} v_1(\bar{n}|\bar{m}) v^*_f (\bar{m}|\bar{f}) \sqrt{p_{\text{in}}^{(r, \eta)}(\bar{m}) p_{\text{in}}^{(r, \eta)}((\bar{f})},
\]

where

\[
p_{\text{in}}^{(r, \eta)}(\bar{m}) = \sum_{\bar{f} \geq \bar{m}} \prod_{i=1}^{M} \left( \frac{i}{\bar{m}} \right)^{n_i} (1 - \eta)^{s_i - n_i} p_{\text{in}}^{(r)}(\bar{f}),
\]

\[
\eta = \frac{\tanh r_i}{\tanh r'_i},
\]

\[
\frac{\tanh r_i}{\tanh r'_i} = \frac{\cosh r_i}{\cosh r'_i} = i.
\]
and \( p^{(c)}_{\text{in}}(i) \) is defined as in theorem 1. We have

\[
p^{(r)}_{\text{in}}(m) p^{(r')}_{\text{out}}(|\tilde{m}|) \left( \frac{\eta'}{\eta} \right)_{[m]}
= \sum_{\tilde{m}} \prod_{i \geq m} \left( \frac{i}{\tilde{m}} \right) \eta^{m_p} (1 - \eta)_{i - m_p} p^{(c)}_{\text{in}}(i) \left( \frac{\eta'}{\eta} \right)_{[i]}
\]

\[
\times \prod_{q=1}^{K} \frac{\cosh r_q'}{\cosh r_q} \left( \frac{\tanh r_q'}{\tanh r_q} \right)_{[i]}
= \sum_{\tilde{m}} \prod_{i \geq m} \left( \frac{i}{\tilde{m}} \right) \eta^{m_p} (a(1 - \eta))_{i - m_p}
\]

\[
\times p^{(c)}_{\text{in}}(i) \left( \frac{\tanh r_p}{\tanh r_p} \right)_{[i]} \prod_{q=1}^{K} \frac{\cosh r_q'}{\cosh r_q}
= \sum_{\tilde{m}} \prod_{i \geq m} \left( \frac{i}{\tilde{m}} \right) \eta^{m_p} (1 - \eta')_{i - m_p} p^{(c)}_{\text{in}}(i)
= P^{(r')}_{\text{in}}(m),
\]

where we used the relation \( \eta' = 1 - (1 - \eta) \epsilon \).

Combining equations (C3) and (C1), we can show the correctness of theorem 2.

**Appendix D. Non-uniform loss**

Consider an \( M \)-mode interferometer (passive linear optical network), if the interferometer is lossy, then the input-output relation is [38–40]

\[
\hat{b} = A\hat{a} + \sqrt{1 - AA^\dagger} \epsilon,
\]

where \( \epsilon \) represents the loss modes which will be traced out, and \( A \) is a complex matrix satisfying \( AA^\dagger \leq I \). The matrix \( A \) can be decomposed (singular decomposition) as \( A = V_1 DV_2 \), where \( V_1 \) and \( V_2 \) are unitary matrices, and \( D = \text{diag} \{ \sqrt{\eta_1}, \sqrt{\eta_2}, \ldots, \sqrt{\eta_M} \} \) with \( \eta_i \in [0, 1] \).

Denote \( \eta_i = \max (\eta_i) \) for \( i = 1, 2, \ldots, M \). A uniform loss layer with transmission rate \( \eta_i \) can be extracted from the lossy passive linear optical network. This can be seen from:

\[
A = \sqrt{\eta_i} I \left( \frac{1}{\sqrt{\eta_i}} \right) V_1 DV_2.
\]

Thus the whole physical process is equivalent to that the input SMSSs first go through a passive linear optical network with a nonuniform loss described by \( A_1 = \left( \frac{1}{\sqrt{\eta_i}} \right) V_1 DV_2 \) and then a uniform loss channel described by \( A_2 = \sqrt{\eta_i} I \).

The probability that an output pattern \( m \) is obtained is thus

\[
P_{\text{out}}(m) = \langle \tilde{m} | \text{Tr}_{2M} \{ \hat{U}_1 \hat{U}_2 \hat{U}_3 \hat{U}_4 \hat{U}_5 \hat{U}_6 \hat{U}_7 \hat{U}_8 | \langle \tilde{0} \rangle \langle 0 | \}
\times \hat{U}_1 \hat{U}_2 \hat{U}_3 \hat{U}_4 \hat{U}_5 \hat{U}_6 \hat{U}_7 \hat{U}_8 | m \rangle.
\]

Since there are both uniform loss and non-uniform loss processes, we need 2\( M \) loss modes. After some calculations, we have

\[
P_{\text{out}}(m) = \sum_{\tilde{n}, \tilde{l}, \tilde{k}, \tilde{j}, \tilde{a}, \tilde{b}, \tilde{v}, \tilde{l}, \tilde{d}} v_1(m | \tilde{n}) v_1^*(\tilde{m} | \tilde{l}) v_2(\tilde{k} | \tilde{d})
\times v_2^*(\tilde{z} | \tilde{g}) \sqrt{p^{(r)}_{\text{in}}(\tilde{d}) p^{(r)}_{\text{in}}(\tilde{g})} \langle \tilde{n}, \tilde{b} | \hat{U}_n | \tilde{i}, \tilde{0} \rangle
\times \langle \tilde{j}, \tilde{0} | \hat{U}_j | \tilde{l}, \tilde{p} \rangle \langle \tilde{i}, \tilde{c} | \hat{U}_i | \tilde{k}, \tilde{d} \rangle \langle \tilde{z}, \tilde{0} | \hat{U}_z | \tilde{v}, \tilde{v} \rangle,
\]

where \( v_2(\tilde{k} | \tilde{d}) = \langle \tilde{k} | \hat{U}_d | \tilde{d} \rangle \) and \( v_2^*(\tilde{z} | \tilde{g}) = \langle \tilde{g} | \hat{U}_z | \tilde{g} \rangle \). Since the passive linear optical network keeps the total photon number in the input and output states unchanged, we have \( |\tilde{m}| = |\tilde{n}| = |\tilde{l}|, |\tilde{k}| = |\tilde{d}|, |\tilde{z}| = |\tilde{g}|, |\tilde{k}| = |\tilde{z}| + |\tilde{i}|, |\tilde{z}| = |\tilde{j}| + |\tilde{v}|, |\tilde{j}| = |\tilde{p}| + |\tilde{l}|, |\tilde{l}| = |\tilde{n}| + |\tilde{d}| \).
Next, take the transmission rate $\eta_p$ of the uniform loss layer as the overall transmission rate to do the post-selection in protocol 1. The probability distribution of the samples obtained after the post-selection is (without considering cut-off)

$$P_{\text{out}}(\bar{m})P_{\text{pos}}(\bar{m}) = \mathcal{N} \sum_{n,l} n_{\bar{m}(\bar{n})} v_1(\bar{m} | \bar{n}) v_1^*(\bar{m} | \bar{l}) v_2(\bar{k} | \bar{d})$$

$$\times v_2^*(\bar{z} | \bar{g}) \sum_{i \in R} \sum_{j \neq i} \sum_{\bar{k} \geq \bar{i} \geq \bar{z}} \sqrt{P_{\text{in}}^{(i)}(\bar{k}) P_{\text{in}}^{(j)}(\bar{z})}$$

$$\times \sqrt{\left(\frac{k}{i}\right)} (\eta_\theta)^i (c(1 - \eta_\theta))^{k-i}$$

$$\times \sqrt{\left(\frac{z}{i}\right)} (\eta_\theta)^j (c(1 - \eta_\theta))^{z-j}$$

$$\times \sqrt{\left(\frac{i}{n}\right)} (\eta_u)^n (1 - \eta_u)^{i-n}$$

$$\times \sqrt{\left(\frac{i}{i}\right)} (\eta_\theta)^i (1 - \eta_\theta)^{i-n},$$

where $\mathcal{N}$ is the normalization factor, $\eta_\theta = (\eta_1/\eta_1, \eta_2/\eta_2, \ldots, \eta_M/\eta_M)$, $(\eta_\theta)^i = \prod_{p=1}^M (\eta_p)^i$, $(\eta_\theta)^j = \prod_{p=1}^M (\eta_p)^j$. From the equation above, it can be seen that post-selection reduces the uniform loss while making the non-uniform loss have less effect on the sampling results. Because, before the post-selection process, the non-uniform loss layer reduces the state $|\bar{k}\rangle$ to $|\bar{i}\rangle$ with the probability

$$P_n(\bar{k}, \bar{i}) = \left(\frac{k}{i}\right) (\eta_\theta)^i (1 - \eta_\theta)^{k-i}.$$  

(D6)

After the post-selection, the state $|\bar{k}\rangle$ is reduced to $|\bar{i}\rangle$ with the probability

$$P'_n(\bar{k}, \bar{i}) = \left(\frac{k}{i}\right) (\eta_\theta)^i (1 - \eta_\theta)^{k-i}.$$  

(D7)

Since $\sqrt{c(1 - (\eta_\theta)_i)} \leq \sqrt{c(1 - (\eta_\theta)_i)}$ for $i \in \{1, 2, 3, \ldots, M\}$, we have $P'_n(\bar{k}, \bar{i}) < P_n(\bar{k}, \bar{i})$. This corresponds to a reduction in the proportion of the output samples that is affected by non-uniform loss. Therefore, post-selection reduces the effect of non-uniform loss on output samples.

### Appendix E. Detector efficiency

Different detection efficiencies between detectors can also lead to a non-uniform loss process for the GBS device except for the detection efficiency is denoted as $\eta_d$ and the detection efficiency of each photon detector is denoted as $\eta^d_i$, where the subscript $i$ refers to the $i$th photon detector. The whole loss process can be described by a unitary operator $\hat{U}_d$ which includes $M$ actual modes and $M$ loss modes. The input-output relationship corresponding to the whole loss process is

$$\hat{U}_d \hat{a}^\dagger \hat{U}_d^\dagger = \Lambda_d \hat{a}^\dagger,$$

where

$$\Lambda_d = \left( \oplus_{i=1}^M \sqrt{\frac{\eta_i}{\eta_i - \eta}} \oplus_{i=1}^M \sqrt{1 - \eta_i} \right),$$

(E2)

and $\eta_i = \eta^d_i$. Similar to the analysis in appendices D and C, the probability distribution of the output sample is

$$P_{\text{out}}(\bar{m}) = \sum_{n,l} v(\bar{m} | \bar{n}) v^*(\bar{m} | \bar{l}) \sum_{i \in R} \sum_{j \neq i} \prod_{p=1}^M \prod_{q=1}^M P_{\text{in}}(i) P_{\text{in}}(j)$$

$$\times \sqrt{\left(\frac{l_p}{n_p}\right)} (\eta_p)^{n_p} (1 - \eta_p)^{l_p - n_p} \sqrt{\left(\frac{j_q}{l_q}\right)} (\eta_q)^{l_q} (1 - \eta_q)^{j_q - l_q},$$

(E3)
The above equation shows that, in this case, we can change the post-selection probability in protocol 1 to

\[
P_{\text{post2}}(|\eta|) = e^{-|\eta|} \prod_{i=1}^{M} \left( \frac{\eta_i}{\eta_i'} \right)^{n_i},
\]

where \( \eta_i' = 1 - (1 - \eta_i) \tanh \frac{r}{\tanh r}. \)

**References**

[1] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* 10th edn (Cambridge University Press)

[2] Preskill J 2018 *Quantum* 2 79

[3] Shor P 1994 *Proc. 35th Annual Symp. on Foundations of Computer Science* pp 124–34

[4] Arute F et al 2019 *Nature* 574 505

[5] Wu Y et al 2021 *Phys. Rev. Lett.* 127 180501

[6] Huang H-L, Wu D, Fan D and Zhu X 2020 *Sci. China Inf. Sci.* 63 180501

[7] Gong M et al 2021 *Science* 372 948

[8] Bruzewicz C D, Chiaverini J, McConnell R and Sage J M 2019 *Appl. Phys. Rev.* 6 021314

[9] Zhong H-S et al 2020 *Science* 370 1460

[10] Arrazola J M et al 2021 *Nature* 591 54

[11] Zhong H-S et al 2021 *Phys. Rev. Lett.* 127 180502

[12] Slussarenko S and Pryde G J 2019 *Appl. Phys. Rev.* 6 041303

[13] Bourassa J E et al 2021 *Quantum* 5 392

[14] McClean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 *New J. Phys.* 18 023023

[15] Farhi E, Goldstone J and Gutmann S 2014 arXiv:1411.4028 [quant-ph]

[16] Bouland A, Fefferman B, Nirkhe C and Vazirani U 2019 *Proc. 35th Annual Symp. on Foundations of Computer Science* pp 124–34

[17] Rahimi-Keshari S, Lund A P and Ralph T C 2015 *Phys. Rev. Lett.* 114 060501

[18] Hamilton C S, Kruse R, Sansonni L, Barkhofs s, Silberhorn C and Jas I 2017 *Phys. Rev. Lett.* 119 170501

[19] Arute F, Brod D J, Quesada N and Killoran N 2020 *Phys. Rev. A* 98 062332

[20] Mclean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 *New J. Phys.* 18 023023

[21] Farhi E, Goldstone J and Gutmann S 2014 arXiv:1411.4028 [quant-ph]

[22] Bouland A, Fefferman B, Nirkhe C and Vazirani U 2019 *Proc. 35th Annual Symp. on Foundations of Computer Science* pp 124–34

[23] Rahimi-Keshari S, Lund A P and Ralph T C 2015 *Phys. Rev. Lett.* 114 060501

[24] Hamilton C S, Kruse R, Sansonni L, Barkhofs s, Silberhorn C and Jas I 2017 *Phys. Rev. Lett.* 119 170501

[25] Arute F, Brod D J, Quesada N and Killoran N 2020 *Phys. Rev. A* 98 062332